A closer look on dementia, elder abuse and multimorbidity in representative samples of the oldest old

Inauguraldissertation

zur

Erlangung des Doktorgrades

der Humanwissenschaftlichen Fakultät

der Universität zu Köln

nach der Promotionordnung vom 18.12.2018

vorgelegt von

Thomas Brijoux

März 2023

Erstgutachterin: Prof. Dr. Susanne Zank

Zweitgutachter: Prof. Dr. Holger Pfaff

Diese Dissertation wurde von der Humanwissenschaftlichen Fakultät der Universität zu Köln im August 2023 angenommen.

Acknowledgement

Dedicated to the many colleagues and supervisors as well as friends and family members who have accompanied me in life, science and gerontology over the last 2 decades. They all shaped me and supported me in developing myself in a way that made this work possible.

Contents

List of Tables and Figures

List of Abbreviations

ADL *Activities of daily living*
BASE *Berlin Ageing Study*
CHIS California Health Interview Survey
EACS Elder Abuse and Emotional Consequences scale
GEDA Gesundheit in Deutschland aktuell
IADL *Instrumental Activities of daily living*
ICD International Statistical Classification of Diseases and Related Health Problems
MSE mean squared error
RKI *Robert Koch Institut*

Summary

Background: Due to the growing number of old aged persons in Germany and other countries the question arises how old aged persons can be described representatively in surveys. In this work old aged people are defined as persons above and including the age of 80 years. Old age describes a population that, due to limited health, a higher prevalence rates of dementia and a substantial part of the population living in nursing homes, is not only hard to survey. Additionally, constructs that are difficult to measure such as elder abuse become of greater importance when other resources diminish which is more common in old age. A theoretical framework that was formed to improve survey quality and enables discussions about possible survey errors is the Total Survey Error framework.

Objective: Using the Total Survey Error framework it is examined how the constructs of elder abuse, dementia and multimorbidity can be assessed in representative samples of the old age.

Results: Regarding representativity the presented papers show, that the inclusion of the nursing home population, people with dementia and proxy informants enhances the representativity of the survey and reduce non-response bias. On the other hand the inclusion of these hard to survey groups can lead to higher measurements errors. Measurement invariance is not established between proxy and self-reports in reports of elder abuse leading to an exclusion of proxy interviews in measurements of elder abuse. By reweighting the resulting non-response bias is reduced.

Discussion: Inclusion of hard to survey groups can increase measurement errors while their exclusion increases sampling, frame and non-response errors. Presented methods that moderate these kinds of errors are reweighting the survey and separate analyses and presentation of subpopulations of the survey. The latter can allow to optimize measurements in subpopulations but may not yield a common estimator for the whole population.

1 Introduction

This dissertation centres around problems of representative research among the oldest old who are in this case defined as people of and above the age of 80 years. While the size of this age group is constantly increasing, more and more people are hard to survey due to various health reasons. Therefore, the difficulties of gerontological research in surveying the oldest old become of a greater societal significance. This thesis puts a special focus on dementia, multimorbidity, and elder abuse as examples for constructs that are difficult to survey.

The thesis comprises five articles. The first paper directly focuses on hard-to-survey groups and weighs advantages and disadvantages of the inclusion of these groups. Two papers focus on experiences of elder abuse and present the Elder Abuse and Emotional Consequences Scale (EACS), report 12-month prevalence rates of different forms of elder abuse, and explore consequences and risk factors of elder abuse. For this thesis, the reported influence of different survey settings on reported elder abuse, like third persons present during the interview, and ways to cope with these limitations are of greater importance. The fourth paper presents different aspects of quality of life and examines relationships between resources and life results in the oldest old. The last paper concentrates on frequent diseases as well as multimorbidity and its impact on functioning, autonomy, and life satisfaction in old age. The theoretical framework of the Total Survey Error is used for a structured reflection around the underlying questions of how to survey the oldest old and which advantages and disadvantages should be weighed against each other.

All articles in this thesis used data of the first sample of the NRW80+ study which will briefly be introduced hereinafter. The NRW80+ study is the first representative study on the quality of life of the oldest old in North Rhine-Westphalia. The gross sample of the NRW80+ study was collected through residents' registration offices. Data was collected through face-to-face interviews conducted in North Rhine-Westphalia between August 2017 and February 2018. The survey instrument centres around psychological, sociological, and ethical constructs while also gathering some medical data. Wherever target persons could not participate in the interview due to health reasons, the interview could be conducted with a proxy informant who answered on behalf of the target person. The realized sample consists of 1863 persons, with 176 interviews being performed with proxy informants. Table 1 displays the realized interviews in different age and sex groups (Brix et al., 2018). Furthermore, the CHAPO framework (Neise et al., 2019; Wagner et al., 2018) was developed as a theoretical framework for ordering and classifying the examined constructs.

Women			Men		
80-84 years	85-89 years	90+ years	80-84 years	85-89 years	90+ years
344	326	266	384	299	244

Table 1 Unweighted realized interviews of the first wave of the NRW 80+ sample

Further details of the sampling, data collection, and achieved sample are published elsewhere (Hansen et al., 2021). An overview of the final weighted sample is presented in Table 2 (ibid.).

To start with, the thesis provides an introduction of terms and concepts that will be used hereafter. At first, the term "old age" and other synonyms for this group are presented and related concepts like the CHAPO framework are introduced. This is followed by a brief introduction of the concepts of dementia, elder abuse, and multimorbidity, complemented by an introduction of the term "representativity" and the Total Survey Error framework. The subsequent chapter uses the Total

Survey Error framework as a background sheet and evaluates difficulties in the examination of people with dementia, elder abuse, and multimorbidity in large scale surveys. Finally, the five papers which are relevant for this thesis are presented; they will also be discussed in the last chapter.

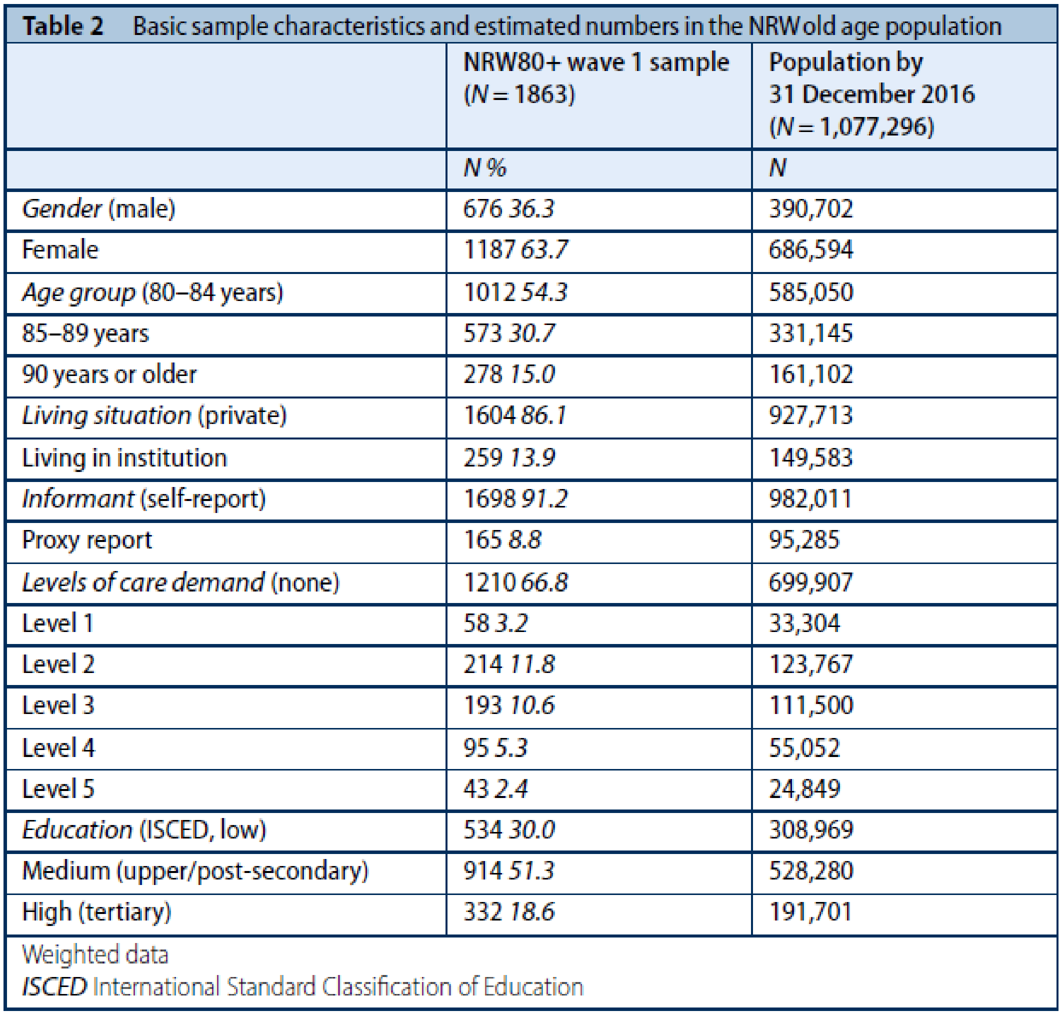


Table 2 Basic sample characteristics and estimated numbers in the NRW old age population

	NRW80+ wave 1 sample (*N* = 1863)	**Population by 31 December 2016 (*N* = 1,077,296)**
	N %	*N*
Gender (male)	676 *36.3*	390,702
Female	1187 *63.7*	686,594
Age group (80–84 years)	1012 *54.3*	585,050
85–89 years	573 *30.7*	331,145
90 years or older	278 *15.0*	161,102
Living situation (private)	1604 *86.1*	927,713
Living in institution	259 *13.9*	149,583
Informant (self-report)	1698 *91.2*	982,011
Proxy report	165 *8.8*	95,285
Levels of care demand (none)	1210 *66.8*	699,907
Level 1	58 *3.2*	33,304
Level 2	214 *11.8*	123,767
Level 3	193 *10.6*	111,500
Level 4	95 *5.3*	55,052
Level 5	43 *2.4*	24,849
Education (ISCED, low)	534 *30.0*	308,969
Medium (upper/post-secondary)	914 *51.3*	528,280
High (tertiary)	332 *18.6*	191,701

Weighted data
ISCED International Standard Classification of Education

Table 2 Description of the weighted NRW 80+ sample (Hansen et al., 2021)

2 Terms and Concepts

In order to talk about representativity in old age, some basic terms and concepts need to introduced first. This chapter introduces the basic concepts and terms for this thesis, provides definitions, and presents scientific literature on the empirical concepts presented. It begins with an introduction of the term “old age”. The constructs of dementia, elder abuse, and multimorbidity which are particularly relevant for this work are subsequently presented. This is followed by a brief introduction into the concept of representativity and a showcase of R-Indicators as a concept for the assessment of representativity. Furthermore, the Total Survey Error framework is introduced as a theoretical framework to describe survey quality.

2.1 Old Age

Globally, the demographic group of elderly people is growing fast (United Nations, 2021). Thus, it is necessary to look at the scientifical meaning of the term “old age” as well as at important concepts that are used to describe old age. Afterwards, the chapter continues with examples and introduces three concepts that are of increasing importance in the field of old age: dementia, elder abuse, and multimorbidity.

2.1.1 Definitions of Old Age

It is well understood that heterogeneity increases in most areas of human behaviour, health, or resources along the life span (Ferrucci & Kuchel, 2021). Generally, with the exception of laboratory markers, the observation that older age cohorts exhibit greater heterogeneity remains valid (Nguyen et al., 2021). Similarly, to this heterogeneity in the population, there is also a heterogeneity of terms and concepts designating the oldest age group. Examples of the labelling of the oldest age group comprise “old age”, “oldest old”, “fourth age”, “disabled elderly”, “advanced old age”, “the very old”, or “very old age group” (Kydd et al., 2020). Kydd et al. (2020) analyse 48 papers on the oldest age groups for the use of terms describing this group. They find the most commonly used terms to be “oldest old” or “oldest old adults” (31 mentions) as well as phrases using the word “very” as in “very old”, “very elderly”, or “very old adults” (12 mentions). Regardless of the term used, the problem remains: Each of the terms can be used with the purpose of ageism. Therefore, Kydd et al. (2020) propose to instead utilise terms that describe age bands of ten years, such as “septuagenarians” or “octogenarians”, as these have a universal meaning and also avoid ageist connotations. However, the groups described in this thesis contain octogenarians, nonagenarian, and centenarians. Naming all three groups individually is linguistically complicated; therefore, the term “oldest old” is preferred in this thesis, as it is generally the most used term.

Degnen (2007) points out that the underlying question of all these different conceptualizations is whether chronological age or some form of functioning and ability should be the most reliable frame of reference. Defining old age via some level of (often impaired) functioning has the advantage of reducing heterogeneity in the referenced group. This, in turn, allows for specifically tailored instruments (Lord, 1968) as well as more precise descriptions, because the reduced heterogeneity leads to reduced standard errors (Bortz & Schuster, 2010; Leonhart, 2009). Reduced standard errors also correspondent to a higher statistical power (Bortz & Schuster, 2010; Kieser, 2018), which is especially important in old age, where study effects are often smaller (Brydges, 2019) and standard deviations are often underestimated (Chen et al., 2013).

Besides these methodological arguments, some gerontologists point out that a definition focusing on biological-medical, psychologic, or social functioning facilitates research questions around the topic of achieving a good or successful age.

> (…) the research question needs to be broadened from a primary focus on outcomes; that is, from: ' *What* is successful ageing?' to include *'How* do people age successfully?' or 'What are the processes that allow for mastery of goals in old age?' We suggest that understanding the *processes* that people use to reach their goals under increasing limitations in resources, be they social, psychological or biological, will lead to additional insights and progress in the field. (Baltes & Carstensen, 1996, p. 399)

However, definitions of old age as the fourth or often (un-)successful age imply categories in which ageing processes have been (un-)successful. In 1997, Rowe and Kahn already proposed a concept of successful ageing known as the McArthur model that contains three main components: 1) low probability of disease and disease-related disability, 2) high cognitive and physical functional capacity, and 3) active engagement with life. While this is the most commonly used definition of dimensions of successful ageing (Plugge, 2021), the choice of dimensions is not self-evident – more than 100 variations of the original model have been proposed (Rowe & Kahn, 2015). For example, Foo et al. (2019) recently proposed epigenetic age, brain age, cardiovascular age, renal age, endocrine age, and musculoskeletal age as dimensions of medical age, and cognitive age and emotional age as dimensions of psychologic age. However, there is no generally accepted list of dimensions. Regardless of the dimensions proposed, the main issue remains: Either these dimensions have to be considered equally important or they have to be ranked (Baltes & Carstensen, 1996).

> A solution to this problem seems to be a flexible definition of success outcomes. Success can be defined by different authorities *(e.g.* individual, peer group, society, scientific theory), by different criteria of assessment *(e.g.* subjective vs. objective) and by different norms *(e.g.* functional, statistical, ideal norm). (Baltes & Carstensen, 1996, p. 404)

In the context of representative research on the very old, it should also be pointed out that Rowe & Kahn have conceived three categories of the old-aged population, namely successful agers (10%), normal agers (80%), and pathological agers (10%). Some empirical studies such as LebenDig (Lang et al., 2019) focus primarily on the fit and successful agers and consequently do not allow for representative statements. Their results can, hence, not easily be applied to the entire population.

An important extension of the McArthur model, especially in the context of research on the very old, was published by Tesch-Römer and Wahl (2017), who "see it as a major limitation that Rowe and Kahn's model is not able to cover the emerging linkage between increasing life expectation and aging with disability and care needs" (ibid., 310). They have expanded the model by including people who grow old with disabilities and care needs. One major finding in empirical research on the oldest old is that people can maintain a high life satisfaction in spite of objectively bad health or living conditions (Puvill et al., 2016), or, as Baltes and Carstensen (1996) put it, "what is considered successful according to functional norms might not meet with ideal norms, nor square with statistical norms." (Baltes & Carstensen, 1996, p. 399).

The alternative approach to defining old age based on functioning is the definition based on chronological age. At first glance, a definition of old age based on chronological age offers a more distinct differentiation between old and non-old age. However, here too, there are many different classifications. Kydd et al. (2020) identify the age groups of 75 and above, 80 and above, 85 and above, 90 and above, 90 to 100, and 100 and above to describe the old-aged population. Among those, the definition of 85 and above is most used, followed by the definition of 80 and above. But although calendrical age is an objective criterion, its meanings can differ for subpopulations. For example, age-specific prevalence rates for diseases like dementia (Alzheimer Europe, 2019), osteoporosis (Burge et al., 2007), or multimorbidity in general (Marengoni et al., 2011) vary between the two sexes. Some kind of cultural sensibility is also evident: An 80-year-old Sudanese woman ages in a very different welfare state and a very different healthcare system than an 80-year-old German woman. However, the data used in this thesis has been collected in North Rhine-Westphalia, Germany, so at least the welfare systems are to some extend comparable among study participants.

In this text, the terms “oldest old” and “old age” are used most frequently based on a demographic definition of old age that includes all persons aged 80 and above. Against the background of this work, which focuses on representative research on the very old, this approach seems to be the most appropriate since functional definitions of old age do not include the whole population.

Old age has already been identified as a life stage that is characterized by high heterogeneity in relevant constructs. Although a dedifferentiation of the covariance of some constructs has been observed (i.e. cognitive and sensory constructs (Lindenberger, 2000), but also brain morphology (Aboud et al., 2019)), ageing cannot be described with just a few constructs, especially in old age; it is rather seen as a multidimensional (Balachandran & James, 2019; Rivadeneira et al., 2021), process that can only be described by numerous scientific disciplines (Hank et al., 2019). The challenges and potentials (CHAPO) model of quality of life in very old age offers a possible ordering of age-related constructs centring around quality of life (Neise et al., 2019; Wagner et al., 2018). It expands Veenhoven’s (2013) model of the four qualities of life to seven different areas. On the horizontal line, an environment-person continuum (from left to right) is intended to distinguish between the liveable environment and the life-ability of a person, while on the vertical axis, life chances (distinguished between “reference standards” above and “actual conditions” below) and life results are presented.

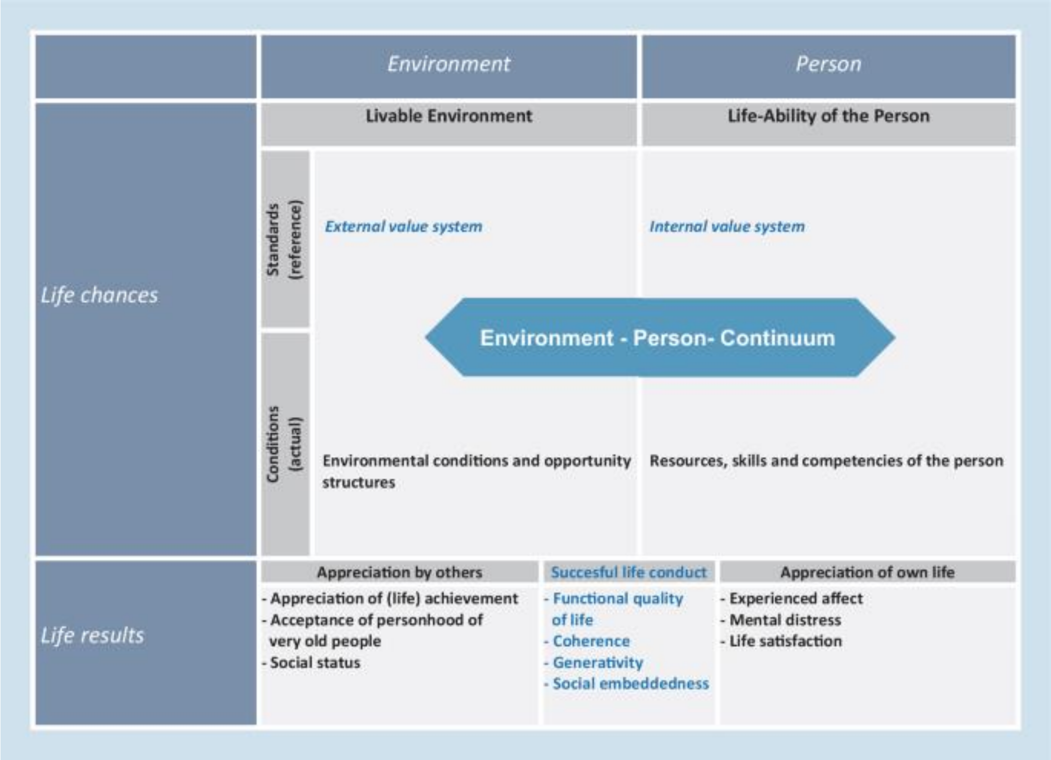


Figure 1 The CHAPO model (Hansen et al., 2021)

Life results can be distinguished between appreciation by others, successful life conduct, and appreciation of one's own life. This classification results in seven areas: 1) reference standards for the liveable environment, such as dominating societal values or images of old age, 2) reference standards for the life-ability of the person, such as the subjective image of age or awareness of age-related change, 3) environmental conditions or opportunity structures, such as the welfare state or living conditions, 4) personal resources, skills and competencies, i.e. physical, psychological or functional health and health competence, 5) appreciation by others, such as societal status of aged persons, 6) successful life conduct that leads to a good embeddedness in society, and 7) appreciation of one's own life, i.e. life satisfaction. However, the presented examples are not exhaustive. For instance, nearly any health-related outcome or disease can be seen as a limiting factor to personal resources and therefore life-chances; at the same time, it can also be considered the result of a life conduct that avoids or accepts risk factors and therefore as a life result. Furthermore, the CHAPO model does not directly describe mechanisms of ageing, but merely presents a possible structure and framework for different research areas and findings (Neise et al., 2019).

2.1.2 Constructs in Old Age

Some constructs that are relevant for the description of old age have already been listed. The topics of dementia, elder abuse, and multimorbidity, which are of greater importance for this work, are described in more detail in the following chapter. While these three topics do not allow for a description of old age in its entirety, they are of increased importance for the concept of representativity and the scope of this thesis.

2.1.2.1 Dementia

Dementia describes a group of diseases that result in reduced cognition and among which Alzheimer's disease is the most common. The Alzheimer Europe report on estimates for dementia prevalence states that 4.1 million people in the 80+ age group suffer from dementia in Europe in 2018 and estimates 4.6 million cases in 2025 (own calculation based on Alzheimer Europe (2019)).

According to the tenth version of the International Statistical Classification of Diseases and Related Health Problems (ICD-10), dementia is defined as:

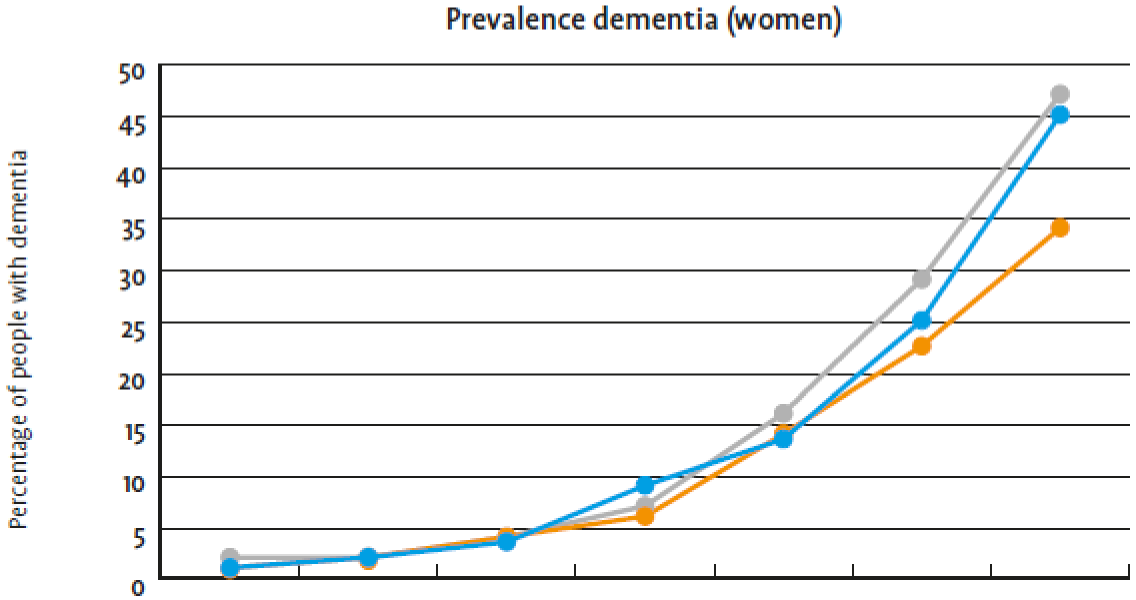
> (…) a syndrome due to disease of the brain, usually of a chronic or progressive nature, in which there is disturbance of multiple higher cortical functions, including memory, thinking, orientation, comprehension, calculation, learning capacity, language, and judgement. Consciousness is not clouded. The impairments of cognitive function are commonly accompanied, and occasionally preceded, by deterioration in emotional control, social behaviour, or motivation. This syndrome occurs in Alzheimer disease, in cerebrovascular disease, and in other conditions primarily or secondarily affecting the brain (World Health Organization, 2019).

For a diagnosis of dementia according to ICD-10, disturbances of memory, thinking, and emotional control must be present over a period of six months. The disorder must not be caused by other diseases (Gutzmann & Zank, 2005).

In the ICD-11, dementia was at first not described as a syndrome but rather as multiple independent diseases of the nervous system (Sathyanarayana Rao et al., 2017). This triggered an intensive international debate as there was a fear of negative consequences for diagnosis and treatment in low- and middle-income countries (Gaebel et al., 2018; Sathyanarayana Rao et al., 2017). Finally, the debate ended in maintaining the concept of a syndrome in the ICD-11 chapter on mental, behavioural or neurodevelopmental disorders and locating the aetiology of the underlying diseases in the ICD-11 chapter on “diseases of the nervous system” (Gaebel et al., 2018; World Health Organization, 2022).

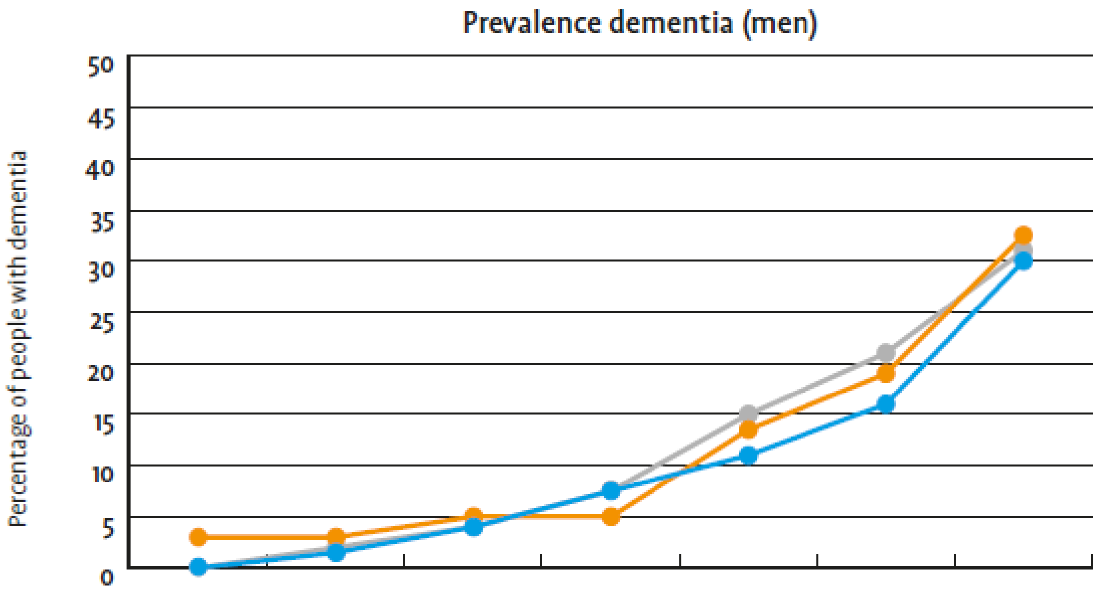
According to ICD-10, dementia describes a complex of symptoms which originate from various diseases. As the discussion on the positioning of dementia in the ICD-11 has shown, there is currently no complete and undisputed classification of these diseases. However, one way to classify different types of dementia is that on a first level primary and secondary dementia are distinguished. Primary dementia diseases are based on brain-organic damage that can be attributed to a disease of the brain itself and are usually not reversible (Doblhammer-Reiter, 2012; Engel, 2006). The most common primary dementias are Alzheimer's dementia, vascular dementia, dementia associated with Parkinson's disease and Lewy body disease, and frontotemporal dementia. In secondary forms of dementia, cognitive symptoms occur as a result of other diseases or injuries that do not primarily affect the brain (Doblhammer-Reiter, 2012). They can be reversible if the underlying disease is cured (for example in the case of malnutrition). The identification of dementia forms and subsequent estimates of their frequency are limited, in part because of the frequent occurrence of mixed forms. However, one possible distribution estimates Alzheimer’s dementia accounts for 60% of primary dementia, with vascular and Lewy body dementia making up 15% and frontotemporal and other types of dementia 5% of primary dementia (Mahlberg & Gutzmann, 2009).

The most important risk factor for dementia is age (Abbott, 2011). In the age group of 80 to 84, the prevalence of dementia is approximately 12%, and increases to approximately 41% in the age group of 90+. The gender-dependent presentation of prevalence (see Figure 2 and Figure 3) also shows that women suffer more frequently from dementia than men. As a result of demographic change, the number of people with dementia in Europe is expected to increase to approximately 4.6 million in 2025 and to approximately 8.1 million in 2050 (Alzheimer Europe, 2019) for those aged 80 and above. Anstey et al. (2019) explore meta-analyses on risk factors for dementia and state highest evidence for cardio-vascular risk factors. Other risk factors are classified into demographic (i.e. education), lifestyle (i.e. diet, smoking), health (i.e. diabetes, depression), medication (i.e. benzodiazepines, statins), and environmental factors (i.e. pesticides).



	60–64	65–69	70–74	75–79	80–84	85–89	90+
AE 2019	0.90	1.50	3.40	8.93	13.05	24.85	44.81
EUROCODE	0.90	1.40	3.80	7.60	16.40	28.50	46.60
EURODEM	0.47	1.10	3.86	6.67	13.50	22.76	34.13

Figure 2 Prevalence of dementia in women in Europe in different age groups (Alzheimer Europe, 2019)



	60–64	65–69	70–74	75–79	80–84	85–89	90+
AE 2019	0.20	1.11	3.10	6.98	10.67	16.29	29.70
EUROCODE	0.20	1.80	3.20	7.00	14.50	20.90	30.80
EURODEM	1.58	2.17	4.61	5.04	12.12	18.45	31.84

Figure 3 Prevalence of dementia in men in Europe in different age groups (Alzheimer Europe, 2019)

The cognitive impairments described as well as the changes in behaviour, emotional control, and motivation increase over the course of the disease. Alzheimer's Disease International (n.d.) describes the following early symptoms of dementia: memory loss, difficulty performing familiar tasks, problems with language, disorientation in time and place, poor or decreased judgement, problems with concentration, planning or organising, misplacing things, changes in mood or behaviour, trouble with images or spatial relationships, and withdrawal from work or social activities. These limitations increase in severity during the medium stage of Alzheimer's disease to a degree where an independent lifestyle is no longer possible (Deutsche Alzheimer Gesellschaft, 2022). The severe dementia stage is characterised by a complete need for care; the patients are bedridden and it is usually no longer possible to communicate with them (Deutsche Alzheimer Gesellschaft, 2019).

The described decline in daily living skills is accompanied by a need for assistance that is conducted by other people, usually both professional as well as private caregivers (Kutzleben et al., 2017). Caring for people with dementia entails great stress for the caring relatives (Engel, 2006; Gutzmann & Zank, 2005; Zank et al., 2007), which usually results in the move of the person with dementia to a nursing home because the home care situation could no longer be maintained. 49% of people with dementia die in nursing homes (Dasch & Lenz, 2022).

2.1.2.2 Elder Abuse

In comparison to dementia, research on experiences of elder abuse is rather new and arose in the 1960s, when research started to focus on marginalised groups (Teaster et al., 2010). In the course of the demographic change in the US and in Western European countries, the perception of elder abuse has shifted; while it used to be considered relevant only for a small group of people, it is now seen as a major health problem (Dong, 2017). Elder abuse has many facets, and understandings and operationalisations on elder abuse range from severe criminal acts (Mulford & Mao, 2017) to perceived ageism (Pillemer et al., 2021). The WHO has developed a definition of elder abuse that encompasses these aspects and is widely accepted. It defines elder abuse as "a single or repeated act or lack of appropriate action, occurring within any relationship where there is an expectation of trust which causes harm or distress to an older person" (World Health Organization & Université de Genève, 2008, p. 1). Elder abuse can be of various forms, such as physical, verbal, psychological / emotional, sexual, and financial abuse as well as intentional or unintentional neglect (ibid.).

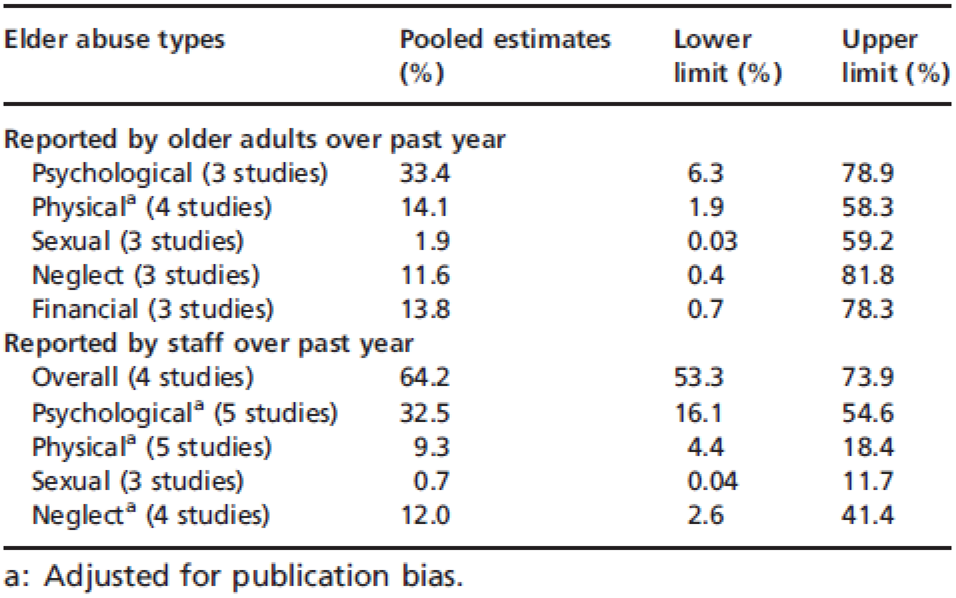
Several meta-analyses estimate the prevalence of elder abuse. Yon et al. (2017) report an annual prevalence of 15% with a corresponding 95% confidence interval from 13% to 19% for elder abuse in community settings. The included studies show a great heterogeneity of 99% in Higgins I^2 and a significant p-value of < 0.001 for the Q-statistic. The reported prevalence of included studies ranges from 0.8% (Marmolejo, 2008) to 79.7% (Silva-Fhon et al., 2015). This wide range and the reported heterogeneity suggest that no single definition of elder abuse was operationalised in the included studies. Indeed, the included studies differ on how narrow or wide they interpret violence. While they agree on a definition of violence or elder abuse as an "act that harms or may harm an elderly person" (Marmolejo, 2008, p. 88), it is not clear where harm begins. Therefore, it is debatable if a meta-analysis should have been performed with all included studies. Yon et al. (2017) further show that sample size was significantly associated with elder abuse and that income (high and medium vs. low) and sampling method show a questionable association ($0.05 < p < 0.1$) with elder abuse. Further meta-analyses on types of abuse reveal the prevalence rates presented in Table 3.

Type of abuse	Annual prevalence (95% confidence interval)
Psychological abuse	11.6% (8.1% - 16.3%)
Financial abuse	6.8% (5.0% - 9.2%)
Neglect	4.2% (2.1% - 8.1%)
Physical abuse	2.6% (1.6% - 4.4%)
Sexual abuse	0.9% (0.6% - 1.4%)

Table 3 Annual prevalence of different types of elder abuse (Yon et al., 2017)

In another meta-analysis that focuses solely on population-based studies, Ho et al. (2017) estimate an annual prevalence of elder abuse of 10.0% (95% CI: 5.2 – 18.6). Again, a high heterogeneity can be observed. Higgins I^2, that reflects on what proportion of the total variance lies between studies (Borenstein et al., 2021, p. 117), is rounded to 100%, and the p-value of the Q-statistic is < 0.001. Meta-regression revealed that the proportion of married older adults is a significant predictor of elder abuse, while a significant influence of the proportion of women in the sample, proportion of sample living with others, proportion of perpetrators of adult children, proportion of perpetrators as spouses, and proportion of abused as females could not be proven. Ho et al. (2017) further perform a meta-analysis on third-party- or caregiver-reported studies that yields results in a prevalence estimate of 34.3%.

A further meta-analysis explores prevalence rates of elder abuse in institutional settings (Yon et al., 2019). Unlike in the meta-analyses presented earlier, no overall estimate of elder abuse in self-reports is calculated here. For self-reports, only individual prevalence estimates for each form of abuse as shown in Figure 4 are calculated.



Elder abuse types	Pooled estimates (%)	Lower limit (%)	Upper limit (%)
Reported by older adults over past year			
Psychological (3 studies)	33.4	6.3	78.9
Physical[a] (4 studies)	14.1	1.9	58.3
Sexual (3 studies)	1.9	0.03	59.2
Neglect (3 studies)	11.6	0.4	81.8
Financial (3 studies)	13.8	0.7	78.3
Reported by staff over past year			
Overall (4 studies)	64.2	53.3	73.9
Psychological[a] (5 studies)	32.5	16.1	54.6
Physical[a] (5 studies)	9.3	4.4	18.4
Sexual (3 studies)	0.7	0.04	11.7
Neglect[a] (4 studies)	12.0	2.6	41.4

a: Adjusted for publication bias.

Figure 4 Prevalence of forms of elder abuse in institutionalised settings (Yon et al., 2019)

In institutional as well as in community-based samples, emotional forms of abuse are the most frequent forms of abuse. However, in institutional settings, physical forms of abuse are more common than in community-based samples; here, they constitute the second most frequent form of abuse, followed by financial abuse, neglect, and sexual abuse.

In staff reports, the overall prevalence rate for elder abuse is 64.2%. Prevalence rates reported by staff are higher than those indicated in self-reports. In both forms of reporting, psychological abuse is the most frequent form, followed by physical abuse, financial abuse, and neglect. The lowest prevalence rates are observed for sexual abuse. The reported range of prevalence rates as well as the indicators for heterogeneity are again high. I^2 ranges from 91.1% to 98.3% in self-reports and from 93.4% to 99.1% in staff reports, and the Q-statistic yields significant p-values for each analysis. The observed heterogeneity makes it clear that different constructs fall under the uniformly used WHO definition of elder abuse. Operationalisations differ in how broadly or narrowly the phrase “causes harm or distress” used by the WHO is understood. Due to this heterogeneity, the construct of elder abuse must be further differentiated. The second and third papers presented in this thesis argue that the very old population is a group that requires a specific focus on elder abuse. To explain this further a brief overview of risk factors and theories of elder abuse, which follows hereinafter.

Two reviews (Johannesen & LoGiudice, 2013; Storey, 2020) provide information about the evidence of known risk factors the victims of for elder abuse. Storey (2020) names eight categories of risk factors: 1. problems with physical health, reduced functioning in activities of daily ling (ADL) or instrumental activities of daily ling (IADL), 2. problems with mental health, which includes problems with cognition, and depression; not only the frequency but also the severity of abuse is associated with mental health issues, 3. problems with substance abuse, especially alcohol abuse, 4. dependencies between victims and perpetrators which can be of financial or care-related nature; they also affect case management because dependency on the perpetrator may affect help-seeking as well as the victim’s ability and willingness to defend oneself, 5. problems with stress and coping since adults reporting higher levels of stress have a greater likelihood of being victims of elder abuse; this also implies consequences for case management, where active and problem-focused coping styles are less common, 6. problems with attitude that affect help-seeking, such as a desire to protect the perpetrator or stoicism, 7. victimisation, where people with prior experiences of victimisation are more likely to become victims of elder abuse, and 8. problems with relationships, especially a poor relationship between victim and perpetrator which increases the likelihood of elder abuse. Moreover, Johannesen and LoGiudice (2013) report ethnicity and the lack of a regular doctor as risk factors. They also report that evidence on age, gender, and education are unclear.

Some of these risk factors are known to be more prevalent in old-aged persons. Therefore, higher prevalence rates for old-aged persons should be expected, although the evidence on age alone is unclear (Johannesen & LoGiudice, 2013). However, it is not the expected higher prevalence rates that cause the need for a unique approach in the research of elder abuse in the oldest old; it is rather the almost fatal combination of lower defence options against possible abuse, reduced options for seeking help and guidance in case of abuse, and difficulties regarding the accessibility of social science studies (Görgen, 2009).

Research on perpetrator risk factors identify perpetrators characteristics as better predictors for elder abuse than victim characteristics (Roberto & Teaster, 2017). If the victim has a need for care, caregiver burden, caregiver inexperience, and psychiatric problems are perpetrator risk factors (Johannesen & LoGiudice, 2013). Generally, drug or alcohol abuse and gambling as well as financial difficulties, personality traits such as a blaming or anti-social personality, trauma and past abuse, ethnicity, and cognitive impairment are known risk factors for perpetrators (ibid.). Johannesen and LoGiudice (2013) report no clear evidence for gender as a risk factor. In the more recent review by Storey (2020), physical health problems, a negative attitude towards older people, and being a victim to family violence are seen as additional risk factors for perpetrators.

Risk factors for victims and perpetrators are linked to the issue of reasons and theories of elder abuse. Researchers have often adapted theories from non-gerontological research on violence to develop theories on elder abuse (Roberto & Teaster, 2017). Without any claim to completeness, three theories of elder abuse are briefly presented below. The most prominent theory of elder abuse is the caregiver-stress hypothesis. It states that "the high levels of stress experienced by the caregiver result in abuse of the dependent adult" (Roberto & Teaster, 2017, p. 24). Research has shown that factors strengthening the victim's dependency on others contribute to all kinds of abuse (ibid.). More recently, research has also stated that the perpetrator's dependency on the victim is a strong contributing factor. "Although it was initially theorized that most elder abuse stemmed from the older adult's dependence on the perpetrator and the resulting caregiver stress, empirical evidence has instead shown the reverse to be true (Storey, 2020, p. 5)". However, in old age, dependencies as well as a need for care become more likely (Statistisches Bundesamt, 2022), which corresponds to an assumed higher prevalence of elder abuse in old age. The second major theory on elder abuse centres around ecological models relying on Bronfenbrenner's (1986) ecological model. Bronfenbrenner's model states that each individual is embedded in a variety of systems "that interact with another and with the individual to influence personal development and life experiences" (Roberto & Teaster, 2017, p. 25). The model distinguishes between microsystems, where individual factors like dementia, substance abuse, or disruptive behaviour of the victim or the perpetrator as well as familial variables like living arrangements or history of abuse can be described, the mesosystem including personals structures like employment status and financial resources, the exosystem comprising social isolation or social support, and the macrosystem, which includes cultural norms or public policy. All systems undergo changes in the latest stage of live, the old age. In the macrosystem, an increased exposure to ageism can be observed (World Health Organization), in the mesosystem, institutional living becomes more likely (Statistisches Bundesamt, 2022), and in the microsystem, personal vulnerabilities like need for care or dementia (Alzheimer Europe, 2019) are more prevalent. The third theoretical framework is the feminist theory, which evolves around unequal economic and social resources between the two genders and may lead to women being more likely to stay in abusive relationships. These inequalities are particularly pronounced in old age (Rosenman & Scott, 2009).

All theories of elder abuse presented here refer to life circumstances or mechanisms that are particularly influenced by the decline in health or economic resources associated with old age. Therefore, a special focus on elder abuse in old age seems to be necessary also from a theoretical perspective.

In their review, Dong et al. (2013) point out that elder abuse has a severe impact on the elderly, including increased mortality, reduced physical health, increased risk of depression and depressive symptoms, anxiety, and post-traumatic avoidance (Dong et al., 2013). It has already been pointed out that these consequences are likely to be more pronounced in old age. This, thus, underlines the need for the examination of elder abuse in old age.

2.1.2.3 Multimorbidity

In old and very old age, an increase in diseases, especially chronic diseases, can be observed (Maresova et al., 2019). For people over 75 years of age, the Robert Koch Institute (RKI) lists high blood pressure, osteoarthritis, and elevated blood lipids as the three most common diseases in women, and high blood pressure, cardiovascular diseases, and osteoarthritis as the most common diseases in men (Robert Koch-Institut, 2015). The higher probability of individual diseases in older adults is accompanied by a higher probability of being affected by several diseases at the same time. The Berlin Ageing Study (BASE) shows that very old people are often affected by several chronic diseases at the same time (Steinhagen-Thiessen & Borchelt, 2010). The simultaneous presence of several diseases is called multimorbidity (Johnston et al., 2019). There is some ambiguity as to the number of diseases required to speak of multimorbidity and the way these diseases are assessed (Fortin et al., 2012; Ho et al., 2021).

In the Berlin Ageing Study (Steinhagen-Thiessen & Borchelt, 2010), 18.6% of men and 27.1% of women aged 70 to 84 were affected by five or more diseases. In the group of people aged 85 and older, this increased to 40.9% (men) and 54.3% (women). While the BASE study defines multimorbidity as the presence of five or more diseases, more recent studies use two or three diseases as the cut-off criterion. In the study “Gesundheit in Deutschland aktuell” (GEDA), multimorbidity is defined as the presence of at least two diseases. Based on this approach, 75.4% of women and 68% of men between the ages of 65 and 74 were classified as multimorbid. At age 75 or older, these numbers increased to 81.7% (women) and 74.2% (men) (Robert Koch-Institut, 2015). This age- and gender-differentiated presentation of prevalence rates already illustrates that age and gender are significant risk factors for multimorbidity (Violan et al., 2014). Multimorbidity is also related to the socioeconomic status (Lampert & Hoebel, 2019). Globally, the number of multimorbid people is increasing, partly due to improved diagnostics and lower mortality rates after strokes or heart attacks (Calderón-Larrañaga et al., 2019). Prevalence rates differ between high-income countries and middle-and low-income countries, with higher prevalence rates in high-income countries. Reasons for this „are not known but methodological factors and differential survival are plausible hypotheses" (Skou et al., 2022, p. 3) (see Figure 5). Figure 5 also indicates that data points are scarce beyond the age of 80 in medium and low income countries and that most studies cannot investigate developments of multimorbidity in old age.

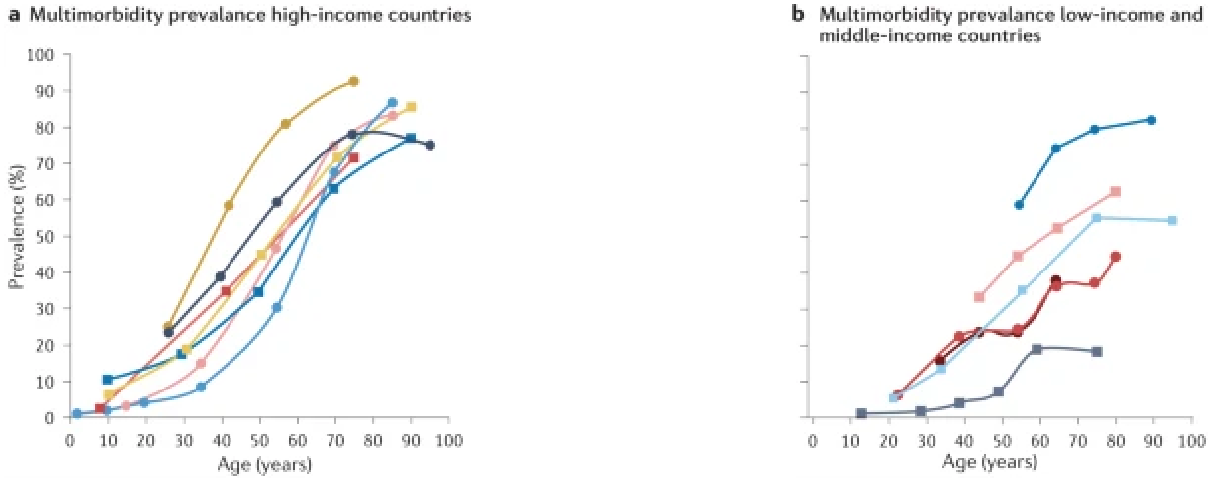


Figure 5 Prevalence of multimorbidity in high- versus middle- and low-income countries (Skou et al., 2022)

Although the concept of multimorbidity is widely used, (On October 20th 2022, Pubmed lists 8326 published articles during the past 10 years), there is no established measurement for multimorbidity. Differences in measurement are thoroughly investigated in a review by Ho et al. (2021) including 566 articles. They include, inter alia, differences concerning the investigated diseases (only eight diseases are included in more than 50% of the studies), the number of diseases investigated which ranges from two to 285, the threshold used for multimorbidity with two or three diseases being the most commonly used threshold, and the used data source which is most often self-report by or interview of the patient (Ho et al., 2021). In spite of the vast amount of published articles with regard to multimorbidity, people with multimorbidity are often excluded in clinical trials, which makes them under-researched in some aspects (Skou et al., 2022). Special focus is therefore needed for this group that is difficult to access.

The effects of multimorbidity are summarised in the systematic reviews by Marengoni et al. (2011) and Makovski et al. (2019). According to these reviews, multimorbidity is associated with functional decline, higher depression and lower life satisfaction, increased mortality, and increased use of the health care system. External resources are of heightened importance in the treatment of chronic diseases (Holzhausen & Scheidt-Nave, 2012). However, these resources are subject to degradation processes in old age. Therefore, it comes as a surprise that the effects of multimorbidity have not been researched with a special focus on the oldest old; this is, hence, a desideratum that this thesis addresses.

2.2 Representativity

A science that relies on empirical evidence to accompany societal changes must first be able to describe these societies. For this purpose, collecting samples that ideally reflect the characteristics of society has become an established practice, and in this context, the term "representativeness" was introduced. This chapter begins by introducing meanings of the term "representativity used within and outside of academia. Afterwards, R-indicators are introduced as a method of measuring representativity, followed by the concept of Total Survey Error.

2.2.1 Definition

The concept of representativity is frequently used in scientific as well as in non-scientific literature but is often not well defined (Schouten, Cobben, & Bethlehem, 2009). An extensive overview of the use of the term is presented by Kruskal and Mosteller (1979a-c and 1980). In four articles, they explore meanings of representativity in non-scientific literature (Kruskal & Mosteller, 1979a), scientific literature excluding statistics (Kruskal & Mosteller, 1979b), scientific literature including statistics (Kruskal & Mosteller, 1979c), as well as the conceptual history of the term (Kruskal & Mosteller, 1980). They summarise nine concepts to which the term "representativity" is often applied:

1. General, usually unjustified acclaim for data
2. Absence of selective forces
3. Miniature of the population
4. Typical or ideal case
5. Coverage of the population
6. Vague term to be made precise
7. Some specific sampling method
8. Permitting good estimation
9. Good enough for a particular purpose

This wide range of meanings results in the criticism that the term "representativity" itself is unspecific und unnecessary (Schnell et al., 2018). In survey methodology, the term was first introduced by Kjaer in 1895 who states that „if a sample was representative with respect to variables for which the population distribution was known, it would also be representative with respect to other survey variables“ (Dumicic, 2011). Although nowadays, this optimism is no longer generally shared (Schnell et al., 2018), the concept of representativity is still used. A modern definition of a "representative sample" is a sample that „ensures external validity in relationship to the population of interest the sample is meant to represent“ (Lavrakas, 2008). Döring and Bortz (2016) define sample representativeness as a measure of how well a sample mirrors the characteristics of a population. They distinguish between a characteristic-specific sample, in which the sample mirrors the population in a given set of variables, and a global representative sample, where the sample mirrors the population in all characteristics (Döring & Bortz, 2016). All these definitions of representativity and representative samples allude to a consistency between sample and population. Moreover, they all contain the problem that an operationalisation for the measurement of representativity is not included or directly apparent.

The problem of measuring representativity has been recently addressed by Schouten et al. (Schouten, Cobben, & Bethlehem, 2009). They distinguish between weak and strong representativity (Schouten, Morren et al., 2009). Strong representativity refers to equal inclusion probabilities for all individuals in the population, while weak representativity means that response probabilities are constant with respect to a set of auxiliary variables (Schouten, Morren et al., 2009). While strong representativity is still not a measurable characteristic of a population, weak representativity can be measured.

Before this is outlined, some differences between the presented definitions need to be delineated first. At first glance, the greatest difference between the understanding of representativity of Schouten, Cobben, & Bethlehem (2009) and Dumicic (2011) as well as Döring and Bortz (2016) seems to be that Schouten et al. define representativity as a characteristic of the population and not of a sample drawn from a population. However, on the basis of the central limit (Field et al., 2014; Leonhart, 2009) theorem, any sample that is representative in the sense of Schouten, Cobben, and Bethlehem (2009) is also representative in the definition given by Lavrakas (2008). Another important differentiation is that, while strong representativity applies to all variables of a sample, weak representativity only refers to a set of variables.

2.2.2 R-indicators

To estimate the representativity of a population, Schouten, Morren et al. (2009) introduce R-indicators as indicators for the representativeness of a sample. They propose the term

$$R(\rho) = 1 - 2 * S(\rho)$$

as an indicator for deviations of representativity with ρ being the response probability and $S(\rho)$ being the standard deviation of the response probabilities defined as:

$$S(\rho) = \sqrt{\frac{1}{N-1} * \sum_{i=1}^{N} (\rho_i - \bar{\rho})^2}$$

with N being the sample size, ρ_i being the response probability of a single person in the population, and $\bar{\rho}$ being the average response probability. As the maximal standard deviation of a dichotomous variable is 0.5, the possible values of R-indicators range from 0 to 1 (Schouten, Morren et al., 2009). This formula indicates that representativity is defined by the variance of response probabilities in the population, with higher variance in response probability inducing less representativity of the population. Response probabilities can be estimated with respect to variables that are known for the population in question. However, this approach results in two mathematical problems. First, the required estimation of response probabilities in the given formulas is not directly possible, becomes mathematical complex, and is not free of bias (Shlomo et al., 2009), which is why, in reality, response propensities are used (ibid.). However, these statistical problems can be solved. The greater problem in measuring representativity is that the auxiliary variables need to be known for the population, which is usually not the case for all variables of interest. Therefore, representativity cannot be estimated or claimed for all variables of interest. For example, life satisfaction is usually not known for every individual of a population and hence, a sample cannot be assessed to be representative for life satisfaction but only for auxiliary variables, such as age, sex, or nursing home status, that are known for the population. By choosing auxiliary variables deliberately, a high value of representativity could even be assumed when the opposite is the case – a process Schnell et al. (2018) seem to fear when they criticise that the “expected applications of these indicators can be viewed as misleading” (Schnell et al., 2018).

Särndal and Lundström (2008) propose a different estimator for the measurement of representativity that ranks alternative auxiliary vectors in their ability to calibrate survey weights and thereby reduce bias. However, a detailed description of this approach is beyond the scope of this text.

2.2.3 Total Survey Error

In the scientific discourse, an assessment and discussion of the quality of surveys is of central importance to understand what a good quality is and how it can be achieved (Faulbaum, 2022).

For a good assessment of survey quality, not only representativeness is crucial, which can be measured with respect to some variables in the form of R-indicators, but also other components, which are presented within the scope of the Total Survey Error (Groves & Lyberg, 2010). In the explanatory framework of the Total Survey Error, two groups of errors are distinguished: sampling errors and non-sampling errors (Biemer, 2010). Sampling errors originate from the selection of a sample instead of the whole population (Biemer & Lyberg, 2003) and are often accompanied by a restriction of representativeness. While Biemer (2010) differentiates sample errors into errors due to a) sampling scheme (e.g. single or multi-stage sampling), b) sample size, and c) estimator choice, Faulbaum (2022) more recently distinguishes between estimation errors and selection errors. Estimation errors broach the issue that in any estimation of a population value by a random sample, sample characteristics can deviate from the true population value. Selection errors occur when assumptions of random sampling are violated and response probabilities are diverse (Faulbaum, 2022).

Non-sampling errors Biemer (2010) are distinguished between a) specification errors, b) nonresponse errors, c) framing errors, d) measurement errors and e) data processing errors (Biemer, 2010). Specification errors occur when the intended construct and the measured construct differ, which is akin to the concept of validity in psychometrics (Groves & Lyberg, 2010). Nonresponse errors comprise unit as well as item nonresponse. Unit nonresponse refers to a sample unit, often an individual, that does not answer any part of a questionnaire, while item nonresponse refers to specific missing values of parts of a questionnaire. Often, information on financial resources is more exposed to this type of error. If the item nonresponse is assumed to be missing at random, nonresponse errors can be dealt with through multiple imputation (Patrician, 2002) or using full information maximum likelihood estimators (Schminkey et al., 2016). Both procedures tend to deliver similar results (Lee & Shi, 2021). Unit nonresponse can be treated by weighting the sample if the desired information is known for the population (Brick, 2013). Frame errors occur when units of the desired sample are omitted or doubled (Biemer, 2010). The omitting of units can easily happen with difficult to survey groups like nursing home residents or prisoners.

Measurement errors are said to be the best researched source of errors (Biemer, 2010). In the framework of classical test theory, an observed variable x_i is usually divided into its true value t_i and an error variable ε_i

$$x_i = t_i + \varepsilon_i,$$

where i is the index of the sampled unit (e.g. a person) und goes from 1 to n. The error variables are assumed to be normally distributed with a mean of zero. Measurement errors are therefore the difference between the true and the observed value of an individual (Faulbaum, 2022). They arise from respondents, interviewers, question wording and question design, and various other interview factors (Biemer, 2010; Faulbaum, 2022). Respondents may intentionally or unintentionally give incorrect answers to questionnaires. Interviewers can influence answers by their speech, appearance, and mannerisms (Biemer, 2010). The interviewer’s influence comprises lower item nonresponse for male interviewers (Bittmann, 2022), while the interview difficulty is perceived lower for female interviewers (Vidovičová & Doseděl, 2018). The questionnaire can cause errors when questions are ambiguous, instructions confusing, and when terms are easily misunderstood (Biemer, 2010). Interview factors that may influence the interview situation are present third persons who can both harm and better answers of respondents with limited communication abilities; such factors are more common in surveys of the oldest old. Differences between proxy information and target person information are also a common kind of measurement error in examinations of the oldest old. Differences of measurement between groups can be explored through an analysis of measurement invariance, which tests if the factorial structures of a questionnaire differs between different groups (Vandenberg & Lance, 2000). In the context of surveys on old age, it can be used to examine differences between proxy and self-reported interviews. If the factorial structure differs between groups, this can be seen as an indication to not pool the measurements of the analysed groups. Four kinds of measurement invariance are distinguished: 1) configural, the equivalence of the model forms, 2) metric, where factor loadings are equal between the analysed groups, 3) scalar, where item intercepts or thresholds are equal between groups, and 4) residual measurement invariance, where items' residuals are equal (Putnick & Bornstein, 2016). The last type of errors, data-processing errors, "includes errors in editing, data entry, coding, assignment of survey weights, and tabulation of the survey data" (Biemer, 2010, p. 824).

While the Total Survey Error is a theoretical framework for the description of all kinds of survey errors, the mean squared error (MSE) is its empirical representation. It is defined as the expected squared difference between an estimate $\hat{\theta}$ and the parameter it is intended to estimate, Ɵ. This can be written as

$$MSE = E\left(\hat{\theta} - \Theta\right)^2.$$

The formula can be converted to show that the mean squared error is composed of one term describing bias and one term describing the variance of a variable.

$$MSE = Bias^2 + Variance.$$

This means that the error is not only reduced by reducing bias, but also by reducing variance in the measurement An overview of the Total Survey Error framework and the MSE as given by Biemer (2010) is presented in Figure 6.

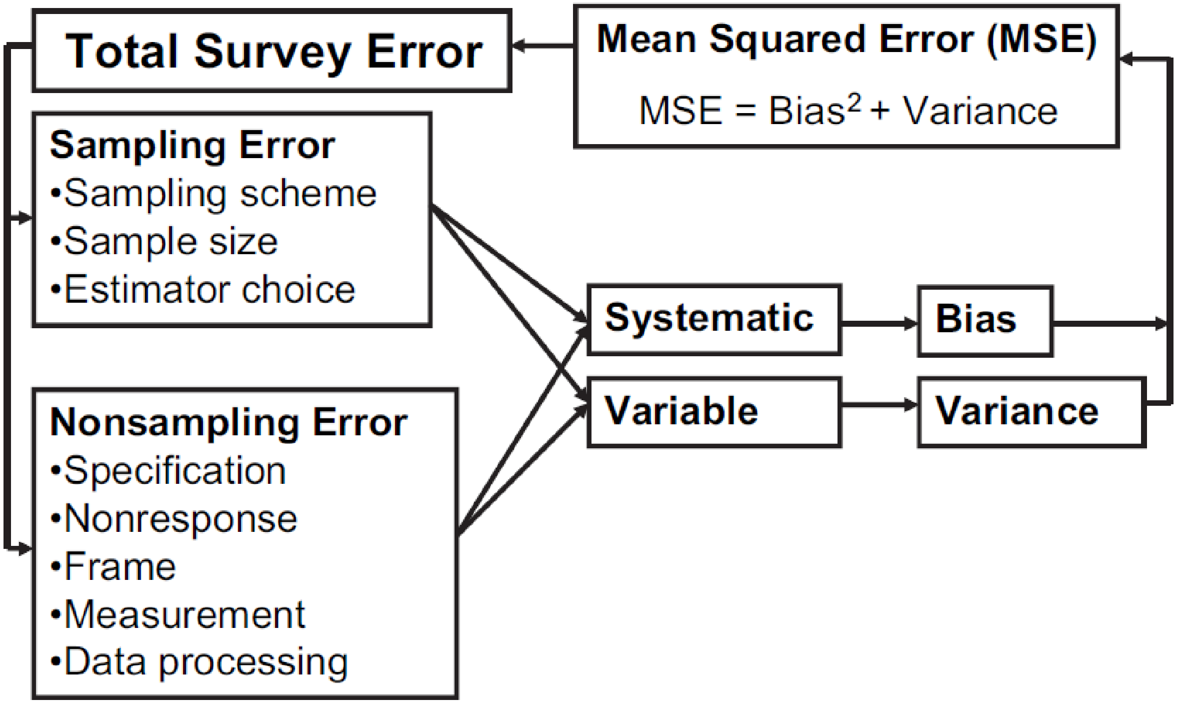


Figure 6 Total Survey Error and MSE (Biemer, 2010)

3 Representativity and Total Survey Error in Old Age

The following chapter a connects the concepts of old age and representative research and discusses challenges in conducting representative research on the oldest old. Problems are discussed on the basis of the three topics of dementia, elder abuse, and multimorbidity, starting with the topic of dementia. The Total Survey Error framework is used as a structure for the three chapters.

3.1.1 Total Survey Error and Dementia

The first challenge in surveying people with dementia concerns the inclusion of people with dementia and therefore evolves around sampling or frame error. In 2000, Riedel-Heller et al. (2000) already reported that prevalence rates of dementia are lower in community-based samples of people aged 75+ when proxy information and the nursing home population are not accounted for. Among face-to-face interviews, there was a prevalence for dementia of 5.3%, while dementia prevalence increased to 6.3% when proxies were included and to 10.5% when nursing home population was additionally sampled. Conversely, this means that sampling errors will be greater if proxies and the nursing home population are not included.

Further challenges arise around the issue of the inclusion of sampled people with dementia. To include people in any survey, informed consent is necessary (Manti & Licari, 2018). Although at the onset of dementia, the mental ability for informed consent of the sampled target person might be reduced, the exclusion of people with dementia is no viable alternative (Holden et al., 2018; Slaughter et al., 2007). Despite newly developed procedures to obtain informed consent from people with dementia (Wied, Haberstroh et al., 2021) which are well reviewed among scientists (Wied, Poth et al., 2021), an increased participation of people with dementia by use of these procedures cannot be observed (Poth et al., 2022). Therefore, obtaining informed consent from people with dementia is still difficult. Depending on when informed consent is obtained, these difficulties can lead to sampling error or unit nonresponse.

Higher item nonresponse among people with dementia is also possible when questions are difficult to understand for them. However, this hypothesis is not well researched. On November 9th 2022, the search term “(response rate) and (alzheimer or dementia) and (survey or representative or population-based) and (old age or oldest old or fourth age)” yielded just 113 hits in Pubmed and the subsequent title screening excluded 105 of these hits. Only one study directly compares missing values in people with dementia and people without dementia; It finds that in subjective health status people with dementia show more missing values than people without dementia (Rodríguez-García et al., 2019). However, Smeding and Koning found already in 2000 that missing values resulting from behavioural disturbances increase in people with frontotemporal dementia which, in turn, intensifies behavioural disturbances. Furthermore, Michalowsky et al. (2020, p. 765) state that the number of missing values “increases with the progression of dementia diseases”.

Specification error might occur when constructs change during the onset of dementia or when subjective perceptions of reality of the person with dementia do not match to others. This is, for example, of particular importance when the research question investigates the topic of living with dementia and the person is not aware of their disease due to anosognosia or other forms of limited awareness of the disease (Alexander et al., 2019; Clare, 2010; Stechl et al., 2007).

Although measurement error is considered to be the best analysed error in the Total Survey Error framework (Biemer, 2010), here, too, research on measurement error in connection with people with dementia has just begun. So far, there is one guideline for “reporting methodological challenges and evaluating potential bias in dementia research” (Weuve et al., 2015). Concerning measurement error, Weuve et al. (2015) point out two major, rather general issues. First, they claim that “there is often no clear gold standard in dementia research” (Weuve et al., 2015, p. 1103), which makes it difficult to assess the validity of measurements. Second, they indicate that the frequently used interval scaling of ordinal scales might lead to an inflation of type 1 error, especially when there are floor and ceiling effects. Furthermore, on the measurement of cognition itself, it is highlighted that the repetition of diagnostic procedures implies training effects and that “Cognitive declines stemming from incipient dementia may thus be impossible to detect because of practice-related improvements on test performance” (Weuve et al., 2015, p. 1104). Because measures that are valid for one group of people may not be valid for another (ibid.), some studies compare measurements in people with dementia to measurements in people without dementia, and find that measurements in people with dementia to be less reliable (Olsen & Bergland, 2017; van Deun et al., 2018). Moreover, they infer that evidence in people with dementia is harder to come by due to limitations in sample size (Hörnsten et al., 2021; Olsen & Bergland, 2017). However, these findings are rather spurious, and no common body of literature can be identified.

In some constructs, for example in pain assessment for people with dementia, it has become an established practice to use not only information from people with dementia themselves, but also from proxy informants. “In regard to pain assessment, a large proportion of people with moderate to severe dementia were unable to complete a self-report pain instrument. Pain was more commonly reported by informal caregivers than the person with dementia themselves.” (Bullock et al., 2019, p. 807). As the cognitive impairment of patients increases, self-reports of experiences of pain people with dementia are considered less valid.(Lautenbacher & Kunz, 2019). Relying on information other than self-reports by people with dementia is not only common in pain assessment but also in examinations of well-being (Martyr et al., 2018) or functional impairment, quality of life, and behavioural problems (Sheehan et al., 2012). Although sometimes, proxies seem to be the only viable source of information about people with severe dementia, the differences in these types of measurements can be a source of measurement error. Hounsome et al. (2011) reports that there is ”a lack of association between patient and proxy ratings“ (Hounsome et al., 2011, p. 390) and shows that different proxies (family carers, institutional carers) differ in their assessment of the patient’s quality of life. Similar effects have been reported by Griffiths et al. (2020), Algar et al. (2016), and Orgeta et al. (2015). These findings should not be the case if all proxy informants referred to the same gold standard: the self-report. However, given the differences between proxy and self-ratings even in milder cases of dementia on the one hand (O'Shea et al., 2020) and an inability to take part in interviews in, the central question is what a valid and reliable measurement in person with severe dementia is and how it can be achieved. The inclusion of proxies as informants seems to be an important step forward (Sheehan et al., 2012), but the ideal, namely that self and proxy ratings complement one another (Burks et al., 2021), is not always naturally achieved. Future research on representativity needs to analyse whether the measurement error originating from different kinds of measurements outweighs the sampling error that derives from excluding people with dementia.

Frame error might occur more frequently in people with dementia as they are part of the nursing home population that is sometimes omitted in surveys (Riedel-Heller et al., 2000). No research can be identified that addresses dementia in connection with data processing errors. It is noteworthy that in Faulbaum's (2022) classification, these errors might be considered coverage errors.

Two areas can be identified for further research. First, the spurious findings on all types of errors that occur in surveys on people with dementia need to be complemented by further empirical studies before a systematic review can summarise these findings. Second, it needs to be evaluated whether the inclusion of people with later stages of dementia via proxies reduces or enhances the mean squared error in the sample.

3.1.2 Total Survey Error and Elder Abuse

In general, compared to research on dementia, there is less empirical evidence on the issues surrounding the survey of old-aged persons on the topic of elder abuse. Concerning sampling errors, the choice of estimator is a first source of error. While many studies report prevalence rates of elder abuse (Ho et al., 2017; Yon et al., 2017; Yon et al., 2019), this dichotomous measurement does not account for different severities and frequencies of abuse. Problems with sample size arise as elder abuse comprises rare as well as frequent acts of abuse. Jia et al. (2021) mention that meta-analyses of rare events tend to be underpowered while Zhou and Shen (2022) as well as Cai et al. (2010) emphasise issues with common estimates in meta-analyses. The third source for sample errors in the Total Survey Error framework is the sampling scheme. Görgen, Herbst, and Rabold (2009) recommend sampling through registration offices for surveys on old-aged persons.

A number of non-sampling errors have been investigated in connection with research on elder abuse. In an analysis of nonresponse in an elder abuse survey with a European sample, De Donder et al. (2013) find that item nonresponse "is influenced by individual characteristics (social status, vulnerability), method effects such as content (sensitivity), the order of the questions (forms of abuse), by type of data collection and the presence of assistance in survey completion" (De Donder et al., 2013, p. 1021). With regard to frame error, the doubling or omitting of sample units, a special focus needs to be laid on the nursing home population that is especially difficult to survey, even more so during the COVID-19 pandemic (Scherpenzeel et al., 2020). Concerning data processing, there seems to have been no specific research in the context of elder abuse. Some research, however, focuses on measurement error. Visschers et al. (2017) find that the answers of perpetrators in an investigation of intimate partner violence show social desirability bias, but the extent of this bias remains small. Hence, the question remains: To what extent can elder abuse be measured through proxy informants who in some cases might be perpetrators themselves? Fang et al. (2022) show moderate to substantial agreement between caregivers' and care recipients' report, and, contradictory to Visschers et al. (2017), caregivers' answers yield higher prevalence rates in all investigated forms of abuse than care recipients' answers. Thus, more studies need to investigate the validity of proxy-based information in elder abuse. Similar questions can be asked about the validity of the participation of people with dementia and of present third persons. Specification error seems to be the most difficult type of error to address. Although various researchers have been stating for decades that research on elder abuse has a theoretical deficit and that operationalised concepts show a wide heterogeneity (Ho et al., 2017; Lachs & Pillemer, 2004; Yon et al., 2017; Yon et al., 2019), the lack of uniform operationalisations still leads to great heterogeneity in meta-analyses. A further investigation of the influence of third persons and proxy informants on interviews and the creation of a better foundation for the theory of elder abuse, especially in old age, seem to be the most eminent research tasks.

3.1.3 Total Survey Error and Multimorbidity

In the field of multimorbidity, sampling processes are more extensively researched than in the field of elder abuse. Two reviews analyse sampling sizes, sampling schemes, and estimator choices (Fortin et al., 2012; France et al., 2012). While France's et al. (2012) review focuses on cohort studies, Fortin et al. (2012) addresses the methodology of the included articles directly. In regard to sampling size and sampling scheme, France et al. (2012) state that no study uses random sampling and that sample sizes are relatively small. Fortin et al. (2012) report differences in sample sizes and recruitment method. Differences also include the number of diagnoses required for multimorbidity, which is the “most important factor on estimating prevalence” (Fortin et al., 2012, p. 142).

Apart from data processing errors, all non-sampling errors of the Total Survey Error framework have been examined with respect to multimorbidity. Different studies examine nonresponse bias on health related outcomes. Lee et al. (2009) directly estimate nonresponse bias through a comparison of the California Health Interview Survey (CHIS) and US Census data and state that the “response rate in CHIS did not result in significant nonresponse bias and did not substantially affect the level of data representativeness” (Lee et al., 2009, p. 1811). However, this result is not reported uniformly – even within the same country. Fakhouri et al. (2020) find that, unlike in previous waves, in the 2017-2018 sample of the National Health and Nutrition Examination, weighting could only reduce but not eliminate nonresponse bias. Halbesleben and Whitman (2013) as well as Locker (2000) also mention weighting as a tool to address nonresponse bias in health-related surveys. Concerning frame error, the nursing home population needs to again be accounted for. Specification error has been discussed in the previously mentioned Berliner Altersstudie. Steinhagen-Thiessen and Borchelt (2010) present differences between diagnoses made by general practitioners and by specifically trained study doctors and report that several diseases in old age are not properly diagnosed by general practitioners. Although the different tools for measuring multimorbidity overlap in large areas, they differ in detail, which provides another opportunity for specification error (Johnston et al., 2019; Starfield & Kinder, 2011). Another form of specification error can occur when different data sources are used for health information:

> “Current methods include interviews, self-reports, medical record reviews, administrative databases, and clinical examinations. Analyses of surveys containing both self-report and objective measurements of health status have documented systematic biases in self-reports according to age, sex and socioeconomic status (Sadana, 2000)” (In:Marengoni et al., 2011, p. 435).

Subsequently, measurement errors can occur when target persons are not aware of their diseases because they have not been diagnosed. Depending on the wording of the question (the NRW80+ study, for example, asks explicitly for treated diseases), misunderstandings can happen easily, especially when drop-off questionnaires are administered.

4 Publications

This thesis contains five papers that address the problem of representativity in old age. The first paper focuses on methodological challenges and idescribes advantages and disadvantages of the inclusion of nursing home residents, the representation of people unable to participate in surveys via proxy interviews, and cognitively impaired persons. It links the other papers, that illustrate the topics of elder abuse, dementia and multimorbidity. The subsequent two articles look at the subject of elder abuse. While the second paper concentrates on the measurement of elder abuse, the third paper focuses on prevalence rates, risk factors, and consequences of elder abuse. The fourth paper describes different aspects of quality of live in old age. The final paper emphasises the relevance of multimorbidity on different quality of life outcomes.

The literature quoted in these papers is presented in the references within the papers and not part of the references presented in chapter 7.

4.1 Challenges and Benefits of Including the Institutionalized, Cognitively Impaired and Unable to Respond in a Representative Survey of the Very Old

The first paper within this thesis focuses on advantages and disadvantages of surveying certain difficult-to-survey groups that are common among the oldest old. It describes the number of missing values in the nursing home population as well as in people with impaired cognition and examines consequences of the exclusion of the nursing home population, people with limited cognitive functioning, and people who are represented by proxy interviews on R-indicators and nonresponse bias. Furthermore, it explores the consequences of the exclusion of these three groups for the estimation of socioeconomic status, functioning, and well-being.

Challenges and Benefits of Including Institutionalized, Cognitively Impaired, and Unable-to-Respond Individuals in a Representative Survey of the Very Old

Roman Kaspar[1], Thomas Brijoux[1], Andrea Albrecht[1], Jaroslava Zimmermann[1], Judith Wenner[2], Jonas Fey[1], Marcella Reissmann[1], Michael Wagner[1,3] & Susanne Zank[1,2]

[1] University of Cologne, Cologne Center for Ethics, Rights, Economics, and Social Sciences of Health, Cologne, Germany

[2] University of Cologne, Faculty of Human Sciences, Cologne, Germany

[3] University of Cologne, Faculty of Management, Economics and Social Sciences, Institute of Sociology and Social Psychology (ISS), Cologne, Germany

Abstract

Institutionalization, cognitive impairment, and the inability to conduct an interview due to health impairment are among the top exclusion criteria for most large-scale social and aging surveys. Reservations about targeting vulnerable groups result from economic or legal restrictions of recruitment and concerns regarding research ethics or the validity of the data obtained. However, failure to include these individuals may lead to substantial bias. Metadata showed that privileged data access and checks against nursing home repositories prevented the undercoverage of institutionalized individuals. Measures to include difficult-to-survey groups led to a marked increase in response rates. Individuals with health impairments substantially contributed to the representativity of the sample. Nonresponse bias was cut in half when compared with a less inclusive study protocol. From a Total Survey Error perspective, reductions in nonresponse bias, low item-nonresponse, and evidence of measurement invariance across self-reports and proxy reports for key outcome variables show significant benefits of including difficult-to-survey groups in estimating characteristics of this population.

Keywords: quality of life, proxy report, total survey error, nonresponse bias, measurement invariance

Challenges and Benefits of Including Institutionalized, Cognitively Impaired, and Unable-to-Respond Individuals in a Representative Survey of the Very Old

1 Introduction

Very old individuals have been described as a difficult-to-survey part of the population (Feskens, 2009). Institutionalization, cognitive impairment, and lack of communicative ability to conduct an interview or questionnaire are among the top reasons for exclusion or nonresponse of the oldest old[1] in most large-scale social surveys (Sakshaug, 2022; Schanze, 2017) and many aging studies (e.g., Davies et al., 2010; Deeg et al., 2002; Schiel et al., 2021). Much of the reservation against targeting more vulnerable clients is well substantiated by economic or legal restrictions and concerns about research ethics or the validity of data obtained. For example, community registers in Germany are prohibited by law from including addresses of nursing home residents (NHR) in samples drawn for commercial survey institutes (BMG §52). Davies et al. (2010) report in the Newcastle 85+ study that increased time was necessary to engage with family and gatekeepers. Lower contact rates have consistently been reported for NHR (Gaertner, Koschollek, et al., 2019; Wagner et al., 2019). Based on data from the Survey of Health, Ageing and Retirement in Europe SHARE, Schanze (2021) found the data quality was lower in individuals with lower socioeconomic and health status. Simultaneously, failure to include those with more prominent health issues or in nonprivate dwellings may lead to a substantial underestimation of, for example, dependencies in the activities of daily living (Kelfve et al., 2013; Schanze & Zins, 2019) or health inequalities (Kelfve, 2019). Vulnerable individuals represent a nonnegligible portion of the population aged 80 years and older. Hence, their inclusion is essential in surveying the older

---

[1] There is no consensus definition of very old age or the oldest old. In this study, we use the cut-off of 80 years or older, also used in the Survey of Health, Ageing and Retirement in Europe SHARE (Andersen-Ranberg et al., 2005).

population. The current lack of experience with including the most vulnerable in survey research is not limited to questions of recruitment and the mode of the interview but also pertains to identifying valid ways to integrate the collected information into the analysis and subsequent reporting. This study exemplifies how including difficult-to-survey groups of older adults could reduce nonresponse bias and provide better population estimates from survey responses.

2 Elements of Survey Quality

The Total Survey Error (TSE) framework gradually arose from the work of scholars wanting to optimize survey design concerning different threats to survey quality (for an overview, see Groves & Lyberg, 2010). An early example is Groves (1987), who jointly considered both nonobservation error (i.e., coverage, sampling, and nonresponse errors) and observation (i.e., measurement) error. In this study, we consider the quality of a survey as a function of the coverage of the targeted population in the sampling frame, the definition of inclusion probabilities for individuals drawn from the sampling frame, the share and selectivity of respondents, and the degree to which the response measures the true value of respondents' characteristics. Specific challenges have been reported regarding most, if not all, of these steps when studying very old age (Schanze, 2019; Schanze & Zins, 2019).

2.1 Sampling Errors

Concerning *adequate coverage of the population of the very old*, a first issue is conceptual and arises from the need to a priori define very old age. While various definitions of very old age have been suggested in different research fields (Degnen, 2007; Foo et al., 2019), very limited official register data exist to define the population of the very old other than by chronological age (e.g., mostly 80 years or older). Even a simple demographic definition (i.e., more than half of one's birth cohort has passed) is subject to historical change and jeopardized by gender-specific mortality rates. Second, community residential register data in this age segment have a higher probability of being outdated (e.g., due to the delayed implementation of death records or relocation), incomplete (e.g., retaining the original address after relocating to a nursing home), or inaccessible. Regarding the latter, legal restrictions may limit information access on particularly vulnerable parts of the population, such as NHR. In Germany, data security laws afford statistic bureaus to mask sociostructural information from residential registries to disable potential reidentification, particularly in small communities with only a few very old individuals. Scherpenzeel et al. (2017) have described differences in sampling frames for social surveys across European countries.

Concerns have also been raised regarding *the definition of inclusion probabilities for individuals drawn from the sampling frame*. Because there are many more women in this age segment today than men and the number of individuals declines rapidly with increasing chronological age, simple random sample data will be limited in its potential to allow for reliable comparison of subpopulations in very old age. For example, oversampling of males is required to conduct meaningful subgroup analyses across gender. However, disproportionate sampling schemes compromise sampling efficiency, because they need to be accounted for in the analysis by using design weights (Aust & Gilberg, 2016). Moreover, they require detailed information on the (age) structure of the targeted population to compute selection probabilities adequately. First, and related to problems of retaining detailed information on the age structure, official data on the number of inhabitants above the age of 80 may be difficult to obtain, particularly in small communities, depending on regional data protection laws or default register procedures. For example, standard reporting, from the state administration to the Federal Bureau of Statistics in Germany, excludes more finegrained age or birth cohorts beyond age 80. Second, the relative scarcity of the oldest old at the level of the primary sampling unit (e.g., communities) poses problems for delivering the required number of oldest-old individuals. For any expected total number of participants, these circumstances effectively increase the number of necessary primary sampling units (PSUs) (hence, reducing the design effect), result in synthetic sampling points or reduce the gross sample because communities may not be able to deliver the required number of oldest-old individuals. They also effectively limit the disproportionality that can be accomplished in a sampling design to safeguard subgroup comparison. Specific and sequential mixed-mode designs have been suggested to survey very old adults (Gaertner, Koschollek, et al., 2019; Gaertner, Lüdtke, et al., 2019). Additionally, separate sampling frames for subpopulations, such as nursing home registers, have been considered (Schneekloth & Müller, 1998). Integrating alternative data sources into multi-frame sampling has been suggested (Sand, 2014). However, such schemes may share the problems reported above for residential register data and challenges related to selecting random samples in dynamic and unique entities, such as nursing homes (preselection, high intraclass correlation).

2.2 Nonresponse and Measurement Errors

Specific challenges have been reported regarding *the number of very old individuals willing to and capable of participating* in survey research and whether this group is *representative of the general population of very old adults*.

First, aging survey response rates have decreased over the last decades. In the German Aging Survey, base sample participation dropped from 50.3% in 1996 to 27.1% in 2014 (Vogel, Klaus, et al., 2020). Mistrust of strangers and lack of interest in issues not affecting them were presumed reasons for the lower motivation of older adults to participate (Kühn et al., 1999). Wagner et al. (2019) have identified specific dropout mechanisms for subgroups of the very old. More specifically, individuals in different settings (e.g., private vs. institutional) and with different levels of functioning drop out of the study at different points of the recruitment process. For example, low contact rates in NHR indicate difficulty in accessing some of these target persons. However, significantly lower refusal rates have been found in NHR that could ultimately be contacted, corroborating the importance of gatekeepers for surveying vulnerable individuals (Davies et al., 2010; Hall et al., 2009).

Second, very old and oldest-old age is characterized not only by more years to live but also by a substantial proportion of time spent with illness and loss of independence (Tesch-Römer & Wahl, 2017). Even aging studies targeting very old adults sometimes fail to include some of the most vulnerable segments of this population, often for practicability or skepticism concerning the quality of the data obtained (Schanze, 2021). Exclusion criteria (e.g., unable to conduct an interview, nonprivate setting, or cognitive impairment) not only limit the generalizability of findings but also cause difficulty in defining eligibility and the computation of response rates with potential adverse effects of the initial overcoverage of the target population. For example, individuals who died between sampling and recruitment are often considered non-eligible cases in younger samples. In contrast, deceased target individuals may be considered eligible and nonrespondents in a population segment characterized by increased age-associated risk of mortality.

Finally, a large body of evidence exists that describes potential differences between age groups regarding *the quality of information obtained from very old and oldest-old adults in standard survey data collection procedures* and suggests an adaptation to procedures and materials (Isakson et al., 2007). Some of these differences pertain to the specific historical background of the oldest old and differences in socialization, language, norms, education, or life experience (e.g., war) that may result in potential misfit with assessment procedures or concepts generally developed more recently with and for a younger audience (Chan et al., 1999). Examples of standard survey assessment where the current practice may not fully map onto the realities of very old individuals are the coding of educational or vocational backgrounds using the current ISCED standardization. Educational systems and access to higher education have changed over time, making it difficult to compare and classify the educational qualifications of today's oldest old. Similarly, assessing income or occupational prestige among mainly retired or institutionalized individuals is difficult and might ultimately lead to their exclusion from the analysis. Simultaneously, overcomplex assessments result in high item-nonresponse or information bias. Measures of social status are nonetheless indispensable for assessing social and health inequalities among oldest old individuals (Darin-Mattsson et al., 2017). Adapting standard instruments, e.g., inquiring about the last (former) occupation instead of the current employment status for coding social class, is a practical and straightforward solution (Grundy & Holt, 2001).

Other differences in responses to survey questions by very old individuals compared with younger age groups pertain to developmental change. Research into measurement invariance across the lifespan has just begun acknowledging that changes in how survey questionnaires are understood and responded to may offer a unique window into understanding intra-individual developmental processes (e.g., response shift, see Edwards & Wirth, 2009; Kaspar et al., 2018). Evidence for a developmental change in late life has been reported for key characteristics (e.g., cognition: Hülür et al., 2015, personality: Mueller et al., 2017, values: Borg, 2019; Reissmann et al., 2021). Measures of functional health, such as activities of daily living (ADL) or instrumental activities of daily living (IADL) scales exhibit some degree of dependency on age and context (e.g., living at home vs. nursing homes) at the item level (Fleishman et al., 2002; LaPlante, 2010).

However, the combined scales were rather stable over time, setting, and age, making ADL/IADL scales an important and comparable measure for functional health among the oldest old (Finlayson et al., 2005).

A third large group of concerns regarding the quality of survey responses obtained from very old individuals refers to compromised sensory, physical, or cognitive abilities more prevalent in this age segment (Schanze, 2021). On the most general level, challenges are reported regarding interview length, attention span, fatigue, or the need for breaks (Davies et al., 2010). Similarly, evidence shows that not all assessment modes are equally suited for very old adults, suggesting that face-to-face (f2f) interviews are the mode most responsive to challenges posed by sensory loss or limited cognitive capacity (Farmer & Macleod, 2011; Isakson et al., 2007). Concerning obtaining retrospective biographical information, the potential combination of a very rich and long biography and memory impairment may result in inconsistent data (El Haj et al., 2015). This may even be true within reasonable timeframes such as “the last 12 months.” These are extensively used in assessing activity, well-being, or social engagement for good reasons, including, but not limited to, expected seasonal variation, celebrations, or transient states such as short-term illness. Findings on the positive and negative affect schedule (PANAS) have consistently shown invariant measurement properties across different age groups and in very old individuals (Kercher, 1992; Mackinnon et al., 1999). However, a modified version of the PANAS that used dichotomized items reported low reliability of the scale in NHR with dementia (Gerritsen et al., 2007).

Nevertheless, the population of the very old includes a substantial number of individuals for whom adaptations of the f2f interview will not succeed in surpassing communication impairment due to substantial physical or cognitive decline. The group of individuals *unable to conduct the interview* (UCI) themselves for health reasons varies as a function of the expected interview burden (e.g., length and complexity). We argue that obtaining a proxy report for these cases may provide helpful information for many aspects of the life of the targeted individual and offer a window to explore (the lack of) specific knowledge (e.g., introspection) or motivation (e.g., response style) of the proxy informant as a substantive-matter research question (Vazire, 2010). Maybe even more importantly, it allows scholars to test empirically to what extent such data could validly be integrated into population estimates.

2.2 Research Aims

This study weighs the advantages and disadvantages of including three difficult-to-survey subgroups (i.e., institutionalized, UCI, and cognitively impaired individuals) in a representative survey on quality of life (QoL) in very old age. More specifically, we first evaluated the success of sampling and recruitment measures taken to better address NHR, individuals with varying degrees of cognitive impairment, and information on those no longer able to conduct the 90-min-long interview themselves (i.e., UCI via proxy interviews). We expected an improvement in the absolute number of realized interviews and response rates relative to studies that adhere to the same inclusive definition of the target population but refrain from going the extra mile to address hard-to-survey subgroups during sampling and recruitment. Moreover, we expected increased representativity of the realized sample, thus reducing the risk of nonresponse bias relative to such survey protocols that fail to specifically address these difficult-to-survey subgroups of older adults.

Next, we evaluated the extent to which integrating responses from these difficult-to-survey groups altered substantive-matter conclusions about socioeconomic resources, health resources, and well-being outcomes of the very old population. We expected that subject-matter responses for these groups might increase insecurity about population parameter estimates, particularly when options to establish measurement invariance (MI) or estimate response bias in subgroups (e.g., proxy interviews) were limited. However, from the viewpoint of the TSE framework, we assumed that the benefits of representing difficult-to-survey subgroups of very old adults in the sample (i.e., reduced nonresponse bias) could outweigh concerns about data quality (e.g., measurement bias from proxy interviews or inconsistent responses from cognitively impaired individuals) when including more vulnerable individuals.

3 Methods

3.1 Participants and Procedures

The data are from a study on QoL and well-being of very old adults conducted in Germany's most populous state, North Rhine-Westphalia (Wagner et al., 2018). A multistage sampling design was employed to define the sample. First, communities were drawn as PSUs based on the number of inhabitants aged 80 years or older (i.e., proportional-to-size selection). A total of 120 PSUs was selected, with large communities contributing multiple PSUs. Next, community offices drew random samples of 400 individuals (i.e., secondary sampling units) per PSU. From the resulting sampling frame of more than 48,000 individuals, a disproportional gross sample of 8,040 individuals was drawn that should result in approximately 1,800 realized interviews. Persons 85 years and older and men were oversampled to allow for robust subgroup analyses according to a priori power analyses. Details about the computation of survey weights are provided in Appendix A. A total of 1,863 computer-assisted personal interviews (CAPI) were conducted at participants' homes to assess a wide array of individual QoL resources (e.g., economic, health) and subjective QoL outcomes (e.g., well-being). The study protocol also included objective testing of handgrip strength and mild cognitive impairment. The mean age of the realized sample at the time of the interview was 87.0 years (*SD* = 4.5; range: 80.1 to 102.9

years). The ethical board of the medical faculty at the University of Cologne approved this study (Protocol #: 17–169).

3.2 Major Design Decisions

The expected length of an interview is a critical determinant of study participation. Results from a pilot study (Brix et al., 2016) indicated that interviews with very old individuals should not exceed 90 min on average. Potential proxy interviews for UCI were not limited a priori to specific groups of informants (e.g., partners and children) or specific content (such as "facts" or easily observable characteristics of the target persons) to prevent over-exclusiveness and test the limits of data collection in this population. However, informants have been explicitly instructed to choose "refuse" or "do not know" categories for questions they felt uncomfortable or unable to answer. Additionally, all interviews conducted with individuals willing to participate and capable of understanding and answering the questions are included in the study, irrespective of their screening test results for mild cognitive impairment and suspected dementia.

3.3 Fieldwork Metadata

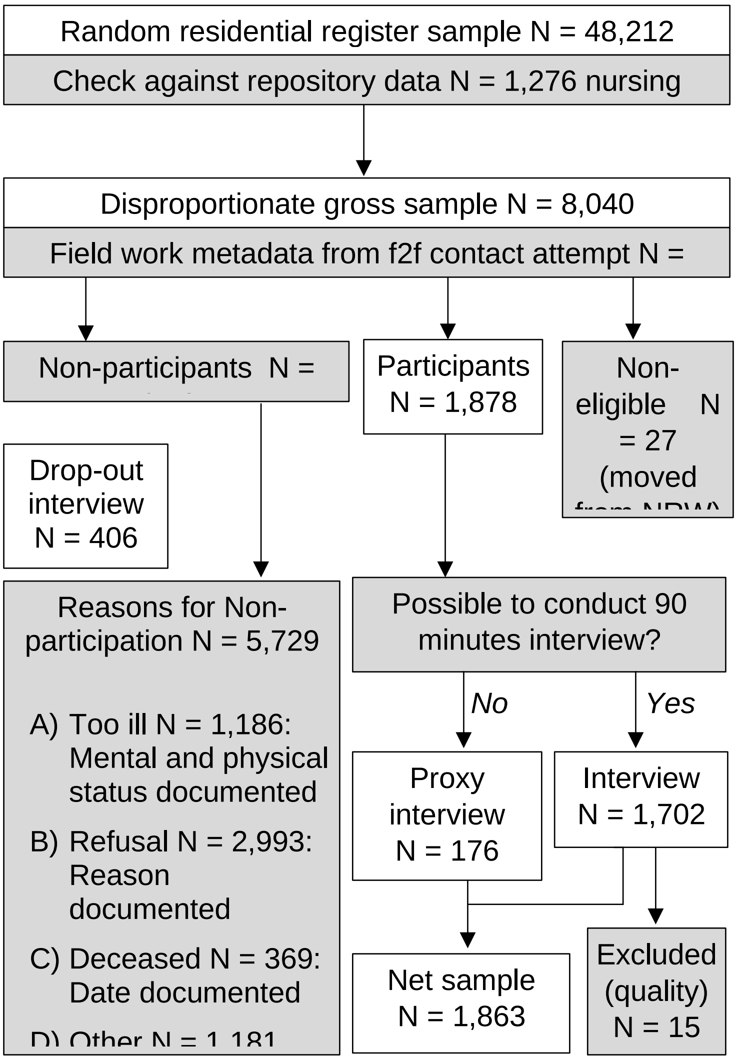
During the initial steps of obtaining the study sample and contacting potential respondents, metadata (e.g., communicative abilities, health status, living conditions) were generated that identified difficult-to-survey individuals, leading to tailored approaches and inclusion strategies (Figure 1).

Detailed reasons for nonparticipation were documented for 2,993 individuals who actively refused to participate. Only in cases when targeted individuals were too ill to conduct the 90-minute interview themselves (i.e., UCI) was an attempt made to conduct a proxy interview, and specific reasons have been documented if no proxy interview could be realized (N = 1,186).

Institutionalization

Random residential register samples from 94 selected communities were received between May and September 2017. In the beginning, data were collected by a commercial survey institute. However, population registration offices were hesitant to provide data on NHR to a commercial survey institute because NHR had just been granted additional protection by novel data protection legislation at that time. In response, the academic staff obtained privileged access to register data directly from the population registration offices under the applicable data protection laws. Sample quality control included comparing the most recent official register data and screening for clear deviations from random sampling (e.g., demographic structure, name- or streetwise selection). A comprehensive repository of 1,276 addresses of nursing homes and care facilities in NRW was used to identify individuals in the samples living in an institution. The share of identified NHR was compared with the community-level census data on the prevalence of institutionalization in the target population. If the data suggested that NHR were systematically excluded, a new and unrestricted random sample was requested.

Figure 1. Flowchart of potential study participants, fieldwork metadata, and interview data.



Cognitive status

Extended documentation of mental or physical health conditions was sought in case the target person was unable to conduct the interview him/herself due to health reasons (i.e., UCI).

Proxy interviews

Documentation of nonparticipation was also extended to include information on why no proxy interview could be conducted in UCI. More specifically, the lack of potential informants and the refusal of available informants were documented. In this study, we defined the difficult-to-survey subgroup of UCI as all individuals who were unable to conduct the interview him/herself due to health reasons and who therefore had to be included in the study via proxy interviews.

3.4 Data from Survey Interviews

Institutionalization

Respondents were asked to report whether they received formal or informal care. Additionally, interviewers rated the respondents' dwellings according to different categories of private, supported, or institutionalized housing. The time between drawing the register samples and interviewing was kept to a minimum (average 112 days, range 12 to 208 days). The community register information on nonprivate dwellings was validated before contacting potential respondents (see the section on fieldwork metadata). Nevertheless, some individuals initially classified as private-dwellings were finally interviewed in institutional settings. Details on the definition of the difficult-to-survey subgroup of very old individuals in nursing homes (NHR) in this study are reported in Appendix B.

Cognitive status

The DemTect has been developed as a brief screening tool for mild cognitive impairment and the early stages of dementia (Kalbe et al., 2004). It comprises subtests targeting immediate/delayed word recall, digit span, number transcoding, and verbal fluency. Favorable psychometric and diagnostic properties for identifying mild cognitive impairment (MCI) have been reported compared with alternative screening tools, such as the Mini-Mental State Examination (Kalbe et al., 2013).

Because most of the empirical evidence regarding the psychometric properties and clinical utility of the DemTect has been reported for clinical populations, a rigorous test of the tool for use in large-scale surveys in the general old age population has been conducted. In a comprehensive feasibility study, the DemTect was implemented in CAPI interviews conducted in a random register sample of participants (N = 291) aged 80 years or above who were living privately and in long-term care facilities. All interviewers participated in a half-day test administration and scoring training. During the assessment, 49 participants were unable to answer the number transcoding task for reasons unrelated to cognitive problems (i.e., problems with reading/vision or writing 16.8%). In these cases, subtest scores were estimated based on the available partial responses in this subtest and performance in all other subtests to minimize the punishment for noncognitive impairment of test-takers. We used age-specific score transformation and derived an ordinal diagnostic group variable (i.e., age-adequate function, MCI, early dementia) based on the cut-off values suggested for persons aged 80 and over (Kessler et al., 2014). In the feasibility study, 26.8% of participants declined to take the test or some subtests, precluding the interpretation of their diagnostic group membership. The DemTect has been evaluated multiple times (Kalbe et al., 2004; Kessler et al., 2010; Perneczky, 2003), and a joint estimator using a Reitsma function (Doebler, 2020) yields a sensitivity of 94.1% and specificity of 91.4%. Based on the published age- and sex-specific prevalence rates

(Doblhammer et al., 2012), 19.5% of the feasibility study sample can be expected to have dementia. A comparison with the DemTect classification yielded a positive predictive value of 72.5% and a negative predictive value of 98.5% of the screening.

In the survey reported here, an identical procedure was followed for the test administration and scoring in interviews with respondents. In the case of proxy interviews, cognitive status was reported with the Global Deterioration Scale (GDS, Reisberg et al., 1982) in seven stages, from 1 (*no cognitive impairment*) to 7 (*most severe*). More recently, Reisberg and colleagues aligned GDS stage 3 to correspond to a clinical presentation of MCI (Reisberg et al., 2011).

For the current analysis, all respondents whose cognitive function was labelled as MCI or dementia according to the diagnostic screening were considered members of the difficult-to-survey subgroup of very old adults with cognitive impairment.

Socioeconomic status

The International Socioeconomic Index of Occupational Status (ISEI-08, Ganzeboom et al., 1992; Ganzeboom & Treiman, 2003) was used to measure socioeconomic status (SES). The ISEI-08 is a metric measure (range 10–90) building mainly on the occupation of the respondents according to the International Standard Classification of Occupations (ISCO-08, Ganzeboom, 2010). The measure also considers the required educational qualification. Given that most participants were retired, we took the last occupation as a basis for the ISEI-08. In case the targeted individual reported no previous occupation, the last occupation of the partner was used instead. ISEI-08 scores in this sample varied between 10 and 89.

Independence in activities of daily living

Adults' self-reported performance on the basic ADL (Katz et al., 1963) and IADL (Fillenbaum, 1988; Lawton & Brody, 1969) was used to measure everyday functioning. In this paper, we use five items representing ADL (e.g., getting dressed, walking) and seven items representing IADL (e.g., preparing meals, handling finances) with response options 0 (*impossible without help*), 1 (*some help needed*), or 2 (*no help needed*). The reliability of the ADL and IADL scales in the current sample was high (MacDonald's ω = 0.92 and 0.93, respectively).

Subjective well-being

The positive affect subscale of the short form of the PANAS (Kercher, 1992) was used as an indicator of subjective well-being (SWB). The frequency of positive affective states (e.g., "enthusiastic," "excited") across the past 12 months was reported on a 5-point response scale from 1 (*never*) to 5 (*very often*). Favorable psychometric properties of this very brief instrument have been reported for age-diverse and old-old samples (Hilleras et al., 1998; Kercher, 1992; Mackinnon et al., 1999). The scale consistency in our sample was satisfying (MacDonald's ω = 0.88).

3.5 Plan of Analysis

We adopted the TSE framework to discuss advantages and disadvantages of including three groups of particularly difficult-to-survey individuals in a representative survey of QoL in the very old. First, we used sampling and fieldwork metadata (e.g., share of nursing home addresses, reasons for nonparticipation) to evaluate the success of tailored strategies to include NHR, individuals with (beginning) cognitive impairment, and UCI and to estimate their effect on representing the population 80 years or older. Here, based on random samples of individuals from 94 communities, we investigated the potential threat of undercoverage of the institutionalized population in the sampling frame due to legal restrictions on data access and the lack of information on nonprivate dwellings in the register data. Based on fieldwork metadata, we compared the reasons for the refusal of residents in private dwellings versus institutional settings and the reasons why no proxy interview could be conducted for most UCI. We then used response rates and representativity (R) indicators (Schouten et al., 2009) to estimate differences in nonresponse bias resulting from the exclusion of specific subgroups of respondents. The response rate was defined based on the AAPOR definition of RR1 to include all eligible cases (AAPOR, 2016; Wagner et al., 2019). R-indicators estimate the representativity of a sample by predicting the nonresponse propensity by auxiliary variables known for the whole population (Schouten et al., 2009). Hence, they represent a measure of the deviation from weak representativity regarding theoretically meaningful threats to survey participation. The population R-indicator is defined as 1 minus two times the standard deviation of the response probability in the population ρ and ranges from 0 to 1, with higher values indicating higher representativity. Because response probabilities are typically unknown, they need to be estimated by regressing observed (non)responses on auxiliary variables. In this study, we used age, sex, community size, and NHR status as proxies for selection mechanisms (e.g., health and social status restrictions). The design-weighted (i.e., s_i/π_i) response-based R-indicator is defined by Schouten et al. (2009) as follows:

$$\hat{R}(\rho) = 1 - 2\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\frac{s_i}{\pi_i}(\hat{\rho}_i - \hat{\bar{\rho}})^2} \quad . \qquad (1)$$

We use the coefficient of variation of response propensities (CV) as a measure of the maximum potential nonresponse bias. The corresponding estimate acknowledges that nonresponse bias is a function of both the relative share of individuals participating in the study (i.e., response rate RR1) and the degree to which the realized sample is weakly representative of the target population:

$$CV(X) = \frac{1-R(X)}{2RR1} \qquad (2).$$

The CV represents a conservative worst-case estimate of nonresponse bias, assuming that nonresponse correlates maximally with the selected auxiliary variable X (de Heij et al., 2015).

In the second step, substantive-matter responses from the realized sample of 1,863 individuals (of which NHR: 211, screened cognitively impaired: 504, and UCI: 176) were considered to estimate potential bias using a multi-group approach to measurement invariance and mode effects. In this study, we restricted our focus to socioeconomic status, independence in ADL/IADL, and SWB as prominent markers of welfare and well-being research characterized using distinct approaches to their measurement. To empirically estimate potential bias introduced by including responses from proxy reports, we used a multi-group factor analytic (MGCFA) approach to test MI of multi-item scales (i.e., PANAS and ADL/IADL) as suggested by Meredith (1993). Cross-group equality constraints regarding factor loadings or/and item intercepts were used to test for metric and scalar MI, respectively. Violations of scalar MI may introduce systematic response bias into the estimating population means unless such group-specific deviations are adequately addressed in estimating factor scores. If the observed or latent group means can be validly compared, bias resulting from excluding proxy information can be directly estimated. The benefit of including difficult-to-survey subgroups of very old individuals was estimated relative to their contribution to representing the target population and the potential nuisance introduced to the assessment of key QoL indicators.

Analyses were performed using SAS 9.4 (SAS Institute, Cary, NC), Mplus 8.6 (Muthén & Muthén 1998–2021), and the *RISQ 2.1 package* (de Heij et al., 2015) for R software (R Core Team, 2017).

4 Results

4.1 Including the Institutionalized Population

Coverage in the sampling frame

In the random register samples, the proportion of NHR identified was almost consistently smaller than expected from the best available census data (Table 1). The overall percentage of NHR in the gross sample obtained from community registers was 8.1%, with 12.0% expected from the census data. However, even in communities that explicitly excluded individuals flagged as NHR in their databases, the obtained samples included 62.8% of all institutionalized individuals expected from the census data. Most remarkably, even in communities that provided ample evidence for using the full sampling frame, we noted a discrepancy of -2.3 percentage points, most likely attributable to a broader definition of institutional settings in census data compared with institutions listed in the nursing home register. Hence, the margin of uncertainty that arises from diverging definitions of the institutionalized population and procedures to identify NHR in the register data is even greater than the estimated degree of potential undercoverage remaining without it (-1.6 percentage points).

Table 1. Undercoverage of the institutionalized population in population register samples.

	Communities with explicit				All Communities			
Origin of	inclusion		exclusion					
information	N[1]	% NHR	N[1]	% NHR	N[1]	N individuals	N NHR	% NHR

Census	41	12.3%	13	12.1%	88	560,221	67,242	12.0%
Sample	41	10.0%	13	7.6%	88	45,809	3,701	8.1%
Difference		-2.3%		-4.5%				-3.9%

Note. [1]Census data are not available for six communities. NHR = nursing home residents.

Survey participation

Based on data from a feasibility study, mechanisms of survey nonresponse have been reported to be different for individuals living in private dwellings versus institutional settings (Wagner et al., 2019). Particularly, potential respondents in nursing homes were less easily accessible (i.e., lower contact rates) due to health impairments. However, once contacted, they show much lower refusal rates than individuals in private dwellings. In this study, reasons for refusing to participate were documented for 2,993 individuals (Figure 1). The main reasons for these were similar in private dwellings and NHR (Table 2). No interest (more than two out of three) and a general refusal to participate in surveys (one out of four or five) were the most common reasons for both groups of potential participants. In NHR, the length of interviews (8.4% vs. 6.8%) and inability to comment on the survey's subject (6.0% vs. 4.4%) were more frequent reasons for refusal than they were for private-dwelling individuals.

Table 2. Reasons for individuals living in private housing or institutions refusing to participate.

	% Total (n = 2,993)	% Private (n = 2,861)	% NHR (n = 132)
Not interested[1n]	69.4	69.6	63.5
General refusal	25.3	25.6	18.8
Not allowed by others	8.2	8.2	8.4
Interview too long	6.8	6.8	8.4
Subject too intimate	7.4	7.4	5.8

Lack of knowledge	4.4	4.4	6.0
Privacy concerns	3.7	3.7	3.4
No time	3.6	3.7	1.5
Participated too often in surveys	0.5	0.4	2.1

Note. Design-weighted data. [1]Multiple reasons for refusal to participate could be given. NHR = nursing home residents.

In the group of participants, the living status may have changed since the sample generation. A total of 151 addresses of participants had been flagged as institutions based on the residential register or nursing home repository data before fieldwork. Based on all available information related to dwellings at the time of the interview, 211 participants were identified as NHR. In some cases, information from different sources was contradictory. For example, 13 participants listed as private-dwelling according to the residential register lived in a nursing home at the time of first contact or at the time of the interview.

4.2 Including Those Unable to Conduct the Interview Via a Proxy Interview

Of all 8,040 individuals contacted, 1,362 were UCI themselves for health reasons (Figure 1). Of these, it was possible to conduct 176 proxy interviews. The design-weighted gross sample estimated that this corresponded to 15.1% of the noncommunicative, very old population that could, in principle, be included in this measure. For 38.4% of the remaining cases, the inability to provide informed consent due to mental illness precluded a proxy interview. In cases where informed consent could have been obtained, the most frequent reasons for not conducting a proxy interview were proxy refusal (24.0%), the target person opposing a proxy interview (23.7%), and no identifiable proxy (12.6%). However, in many cases, the reasons why the proxy interview could not be realized remain unknown (35.6%). The number of contacts needed to obtain a proxy interview in this population was estimated to be significantly higher than for self-reports (M = 3.0, 95% CI = [2.6 to 3.5] vs. M = 2.4, 95% CI = [2.3 to 2.5]).

About half (46.2%) of the interviews conducted with a proxy informant were with the children of the target persons, mostly daughters (64.0%). About one in five proxy interviews was conducted with spouses (18.7%). Proxy informants were predominantly female (70.2%). The percentage of self-reports decreased across age groups (Table 3). However, a similar number of proxy interviews were conducted in all three age groups. UCI were more often living in institutions, female, and less healthy than self-reporting target persons. For example, UCI were treated for significantly more health conditions. The possibility of conducting proxy interviews appears particularly relevant for representing NHR, as their share is four times higher in proxy reports than in self-reports (43.3% versus 11.0%). In line with our expectation, the availability of information differed between self- and proxy reports for different reasons. Proxy informants, for example, reported more often that they did not know the answer to a question than respondents in self-reports. However, refusal rates regarding specific questions were substantially higher in self-reports than in proxy interviews. Despite these plausible limitations, information on more than 90% of all questions posed is available to investigate the QoL of UCI.

4.3 Including Individuals with Cognitive Impairment in the Survey

Regarding the cognitive status of the participants, 3.7% of respondents (self-reports) declined to participate in the cognitive screening test, and 14.3% of participants did not complete all DemTect subtests. GDS-Scale scores were available for 96.1% of individuals represented by proxy interviews. Few proxy informants answered “do not know” or refused to rate the cognitive status of the target person (1.2% and 2.8%, respectively). Overall, cognitive status could be assessed and interpreted for 83.2% of participants. The assessment of cognition was more likely to be available in the youngest age group (85.5%, $\chi^2 = 7.03$, df = 2, $p = .030$) and in proxy interviews (96.1%, $\chi^2 = 9.09$, df = 1, $p = .003$). In contrast, the availability of information on cognitive status was independent of gender ($\chi^2 = 1.90$, df = 1, $p = .168$) and living situation ($\chi^2 = 1.28$, df = 1, $p = .258$).

The cognitive function of most individuals capable of self-report was screened as age-adequate (73.7%), whereas 16.4% had MCI and 9.8% had suspected dementia (Table 3). In UCI, however, the most common cognitive status was suspected dementia (73.5%), 7.4% were classified as MCI, and 19.1% had age-adequate cognitive functioning. Thus, in line with our expectations, including proxies as informants led to a better representation of people with dementia in the data. Overall, estimates from the total sample yielded a prevalence of 16.3% for dementia and 15.5% for MCI in the population 80 years or older.

Table 3. Characteristics of respondents able versus unable to answer the interview themselves.

	Self-report (N = 1,687) % or M [95% CI]	Proxy report (N = 176) % or M [95% CI]	Test[1]
Gender			
- Male	37.0%	28.3%	χ^2=3.63, df=1,
- Female	63.0%	71.7%	p = .057
Age group			
- 80–84 yrs	56.3%	33.0%	χ^2=44.90, df=2,
- 85–89 yrs	30.6%	32.4%	p < .001
- 90 yrs or older	13.1%	34.6%	
Setting			
- Private dwelling	89.0%	56.7%	χ^2=66.86, df=1,
- Institution	11.0%	43.3%	p < .001

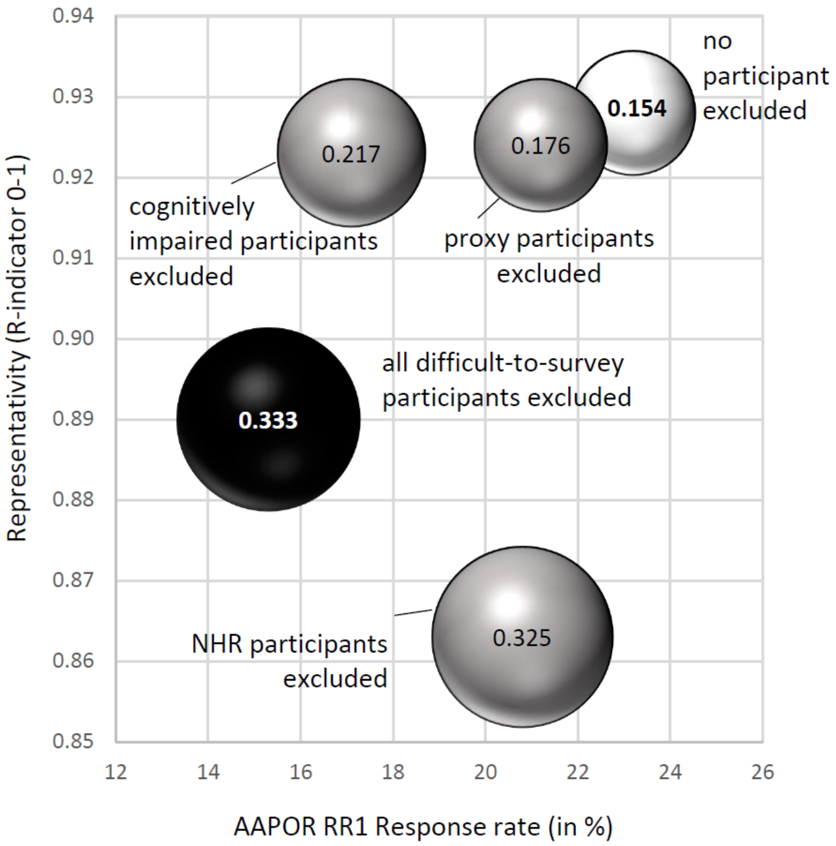
Treated health conditions (0–19)[2]	3.4 [3.2 – 3.6]	4.2 [3.7 – 4.8]	χ^2=11.71, df=1, *p* < .001
Cognitive status[3]			
- Age-adequate function	73.7%	19.1%	χ^2=153.55, df=2, *p* < .001
- MCI	16.4%	7.4%	
- Suspected dementia	9.8%	73.5%	
Response			
- Do not know (%)	1.5 [1.3 – 1.8]	5.8 [4.7 – 6.9]	F=58.1, df=1/119, *p* < .001
- Refuse to answer (%)	2.0 [1.6 – 2.4]	0.4 [0.2 – 0.7]	F=40.8 df=1/119 *p* < .001

Note. Calibration-weighted data. [1]Wald test for predicting proxy interview by participant characteristics or F-test for predicting levels of "do not know" answers and refusals by interview type. All tests consider the complex sampling design and use unbiased standard errors (Taylor linearization). [2]A list of 19 currently treated health conditions (e.g., heart disease, hypertension, respiratory or lung disease, diabetes). [3]A classification based on DemTect scores from self-reports and GDS ratings from proxy reports. 95% CI = 95% confidence interval.

4.4 Consequences for Representativity and Unit-Nonresponse-Bias

Failure to allow for proxy interviews in case the target persons were willing to participate but unable to conduct the interview themselves would have "lost" 176 participants and led to a decrease of 2.2 percentage points in the response rate. Similarly, the inclusion of the 211 individuals interviewed in nonprivate settings accounted for a 2.7 percentage point increase in the response rate. Restricting the sample to the 1,359 screened participants with uncompromised cognitive function (i.e., using the MCI classification as an exclusion criterion) would have resulted in a 6.3 percentage point drop in the response rate. The measures taken in this study to assure the best-possible inclusion of these difficult-to-survey groups increased the response rate from 15.3% to 23.2%.

Figure 2. Maximum nonresponse bias (CV) resulting from the failure to include difficult-to-survey groups of very old individuals. CV = Coefficient of Variation. Bubble size is proportional to the size of the CV. Bubble location represents a unique combination of the response rate and representativity of the retained sample.



Although an increase in survey participation represents an important step toward unbiased population estimates, a higher response rate is insufficient. It should be accompanied by increased representativity of the realized sample to reduce survey nonresponse bias efficiently. The analysis of response propensities for the sample showed only a minor predictive value of age, gender, living in an institution, or regional characteristics (BIK). Thus, a resulting R indicator of 0.928 with a 95% confidence interval [0.908 to 0.948] underscored the high representativity of the realized sample regarding these potential threats to participation. Additionally, an estimate of 0.154 [0.112 to 0.196] for the coefficient of variation suggested that even under a worst-case scenario regarding the characteristics of very old individuals not included in the study, population estimates will not exhibit more than a maximum of 15.4 percent nonresponse bias (Figure 2). In contrast, the failure to include the 640 individuals who showed signs of cognitive decline, lived in institutions, or were unable to conduct the interview themselves for health reasons would have resulted in a lower response rate, significantly worse representativity of the realized sample (R = 0.89, 95% CI = [0.881 to 0.905]), and more than twice the insecurity regarding the maximum potential nonresponse bias (CV = 0.333, 95% CI = [0.296 to 0.370]) in population parameter estimations.

4.5 Consequences for Substantive-Matter Conclusions

Population estimates for socioeconomic and health resources and SWB showed substantial differences across difficult-to-survey subgroups of very old individuals (Table 4). Hence, not including specific subpopulations may lead to systematic bias in estimates for the population 80 years or older.

Table 4. (Sub-)Population estimates of QoL resources and outcomes in very old age.

	Socioeconomic Status (ISEI, 10–90) M [95% CI][1]	Functional health (I/ADL, 0–2) M [95% CI] [1]	Well-being (PANAS, 1–5) M [95% CI] [1]
Full sample	40.84 [39.38–42.30]	1.54 [1.50–1.58]	3.26 [3.16–3.34]
Setting			
- Private dwelling	41.41 [39.88–42.94]	1.65 [1.61–1.68]	3.31 [3.17–3.34]
- Institution	37.08 [33.82–40.33]	0.88 [0.80–0.96]	2.91 [2.73–3.08]

Cognitive status			
- Age-adequate	44.10 [42.28–45.93]	1.73 [1.70–1.76]	3.41 [3.32–3.50]
- MCI or suspected dementia	35.38 [32.89–37.86]	1.20 [1.11–1.28]	2.94 [2.80–3.08]
Mode of survey participation			
- Self-report	41.31 [39.78–42.83]	1.64 [1.61–1.67]	3.32 [3.22–3.41]
- Proxy report	35.89 [32.48–39.31]	0.52 [0.44–0.60]	2.60 [2.42–2.78]

Note. Calibration-weighted data. [1]95% confidence intervals are based on standard errors that consider the complex sampling design (Taylor linearization).

Socioeconomic status

In difficult-to-survey subgroups, ISEI-08 scores could not be computed for 5.6% of cognitively impaired individuals, 6.0% of those represented by proxy interviews, and 8.2% of those living in institutions. In contrast, missing rates for SES ranged between 2.1% and 3.1% in the remaining participants. Estimates for SES were substantially lower in all three subgroups (Table 4). In individuals with limited cognitive functioning and UCI, ISEI-08 scores were significantly (as judged from nonoverlapping confidence intervals) lower than those in the respective self-report and age-adequate cognition groups were. Hence, a substantial bias suggesting better socioeconomic resources for the very old population would result from failing to include these difficult-to-survey groups.

ADL/IADL independence

The data availability regarding functional status was high (92.8% to 100%) in both self- and proxy reports. Interestingly, more “do not know” answers were encountered with more complex tasks of everyday living (e.g., housework) in self-reports and private settings. This suggests that sharing chores was more common, leaving respondents unsure if they could perform this task when needed. Unsurprisingly, “do not know” answers were more common in individuals screened as cognitively impaired, although missing rates did not exceed 2.6% for any item in this subgroup.

The estimates for functional health were reported to be significantly lower in UCI than in interviewed participants (0.50 versus 1.6, Table 4). However, loglikelihood ratio tests for MI showed that proxy- and self-reports were not comparable already at the level of metric invariance (Δ-2LL$_{corr}$ = 22.05, Δdf = 2, p <.001). Therefore, the degree of potential response bias resulting from proxy informants answering items differently cannot be estimated empirically. The comparison of group means and even the inclusion of UCI in covariance structure analysis are severely limited.

Subjective well-being

The availability of information on affective states was comparable in self-reports and UCI participants on all five items and ranged between 91.7% (“alert”) and 96.6% (“enthusiastic”). However, there was more item-level nonresponse due to refusal in self-reports (1.3% to 1.4%) compared with less than four per thousand for all items in proxy reports. Refusal rates were slightly higher on average in NHR (1.8%) and those with beginning cognitive impairment (1.0%). However, “do not know” answers were responsible for most of the item-nonresponse in all respondents.

The tests for metric and scalar MI found no evidence of a) a different conceptual understanding of positive affect in self- compared with proxy reports or b) systematic differences between groups in the interpretation of the 5-point response scale, respectively (Table 5). Thus, full scalar MI was established for the PANAS scale, allowing for uncompromised integration of information on UCI in population estimates. The level of positive affect was significantly lower in proxy interviews than in self-reports (M = 2.60 and M = 3.32, respectively; Table 4). Thus, failure to include information on UCI would have led to a positively biased estimate of affective well-being in the 80+ population. However, a mean population value of 3.26 [3.16 – 3.34] was estimated for positive affect in the combined sample, suggesting that the magnitude of bias due to the exclusion of UCI would be minor (i.e., 0.06 scale points). Similar risks apply when failing to include individuals living in institutions and those with (beginning) cognitive impairment, as levels of well-being were also significantly lower in these subgroups. Although proxy interviews appear to hold merit in representing individuals with pronounced cognitive decline, questions may linger regarding the validity of information collected via self-reports in less cognitively impaired individuals (e.g., MCI or suspected dementia). However, the "known truth," against which responses might unequivocally be validated, is generally limited in most social science surveys. In this sample, discrepancies between reported and registered birth dates were observed to a similar degree in those with age-adequate cognitive function compared with those classified MCI or suspected dementia (6.0% and 5.8%, respectively).

Table 5. Nested model comparison of configural, metric, and scalar MI for the PANAS scale in self- and proxy report subgroups.

Degree of invariance (equality constraint)	Absolute model fit	Relative fit indices	LR difference test[1]
Configural MI (factor model)	χ^2 = 79.0, df = 10, *p* < .001	RMSEA [CI90] = 0.087 [0.070–0.106], *p* < .001	
Metric MI (loadings)	χ^2 = 91.4, df = 14, *p* < .001	RMSEA [CI90] = 0.078 [0.063–0.094], *p* = .001	Δ-2LL$_{corr}$ = 2.44, Δdf = 4, *p* = .656
Scalar MI (loadings and intercepts)	χ^2 = 103.2, df = 18, *p* < .001	RMSEA [CI90] = 0.072 [0.059–0.086], *p* = .003	Δ-2LL$_{corr}$ = 13.00, Δdf = 8, *p* = .112

Note. [1]Likelihood-ratio model comparison for robust maximum likelihood estimation using scaling correction factors.

5 Discussion

This study drew on metadata from sampling and fieldwork. It also used the responses from a large-scale survey on QoL to weigh the advantages and disadvantages of targeting and including difficult-to-survey subgroups of very old individuals. After comprehensive piloting, a study protocol was employed to maximize the inclusion of those with limited cognitive function, NHR, and UCI to optimize the representativity of the data and the precision of estimates. The results show that the failure to include these individuals would have resulted in much lower response rates and worse representativity. Hence, a potential maximum nonresponse bias is more than two times the size achieved with the full sample. Increasing response rates may be critical to guarantee that there is enough power for subgroup comparison, given the very uneven proportion of men and women or different birth cohorts (i.e., oldest-old) in this population segment and the accumulated risk to survey participation due to communication or care needs. Results regarding NHR showed that the omission of a relatively small difficult-to-survey group had minor adverse effects on response rates but a substantial impact on potential nonresponse bias. While this finding may be surprising upfront, it is a direct consequence of the fact that this characteristic was deemed a particularly important aspect of representativity analysis and has been explicitly considered a predictor of unit nonresponse in this study. Because of the scarcity of characteristics of the oldest old available from the register data for representativity analysis, our promising results regarding the identification of NHR in community register data, and the expectation that NHR status serves as a good proxy for threats to participation because of gatekeepers and poor health status, we think this emphasis is well-justified.

The findings of this study underscore the need to include difficult-to-survey subgroups, such as the institutionalized and communication impaired, to represent the population of individuals aged 80 years or older. Undercoverage of NHR in random register samples may be avoided using privileged data access of public-law institutions (e.g., universities) and dedicated quality control of received samples. This assertion is confirmed by more recent experiences from the second wave of the study. Here, all register samples were obtained directly by the university, and the proportion of identified nursing home addresses was even closer to the expected figure. According to the census data, the percentage of NHR in these communities was an expected 11.7%, while 8.6% were identified in the sample. Acknowledging that a difference of about 2.3 percentage points could be attributed to a broader definition of this group in the census, only a minor potential undercoverage of -0.8 percentage points results from the sampling procedure employed. Thus, we expect that alternative strategies, such as dual-frame sampling, may be challenged by considerable uncertainty in defining the institutionalized subpopulation and a loss of efficiency from overlapping target populations in the nursing home and community register samples.

The cognitive status could be measured for 83.2% of the sample and resulted in a prevalence estimation of 163 per thousand, a slight underestimation of the actual dementia prevalence of 194 per thousand known from the health insurance data. This prevalence is also achieved by including persons with dementia (PWD) via proxy informants. Studies not including UCI via proxy interviews, such as the SHARE study, report an even lower prevalence, of 104 per thousand, in an even older (85+ years) population (Ferreira et al., 2020). Although proxy interviews for UCI come with difficulties in the comparability of measurement, they represent a keystone for the representation of PWD in survey samples.

Our results also showed that substantive-matter (here: socioeconomic status, functional health, and SWB) research could profit from including difficult-to-survey subgroups in multiple ways. First, we presented evidence that information on the well-being of UCI can validly be integrated, allowing for a less biased population estimate where subgroup levels differ from those generally observed in less inclusive studies. Second, differences observed in difficult-to-survey subgroups regarding socioeconomic status or independence in ADL/IADL agree with findings from studies targeting more specific subpopulations or studies with similar inclusive approaches (Kelfve et al., 2013). The identified tendency for lower socioeconomic status in these groups underlines the importance of including them in research on social and health inequalities.

5.1 Limitations and Future Directions

Some limitations of the current study warrant attention. First, while consulting both interview data and fieldwork metadata may ultimately result in the best-possible characterization of participants' situation at the time of the interview (e.g., concerning NHR status) and benefits data interpretation, such updates may not be easily or timely integrated into the computation of survey weights. In this study, using updated information on NHR status instead of outdated residential register information may have contributed to more efficient calibration weights. Second, the assessment of cognitive function in UCI relied on proxy GDS ratings instead of objective testing. We acknowledge that although some degree of age-adequate functioning and MCI may be compared across instruments, this option may not extend to more severe levels of cognitive impairment (i.e., higher GDS scores). Third, and more conceptually, the usefulness of the R indicator approach critically depends on selecting relevant predictors of participation. Considering institutionalization as a proxy for adverse conditions in very old age may allow for a rather strict test of the risks of noninclusion of function-impaired groups to the representativity of our old age sample. When exploring survey error related to the exclusion of other difficult-to-survey subgroups (e.g., migrants), different predictors of response propensity (e.g., language proficiency) may be deemed relevant. In this situation, identifying bias (e.g., cultural bias) in survey responses using MI testing may be warranted. Therefore, we expect that this example of considering both sampling errors and measurement errors related to the inclusion of difficult-to-survey older adults will be useful beyond research on aging. Finally, the measures taken to reduce unit nonresponse in the difficult-to-survey subpopulations studied here relied on structural resources (e.g., detailed structural data available from registers, communities able/willing to draw samples, institutions granting access to NHR) or social resources such as the availability of proxy informants. Because such resources are certainly not randomly distributed in the population of 80+, generalization of findings concerning the characteristics of those living in institutions, those unable to answer the interview, and those with impaired cognitive functioning remain limited. Similarly, although the difficult-to-survey groups of very old adults studied here are very likely to show lower survey participation also in other countries, their prevalence and characteristics may diverge, and different structural (e.g., digital infrastructure, data protection law, see Scherpenzeel, 2017 for an overview of sampling frames for European social surveys) as well as social resources (e.g., caring relatives) may be available for inclusive survey research.

5.2 Conclusion

Taken together, this study provided ample evidence that the benefits of tailored measures to reduce unit nonresponse in difficult-to-survey groups and the integration of responses obtained from NHR, the cognitively impaired, and for UCI via proxy interviews may outweigh the undisputed challenges along this road. However, evaluations should be increasingly based on more substantiated research hypotheses that may well go beyond the estimation of population point parameters focused on in this study and consider bias in the covariance structure.

The strong discrepancy in reported levels of I/ADL independence between self- and proxy reports, with evidence of the lack of MI across informants, warrants further investigation. The authors hope that the availability of such information may help shed light on mechanisms that could ultimately promote independence in daily living, particularly for those very old individuals who already receive support and assistance or those less aware of existing threats to their independence.

Acknowledgment

The authors declare that they have no conflicts of interest. Survey materials and a de-identified dataset of survey responses are available as scientific use files at the GESIS – Leibniz Institute for the Social Sciences data repository (Zank et al., 2020). The de-identified metadata and analytical code used in this study are available at the GESIS – Leibniz Institute for the Social Sciences data repository (Kaspar et al., 2023). The project is part of the Key Research Area "Aging and Demographic Change" at the Cologne Center for Ethics, Rights, Economics, and Social Sciences of Health (ceres) at the University of Cologne. The members of the project board are Michael Wagner and Susanne Zank. This work was supported by a grant from the Ministry of Culture and Science of the German State of North Rhine-Westphalia (323-8.03-125240).

References

Andersen-Ranberg, K., Petersen, I., Robine, J-M., & Christensen, K. (2005). Who Are the Oldest-Old? In A. Börsch-Supan, A. Brugiavini, H. Jürges, J. Mackenbach, J. Siegrist,

& G. Weber (Eds.), *Health, Ageing and Retirement in Europe: First Results from the Survey of Health, Ageing and Retirement in Europe* (pp. 35-40). Mannheim Research Institute for the Economics of Aging.

Aust, V., & Gilberg, R. (2016). *Überlegungen zur Gemeindestichprobe der geplanten Erhebung NRW80+: Kurzbericht*. infas Institut für Sozialforschung.

Borg, I. (2019). Age- and gender-related differences in the structure and the meaning of personal values. *Personality and Individual Differences*, *138,* 336–343. https://doi.org/10.1016/j.paid.2018.10.013

Brix, J., Steinacker, G., Schneekloth, U., & Stadler, M. (2016). *Machbarkeitsstudie NRW80+: Expertise zur Erreichbarkeit einer Personenstichprobe Hochaltriger*. München. TNS Infratest Sozialforschung.

Chan, K.-Y., Drasgow, F., & Sawin, L. L. (1999). What is the shelf life of a test? The effect of time on the psychometrics of a cognitive ability test battery. *Journal of Applied Psychology, 84*(4), 610–619. https://doi.org/10.1037/0021-9010.84.4.610

Darin-Mattsson, A., Fors, S., & Kåreholt, I. (2017). Different indicators of socioeconomic status and their relative importance as determinants of health in old age. *International Journal for Equity in Health, 16*(1). https://doi.org/10.1186/s12939-017-0670-3

Davies, K., Collerton, J. C., Jagger, C., Bond, J., Barker, S. A. H., Edwards, J., Hughes, J., Hunt, J. M., & Robinson, L. (2010). Engaging the oldest old in research: Lessons from the Newcastle 85+ study. *BMC Geriatrics*, *10,* 64. https://doi.org/10.1186/1471-2318-10-64

Deeg, D. J., van Tilburg, T., Smit, J. H., & Leeuw, E. D. de (2002). Attrition in the Longitudinal Aging Study Amsterdam. *Journal of Clinical Epidemiology, 55*(4), 319–328. https://doi.org/10.1016/S0895-4356(01)00475-9

Degnen, C. (2007). Minding the gap: The construction of old age and oldness amongst peers. *Journal of Aging Studies*, *21*(1), 69–80. https://doi.org/10.1016/j.jaging.2006.02.001

De Heij, V., Schouten, B., & Shlomo, N. (2015). *RISQ manual 2.1 Tools in SAS and R for the computation of R-indicators, partial R-indicators and partial coefficients of variation*. http://hummedia.manchester.ac.uk/institutes/cmist/risq/RISQ-manual-v21.pdf

Doblhammer, G., Schulz, A., Steinberg, J., & Ziegler, U. (2012). *Demografie der Demenz* (1. Aufl.). *Programmbereich Gesundheit*. Huber. http://www.socialnet.de/rezensionen/isbn.php?isbn=978-3-456-85050-4

Doebler, P. (2020). *mada: Meta-analysis of diagnostic accuracy* (Version R package version 0.5.10) [Computer software]. https://CRAN.R-project.org/package=mada

Edwards, M. C., & Wirth, R. J. (2009). Measurement and the Study of Change. *Research in Human Development*, *6*(2-3), 74–96. https://doi.org/10.1080/15427600902911163

El Haj, M., Antoine, P., Nandrino, J. L. & Kapogiannis, D. (2015). Autobiographical memory decline in Alzheimer's disease, a theoretical and clinical overview. *Ageing Research Reviews, 23*(Pt B), 183–192. https://doi.org/10.1016/j.arr.2015.07.001

Farmer, M., & Macleod, F. (2011). *Involving disabled people in social research: Guidance by the Office for Disability Issues*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/321254/involving-disabled-people-in-social-research.pdf

Ferreira, R. G., Brandão, M. P., & Cardoso, M. F. (2020). An update of the profile of older adults with dementia in Europe: Findings from SHARE. *Aging & Mental Health*, *24*(3), 374–381. https://doi.org/10.1080/13607863.2018.1531385

Feskens, R. C. W. (2009). *Difficult Groups in Survey Research and the Development of Tailor-Made Approach Strategies* [Dissertation]. Utrecht, Utrecht University.

Fillenbaum, G. G. (1988). *Multidimensional Functional Assessment of older Adults: The Duke Older Americans Resources and Services Procedures*. Duke University Center for the Study of Aging and Human Development.

Finlayson, M., Mallinson, T., & Barbosa, V. M. (2005). Activities of daily living (ADL) and instrumental activities of daily living (IADL) items were stable over time in a longitudinal study on aging. *Journal of Clinical Epidemiology*, *58*(4), 338–349. https://doi.org/10.1016/j.jclinepi.2004.10.008

Fleishman, J. A., Spector, W. D., & Altman, B. M. (2002). Impact of differential item functioning on age and gender differences in functional disability. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *57*(5), S275-84. https://doi.org/10.1093/geronb/57.5.S275

Foo, H., Mather, K. A., Thalamuthu, A., & Sachdev, P. S. (2019). The many ages of man: Diverse approaches to assessing ageing-related biological and psychological measures and their relationship to chronological age. *Current Opinion in Psychiatry, 32*(2), 130–137. https://doi.org/10.1097/YCO.0000000000000473

Gaertner, B., Koschollek, C., Grube, M. M., Lüdtke, D., Fuchs, J., Scheidt-Nave, C., Gößwald, A., & Wetzstein, M. (2019). Including nursing home residents in a general population health survey in Germany. *Survey Methods: Insights from the Field (SMIF).* Advance online publication. https://doi.org/10.13094/SMIF-2019-00010

Gaertner, B., Lüdtke, D., Koschollek, C., Grube, M. M., Baumert, J., Scheidt-Nave, C., Gößwald, A., Fuchs, J., & Wetzstein, M. (2019). Effects of a sequential mixed-mode design on participation, contact and sample composition – Results of the pilot study “IMOA – Improving Health Monitoring in Old Age”. *Survey Methods: Insights from* *the Field (SMIF).* Advance online publication. https://doi.org/10.13094/SMIF-2019-00016

Ganzeboom, H. B. G. (2010). *A New International Socio-Economic Index [Isei] of Occupational Status for the International Standard Classification of Occupation 2008 [Isco-08] Constructed with Data from the Issp 2002-2007. 2010.* http://www.harryganzeboom.nl/Pdf/2010%20-%20Ganzeboom-ISEI08-ISSP-Lisbon-(paper).pdf

Ganzeboom, H. B. G., Graaf, P. M. de, & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research, 21*(1), 1–56. https://doi.org/10.1016/0049-089X(92)90017-B

Ganzeboom, H. B. G., & Treiman, D. J. (2003). Three Internationally Standardised Measures for Comparative Research on Occupational Status. In J. H. P. Hoffmeyer-Zlotnik & C. Wolf (Eds.), *Advances in Cross-National Comparison* (pp. 159–193). Springer US. https://doi.org/10.1007/978-1-4419-9186-7_9

Gerritsen, D. L., Steverink, N., Ooms, M. E., Vet, H. C. W. de, & Ribbe, M. W. (2007). Measurement of overall quality of life in nursing homes through self-report: The role of cognitive impairment. *Quality of Life Research, 16*(6), 1029–1037. https://doi.org/10.1007/s11136-007-9203-7

Groves, R. M. (1987). Research on survey data quality. *Public Opinion Quarterly, 51*(4 PART 2), S156-S172. https://doi.org/10.1093/poq/51.4_PART_2.S156

Groves, R. M., & Lyberg, L. (2010). Total Survey Error: Past, Present, and Future. Public Opinion Quarterly, 74(5), 849–879. https://doi.org/10.1093/poq/nfq065

Grundy, E., & Holt, G. (2001). The socioeconomic status of older adults: How should we measure it in studies of health inequalities? *Journal of Epidemiology and Community Health, 55*(12), 895–904. https://doi.org/10.1136/jech.55.12.895

Hall, S., Longhurst, S., & Higginson, I. J. (2009). Challenges to conducting research with older people living in nursing homes. *BMC Geriatrics*, *9*, 38. https://doi.org/10.1186/1471-2318-9-38

Hilleras, P. K., Jorm, A. F., Herlitz, A., & Winblad, B. (1998). Negative and Positive Affect among the Very Old. *Research on Aging*, *20*(5), 593–610. https://doi.org/10.1177/0164027598205003

Hülür, G., Ram, N., Willis, S. L., Schaie, K. W., & Gerstorf, D. (2015). Cognitive dedifferentiation with increasing age and proximity of death: Within-person evidence from the Seattle Longitudinal Study. *Psychology and Aging*, *30*(2), 311–323. https://doi.org/10.1037/a0039260

Isaksson, U., Santamäki-Fischer, R., Nygren, B., Lundman, B., & Åström, S. (2007). Supporting the Very Old When Completing a Questionnaire. *Research on Aging*, *29*(6), 576–589. https://doi.org/10.1177/0164027507305924

Kalbe, E., Calabrese, P., Fengler, S., & Kessler, J. (2013). Demtect, PANDA, EASY, and MUSIC: Cognitive screening tools with age correction and weighting of subtests according to their sensitivity and specificity. *Journal of Alzheimer's Disease: JAD*, *34*(4), 813–834. https://doi.org/10.3233/JAD-122128

Kalbe, E., Kessler, J., Calabrese, P., Smith, R., Passmore, A. P., Brand, M., & Bullock, R. (2004). Demtect: A new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. *International Journal of Geriatric Psychiatry*, *19*(2), 136–143. https://doi.org/10.1002/gps.1042

Kaspar, R., Gabrian, M., Brothers, A., Wahl, H.-W., & Diehl, M. (2019). Measuring Awareness of Age-Related Change: Development of a 10-Item Short Form for Use in Large-Scale Surveys. *The Gerontologist*. https://doi.org/10.1093/geront/gnx213

Kaspar, Roman, Brijoux, Thomas, Albrecht, Andrea, Zimmermann, Jaroslava, Wenner, Judith, Fey, Jonas, Reissmann, Marcella, . . . Zank, Susanne (2023). Replication data: Challenges and benefits of including the institutionalized, cognitively impaired and unable to respond in a representative survey of the very old. GESIS Datenarchiv, Köln. Data file version 2.0.0, https://doi.org/10.7802/2536.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychosocial function. *JAMA: The Journal of the American Medical Association*, *185*, 914–919. https://doi.org/10.1001/jama.1963.03060120024016

Kelfve, S. (2019). Underestimated Health Inequalities Among Older People-A Consequence of Excluding the Most Disabled and Disadvantaged. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *74*(8), 125–134. https://doi.org/10.1093/geronb/gbx032

Kelfve, S., Thorslund, M., & Lennartsson, C. (2013). Sampling and non-response bias on health-outcomes in surveys of the oldest old. *European Journal of Ageing*, *10*(3), 237–245. https://doi.org/10.1007/s10433-013-0275-7

Kercher, K. (1992). Assessing Subjective Well-Being in the Old-Old: The PANAS as a Measure of Orthogonal Dimensions of Positive and Negative Affect. *Research on Aging*, *14*(2), 131–168. https://doi.org/10.1177/0164027592142001

Kessler, J., Calabrese, P., & Kalbe, E. (2010). DemTect-B: ein Aquivalenztest zum kognitiven Screening DemTect-A [DemTect-B: A parallel test version to the cognitive screening instrument DemTect-A]. *Fortschritte der Neurologie-Psychiatrie*, *78*(9), 532–535. https://doi.org/10.1055/s-0029-1245452

Kessler, J., Fengler, S., Kaesberg, S., Müller, K., Calabrese, P., Ellwein, T., & Kalbe, E. (2014). DemTect 40- und DemTect 80+: Neue Auswertungsroutinen für diese Altersgruppen [DemTect 40- and DemTect 80+: New scoring routines for these age groups]. *Fortschritte der Neurologie-Psychiatrie*, *82*(11), 640–645. https://doi.org/10.1055/s-0034-1385278

Kühn, K., Porst, R., & Zentrum für Umfragen, Methoden und Analysen -ZUMA- (Eds.). (1999). *ZUMA-Arbeitsbericht 99/03. Befragung alter und sehr alter Menschen.: Besonderheiten, Schwierigkeiten und methodische Konsequenzen. Ein Literaturbericht.* ZUMA.

LaPlante, M. P. (2010). The classic measure of disability in activities of daily living is biased by age but an expanded IADL/ADL measure is not. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *65*(6), 720–732. https://doi.org/10.1093/geronb/gbp129

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist*(9), 179–186.

Mackinnon, A., Jorm, A. F., Christensen, H., Korten, A. E., Jacomb, P. A., & Rodgers, B. (1999). A short form of the Positive and Negative Affect Schedule: Evaluation of factorial validity and invariance across demographic variables in a community sample. *Personality and Individual Differences*, *27*(3), 405–416. https://doi.org/10.1016/S0191-8869(98)00251-7

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543. https://doi.org/10.1007/BF02294825

Mueller, S., Wagner, J., & Gerstorf, D. (2017). On the role of personality in late life. In J. Specht (Ed.), *Personality development across the lifespan* (pp. 69–84). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-804674-6.00006-5

Perneczky, R. (2003). Die Eignung einfacher klinischer Tests für die Erkennung der leichten kognitiven Störung und der leichtgradigen Demenz. *Aktuelle Neurologie*, *30*(3), 114–117. https://doi.org/10.1055/S-2003-38269

R Core Team. (2017). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.r-project.org/

Reisberg, B., Ferris, S. H., Leon, M. J. de, & Crook, T. (1982). The Global Deterioration Scale for assessment of primary degenerative dementia. *The American Journal of Psychiatry*, *139*(9), 1136–1139. https://doi.org/10.1176/ajp.139.9.1136

Reisberg, B., Jamil, I. A., Khan, S., Monteiro, I., Torossian, C., Ferris, S., Sabbagh, M., Gauthier, S., Auer, S., Shulman, M. B., Kluger, A., Franssen, E., & Wegiel, J. (2011). Staging Dementia. In A. Kumar, C. L. E. Katona, & M. T. Abou-Saleh (Eds.), *Principles and practice of geriatric psychiatry* (3rd ed., pp. 162–169). Wiley. https://doi.org/10.1002/9780470669600.ch31

Reissmann, M., Storms, A., & Woopen, C. (2021). Individual values and spirituality and their meaning for affective well-being and engagement with life in very old age [Individuelle Werte und Spiritualität und ihre Bedeutung für affektives Wohlbefinden und aktive Verbundenheit mit dem Leben in der Hochaltrigkeit]. *Zeitschrift Für Gerontologie Und Geriatrie*, *54*(2), 85–92. https://doi.org/10.1007/s00391-021-01974-9

Sakshaug, J. W. (2022). Reducing Nonresponse and Data Linkage Consent Bias in Large-Scale Panel Surveys. *Forum for health economics & policy.* online first https://doi.org/10.1515/fhep-2021-0060

Sand, M. (2014). *Dual-Frame-Telefonstichproben: Entwicklung, Handhabung und Gewichtung. GESIS-Technical Reports: Vol. 2*. GESIS – Leibniz-Institut für Sozialwissenschaften.

Schanze, J.-L. (2017). *Report on sampling practices for the institutionalized population in social surveys.* Deliverable 2.1 of the SERISS project funded under the European Union's Horizon 2020 research and innovation programme (GA No: 654221).

Schanze, J.-L. (2019). Schwer befragbar und vernachlässigbar? Die Bevölkerung in Gemeinschaftsunterkünften. *Informationsdienst Soziale Indikatoren* (62), 13–20. https://doi.org/10.15464/isi.62.2019.13-20

Schanze, J.-L. (2021). Response behavior and quality of survey data: Comparing elderly respondents in institutions and private households. *Sociological Methods & Research,* 004912412199553. https://doi.org/10.1177/0049124121995534

Schanze, J.-L., & Zins, S. (2019). *Undercoverage of the elderly institutionalized population: The risk of biased estimates and the potentials of weighting institutionalized population and not the risk of biased estimates: A simulation study using cross-national survey data.* Survey Methods: Insights from the field (SMIF). https://surveyinsights.org/?p=11275

Scherpenzeel, A., Maineri, A. M., Bristle, J., Pflüger, S.-M., Mindarova, I., Butt, S., . . . Luijkx, R. (2017). *Report on the use of sampling frames in European studies.* Deliverable 2.1 of the SERISS project funded under the European Union's Horizon 2020 research and innovation programme (GA No: 654221).

Schiel, S., Ruiz Marcos, J., & Fahrenholz, M. (2021). *Methodenbericht Deutscher Alterssurvey (DEAS): Schriftliche Panelbefragung 2020.* Bonn: infas - Institut für Angewandte Sozialwissenschaft GmbH.

Schneekloth, U., & Müller, U. (1998). *Hilfe- und Pflegebedürftige in Heimen: Endbericht zur Repräsentativerhebung im Rahmen des Forschungsprojekts „Möglichkeiten und Grenzen selbständiger Lebensführung in Einrichtungen"* (2nd ed.). *Schriftenreihe des BMFSFJ: Vol. 147.2*. Kohlhammer.

Schouten, B., Cobben, F., & Bethlehem, J. (2009). Indicators for the representativeness of survey response. *Survey Methodology, 35*(1), 101–113.

Tesch-Römer, C., & Wahl, H.-W. (2017). Toward a More Comprehensive Concept of Successful Aging: Disability and Care Needs. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *72*(2), 310–318. https://doi.org/10.1093/geronb/gbw162

The American Association for Public Opinion Research. (2016). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* (9th ed.). AAPOR.

Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, *98*(2), 281–300. https://doi.org/10.1037/a0017908

Vogel, C., Klaus, D., Wettstein, M., Simonson, J., & Tesch-Römer, C. (2020). German ageing survey (DEAS). In D. Gu & M. E. Dupre (Eds.), Encyclopedia of Gerontology and Population Aging (pp. 1–9). Springer International Publishing. https://doi.org/10.1007/978-3-319-69892-2_1115-1

Wagner, M., Kuppler, M., Rietz, C., & Kaspar, R. (2019). Non-response in surveys of very old people. *European Journal of Ageing, 16*(2), 249–258. https://doi.org/10.1007/s10433-018-0488-x

Wagner, M., Rietz, C., Kaspar, R., Janhsen, A., Geithner, L., Neise, M., Kinne-Wall, C., Woopen, C., & Zank, S. (2018). Quality of life of the very old: Survey on quality of life and subjective well-being of the very old in North Rhine-Westphalia (NRW80+). *Zeitschrift Für Gerontologie Und Geriatrie, 51*(2), 193–199. https://doi.org/10.1007/s00391-017-1217-3

Zank, S., Woopen, C., Wagner, M., Rietz, C., & Kaspar, R. (2020). *Lebensqualität und Wohlbefinden hochaltriger Menschen in NRW (Repräsentativbefragung NRW80+)*. GESIS Datenarchiv, Köln. ZA7558 Datenfile Version 1.0.0. https://doi.org/10.4232/1.13527

Appendix A

Computation of Survey Weights

A computation of survey weights was done in two steps. First, design weights were calculated for all individuals (N = 8,040) from the gross sample of potential participants as the inverse of the inclusion probabilities. The inclusion probabilities for the individuals in the sampling frame were defined in the two-stage sampling design as the product of probabilities for selecting communities (proportional-to-size pps sampling of primary sampling units) and the probabilities for simple random selection of individuals within PSUs. The design weights also corrected for the deliberately disproportionate inclusion probabilities for the design groups defined by age group (80–84, 85–89, 90+ years) and gender (Table A1) in selecting the gross sample of potential participants to be contacted. Disproportionate sampling was necessary for two reasons. First, the older age groups and men represent small proportions of the very old population. Consequently, a simple random sampling would have resulted in only a few respondents in these design groups and compromised the statistical power for comparing population subgroups. Second, a feasibility study showed different response rates concerning the age and sex of the targeted participants. The analysis of the gender and age population structure also revealed that it would be practically impossible to sample enough individuals in the rare population groups (i.e., men 90+) to achieve equal cell sizes of N = 300 for all design groups in the projected realized (net) sample of approximately 1,800 study participants. Hence, a less extreme oversampling of older age groups and men was conducted that would still result in the high statistical power of the subgroup comparison. The total design weights were calculated by multiplying the probability of the community to be drawn into the sample, the probability of the person to be drawn into the community sample, and an adjustment that considers the deliberate oversampling of specific design groups during the last step of selecting individuals for the gross sample. The resulting design weights ranged from 0.278 to 1.738, and the efficiency of weighting was 84.7%.

Table A1. Gross sample selection probabilities by design group.

Design group	N Sampling frame	N Gross sample	Probability
Male, 80–84 y	10,699	1,407	0.131
Male, 85–89 y	5,150	1,179	0.229
Male, 90+ y	1,739	1,005	0.578
Female, 80–84 y	15,668	1,608	0.103
Female, 85–89 y	9,369	1,501	0.160
Female 90+ y	5,512	1,340	0.243
Total	48,137	8,040	1.00

Second, calibration weights have been computed for all individuals from the realized sample (N = 1,863) to correct for selective nonresponses. The recalibration of the design weights was conducted using an iterative process regarding the known population distributions of household size, institutional versus private dwelling, marital status, administrative district, community size and type (BIK-10), age, and gender. The resulting calibration weights ranged from 0.2 to 1.942, and the weighting efficiency was 72.4%. A comparison of sample and population distributions of key population characteristics at different levels of weighting is reported in Hansen et al., 2021.

Appendix B

Definition of the Difficult-to-Survey Subgroup Nursing Home Residents (NHR)

During the initial steps of obtaining samples of secondary sampling units (SSU) from community registers, nursing home residents (NHR) in the sample were identified based on their primary address of residence and a comprehensive repository of 1,276 addresses of nursing homes and care facilities in North-Rhine Westphalia (NRW).

Additional information on dwelling status and care was collected during contacting and interviewing to validate or update preliminary register-based information from sampling. More specifically, interviewers rated the housing situation during contact for all potential participants in the gross sample, using the categories “typical private apartment or house,” “nursing home,” “nursing home facility (e.g., hospice),” “residential care group,” “multigenerational house,” “senior residence,” “retirement home,” and “assisted living apartment or house.” In addition, interviewers assessed whether the address was likely to be part of an institution after an interview had been realized. During interviews, the interviewee (i.e., target or proxy person) answered questions about the time of relocation and the need for full inpatient care.

For this study, the definition of NHR at the time of the interview considered the most current and consistent information from sampling, fieldwork, and interviews. To this end, a hierarchy of available information on institutionalization was defined. Priority was given to trained interviewers’ assessments of the housing situation at the time of the interview. More specifically, NHRs were defined as all individuals living in a “nursing home,” “nursing home facility (e.g., hospice),” or “residential care group.” Only if this information was unavailable (e.g., due to interviews being conducted elsewhere), or the target person lived in a retirement home or a seniors’ residence in which both private and institutional dwellings are often possible, information from residential registers and contact during fieldwork was considered. In the case of inconsistent information from residential registers and fieldwork, information on the recent relocation from the interview was considered. Fieldwork information from the time of first contact was prioritized over residential register information if the target person had relocated within the last 3 years. In case the target person did not move within the last 3 years or no information on relocation was available, interview information on receiving full inpatient care was considered to define NHR status.

4.2 Development of the Elder Abuse and Emotional Consequences Scale (EACS)

This paper introduces the Elder Abuse and Emotional Consequences Scale and presents a set of 14 out of its 16 original items that measure the frequency of elder abuse. The items describe emotional abuse in the forms of intimidation, shaming and blaming, and paternalism as well as financial abuse, physical abuse, and neglect and include frequent and very rare forms of elder abuse. Furthermore, the paper examines possible measurement errors in the measurement of elder abuse that occur in different interview settings common in old age, i.e. interviews where third persons are present as well as proxy-interviews.

GeroPsych

Development of the Elder Abuse and Emotional Consequences Scale (EACS)

Michael Neise, Thomas Brijoux, and Susanne Zank

Online First Publication, August 25, 2022. http://dx.doi.org/10.1024/1662-9647/a000297

CITATION
Neise, M., Brijoux, T., & Zank, S. (2022, August 25). Development of the Elder Abuse and Emotional Consequences Scale (EACS). *GeroPsych*. Advance online publication. http://dx.doi.org/10.1024/1662-9647/a000297

Full-Length Research Report

Development of the Elder Abuse and Emotional Consequences Scale (EACS)

Michael Neise[1], Thomas Brijoux[2], and Susanne Zank[1]

[1]Rehabilitative Gerontology, University of Cologne, Germany
[2]Cologne Center for Ethics, Rights, Economics, and Social Sciences of Health, University of Cologne, Germany

Abstract: Little data exist on elder abuse (EA) among the very old (>80 years), and instruments have not been developed to identify EA within this group. This study presents a survey instrument: the Elder Abuse and Emotional Consequences Scale (EACS). We conducted interviews in a representative sample of the oldest old and did confirmatory factor analysis (CFA) based on n = 988 persons to confirm the assumed factor structure of the EACS. The results confirm a six-dimensional structure and good model fit of the questionnaire. A preliminary examination confirms the reliability and validity of the dimensions. The EACS is thus a multidimensional assessment of EA that can be used in interviews with the very old.

Keywords: elder abuse, very old, survey, personal interviews, psychometric information

Background

Elder abuse (EA) is a widespread phenomenon that challenges societies with an increasingly aging population (Jagielska et al., 2015). The World Health Organization (WHO) defines the phenomenon as "a single or repeated act, or lack of appropriate action, occurring within any relationship where there is an expectation of trust, which causes harm or distress to an older person" (Perel-Levin & WHO, 2008, p. 5). Research commonly distinguishes between five forms of EA: physical, psychological/emotional, financial, sexual, and neglect (Pillemer et al., 2016).

Meta-analyses have estimated the annual prevalence of EA to be between 10% (Ho et al., 2017) and 15% (Yon et al., 2017) on an international level. The results show substantial variation in prevalence estimates of EA, which may be attributed to differences in the conceptual basis, the method of operationalization, and different sample compositions (Neise & Zank, 2019). Most studies include people who are 60 years or older. Two representative German studies include people up to the age of 85 (Goergen, 2010; Fraga et al., 2014). However, data on the phenomenon of EA in the very old (>85 years) in Germany are incomplete, and reliable figures are lacking.

Surveys of the oldest old represent a fundamental challenge (Hall et al., 2009; Wagner et al., 2019) and appear to be even more challenging for sensitive topics such as experiences of EA. The first challenge derives from the heterogeneity of the population regarding cognitive ability. While most individuals above the age of 80 are cognitively fit, some require assistance by third persons in the interview setting, and others can only be represented via proxy interviews (Wagner et al., 2017). Including people with dementia in surveys is a challenge in itself and often leads to the exclusion of this group (Cridland et al., 2016). It is unknown to what extent the presence of third persons in an interview situation affects the responses, but one can easily imagine that the target persons do not like to reveal abuse actions as these are often accompanied by feelings of shame (Dow et al., 2020). Third persons or proxy informants may also be perpetrators themselves. Therefore, third persons and proxy informants present during the interview represent a second challenge.

A third challenge affects the construct validity of the scale regarding institutionalized settings. While most old-aged individuals live in private settings, a substantial part lives in institutionalized care settings. About 600,000 people over the age of 80 who need care in Germany live in nursing homes (Destatis, 2020). They might be particularly exposed to different (additional) forms of abuse like mechanical or pharmacological custodial measures. The final challenge relates to the time constraint in social surveys, especially in interviews with the very old, which requires using a brief survey instrument to capture EA.

GeroPsych (2022), 1–11
https://doi.org/10.1024/1662-9647/a000297

These challenges limit the prospects of retrieving exact prevalence rates of EA. However, there are decreasing protective resources against EA incidents among the oldest old and a lack of coverage of the oldest old in social science (Goergen, 2010), which is why EA must become a part of social surveys. Yet, existing instruments for measuring EA do not sufficiently address these concerns.

Several instruments are available that were developed to measure elder abuse (Schofield & Mishra, 2003; Sooryanarayana et al., 2013; Schofield, 2017; Simmons et al., 2020). Other instruments originate from research on interpersonal violence and have been adapted to measure EA (Straus, 1979; Straus et al., 1996; Laurie et al., 2012). However, none of these available instruments was developed for interviewing the very old (>85 years). Among the most commonly used scales to assess EA are the Conflict Tactics Scale (CTS; Straus, 1979), which was originally developed to assess abuse in intimate relationships, and the Conflict Tactics Scale 2 (CTS2; Straus et al., 1996). The scales include different dimensions of family violence and some forms of abuse that are also relevant to mapping EA in very old age, e.g., psychological or physical abuse. However, they do not sufficiently differentiate between all previously named forms of EA in very old age, and specifically they lack differentiation in the description of psychological abuse (Schofield & Mishra, 2003). Another EA instrument – the Vulnerability to Abuse Screening Scale (VASS; Schofield & Mishra, 2003) – was used to measure EA in the dimensions of vulnerability, dependence, dejection, and coercion, albeit not in the conceptually adopted forms of the WHO. In addition, the VASS and other scales target specific groups like women (ibid.; Laurie et al., 2012) or hospitalized adults (Simmons et al., 2020).

These scales and other available survey instruments, however, do not cover the aforementioned forms of EA in oldest age, are too extensive for surveys of very old people, or do not apply to specific target groups that are increasingly represented in very old age, such as individuals in inpatient care or people with cognitive impairments. This paper addresses the need for developing a short scale to measure EA which can be used in the aforementioned described heterogeneous living conditions.

Aim

This paper presents a new questionnaire for the multidimensional assessment of EA in the very old. It aims to assess the EA phenomenon in interview contexts with the target group within less than 5 minutes. In this context, the instrument should also be applicable for interviewing people with mild cognitive impairment (MCI) and people in institutional care settings. We describe the questionnaire structure, conduct construct validity, and illustrate specific data collection challenges.

Development of the Dimensions of the EACS

Below, we describe the conception of the Elder Abuse and the Emotional Consequences Scale (EACS). We also address the forms of EA of central interest among the very old and the relevance of a stronger internal differentiation of the psychological dimension of EA in very old age.

First, regarding the differentiation of the psychological form, very old people, especially those dependent on others, might be confronted with infantilizing attitudes that may – or may not – be benevolent at their core but tend to undermine the autonomy of the older person concerned and patronize them (Sánchez-Izquierdo et al., 2019). Empirical studies suggest that infantile patronizing behavior in this context is a specific subaspect of psychological abuse in care relationships (Goergen, 2010). Second, for the conceptualization of EA, intimidating and shaming behavior can be distinguished as psychological abuse (Conrad et al., 2011). Third, experiences of neglect are among the more common forms of EA (Dong et al., 2007). Given the increasing likelihood of the very old to need long-term care and care support (Kingston et al., 2018), consideration of experiences of neglect in old age is of central importance. Experiences of neglect can relate to both care-related and psychosocial needs (Goergen, 2010).

Further, financial exploitation is among the most frequently cited forms of EA (Yon et al., 2017). Risk factors favoring financial exploitation are related to health condition (poor health, cognitive decline, need for care and support) (Peterson et al., 2014). Because very old age is often associated with poorer health conditions (Cho et al., 2011), financial exploitation is more likely for the oldest old and should therefore be considered when measuring EA in this group.

Another form of EA in very old age that needs to be differentiated more strongly is custodial measures. These are applied primarily in cases of advanced need for care or in institutional care contexts and are often administered without the consent of those affected, from a seemingly benevolent, protective stance (Gastmans & Milisen, 2006). In this context, deprivation of liberty may take the form of sedating medication or mechanical measures such as bed height adjustment or restraint belts. Such actions can trigger fears and negative emotions in those affected (ibid.) and should also be conceptually understood as a partial aspect of EA in the very old.

While physical and sexualized forms are comparatively less prevalent than other forms of EA (Yon et al., 2017), these forms are nonetheless part of the overall conceptualization of EA (Pillemer et al., 2016).

Another significant factor for the conceptualization and operationalization of EA in the very old is that all forms of EA are associated with negative emotional consequences (Podnieks & Thomas, 2017).

To obtain initial indications of the construct validity of the developed scale, we used selected evidenced risk factors of EA (Schofield & Mishra, 2003; Johannesen & LoGiudice, 2013; Pillemer et al., 2016). These include the cognitive and mental state of the affected older person, the quality of social relationships, and depression. Living in a nursing home is also considered a risk factor for EA (Yon et al., 2019).

Methods

This study is based on data from the study "Quality of Life and Subjective Well-Being of the Very Old in North Rhine-Westphalia" (NRW80+; Wagner et al., 2017). North Rhine-Westphalia is the most populous state in Germany and has a similar demographic structure to the entire country. The data were collected from August 2017 to February 2018. This survey aims to obtain a representative picture of the very old regarding their quality of life, subjective well-being, and living situation.

Design and Procedure

The overall sample consists of randomly drawn cases from the registers of the municipal registration office. Only individuals who were 80 years of age or older were utilized. The drawing was stratified by age group (80–84, 85–89, 90 +) and sex. Deliberate care was taken to include a disproportionate number of men aged 90+ in the random sample to ensure adequate inclusion of this less common segment of the elderly population. In the second step, we contacted the selected cases by mail through a survey institute and informed them about the study. Finally, trained interviewers contacted the selected individuals personally, and if informed consent was obtained, they conducted a 90-minute face-to-face interview. If individuals were unable to participate in the 90-minute interview, we arranted for an interview with a proxy informant.

Participants

The total sample comprised $n = 1{,}863$ persons aged 80 years or older, of whom 1,687 individuals and 176 proxies were interviewed. All interviews were computer-assisted personal interviews. The average age was 85 years; 36.3% were male, 13.9% were living in institutional settings, and 39.6% had another person present at the time of the interview (including relatives). Survey weights were used to correct for possible bias in age, sex, marital status, household size, institutionalization, and regional characteristics (Hansen et al., 2021).

Measures

In its first, unvalidated form, the EACS included 16 items answered on a 5-point scale (0 = *never* to 4 = *very often*), describing at least mild acts of elder abuse. In eight theoretically assumed dimensions of EA (intimidation, shaming and blaming, paternalism, neglect, financial exploitation, physically abusive behavior, unwanted custodial measures, sexualized abuse), the items ask about self-rated experiences of abuse and their emotional consequences in the past 12 months. The item formulation involved four steps: First, we reviewed existing instruments for measuring abuse in old age based on relevant work and reviews (Fulmer et al., 2004; Cohen, 2011; Conrad et al., 2011; Phelan & Treacy, 2011; Abolfathi Momtaz et al., 2013; Van Royen et al., 2020) and identified common dimensions of EA. Second, we discussed specific forms of abuse in old age regarding increasing dependency need for care and institutionalization with experts ($n = 10$). The results revealed eight theoretically described dimensions of abuse. Third, for each dimension of EA, we formulated and prediscussed two acts of abuse and their possible emotional consequences with peers. We paid special attention to the comprehensibility of the formulations and a description of the act of EA. Therefore, the items describe mild acts of abuse. For example, the selected item text in the physical behavior dimension referred to unpleasant touching in combination with its emotional consequence. Finally, we did a pilot study with $n = 62$ very old individuals, testing the questionnaire for its applicability in different settings (private, institutional) and for specific target groups (e.g., people with mild cognitive impairments) to be finally used in the survey study. The English version of the questionnaire was translated from German into English and backtranslated into German by two independent native speakers.

In addition to the EACS, we also considered other variables in the study to describe the sample and to investigate the expected characteristics related to the dimensions of EACS (described below). To map the respondents' personality-related state, which is considered a potential risk factor for abuse experiences (Johannesen & LoGiudice, 2013; Pillemer et al., 2016), we employed the mean of three items ("I lose my temper too easily, I argue with other people too much," "I feel irritated or annoyed too easily") of the Aggressive behavior subscale of the Inventory of

Interpersonal Problems, which were answered on a 4-point scale (1 = *not at all strongly* – 4 = *very strongly*). A satisfactory internal consistency of Cronbach's alpha = 0.85 was obtained for the entire subscale (Barkham et al., 1996).

We assessed depressive symptoms using the short version of the Depression in Old Age Scale (DIA-S4; Heidenblut & Zank, 2020). The DIA-S4 consists of 4 dichotomous items (0 – *no*, 1 – *yes*). To determine the depressive level, we calculated the sum of the four items.

We assessed cognitive health status using the DemTect, a screening tool that differentiates between age-adequate cognitive status, mild cognitive impairment (MCI), and dementia. This test includes assessing various cognitive functions such as memory impairment, number transcoding, semantic fluency, and working memory (Kalbe et al., 2004). We used age-adequate norm values for the population 80+ to classify the three aforementioned groups (Kessler et al., 2014).

We measured the quality of social relationships by the number of social contacts with whom they felt very close. Other sociodemographic data considered were age (in years), sex (male, female), and residence in an inpatient care facility (private residence, nursing home).

To control for methodological challenges related to the particularly sensitive interview content (EA), the interviewers reported whether other individuals were present in the interview situation and whether the interview was conducted with the target or a proxy person.

Data Analysis

The analysis procedure for answering the research questions is based on a step-by-step approach. First, we used an analysis of the measurement invariance to examine the influence of different interview settings (third persons present, proxy interviews). We applied analyses of measurement invariance to uncover differences between groups in a proposed factorial structure, distinguishing different degrees of measurement invariance in the literature. In this context, evidence of strong factorial invariance was considered sufficient to assume the same measurement structure between different groups and thus the measurement of the same construct (Geiser, 2010; Widaman & Reise, 1997). We assumed strong factorial invariance if the patterns of factor loadings in the respective (EA) dimensions and the respective level (intercepts) did not differ significantly. We used the chi-square difference test to detect strong factorial invariance. Since the EA construct was not normally distributed, we used a robust estimation procedure – maximum likelihood estimation with robust standard errors – to determine the chi-square difference characteristic (Asparouhov & Muthén, 2013).

Second, after analyzing the measurement invariance in the multigroup comparison, we used descriptive statistics to describe the design-weighted analysis sample. The criterion variables (experiences of abuse) were simultaneously illustrated at the item level with details of the individual expressions.

Third, we calculated two confirmatory factor analysis (CFA) models – one with eight theoretically assumed dimensions of abuse and one with six theoretically assumed dimensions of abuse. We compared the two models on relevant parameters to determine model quality ($\chi^2_{(\mathrm{Df,\ SCF})}$, RMSEA (CI90%), CFI/TLI, SRMR) and internal consistency reliability (McDonald's ω) at the dimensional level, identifying the better model.

Finally, we performed correlation analyses (Spearman's rank correlation coefficient ρ) between the output factor scores and the theoretically assumed related variables at the dimensional level, estimating missing values in criterion variables using full-information maximum likelihood (FIML) procedures. The latter is considered a method to replace missing values to keep the loss of information low (Lee & Shi, 2021). Data analysis was performed using SPSS version 27, Excel 365, and MPLUS 8.0.

Results

The factorial structure of the EACS differs between interviews conducted with targeted individuals and those conducted with proxy persons. Strong factorial measurement invariance between groups cannot be assumed (Satorra-Bentler scaled $\chi^2_{(df=7)} = 25.79$, $p = .001$). Therefore, $n = 176$ cases assessed by proxy persons were excluded from further analysis. Comparisons within the remaining interviews of target persons with third persons present during the interview and interviews without third persons present also showed significant differences, so strong factorial measurement invariance could not be assumed (Satorra-Bentler scaled $\chi^2_{(df=7)} = 29.19$, $p \leq .001$). Consequently, an additional $n = 699$ interviews with third persons present were excluded from the final analysis sample.

The results presented below are based on the analysis sample with a total of $n = 988$ interviews conducted with the target persons and with no other persons present during the interview. Table 1 provides an overview of the analysis sample.

The (weighted) analysis sample includes individuals who are on average 85 years old; 70% are female and 13.5% are living in a nursing home. Nearly 76% of the sample has age-appropriate cognitive status.

Table 2 provides an overview of the items of the EACS and their corresponding frequencies. The results indicate

Table 1. Descriptive statistics of the analysis sample (N = 988)

Construct	Category	Number of cases (n)[1]	Mean (SD)/%[2]
Age		988	84.97 (4.04)
Sex	Female	548	69.8%
	Male	440	30.2%
Cognition (DEMTECT)	Normal	612	75.8%
	MCI	132	15.7%
	Dementia	70	8.5%
Nursing home	No	873	86.5%
	Yes	115	13.5%
Personality-related state (aggressive behavior)		987	1.36 (0.48)
Depression (DIA-S4)		934	0.87 (1.12)
Relationship quality		988	1.66 (1.34)

Note. [1]Number of cases considered, unweighted, [2]weighted results.

Table 2. Elder Abuse and Emotional Consequences Scale (EACS) (n = 977[1])[2]

Introduction

No matter how well one gets along with family, friends, neighbors, or other caregivers, such as healthcare professionals, minor or major disputes or disagreements can occur that entail emotional consequences. Below, we are interested in how often – in the past 12 months – you have experienced the following emotions during disputes with familial persons close to you or with whom you have regular contact.

How often have you experienced that someone …

Item no.	Item	Never (0)	Seldom (1)	Sometimes (2)	Often (3)	Very Often (4)
1	… raised their voice against you so that you felt upset or insecure?	66%	20.6%	10.6%	2.4%	0.4%
2	... became offensive toward you so that you felt upset or insecure?	74.6%	15%	9.1%	0.8%	0.4%
3	... talked about your weakness in front of you and other people so that you felt embarrassed?	84.9%	10.2%	4.4%	0.6%	–
4	... blamed you for an incident or circumstance so that you felt bad or upset?	85.3%	10%	3.9%	0.8%	–
5	… disregarded your opinion so that you didn't feel like you were being taken seriously?	70.5%	17%	10.3%	1.8%	0.4%
6	… persuaded you to relinquish a right or wish so that you felt patronized at this moment?	82%	10.6%	5.5%	1.6%	0.4%
7	… didn't provide you with the support you needed so that you felt helpless? (*Interview instructions: Only provide these examples upon request, e.g., going to the toilet, washing, getting dressed*)	82%	9.4%	6%	1.8%	0.7%
8	… didn't offer their time so that you felt unwanted or not cared about?	80.4%	10.6%	6.8%	1.7%	0.6%
9	… used your money or possessions for their own purposes so that you felt exploited?	92.2%	3.2%	2.8%	1.3%	0.4%
10	… let themselves be kept by you (e.g., by lack of contribution to shared expenses) so that you felt taken advantage of?	92.2%	3.6%	2.8%	0.9%	0.4%
11	… touched you firmly or roughly so that it felt unpleasant?	95.5%	2.8%	1.3%	0.2%	0.3%
12	… manhandled you or physically treated you in a cruel way so that it felt unpleasant?	94.4%	3.2%	1.9%	0.2%	0.3%
13	… restricted you in your mobility so that you felt frustrated or angry? (*Interview instructions: Only provide these examples upon request, e.g., by tying somebody up, elevating the edge of the bed, or being locked in a room, apartment, or house*)	95.4%	2.6%	1.7%	0.3%	–
14	... gave you pills or medication without your consent that made you feel tired?	98.8%	0.7%	0.6%	–	–
15	… behaved indecently toward you so that you felt ashamed or distraught	97.5%	1.9%	0.7%	–	–
16	... sexually harassed you verbally or physically so that you felt ashamed or distraught?	99.5%	0.5%	<0.1%	–	–

Note. [1]n = 11 cases have missing values on all items of the elder abuse construct, [2]Weighted results.

Table 3. EACS characteristics (model goodness, factor loadings, reliability data) (n = 988^1)

Item no.	Dimensions of abuse	Standardized factor loadings	Composite reliability (McDonald's ω)
1	Intimidation	0.794	0.798
2		0.840	
3	Shaming and blaming	0.741	0.708
4		0.715	
15		0.391	
5	Paternalism	0.796	0.797
6		0.839	
7	Neglect	0.823	0.823
8		0.849	
9	Financial exploitation	0.878	0.865
10		0.868	
11	Physical behavior	0.914	0.907
12		0.909	
Model fit		χ^2 = 113.396 (df = 77, scaling correction factor = 2.3750), RMSEA: 0.022 (CI 90% 0.012–0.030), CFI/TLI: 0.984/0.978, SRMR: 0.034	

Note. ^1n = 988 estimated with full information maximum likelihood (FIML) considering age and sex as external factors.

Table 4. EACS Dimensions with assumed interrelated characteristics^{1,2,3}

Dimensions	Cognitive decline	Nursing home	Personality-related state (aggressive behavior)	Depression	Relationship quality
Intimidation	0.031	0.032	0.261***	0.140***	−0.087***
Shaming and blaming	0.069**	0.032	0.225**	0.122***	−0.118**
Paternalism	0.084**	0.081**	0.244**	0.170***	−0.106**
Neglect	0.082**	0.122**	0.175**	0.155**	−0.092**
Financial exploitation	0.066**	0.020	0.109**	0.098**	−0.137**
Physical behavior	0.056*	0.090**	0.084**	0.072**	−0.034

Note.^{1} F scores of the EACS Dimensions correlated with assumed interrelated characteristics, ^{2}Spearman's rank correlation coefficient, ^{3}*$p \leq .050$; **$p \leq .010$, ***$p \leq .001$.

higher frequencies particularly for questions that can be attributed to experiences of psychological abuse and neglect, and lower frequencies for questions about experiences of physical abuse. For example, on question 8, 20% of respondents state that, in the past year, they have experienced someone not making time for them, making them feel unwanted, or not cared about. In contrast, on question 11, 4.5% state they were touched firmly or roughly, making them feel uncomfortable.

Table 3 illustrates the results of the confirmatory factor analysis and the results of the reliability analyses of the subscales according to theoretically applied dimensions. The assumed dimension on unwanted custodial measures (items 13 and 14) was excluded from this analysis because of insufficient reliability characteristics (McDonald's ω = 0.097). The hypothesized dimension on sexualized abuse experiences (items 15 and 16) was also excluded from the analysis because of insufficient reliability data (McDonald's ω = 0.037). Because of the content-related proximity of item 15 to items 3 and 4, this item was assigned to a dimension in the final confirmatory factor analysis (Table 3).

As an indication of construct validity, we reviewed the underlying dimensions of EA for commonly known (cognition, mental state, depression, relationship quality) and nursing home as a theoretically discussed risk factor. The results are shown in Table 4.

The results show significant positive associations of mental status (overly aggressive) and depression with all dimensions of the EACS ($r \geq .072$, $p \leq .01$). Good relationship quality with the social environment is significantly negatively related to all EACS dimensions ($r \leq -.087$, $p \leq .01$), except for the dimension on physical behavior ($r = -.034$, $p > .05$). Cognitive decline is significantly positively related to all EACS dimensions ($r \geq .056$, $p \leq .05$), except for the EACS dimension on intimidation ($r = .031$, $p > .05$). Living in a nursing home is significantly positively related to the shaming and blaming, neglect, and physical behavior on the EACS ($r \geq .081$, $p \leq .01$).

Discussion

The EACS

The EACS is a new survey instrument for assessing EA and its emotional consequences in old-aged individuals. The analysis of the available data demonstrates that the chosen six-dimensional data structure can be satisfactorily represented, and that the dimensions show adequate reliability. The scale is based on a broad understanding of EA. This is also reflected in the high prevalence of abuse (Brijoux et al., 2021) and the high frequencies in the dimensions. It may, thus, enable future research to discover the dark field of EA in very old individuals. At the same time, the instrument is applicable both in home and institutional settings, can be used when interviewing people with mild cognitive impairment or in the early stages of dementia, and comprises an average interview duration of less than 3 minutes.

The scale could not reliably capture the dimensions of sexualized abuse and unwanted custodial measures. Depending on the dimension, we can assume different hypotheses. Unwanted custodial measures are used more frequently for people who are seriously ill and in need of care (Walther, 2007). At the same time, there is an increased likelihood that this group of people can no longer be interview participants and therefore requires proxy information. These were excluded from this analysis, thus reducing the number of people who might have experienced this form of EA in our sample. A similar argument can be made for sexual forms of abuse. A comparison of two meta-analyses shows that this form of abuse is detected more frequently in institutional settings than in domestic contexts (Yon et al., 2017, 2019). At the same time, after we excluded proxy informants, the already very rarely reported form occured even less frequently in the sample analyzed. Because of the very low frequency, an adequate statement about the reliability is not sufficiently possible. In addition, in the case of the fifteenth item (Tables 2 and Table 3), there is a greater proximity to the dimension shaming and blaming. Consequently, by removing the two theoretical dimensions, we reduced the mapping of the overall construct. For a reliable consideration of these partial aspects of the overall construct, it appears more favorable to use other methods such as observational studies (custodial measures), drop-off questionnaires, or qualitative approaches (sexualized forms).

Validity of the EACS

Because of time constraints when interviewing the target group, we could not assess concurrent validity with a second EA questionnaire, which would be the gold standard for validating a new instrument. Instead, we used construct validity as the first indication of the validity of the EACS by investigating correlations with expected evident risk factors of EA. A similar approach was already taken in the representative Women's Health Australia Study (Schofield & Mishra, 2003). The expected correlation measures were cognitive decline, personality-related state (aggressive behavior), depression, and relationship quality, which were correlated with the dimensions of the EACS. The results show significant correlations in the expected directions in respect of the EACS dimensions, for which there is already some evidence. For example, Acierno et al. (2010) support the present findings on relationship quality with the relevant EACS dimensions by showing that low social support is associated with increased risk of emotional and physical EA. The association of depression with experiences of physical and emotional abuse, neglect, and financial exploitation demonstrated here has also been supported in several studies with older people (Aylaz et al., 2019; Cisler et al., 2012; Weissberger et al., 2020). The association of respondents' aggressive tendencies in dealing with others regarding psychological abuse experiences, financial exploitation, and neglect is supported by a study by Li and Dong (2018), who found that impulsive personality traits such as neuroticism are associated with the same forms of abuse. Finally, another study examining the relationship between cognition and abuse in old age shows an increased risk for physical and emotional abuse, caregiver neglect, and financial exploitation among individuals with cognitive impairments (Dong et al., 2011), underlining the association of cognitive impairment with specific dimensions of abuse demonstrated here. Thus, the EACS can similarly map proven relations of other studies and shows first indications for a valid representation of different dimensions of EA, which allows for a differentiated examination of the EA construct.

Further, other researchers have discussed the EACS dimensions as applicable across the board (Sengstock et al., 1990). Significant correlations of some EACS dimensions were found with living in an institutional setting (paternalism, neglect, and physical behavior). Comparable findings in other studies are relatively rare, as many studies examine residential settings separately concerning EA (Yon et al., 2017, 2019). In addition, the associations found in the present study should be interpreted with caution and cannot be causally attributed to the institutional setting; rather, they can be explained by the fact that more people in need of assistance live in institutional settings (Peng & Wu, 2015). Regarding the first relationship between paternalism and institutional setting, we can assume that more needy individuals in the institutional care context are more likely to be confronted with paternalistic acts of care, presumably fed by an overprotective attitude of the caregiver in question (Fernández-Ballesteros et al., 2019). The second relationship between neglect and institutional setting could be explained by a higher number of individuals with cognitive and functional impairments in institutional settings. Fulmer et al. (2005) found poorer cognitive and functional status to be significantly related to more neglect experiences among older people. The third significant association of more physical abuse experiences in the institutional setting is supported by the meta-analysis of Yon et al. (2019), who demonstrated a high prevalence of physical abuse in the institutional setting. One possible explanation is the increasing impairment of older individuals in the institutional setting, accompanied by behavioral problems to which caregivers may (inappropriately) respond (Lindbloom et al., 2007).

Methodological Challenges

The results of the present research address two relevant methodological challenges in the context of data generation. These challenges relate to the interviewee and the interview situation. The examination to determine strong factorial measurement invariance reveals a significant difference in the assessment of the EACS in target interviews compared to the assessment by proxies, suggesting a different understanding of the construct in these groups. Consequently, the two groups should not be analyzed together, so we excluded the interviews with proxies from this study. The reasons for the different evaluation can be multifaceted. One possible explanation is that the proxy interviews primarily contained information about particularly impaired individuals. While the ability to provide information is significantly limited in individuals with severe cognitive impairments, and information must often be obtained by proxy informants (Zimmerman & Magaziner, 1994), at the same time, studies on other construct areas such as functionality or emotional well-being show that they are poorly assessed by proxies (Selai & Trimble, 1999). In addition, studies on the prevalence of EA in institutional settings overall show higher levels of psychological, physical, and sexual abuse when studies were conducted with caregivers compared to studies conducted with older individuals themselves (Yon et al., 2019). However, it is impossible to conclusively elucidate the reasons for the differences in EA assessment at this point.

The comparison of EA between the groups with and without third persons present also shows a significant difference in the evaluation of the EA construct. Individuals with third persons present during the interview also rate the construct differently. Thus, a joint evaluation was not advisable here, and the individuals in question were excluded from the analysis. Recommendations on methodological challenges in the interpersonal violence survey indicate that reporting abuse experiences can be shameful, tabooed, or filled with fears (Fraga, 2016). For example, an empirical study on the recording of domestic violence against women in Nicaragua shows that significantly less violence is reported when other people were present in interview situations (Ellsberg et al., 2001). Thus, the findings on the lack of strong factorial measurement invariance indicate that it is not advisable to consider proxies or present third persons in a joint evaluation.

Limitations

The EACS is based on a very broad understanding of abuse to cover as many acts of abuse as possible. However, this means the differentiation and exact identification of more severe acts of abuse are impossible with the instrument, even though it can be assumed that more severe acts of abuse are also being implicitly recorded. At the same time, the final instrument does not address sexualized forms of abuse and custodial measures because of insufficient reliability.

Although the EACS has comparatively high expressions at the level of the individual items, we must also assume that the actual occurrence of abuse is underestimated. Excluding proxies because of the undetectable measurement invariance of the EA construct also excluded individuals who are no longer able to provide information and may be particularly affected by certain acts of abuse.

Implications

The EACS is an EA assessment tool that covers a wide range of forms of abuse, is broadly consistent with those in the WHO definition, and permits greater differentiation, especially regarding experiences of psychological abuse. The instrument is applicable in different residential settings and can also be used with individuals who have cognitive impairments but can still be interviewed. Furthermore, the broad understanding of EA for uncovering the phenomenon in research contexts can be seen as a strength. The instrument is particularly suitable for use in research contexts in which the dark field of EA is to be addressed or an internally differentiated examination of different forms of EA is required. Although the EACS provides preliminary evidence of validity in mapping experiences of abuse, further research should be conducted to uncover concurrent validity with the target population.

Even though the instrument has now been developed and tested for interviewing the very old, its applicability in other target groups (younger old people, old people in need of care) should also be examined.

Furthermore, because of the low-threshold questioning method and the reference to the emotional state of the affected person, the instrument can provide a meaningful approach to detecting cases of EA in practice. But possible consequences for detected cases are permissible only to a limited extent and should be supported by further examinations because of the broad conception of EA. The practical usefulness of this rapidly deployable instrument should therefore also be examined.

References

Acierno, R., Hernandez, M. A., Amstadter, A. B., Resnick, H. S., Steve, K., Muzzy, W., & Kilpatrick, D. G. (2010). Prevalence and correlates of emotional, physical, sexual, and financial abuse and potential neglect in the United States: The National Elder Mistreatment Study. *American Journal of Public Health, 100*(2), 292–297. https://doi.org/10.2105/AJPH.2009.163089

Asparouhov, T., & Muthén, B. (2013). *Computing the strictly positive Satorra-Bentler chi-square test in Mplus* (Mplus Web Notes, No. 12). https://www.statmodel.com/examples/webnotes/SB5.pdf

Aylaz, R., Pekince, H., Işik, K., Aktürk, Ü., & Yildirim, H. (2020). The correlation of depression with neglect and abuse in individuals over 65 years of age. *Perspectives in Psychiatric Care, 56*(2), 424–430. https://doi.org/10.1111/ppc.12451

Barkham, M., Hardy, G. E., & Startup, M. (1996). The IIP-32: A short version of the Inventory of Interpersonal Problems. *British Journal of Clinical Psychology, 35*(1), 21–35. https://doi.org/10.1111/j.2044-8260.1996.tb01159.x

Brijoux, T., Neise, M., & Zank, S. (2021). Elder abuse in the oldest old: Prevalence, risk factors and consequences. *Zeitschrift für Gerontologie und Geriatrie, 54*(Suppl 2), 132–137. https://doi.org/10.1007/s00391-021-01945-0

Cho, J., Martin, P., Margrett, J., Macdonald, M., & Poon, L. W. (2011). The relationship between physical health and psychological well-being among oldest-old adults. *Journal of Aging Research*, 3, 605041. https://doi.org/10.4061/2011/605041

Cisler, J. M., Begle, A. M., Amstadter, A. B., & Acierno, R. (2012). Mistreatment and self-reported emotional symptoms: Results from the National Elder Mistreatment Study. *Journal of Elder Abuse & Neglect, 24*(3), 216–230. https://doi.org/10.1080/08946566.2011.652923

Cohen, M. (2011). Screening tools for the identification of elder abuse. *Journal of Clinical Outcomes Management, 18*(6), 261–270.

Conrad, K. J., Iris, M., Ridings, J. W., Langley, K., & Anetzberger, G. J. (2011). Self-report measure of psychological abuse of older adults. *The Gerontologist, 51*(3), 354–366. https://doi.org/10.1093/geront/gnq103

Cridland, E., Phillipson, L., Brennan-Horley, C., & Swaffer, K. (2016). Reflections and recommendations for conducting in-depth interviews with people with dementia. *Qualitative Health Research, 26*(13), 1774–1786. https://doi.org/10.1177/1049732316637065

Destatis. (2020). *Pflegestatistik. Pflege im Rahmen der Pflegeversicherung – Deutschlandergebnisse – 2019* [Care Statistics. Care in Framework of long-term care insurance – Germany results – 2019]. Statistisches Bundesamt.

Dong, X., Simon, M., & Gorbien, M. (2007). Elder abuse and neglect in an urban Chinese population. *Journal of Elder Abuse & Neglect, 19*(3–4), 79–96. https://doi.org/10.1300/J084v19n03_05

Dong, X., Simon, M., Rajan, K., & Evans, D. A. (2011). Association of cognitive function and risk for elder abuse in a community-dwelling population. *Dementia and Geriatric Cognitive Disorders, 32*(3), 209–215. https://doi.org/10.1159/000334047

Dow, B., Gahan, L., Gaffy, E., Joosten, M., Vrantsidis, F., & Jarred, M. (2020). Barriers to disclosing elder abuse and taking action in Australia. *Journal of Family Violence, 35*, 853–861. https://doi.org/10.1007/s10896-019-00084-w

Ellsberg, M., Heise, L., Pena, R., Agurto, S., & Winkvist, A. (2001). Researching domestic violence against women: Methodological and ethical considerations. *Studies in Family Planning, 32*(1), 1–16.

Fernández-Ballesteros, R., Sánchez-Izquierdo, M., Olmos, R., Huici, C., Casado, J. M. R., & Jentoft, A. C. (2019). Paternalism vs. autonomy: Are they alternative types of formal care? *Frontiers in Psychology, 10*, 1–4. https://doi.org/10.3389/fpsyg.2019.01460

Fraga, S., Lindert, J., Barros, H., Torres-González, F., Ioannidi-Kapolou, E., Melchiorre, M. G., Stankunas, M., & Soares, J. F. (2014). Elder abuse and socioeconomic inequalities: A multilevel study in 7 European countries. *Preventive Medicine, 61*, 42–47. https://doi.org/10.1016/j.ypmed.2014.01.008

Fraga, S. (2016). Methodological and ethical challenges in violence research. *Porto Biomedical Journal, 1*(2), 77–80. https://doi.org/10.1016/j.pbj.2016.04.005

Fulmer, T., Guadagno, L., Dyer, C. B., & Connolly, M. T. (2004). Progress in elder abuse screening and assessment instruments. *Journal of the American Geriatrics Society, 52*(2), 297–304. https://doi.org/10.1111/j.1532-5415.2004.52074.x

Fulmer, T., Paveza, G., Weerd, C. V., Fairchild, S., Guadagno, L., Bolton-Blatt, M., & Norman, R. (2005). Dyadic vulnerability and risk profiling for elder neglect. *The Gerontologist, 45*(4), 525–534.

Gastmans, C., & Milisen, K. (2006). Use of physical restraint in nursing homes: Clinical-ethical considerations. *Journal of Medical Ethics, 32*(3), 148–152. https://doi.org/10.1136/jme.2005.012708

Geiser, C. (2010). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung* [Data analysis with Mplus. An application-oriented introduction]. VS Verlag.

Goergen T. (Ed.). (2010). *Sicherer Hafen oder gefahrvolle Zone? Kriminalitäts- und Gewalterfahrungen im Leben alter Menschen* [Safe harbour or dangerous zone: Experiences of crime and violence in the lives of the elderly]. Verlag für Polizeiwissenschaft.

Hall, S., Longhurst, S., & Higginson, I. J. (2009). Challenges to conducting research with older people living in nursing homes. *BMC Geriatrics, 9*, 38. https://doi.org/10.1186/1471-2318-9-38

Hansen, S., Kaspar, R., Wagner, M., Woopen, C., & Zank, S. (2021). The NRW80+ study: Conceptual background and study groups. *Zeitschrift für Gerontologie und Geriatrie, 54*, 76–84. https://doi.org/10.1007/s00391-021-01970-z

Heidenblut, S., & Zank, S. (2020). Screening for depression in old age with very short instruments: The DIA-S4 compared to the GDS5 and GDS4. *Gerontology and Geriatric Medicine, 6*, 233372142098132. https://doi.org/10.1177/2333721420981328

Ho, C. S. H., Wong, S. Y., Chiu, M. M., & Ho, R. C. M. (2017). Global prevalence of elder abuse: A meta-analysis and meta-regression. *East Asian Archives of Psychiatry, 27*(2), 43–55.

Jagielska, K., Lukasik, J. M., & Pikula, N. G. (2015). *Violence against the elderly. Challenges – research – action*. European Association of Schools of Social Work.

Johannesen, M., & LoGiudice, D. (2013). Elder abuse: A systematic review of risk factors in community-dwelling elders. *Age and Ageing, 42*, 292–298. https://doi.org/10.1093/ageing/afs195

Kalbe, E., Kessler, J., Calabrese, P., Smith, R., Passmore, A. P., Brand, M., & Bullock, R. (2004). DemTect: A new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. *International Journal* *of Geriatric Psychiatry, 19*(2), 136–143. https://doi.org/10.1002/gps.1042

Kessler, J., Fengler, S., Kaesberg, S., Müller, K., Calabrese, P., Ellwein, T., & Kalbe, E. (2014). DemTect 40– and DemTect 80+: New scoring routines for these age groups. *Fortschritte der Neurologie Psychiatrie, 82*(11), 640–645. https://doi.org/10.1055/s-0034-1385278

Kingston, A., Comas-Herrera, A., & Jagger, C. (2018). Forecasting the care needs of the older population in England over the next 20 years: Estimates from the Population Ageing and Care Simulation (PACSim) modeling study. *The Lancet Public Health, 3*(9), E447–E455. https://doi.org/10.1016/S2468-2667(18)30118-X

Laurie, A. R., Showalter, J., Pratt, T., Ballentine, N. H., Chinchilli, V. M., & McCall-Hosenfeld, J. S. (2012). Validity of the brief inpatient screen for intimate partner violence among adult women. *Journal of Women's Health, 52*(7), 679–699. https://doi.org/10.1080/03630242.2012.724763

Lee, T., & Shi, D. (2021). A comparison of full information maximum likelihood and multiple imputation in structural equation modeling with missing data. *Psychological Methods, 26*(4), 466–485. https://doi.org/10.1037/met0000381

Li, M., & Dong, X. (2018). Association between personality of victim and elder abuse. *Innovation in Aging, 2*(Suppl 1), 917. https://doi.org/10.1093/geroni/igy031.3412

Lindbloom, E. J., Brandt, J., Hough, L. D., & Meadows, S. E. (2007). Elder mistreatment in the nursing home: A systematic review. *Journal of the American Medical Directors Association, 8*, 610–616.

Momtaz, Y. A., Hamid, T. A., & Ibrahim, R. (2013). Theories and measures of elder abuse. *Psychogeriatrics, 13*(3), 182–188. https://doi.org/10.1111/psyg.12009

Neise, M., & Zank, S. (2019). Gewalterfahrungen älterer Menschen im sozialen Nahraum–Befunde und Herausforderungen [Experiences of violence by older people in the social environment – Findings and challenges]. In K. Hank, F. Schulz-Nieswandt, M. Wagner, & S. Zank (Eds.), *Alternsforschung. Handbuch für Wissenschaft und Praxis* (pp. 459–490). Nomos Verlag.

Peng, R., & Wu, B. (2015). Changes of health status ad institutionalized among older adults in China. *Journal of Aging and Health, 27*(7), 1223–1246. https://doi.org/10.1177/0898264315577779

Perel-Levin, S., World Health Organization. (2008). *Discussing screening for elder abuse at primary health care level by Silvia Perel-Levin*. World Health Organization. https://apps.who.int/iris/handle/10665/43523

Peterson, J. C., Burnes, D. P. R., Caccamise, P. L., Mason, A., Henderson, C. R., Wells, M. T., Berman, J., Cook, A. M., Shukoff, D., Brownell, P., Powell, M., Salamone, A., Pillemer, K. A., & Lachs, M. S. (2014). Financial exploitation of older adults: A population-based prevalence study. *Journal of General Internal Medicine, 29*(12), 1615–1623. https://doi.org/10.1007/s11606-014-2946-2

Phelan, A., & Treacy, M. P. (2011). *A review of elder abuse screening tools*. NCPOP, School of Nursing, Midwifery and Health Systems, University College Dublin.

Pillemer, K., Bumes, D., Riffin, C., & Lachs, M. S. (2016). Elder abuse: Global situation, risk factors, and prevention strategies. *The Gerontologist, 56*(S2), 194–205. https://doi.org/10.1093/geront/gnw004

Podnieks, E., & Thomas, C. (2017). The consequences of elder abuse. In X. Dong (Ed.), *Elder abuse: Research, practice and policy* (pp. 109–123). Springer.

Sánchez-Izquierdo, M., Santacreu, M., Olmos, R., & Fernández-Ballesteros, R. (2019). A training intervention to reduce paternalistic care and promote autonomy: A preliminary study. *Clinical Interventions in Aging, 14*, 1515–1525. https://doi.org/10.2147/CIA.S213644

Schofield, M. J., & Mishra, G. D. (2003). Validity of self-report screening scale for elder abuse: Women's Health Australia Study. *The Gerontologist, 43*(1), 110–120. https://doi.org/10.1093/geront/43.1.110

Schofield, M. J. (2017). Screening for elder abuse: Tools and effectiveness. In X. Dong (Ed.), *Elder abuse: Research, practice and policy* (pp. 161–199). Springer.

Selai, C., & Trimble, M. R. (1999). Assessing quality of life in dementia. *Aging & Mental Health, 3*(2), 101–111.

Sengstock, M. C., McFarland, M. R., & Hwalek, M. (1990). Identification of elder abuse in institutional settings: Required changes in existing protocols. *Journal of Elder Abuse & Neglect, 2*(1/2), 31–50. https://doi.org/10.1300/J084v02n01_03

Simmons, J., Wiklund, N., Ludvigsson, M., Nägga, K., & Swahnberg, K. (2020). Validation of REAGERA-S: A new self-administered instrument to identify elder abuse and lifetime experiences of abuse in hospitalized older adults. *Journal of Elder Abuse & Neglect, 32*(2), 173–195. https://doi.org/10.1080/08946566.2020.1737614

Sooryanarayana, R., Choo, W.-Y., & Hairi, N. N. (2013). A review on the prevalence and measurement of elder abuse in the community. *Trauma, Violence, & Abuse, 14*(4), 316–325. https://doi.org/10.1177/1524838013495963

Straus, M. A. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics (CT) Scales. *Journal of Marriage and Family, 41*(1), 75–88. https://doi.org/10.2307/351733

Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The revised Conflict Tactics Scales (CTS2): Development and preliminary psychometric data. *Journal of Family Issues, 17* (3), 283–316. https://doi.org/10.1177/019251396017003001

Van Royen, K., van Royen, P., De Donder, L., & Gobbens, R. J. (2020). Elder abuse assessment tools and interventions for use in the home environment: A scoping review. *Clinical Interventions in Aging, 15*, 1793–1807. https://doi.org/10.2147/CIA.S261877

Wagner, M., Rietz, C., Kaspar, R., Janhsen, A., Geithner, L., Neise, M., Kinne-Wall, C., Woopen, C., & Zank, S. (2017). Quality of life of the very old: Survey on quality of life and subjective well-being of the very old in North Rhine-Westphalia (NRW80+). *Zeitschrift für Gerontologie und Geriatrie, 51*, 193–199. https://doi.org/10.1007/s00391-017-1217-3

Wagner, M., Kuppler, M., Rietz, C., & Kaspar, R. (2019). Non-response in surveys of very old people. *European Journal of Ageing, 16*, 249–258. https://doi.org/10.1007/s10433-018-0488-x

Walther, G. (2007). Restraints in long-term care nursing homes for the elderly: Legal aspects and alternatives. *Ethik in der Medizin, 19*, 289–300.

Weissberger, G. H., Mosqueda, L., Nguyen, A. L., Samek, A., Boyle, P. A., Nguyen, C. P., & Han, S. D. (2020). Physical and mental health correlates of perceived financial exploitation in older adults: Preliminary findings for the Finance, Cognition, and Health in Elders Study (FINCHES). *Aging & Mental Health, 24*(5), 740–746. https://doi.org/10.1080/13607863.2019.1571020

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). American Psychological Association. https://doi.org./10.1037/10222-009

Yon, Y., Mikton, C. R., Gassoumis, Z. D., & Wilber, K. H. (2017). Elder abuse prevalence in community settings: A systematic review and meta-analysis. *The Lancet Global Health, 5*, e147–e156. https://doi.org/10.1016/S2214-109X(17)30006-2

Yon, Y., Ramiro-Gonzalez, M., Mikton, C. R., Huber, M., & Sethi, D. (2019). The prevalence of elder abuse in institutional settings: A systematic review and meta-analysis. *European Journal of Public Health, 29*(1), 58–67. https://doi.org/10.1093/eurpub/cky093

Zimmerman, S. I., & Magaziner, J. (1994). Methodological issues in measuring the functional status of cognitively impaired nursing home residents: The use of proxies and performance based measures. *Alzheimer Disease and Associated Disorders, 8*(Suppl 1), S281–S290.

History
Received January 31, 2022
Accepted June 13, 2022

Conflict of Interest
The authors declare no conflict of interest.

Editorial Note
The acting editor was Noel H. Ballentine.

Funding
Open access publication enabled by University of Cologne.

ORCID
Michael Neise
https://orcid.org/0000-0001-7548-591X

Michael Neise
Rehabilitative Gerontology
University of Cologne
Herbert-Lewin-Str. 2
50931 Köln
Germany
michael.neise@lvr.de

4.3 Elder abuse in the oldest old: prevalence, risk factors and consequences

The third paper included in this thesis focuses on estimating the 12-month prevalence rate and examines risk factors and consequences of elder abuse in the oldest old. Relevant for the context of this thesis, the article discusses the difficulties that arise for representativity when parts of the sample cannot be analysed together and presents a method for addressing these difficulties by recalibrating the survey weights.

Z Gerontol Geriat 2021 · 54 (Suppl 2):S132–S137
https://doi.org/10.1007/s00391-021-01945-0
Received: 17 March 2021
Accepted: 25 June 2021

Elder abuse in the oldest old: prevalence, risk factors and consequences

Thomas Brijoux[1] · Michael Neise[2,3] · Susanne Zank[2]

[1] ceres—cologne center for ethics, rights, economics, and social sciences of health, Universität zu Köln, Köln, Germany
[2] Lehrstuhl für Rehabilitationswissenschaftliche Gerontologie, Humanwissenschaftliche Fakultät, Universität zu Köln, Köln, Germany
[3] Landschaftsverband Rheinland, Köln, Germany

Abstract

Background: Experiences of abuse in relationships with an expectation of trust are a common phenomenon among older people and is called elder abuse (EA). This can take various forms, such as physical, verbal, emotional, psychological, financial, sexual abuse or neglect. Due to their high vulnerability and difficulties in receiving support, people aged over 80 years old have been pointed out as a group that needs special focus in research.
Objective: Prevalence, risk factors and consequences of EA for different aspects of quality of life are explored among the oldest old.
Material and methods: Computer-assisted personal interviews were conducted in a representative sample of the oldest old in North Rhine-Westphalia (Germany). 988 self-report interviews without third persons present of the NRW80+ study are used to assess EA with the help of the elder abuse and emotional consequences scale (EACS). The EACS describes EA in six dimensions that give a broad understanding of EA.
Results: Prevalence of experiences of EA within the last 12 months was 54.1%. In logistic regression, multimorbidity, lower functioning, age below 90 years, smaller social network size, and aggressive behaviorwere significant risk factors for EA. People experiencing EA showed less life satisfaction and autonomy and increased loneliness and depressive symptoms.
Conclusion: EA is prevalent among the oldest old. Serious consequences of EA on life results can be shown with a broad operationalization of EA. Future research should focus on a deeper understanding of reasons for EA and reflect on the relationship between and the perspectives of perpetrators and victims.

Keywords
Elder abuse · Old age · Representative survey · Elder mistreatment · Loneliness

Supplementary Information

The online version of this article (https://doi.org/10.1007/s00391-021-01945-0) contains supplementary material, which is available to authorized users.

Scan QR code & read article online

Background

The group of the oldest old (aged over 80 years) is one of the fastest growing populations in Germany [3]. A widespread phenomenon among older people in general (60+ years) is called elder abuse (EA). The most common working definition of EA, which is used in most studies and adopted by the World Health Organization (WHO) describes EA as:

a single or repeated act or lack of appropriate action, occurring within any relationship where there is an expectation of trust which causes harm or distress to an older person. Elder abuse has serious consequences for the health and well-being of older people and can be of various forms: physical, verbal, psychological/emotional, sexual and financial. It can also simply reflect intentional or unintentional neglect [31, p. 1].

The occurrence of EA has been investigated in many national and international studies. Two different meta-analyses reported prevalence in community settings of 10% [13] and 15%, respectively. [33]. Within these two meta-analyses, the range of reported prevalence rates was between 0.8% [20] and 79.7% [26], which indicates heterogeneous definitions and operationalizations of EA that differ between analyzing solely illegal acts and addressing subjective experiences of abuse [8]. In a representative German study, a prevalence rate of 25.6% for psychological aggression and 1.5% for physical violence in community-dwelling people aged between 60 years and 85 years old was reported [8]. In contrast, EA in nursing homes is seldomly explored through the perspective of the victims, which might be due to the high sensitivity of the topic as well as difficult access to and limited communicative abilities of this group [8]. Moreover, a representative prevalence rate of EA in the oldest old (80+ years) is lacking on the international and national levels.

As prevalence rates are missing for the oldest old and consequences of EA are presumably more devastating for them [8], special attention for this group is needed. Firstly, higher need for care, reduction in the size of social networks, higher prevalence rates for dementia, and frailty are reported in this age group [23]. Therefore, psychological forms of violence, like intimidation, paternalism as well as shaming and blaming, should be investigated [21]. Other forms of abuse, like sexual abuse, are rarely reported in old age [33]. Secondly, the victim's perspective is crucial when EA happens in social relationships with a high dependency between victim and perpetrator. Victim self-reports are of utmost importance when victims have less ability to defend themselves against intentional or unintentional violence, when they have fewer possibilities to reach for help and support, and when their access to social studies is limited [8], all of which is the case for the oldest old.

Risk factors for EA abuse have been explored in two recent systematic reviews [16, 28]. Risk factors for the victims can be categorized into sociodemographic, physical, and psychological aspects [16, 28]. Sociodemographic risk factors with consistent evidence are ethnicity and lower education [28]. Contradicting evidence exists for age > 75 years, gender (female), marital status, higher income, and living alone, which are protective factors in some studies and risk factors in others. Psychological constructs like cognitive impairment, aggressive behavior, loneliness, personality traits (e.g. antisocial personality), self-neglect, and stress-related coping processes are associated with EA [16, 28]. Health-related risk factors comprise various measures of functional impairments like ADL and IADL, frailty, multimorbidity, and incontinence as well as psychiatric illnesses (depression, alcohol abuse, past abuse) [16, 28].

EA affects central aspects of quality of life (QoL) in very old age. In a review of 25 cross-sectional studies, Dong et al. [4] pointed out that EA has a detrimental impact on the psychological well-being of the aged population. In addition, they reported that depression is one of the most prevalent psychological consequences of EA. Recently, Wang and Dong [30] showed that community-dwelling older people who report a greater degree of self-perceived loneliness have a greater probability of having experienced EA, especially psychological forms of abuse. A further aspect of quality of life that is influenced by EA is autonomy, often by use of custodial measures.

The first objective of this study is to estimate the prevalence of various facets of EA in very old age. In a second step, risk factors for experiencing EA are identified. Thirdly, we investigate the effects of EA on central outcomes of QoL in very old age.

Study design

The NRW80+ dataset is used to examine EA in the oldest old. The sample consists of $n = 1863$ persons over 80 years and includes $n = 1687$ self-report interviews and $n = 176$ proxy interviews. Details about the recruitment process, a description of the proxy interviews, and a general description of the participants of the NRW80+ study are presented elsewhere [10].

The challenges and potentials (CHAPO) model describes the theoretical framework of this study [29]. The definition of EA already implies a connection between life conditions and emotional results which are also contemplated in the broader CHAPO framework. With respect to this model, we have investigated the role of disposable resources, skills, and competencies relating to experiences of EA and linked these experiences to the model-inherent life results depression, loneliness, life satisfaction, and autonomy.

Variables

The elder abuse and emotional consequences scale (EACS) is used to describe the prevalence of EA within the last 12 months [22]. The EACS is a low-threshold instrument that is designed for use in large representative surveys. It comprises 13 items, describing different actions of EA and their emotional consequences for the victim (e.g. "how often have you experienced that someone raised their voice against you so that you felt upset or insecure"). The frequencies of these actions are graded in five categories from never to very often. We constructed a dichotomous outcome variable for EA: if the interviewed person rated any of the 13 items as seldom or more frequent, this was regarded as EA. The 13 items represent the 6 dimensions: intimidation, shaming and blaming, paternalism, neglect, financial exploitation, and physical abuse.

Sociodemographic risk factors analyzed were gender, age group, nursing home residency, education [32], social network size, and income [2].

Potential psychological risk factors that were analyzed were aggressive and offensive behavior [1] and cognition. Cognition was measured by means of the DemTect [17, 18].

Physical risk factors explored were multimorbidity, frailty, and functional health. The number of treated diseases was used as an indicator of multimorbidity. Frailty was analyzed in accordance with the description by Zimmermann et al. [34], which categorizes people as frail, pre-frail, or non-frail. The instrumental activities of daily living (IADL) subscale of the older Americans resources and services questionnaire was used to describe the functional status [6].

Potential consequences of EA were examined in the form of depressive symp-

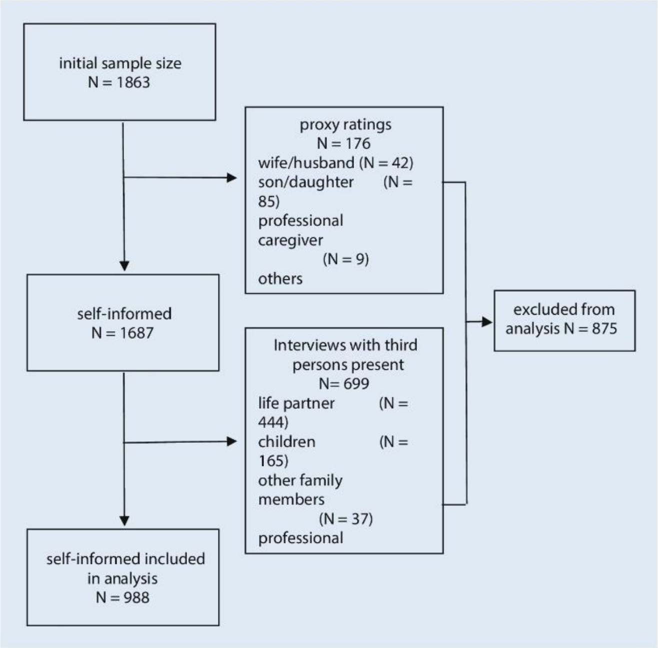


Fig. 1 ▲ Self-report interviews, proxy interviews, interviews conducted in the presence of others

toms, loneliness, autonomy, and life satisfaction. As an indicator for depressive symptoms, we used the short form of the depression in old age scale (DIA-S4) [12]. The DIA-S4 is a screening tool for depression consisting of four questions that are part of the DIA-S [11]. Loneliness [5], autonomy and life satisfaction [19] were measured with one item that has been established in the socioeconomic panel and the European social survey. A more detailed description of the used variables can be obtained in Supplement 1.

Sample and bias

With use of survey weights, the NRW80+ sample is representative in relation to age, sex, and nursing home status. Detailed analyses that are published elsewhere [22] point out limitations in interviews in which third persons were present or which were conducted with proxy informants. In this article, only self-ratings from interviews in which the interviewee and the interviewer were alone were analyzed. Therefore, 176 proxy ratings and 699 interviews in which third persons were present were excluded from this analysis. The resulting sample consisted of 988 cases (see ◘ **Fig. 1**).

The exclusion of proxy interviews and interviews where third persons were present lead to a selective and biased subsample. This bias was reduced by a recalculation of sample weights via logistic regression [7] that gives more weight to underrepresented groups. Consequences of weighting the sample are shown in Supplement 2.

Statistical methods

Frequency of EA and its dimensions were analyzed by descriptive statistics. Confidence intervals for binary variables are 95% Clopper-Pearson intervals. Odds ratios and Cohens D are presented as measures of effect size. Logistic regression is used to examine risk factors of EA. Tests for multicollinearity show no problems. Multiple imputation [15] was used to address missing values, with Rubins formula [25] for the calculation of standard deviations, 20 datasets were imputed [9]. The Holm-Bonferroni for multiple tests is used [14]. Calculations were carried out by means of SPSS (IBM Corp. Released 2020. IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY, USA) Version 27 and Microsoft Excel 365.

Results

Prevalence of elder abuse

About half of the sample ($N = 534$; 54.1%; Confidence Interval (CI): 51.0–57.2%) have experienced a form of EA at least once during the last 12 months. The most frequent form of EA is intimidation with a prevalence of 39.2% (CI: 36.2–42.3%). Other frequent forms are paternalism (32.6%; CI: 29.8–35.6%), neglect (27.0%, CI: 24.4–29.9%), and shaming and blaming (23.1%, CI: 20.6–25.9%). Financial exploitation (10.9%, CI: 9.1–13.0%) and physical behavior (8.8%, CI: 7.2–10.8%) occurred less often. Frequent experience of EA (i.e. often or very often) is rare (i.e. less than 4% of the sample) in all facets of EA (◘ **Table 1**).

Risk factors for elder abuse

Multivariate logistic regression identifies higher multimorbidity (OR: 1.13, $p < 0.001$), social network size (OR: 0.75, $p < 0.001$), and higher levels of aggression of the victim (OR: 2.53, $p < 0.001$) as significant risk factors for EA. No significant influence could be found for cognitive status, education, gender, income, IADL, age group, and frailty. The mean of Nagelkerkes R^2 over all imputed datasets is 0.16. All regression coefficients and their respective odds ratios and confidence intervals are shown in Supplement 3.

Quality of life

People who are affected by EA show worse values in all analyzed QoL outcomes. They show more depressive symptoms (1.14 vs. 0.69, $p < 0.001$), perceive more loneliness (1.49 vs. 1.32, $p < 0.001$) and lower autonomy (3.41 vs. 3.63, $p < 0.001$), and have a lower life satisfaction (7.67 vs. 8.25,

Table 1 Prevalence and frequency of different EA dimensions

EA-Dimension	Never (in %)	Seldom (in %)	Sometimes (in %)	Often (in %)	Very often (in %)
Intimidation	60.8	26.6	10.8	1.4	0.5
Shaming and blaming	76.9	21	1.8	0.3	0
Paternalism	67.4	22.4	7.7	2.2	0.4
Neglect	73	17.8	5.8	2.2	1.2
Financial exploitation	89.1	6.8	2.8	1	0.4
Physical behavior	91.2	5.3	3	0.2	0.3

Table 2 Dimensions of quality of life and EA

		Mean	SD	T	p-value[a]	Cohens D
Depression	No EA ($N=453$)	0.69	1.00	6.32	< 0.001***	0.43
	EA ($N=535$)	1.14	1.17			
Loneliness	No EA ($N=453$)	1.32	0.62	–3.66	< 0.001***	0.20
	EA ($N=535$)	1.49	0.82			
Autonomy	No EA ($N=453$)	3.63	0.64	–4.73	< 0.001***	–0.26
	EA ($N=535$)	3.41	0.81			
Life satisfaction	No EA ($N=453$)	8.25	1.73	–4.82	< 0.001***	–0.42
	EA ($N=535$)	7.67	2.02			

**p* 0.05, ***p* < 0.01, ****p* < 0.001
[a] After the Holm-Bonferroni correction all *p*-values remain significant on the 5% level

$p < 0.001$) (◘ **Table 2**). Effect sizes range between small and medium effects.

Discussion

During the last 12 months, 54.1% of persons aged 80 years and older have experienced some kind of EA. This prevalence is higher compared to other studies for the elderly population in Germany [8]; however, the comparison of different prevalences of EA can be misleading when the difference in operationalizations is not accounted for. In this sample, the low-threshold approach in the wording of the questions and the emphasis of the EACS on emotional forms of abuse are expected to have contributed to higher prevalence rates. The higher age of the informants, which is associated with other risk factors like multimorbidity, also contributes to the higher prevalence. In old age, emotional and psychological abuse are the most common forms of abuse, which resembles the results of known the meta-analysis and reviews [13, 27, 33] in which emotional abuse was also identified as the most frequent form of EA in younger populations. Even though many very old people experience EA, it is nevertheless a rare event for most of them. In each dimension, less than 4% of the people are affected "often" or "very often" by EA.

In the regression model, multimorbidity, and the victim's tendency for aggressive interactions are significant risk factors for EA, while a greater social network is a protective factor. An influence of cognition, nursing home status, IADL, education, gender, age group and income could not be shown in the multivariate logistic regression model. While evidence for the effect of gender, education, age group and income has already been inconsistent in existing literature [16, 28], reasons for the missing effect of nursing home status and cognition are not obvious. In a subsequent bivariate sensitivity analysis of our data, nursing home status has shown a significant association with EA, but this effect was moderated by cognition, functional health, and multimorbidity, which are also more common in nursing homes. Therefore, the effect was attenuated in the full model. The same applies to the constructs related to physical health, where IADL is not significant when multimorbidity is included in the model.

Risk factors for the abusers, like caregiver burden, especially for caregivers of people with dementia or substance abuse [16], cannot be investigated in this article. Likewise, risk factors lying within the victim-perpetrator relationship cannot be explored in this article. Therefore, the overall low amount of explained variance was expected.

On the basis of the CHAPO model, life events, such as experiences of EA, are associated with more generalized life results. In our sample and understanding, EA is associated with higher depression, increased loneliness, a reduction of autonomy, and life satisfaction. For depression, loneliness, and life satisfaction, the observed effect sizes are even higher than reported effect sizes in other populations [4, 24], supporting the proposition that the impact of EA is more severe for vulnerable groups [8]. Autonomy and loneliness are also significantly associated with EA. These two life results are more strongly affected by interactions with others, i.e. victim-perpetrator interaction in the case of EA. For a better understanding of underlying processes, the inclusion of the perspective of potential perpetrators in studies seems necessary. This could be done in future qualitative analyses but is not possible within the NRW80+ dataset. For loneliness, it might also be true that old people who have a smaller social network may feel more alone when they experience abuse in one or more of their few relationships. Causality might also work in the other direction, meaning that loneliness makes old people easier to be taken advantage of.

Based on associations found between aspects of QoL and EA, we conclude that our choice of a broad understanding and low-threshold operationalization of EA leads to results that are relevant for the oldest old and should be pursued in future research, especially in the context of quality of life.

Two limitations of the current study should be considered. Firstly, the exclusion of proxy interviews limits the representativity of the sample. The resulting bias has been addressed by a recalculation of sample weights that gives more importance to cases which are underrepresented; however, a complete elimination of this bias is not possible. Nevertheless, compared to surveys that do not include people with dementia and focus entirely on community-dwelling older adults, the use of adapted sample weights represents an improvement. Secondly, testing causality was not possible within this cross-sectional dataset and results should be replicated. As control group designs are no ethically possible alternative, longitudinal studies need to be performed.

Practical implications

- Experiences of EA are a common phenomenon in the group of older people and refer most often to psychological abuse; however, for most of the affected individuals, these actions of EA remain rare events.
- EA is associated with higher depression and loneliness as well as lower life satisfaction and autonomy.
- Longitudinal surveys are needed to unveil mechanisms that lead to EA.
- To analyze mechanisms of EA, studies are needed that reflect on the relationship between perpetrators and victims.

Corresponding address

Thomas Brijoux
ceres—cologne center for ethics, rights, economics, and social sciences of health, Universität zu Köln
Albertus Magnus Platz, 50923 Köln, Germany
thomas.brijoux@uni-koeln.de

Acknowledgements. The project NRW80+ belongs to the Key Research Area "Aging and Demographic Change" (Speaker: Susanne Zank) of the Cologne Center for Ethics, Rights, Economics, and Social Sciences of Health (ceres) which is directed by Christiane Woopen at the University of Cologne. Members of the project board are Michael Wagner, Christiane Woopen and Susanne Zank. NRW80+ is funded by the Ministry of Innovation, Science and Research, North Rhine-Westphalia.

Funding. Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest. T. Brijoux, M. Neise and S. Zank declare that they have no competing interests.

Ethical standards. All procedures performed in studies involving human participants or on human tissue were in accordance with the ethical standards of the institutional and/or national research committee (Medical Faculty of the University of Cologne) and with the 1975 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

1. Barkham M, Hardy GH, Startup M (1996) The IIP-32: a short version of the inventory of interpersonal problems. Br J Clin Psychol 35:21–35
2. Destasis (2018) Wirtschaftsrechnungen. Laufende Wirtschaftsrechnungen. Einkommen, Einnahmen und Ausgaben privater Haushalte. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Einkommen-Einnahmen-Ausgaben/Publikationen/Downloads-Einkommen/einnahmen-ausgaben-privater-haushalte-2150100177004.pdf?__blob=publicationFile. Accessed 19 July 2021
3. Destasis (2019) Bevölkerung im Wandel. Annahmen und Ergebnisse der 14 koordinierten Bevölkerungsvorausberechnung. https://www.destatis.de/DE/Presse/Pressekonferenzen/2019/Bevoelkerung/pressebroschuere-bevoelkerung.pdf?__blob=publicationFile. Accessed 19 July 2021
4. Dong X, Chen R, Chang E, Simon M (2013) Elder abuse and psychological well-being: a systematic review and implications for research and policy—a mini review. Gerontology (59):132–142. https://doi.org/10.1159/000341652
5. European Social Survey (2019) „Deutschland in Europa". Deutsche Teilstudie im Projekt „European Social Survey" Welle 9 2018 / 2019. https://www.europeansocialsurvey.org/docs/round9/fieldwork/germany/ESS9_questionnaires_DE.pdf. Accessed 19 July 2021
6. Fillenbaum GG (1988) Multidimensional functional assessment of older adults. The duke older americans resources and services procedures. Duke University Center for the Study of Aging and Human Development, New Jersey
7. Gabler S, Kolb J-P, Sand M, Zins S (2015) Gewichtung. Mannheim. GESIS Survey Guidelines. GESIS – Leibniz-Institut für Sozialwissenschaften, https://doi.org/10.15465/gesis-sg_007
8. Goergen T (2009) "Sicherer Hafen" oder "gefahrvolle Zone"? Kriminalitäts- und Gewalterfahrungen im Leben alter Menschen. https://www.bmfsfj.de/blob/94186/727885d90cc7178a738226f533bce61d/kriminalitaets-gewalterfahrungen-leben-alter-menschen-langfassung-data.pdf. Accessed 19 July 2021
9. Graham JW, Olchowski AE, Gilreath TD (2007) How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci 8:206–213. https://doi.org/10.1007/s11121-007-0070-9
10. Hansen S, Kaspar R, Wagner M, Woopen C, Zank S (2021) The NRW80+ study: Conceptual background and design decisions (in press). Z Gerontol Geriat 54. (In press)
11. Heidenblut S, Zank S (2010) Entwicklung eines neuen Depressionsscreenings für den Einsatz in der Geriatrie. Z Gerontol Geriat 43:170–176
12. Heidenblut S, Zank S (2020) Screening for depression in old age with very short instruments: the DIA-S4 compared to the GDS5 and GDS4. Gerontol Geriatr Med 6:1–9. https://doi.org/10.1177/2333721420981328
13. Ho CSH, Wong SY, Chui SY, Ho RCM (2017) Global prevalence of elder abuse: a meta-analysis and meta-regression. East Asian Arch Psychiatry 27:43–55
14. Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6(2):65–70
15. Jakobsen JC, Gludd C, Wetterslev J, Winkel P (2017) When and how should multiple imputation be used for handling missing data in randomised clinical trials—a practical guide with flowcharts. BMC Med Res Methodol 17:162. https://doi.org/10.1186/s12874-017-0442-1
16. Johannesen M, LoGuidice D (2013) Elder abuse: a systematic review of risk factors in community-dwelling elders. Age Ageing 42:292–298. https://doi.org/10.1093/ageing/afs195
17. Kalbe E, Kessler J, Calabrese P, Smith R, Passmore AP, Brand M, Bullock R (2004) DemTect: a new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. Int J Geriatr Psychiatry 19:136–143
18. Kessler J, Fengler S, Kaesberg S, Müller K, Calabrese OP, Ellwein T, Kalbe E (2014) DEmTect40– und DemTect80+: Neue Auswertungsmethoden für diese Altersgruppen. Fortschr Neurol Psychiatr 82:640–645
19. Kantar Public (2019) SOEP-Core-2018: Personenfragebogen, Stichproben A-L3 + N. SOEP Survey Papers 608: Series A. DIW/SOEP, Berlin
20. Marmolejo II (2008) Elder abuse in the family in Spain. Valencia: Queen Sofi a Center for the Study of Violence, 2008. http://www.inpea.net/images/Spain_Report_2008_Elder.pdf. In: Yon, Y., Mikton, C.R., Gassoumius, Z.D., Wilber, K.H. (2017) Elder abuse prevalence in community settings: a systematic review and meta-analysis. Lancet Glob Health 2017(5):e147–e156
21. Neise M, Zank S (2019) Gewalterfahrungen älterer Menschen im sozialen Nahraum – Befunde und Herausforderungen. In: Hank K, Schulz-Nieswand F, Wagner M, Zank S (eds) Alterns-

forschung. Handbuch für Wissenschaft und Praxis. Nomos, Baden-Baden, pp 459–490

22. Neise M, Brijoux T, Zank S (in preparation) The elder abuse and emotional consequences scale (working title)
23. Robert-Koch-Institut (2015) Gesundheit in Deutschland Kapitel 8. https://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GesInDtld/gesundheit_in_deutschland_2015.pdf?__blob=publicationFile. Accessed 19 July 2021
24. Roepke-Buehler SK, Simon M, Dong X (2015) Association between depressive symptoms, multiple dimensions of depression, and elder abuse: a cross-sectional, population-based analysis of older adults in Urban chicago. J Aging Health 27(6):1003–1025. https://doi.org/10.1177/0898264315571106
25. Rubin D (1987) Multiple imputation for nonresponse in surveys. Wiley,
26. Silva-Fhon JR, Del Rio-Suarez AD, Motta-Herrera SN, Fabricio-Wehbe SCC, Partezani-Rodrigues RA (2015) Domestic violence in older people living in the district of Brena, Peru. Revista Fac Med 63:367–375
27. Sooryanarayana R, Choo W-Y, Hairi NH (2013) A review on the prevalence and measurements of elder abuse in the community. Trauma Violence Abus 14(4):316–325
28. Storey JE (2020) Risk factors for alder abuse and neglect: a review of the literature. Aggress Violent Behav 50:101339
29. Wagner M, Rietz C, Kaspar R, Jahnsen A, Geithner L, Neise M, Kinne-Wall C, Woopen C, Zank S (2018) Quality of life of the very old. Survey on quality of life and subjective well-being of the very old in North Rhine-Westphalia (NRW80+). Z Gerontol Geriat 51:193–199
30. Wang B, Dong X (2018) The association between loneliness and elder abuse: findings from a community-dwelling Chinese aging population. Innov Aging 2(S1):282
31. World Health Organization (2008) A global response to elder abuse and neglect: building primary health care capacity to deal with the problem worldwide: main report. https://www.who.int/ageing/publications/ELDER_DocAugust08.pdf. Accessed 19 July 2021
32. Wiest M, Richter M, Krauel F, Maurer S, Henning G, Lejeune C, Rabe C, Engstler H (2014) German Ageing Survey, Deutscher Alterssurvey (DEAS): Documentation of instruments and variables 1996–2011 (pp. 413). Berlin: DZA . http://www.dza.de/fileadmin/dza/pdf/fdz/DEAS_Instrumentendokumentation.pdf. Accessed 19 July 2021
33. Yon Y, Mikton CR, Gassoumius ZD, Wilber KH (2017) Elder abuse prevalence in community settings: a systematic review and meta-analysis. Lancet Glob Health (5):e147–e156. https://doi.org/10.1016/S2214-109X(17)30006-2.
34. Zimmermann J, Hansen S, Wagner M (2021) Home environment and frailty in very old adults (in press). Z Gerontol Geriatr 54

Zusammenfassung

Gewalterfahrungen in der Hochaltrigkeit: Prävalenz, Risikofaktoren und Auswirkungen

Hintergrund: Erfahrungen von Gewalt in Beziehungen mit einer Vertrauenserwartung sind ein häufiges Phänomen und werden Elder Abuse (EA) genannt. EA kann in verschiedenen Formen wie physischem, verbalem, emotionalem, psychischem, finanziellem oder sexuellem Missbrauch sowie Vernachlässigung vorkommen. Aufgrund ihrer hohen Verletzbarkeit und erschwerter Hilfeinanspruchnahme im Fall von Viktimisierungserfahrungen wurden Menschen über 80 Jahren als Risikogruppe identifiziert.

Fragestellung: Dieser Artikel untersucht Prävalenz, Risikofaktoren und Auswirkungen von EA auf verschiedene Dimensionen von Lebensqualität in der Hochaltrigkeit.

Material und Methode: Computergestützte persönliche Interviews wurden in einer repräsentativen Stichprobe der Hochaltrigenbevölkerung in Nordrhein-Westfalen durchgeführt. In dieser Analyse werden nur Selbstberichte verwendet, bei denen keine dritten Personen in der Interviewsituation präsent waren. EA wird mit der Elder Abuse and Emotional Consequences Skala (EACS) gemessen. Die EACS unterscheidet zwischen 6 Formen von EA, denen ein breites Verständnis von EA zugrunde liegt.

Ergebnisse: Die Prävalenz von Gewalterfahrungen innerhalb der letzten 12 Monate beträgt 54.1 %. Multimorbidität, Beeinträchtigung in instrumentellen Aktivitäten des täglichen Lebens, Alter unter 90 Jahre, ein kleineres soziales Netzwerk und aggressives Verhalten waren signifikante Risikofaktoren für EA. Menschen, die von EA betroffen waren, zeigten weniger Lebenszufriedenheit und Autonomie, mehr depressive Symptome und waren häufiger von Einsamkeit betroffen.

Schlussfolgerung: Erfahrungen von EA sind in der Hochaltrigenbevölkerung verbreitet. Mit der gewählten weiten Operationalisierung von EA können bedeutende Konsequenzen für die Lebensqualität der von EA Betroffenen aufgezeigt werden. Zukünftige Forschung sollte die Gründe für Gewalthandlungen besser erforschen und die Beziehung wie auch die Perspektiven von Tätern und Opfern berücksichtigen.

Schlüsselwörter

Elder Abuse · Hochaltrigkeit · Repräsentative Studie · Misshandlung · Einsamkeit

Supplement 1

Variable description:

The Elder Abuse and Emotional Consequences Scale (EACS) is used to describe the prevalence of EA within the last twelve months [23]. The EACS is a low-threshold instrument that is designed for use in large representative surveys. It comprises 13 items, describing different actions of EA and their emotional consequences for the victim (e.g. "How often have you experienced that someone raised their voice against you so that you felt upset or insecure"). The frequencies of these actions are graded in five categories from never to very often. Cronbach's Alpha for the whole scale is 0.87. We constructed a dichotomous outcome variable for EA: If the interviewed person rated any of the 13 items as seldom or more frequent, this was regarded as EA. The 13 items represent the six dimensions: Intimidation, shaming and blaming, paternalism, neglect, financial exploitation, and physical abuse. The intention is to link the construct closely to feelings of the victims and not only to actions of potential perpetrators. Therefore, the items include aspects like felt neglect and or not limited to any specific settings like nursing homes.

Sociodemographic risk factors analyzed were gender, age group, nursing home residency, education, social network size, and income. Education was assessed in three categories established by the German Ageing study [33]. Income was measured as household net income and is analysed in six categories that were established in the household budget surveys [2].

Potential psychological risk factors that were analysed were aggressive and offensive behaviour and cognition. Three items from the "too aggressive" subscale of the Inventory of Interpersonal Problems were used to measure aggressiveness [1].

Cognition was measured by means of the DemTect. The DemTect is a screening instrument for early dementia and mild cognitive impairment with good classification for dementia (100% sensitivity and 92% specificity) of and mild cognitive impairment (80% sensitivity and 92% specificity) [18]. It provides age-specific scores for the oldest old [19]. The Cut-Off scores 13 and 9 were used to differentiate between normal ageing, mild cognitive impairment and dementia [18].

Physical risk factors explored were multimorbidity, frailty, and functional health. The number of treated diseases was used as an indicator of multimorbidity. Frailty was analysed in accordance with the description by Zimmermann et al. [35], which categorizes people as frail, pre-frail, or non-frail. The "instrumental activities of daily living" (IADL) subscale of the Older Americans Resources and Services Questionnaire was used to describe the functional status [7].

Potential consequences of EA were examined in the form of depressive symptoms, loneliness, autonomy, and life satisfaction. As an indicator for depressive symptoms, we used the short form of the Depression in old Age Scale (DIA-S4) [13]. The DIA-S4 is a screening tool for depression consisting of four questions that are part of the Depression in old Age Scale [12]. Loneliness [6], autonomy and life satisfaction [20] were measured with one-item that have been established in the socio-economic Panel and the European social survey.

Loneliness was measured with the item "How often have you felt lonely during the last week?" which was also used in the seventh wave of the European Social Survey [6] and could be answered in four points. A single item was used to measure autonomy ("Do you arrange your life according to your own ideas?") on a four-point scale. Life satisfaction was measured with an eleven-point item ranging from zero to ten ("All in all, how satisfied are you currently with your life?"), that has been used similarly in the socio-economic panel [20].

Sample description

Construct	Category	NRW 80+ Full Sample (N = 1863)[1]	EA Sample (N = 988) (using survey weights) [2]	EA Sample (analysis weights) (N = 988)[3]
Gender	Male	36.3%	30.2%	35.6%
	Female	63.7%	69.8%	64.4%
Cognition	Healthy	68.1%	75.8%	71.1%
	Mild cognitive Impairment	15.5%	15.7%	15.8%
	Dementia	16.4%	8.5%	13.1%
Age group	80-84	54.3%	54.1%	55.2 %
	85-89	30.8%	33.0%	31.4%
	90+	14.9%	12.9%	13.3%
Education	Low	27.8%	24.8%	28.5%
	Middle	53.0%	53.8%	51.4%
	High	17.8%	21.3%	20.1%
Household Income	below 1300 €	16.4%	18.4%	25.8%
	1300 € - 1700 €	14.1%	15.6%	20.7%
	1700 € - 2600 €	23.5%	22.7%	29.8%
	2600 € - 3600 €	11.8%	11.4%	14.4%
	3600 € - 5000 €	4.4%	4.4%	6.0%
	5000 € - 18000 €	2.9%	2.6%	3.3%
Nursing Home	No	86.1%	86.5%	86.7%
	Yes	13.9%	13.5%	13.3%
Frailty	Non-Frail (Ref.)	25.2%	31.2%	20.6%
	Pre-Frail	59.0%	57.0%	42.2%
	Frail	15.8%	11.8%	11.7%
IADL		1.4 (0.69)	1.58 (0.54)	1.44 (0.62)
Multimorbidity	#Diseases (SD)	3.61 (2.36)	3.36 (2.30)	3.42 (2.32)
Loneliness		1.35 (0.67)	1.38 (0.70)	1.41 (0.74)
Depression		0.93 (1.13)	0.87 (1.12)	0.91 (1.12)
Life Satisfaction		7.78 (2.02)	7.98 (1.89)	7.95 (2.10)
Autonomy		3.43 (0.86)	3.59 (0.69)	3.52 (0.83)

1 Sample description of the NRW 80+ dataset using original weights. 2 Sample after exclusion of proxy interviews and interviews with third persons present using original weights. 3 Sample as analysed using new sample weights.

Supplement 2 Risk factors for EA

		Overall Model			
Variable		β	OR	OR CI	p-value[1]
Age group	80 - 84 (Ref.) (n = 389)				
	85 – 89 (n = 363)	-0.17	0.85	0.62 - 1.16	0.29
	90+ (n = 236)	-0.49	0.61	0.39 - 0.96	0.03*
Gender	Men (Ref.) (n = 440)				
	Women (n = 548)	0.12	1.13	0.81 - 1.55	0.48
Education	Low (Ref.) (n = 226)				
	Medium (n = 535)	0.34	1.40	0.96 - 2.05	0.08
	High (n = 228)	0.44	1.55	0.86 - 2.81	0.15
Nursing Home	No (Ref.) (n = 873)				
	Yes (n = 115)	0.34	1.41	0.83 - 2.4	0.21
Household Income	below 1300 € (Ref.) (n = 210)				
	1300 € - 1700 € (n = 212)	-0.09	0.92	0.56 - 1.5	0.73
	1700 € - 2600 € (n = 275)	0.42	1.52	0.96 - 2.4	0.08
	2600 € - 3600 € (n = 147)	0.44	1.56	0.85 - 2.88	0.15
	3600 € - 5000 € (n = 72)	0.24	1.27	0.49 – 3.29	0.62
	5000 € - 18000 € (n = 72)	0.28	1.32	0.40 – 4.33	0.64
Social network	No. of persons	-0.29	0.75	0.66 – 0.85	<0.001***
Personality	Aggression	0.93	2.53	1.83 - 3.49	<0.001***
Cognition	Normal (Ref.) (n = 723)				
	MCI (n = 165)	-0.07	0.94	0.61 - 1.45	0.77
	Dementia (n = 100)	-0.06	0.94	0.54 - 1.65	0.83
Multimorbidity	No. of Diseases	0.13	1.13	1.06 - 1.21	<0.001***
Frailty	Non-Frail (Ref.) (n = 281)				
	Pre-Frail (n = 564)	0.21	1.23	0.8 - 1.90	0.34
	Frail (n = 144)	0.52	1.68	0.89 - 3.2	0.11
Functional Health	IADL	-0.37	0.69	0.49 - 0.97	0.03*
Constant		-1.38	0.25	0.09 – 0.71	<0.01**
Overall R^2					0.16

*p.0.05, **p< 0.01, ***p< 0.001

[1] The shown p-values are not adjusted for multiple testing. Using the Holm-Bonferroni adjustement only social network, personality und multimorbidity are significant at the 5% level.

4.4 Multidimensional Assessments of QoL in Very Old Age: A Population-Based Cross-Sectional Study

The fourth paper presents quality of life as a multidimensional construct and presents different aspects of life results as well as it examines predictors of life results. It follows a different approach in pursuing representativity. Interviews conducted with proxies and interviews conducted with target persons are presented separately here and only the combination of both parts constitutes a representative account of the old-aged population.

Multidimensional Assessments of Quality of Life in Very Old Age: A Population-Based Cross-Sectional Study

Roman Kaspar[1], Andrea Albrecht[1], Thomas Brijoux[1], Jonas Fey[1], Luise Geithner[1], Veronica Oswald[1], Marcella Reissmann[1], Michael Wagner[1,2], Judith Wenner[3], Susanne Zank[1,3] & Jaroslava Zimmermann[1]

[1]University of Cologne, Cologne Center for Ethics, Rights, Economics, and Social Sciences of Health, Germany

[2]University of Cologne, Faculty of Management, Economics and Social Sciences, Institute of Sociology and Social Psychology (ISS), Germany

[3]University of Cologne, Faculty of Human Sciences, Rehabilitative Gerontology, Germany

Corresponding author

Correspondence concerning this article should be addressed to Roman Kaspar, Cologne Center for Ethics, Rights, Economics, and Social Sciences of Health, Albertus-Magnus-Platz, 50923, Cologne, Germany. Email: roman.kaspar@uni-koeln.de

Acknowledgements

This work was supported by a grant of the Ministry of Culture and Science of the German State of North Rhine-Westphalia (323-8.03-125240). The funding agency played no role in the design, execution, analysis and interpretation of data, or writing of the study.

The project is part of the Key Research Area “Aging and Demographic Change” at the Cologne Center for Ethics, Rights, Economics, and Social Sciences of Health (ceres) at the University of Cologne. Members of the project board are Michael Wagner and Susanne Zank.

RK drafted the study, conducted the analyses and wrote the original manuscript. JW, JZ, LG, MR, RK and TB contributed parts of the manuscript, revised and edited the manuscript. All authors read and approved the final manuscript.

We thank Sylvia Hansen, Anna Storms, Michael Neise, Christian Rietz, Wiebke Schmitz and Christiane Woopen for their involvement in planning and conducting the NRW80+ study. We are indebted to Nicole Schäfer for her help in preparing this manuscript.

Abstract

Social survey data on those aged 80 years or older is sparse. Based on a representative sample, this paper validates a multidimensional model proposed for understanding of quality of life (QoL) in very old age. Towards this goal, this paper estimated levels and heterogeneity of personal and environmental resources, well-being, autonomy and perceived appreciation by society in the population of the very old in Germany. Next, the contribution of personal and environmental resources to QoL outcomes and overlap between these outcomes was estimated using a multivariate approach. Results were based on a representative survey on QoL of the very old in North Rhine-Westphalia, Germany's most populous state. The survey included comprehensive information on socio-demographics, health, social and personal QoL resources including objective testing of cognitive function. The sample comprised 1,863 individuals (mean age 86.5 years, range 80-102 years), including 211 individuals residing in non-private dwellings. Proxy interviews were conducted for 176 individuals that were willing to participate but could not conduct the interview themselves due to limited ability to communicate (PLC). Pronounced differences were found for PLC with respect to environmental and personal resources and QoL outcomes. Pronounced differences were also found both with respect lower observed levels of QoL outcomes (e.g., autonomy) and predictors of QoL outcomes (e.g., effect of negative "external" appreciation on subjective well-being). Contrary to the deficit-orientated model of old age, a high degree of autonomy was observed. However, substantial and consequential negative "external" appreciation of very old age was also apparent.

Key words: representative survey, nursing home residents, welfare, well-being, proxy interviews

Multidimensional Assessments of Quality of Life in Very Old Age: A Population-Based Cross-Sectional Study

Different conceptualizations and assessments of quality of life (QoL) have been brought forward in the last decades, most prominently by the World Health Organization (WHO) (1997) that defines QoL as “individuals' perceptions of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept incorporating in a complex way the persons' physical health, psychological state, level of independence, social relationships, personal beliefs and their relationships to salient features of the environment” (WHO, 1997, p. 11). This definition highlights important features of QoL, mainly that it is inherently multidimensional and that living conditions are to be evaluated against the backdrop of individual and societal values and expectations. In this broad understanding of QoL, it is not a single feature (e.g. life satisfaction) that can account for the multiple facets of QoL. Empirical assessments of QoL need to depict these different facets individually and consider simultaneously the relationship between them.

This is even more important in the context of old age which is characterized by the paradoxical empirical observations of high subjective well-being and good subjective health despite the fact that multimorbidity or cognitive and functional limitations are highly prevalent (Schilling, 2006; Wettstein et al., 2016). Studies on QoL in the very old have shown strong psychological resilience despite adverse trajectories in health, function and participation (Jopp et al., 2008; Jyväkorpi et al., 2018). Moreover, QoL in the oldest old has been described as geared towards specific ways to negotiate demands for autonomy, identity and well-being in light of limited health and functioning (Botes et al., 2019; Cho et al., 2011; Kruse, 2021), some of which are more highly appreciated in society than others (e.g. informal caregiving, generativity, technology use).

Conceptual Model of Quality of Life

Within the scope of the large-scale representative survey NRW80+, the Challenges and Potentials Model of QoL in Very old Age (CHAPO, Figure 1) was developed as an interdisciplinary framework and a grid of operationalization (Wagner et al., 2018). CHAPO explicitly considers different dimensions of QoL and suggests operationalizations with a specific focus on old age within each of these dimensions. The dimensions are developed based on Veenhoven's (2000) "Four Qualities of Life Model". His differentiation between QoL 'chances' and 'results' and between 'inner' and 'outer' qualities leads to the four dimensions of QoL, namely, livability of the environment (CHAPO: livable environment), life-ability of the individual (CHAPO: life-ability of the person), external utility of life (CHAPO: appreciation by others), inner appreciation of life (CHAPO: appreciation of own life). The CHAPO model extends Veenhoven's (2000) model by (1) explicitly considering values held by the person and the environment and (2) proposing 'successful life conduct' as a third QoL result characterized by person-environment fit and congruency of values of the individual and their environment. CHAPO achieves specificity with respect to very old age by the selection of QoL markers within conceptual domains that allow for testing competing mechanisms proposed by major ageing theories about how QoL is generated or retained in very old age (Erikson, 1998; Heckhausen et al., 2013; Neise et al., 2019; Wahl et al., 2012).

Life Results

In line with the distinction between inner valuation and external worth (utility) of an individual's life as two kinds of life results (Veenhoven, 2000), we consider an individual's life to be evaluated not only by themselves, but also by their social environment. As manifestations of such external evaluations, the environment's explicit attitudes on and behaviors towards older individuals give them an impression of their *appreciation by others* (Neise et al., 2019). In a performance-oriented society holding generalized prejudice against old age, very old individuals are expected to be confronted with particularly negative evaluations such as constituting an unproductive burden to society (Fürstenberg, 2013). Although evidence for adverse effects of age stereotype on developmental outcomes (including health, well-being and even mortality) have been reported before (Dionigi, 2015; Levy et al., 2002; Levy, 2009), knowledge about the individual factors contributing to such adverse perceptions of external appraisal and potential protective effects remain limited.

As an indicator of subjective appreciation of own life, *positive affect* in very old individuals has been studied extensively. Maximizing positive affect and minimizing suffering are overarching motives of hedonic concepts of QoL. In contrast to rational evaluative judgements such as life satisfaction, however, affective well-being may also result from less conscious processing of goal approach and may signal need satisfaction and successful adaptation to ageing, for example in the competence press model by Nahemow and Lawton (1973) and late-life development (Erikson, 1998). In light of diminishing options for agency and control over one's life and increasing likelihood of chronic disease and unrecoverable loss at the end of life, Tesch-Römer and Wahl (2017) focus on SWB as a marker of successful ageing in very old age. Similarly, in their discussion of an ecological theory of ageing, Wahl et al. (2012) propose autonomy, identity and well-being as developmental outcomes, highlighting the role of adaptive and supportive environments and processes of belonging to achieve these late-life outcomes.

In addition, CHAPO explicitly refers to the concept of *successful life conduct* as a behavioral QoL result. Acknowledging both social-normative and personal goals, we suggest that successful life conduct is characterized by a specific quality of the person-environment constellation (e.g. fit) or congruence between actual behavior and multi-dimensional stipulations held by individuals and society about the good life in very old age (e.g. autonomy, generativity, ego integrity). In this study, we consider *perceived autonomy over one's life* as an indicator of successful life conduct. According to WHO (2002), "*autonomy is the perceived ability to control, cope with and make personal decisions about how one lives on a day-to-day basis, according to one`s own rules and preferences" (WHO, 2002, p. 13).* In ageing policy, autonomy of the individual is a key characteristic of responsible citizenship (WHO, 2002). As self-regulation implies autonomy of the individual, it is a widespread presumption and implicit outcome in ageing theories (Baltes & Baltes, 1990; Carstensen, 1991; Heckhausen, 1997).

Research into the mutual relationship between different aspects of QoL results in very old age is limited. Reliable data on many facets of QoL in the general population of the very old, including those in institutional settings, remains sparse. Especially data that allows for a combined analyses of how individual and environmental chances impact on different QoL results is rare. Wettstein et al. (2015) for example have shown substantial overlap (i.e. within-person correlation across time) between indicators of positive activation (positive affect, depressive symptoms, mastery) but a rather independent pattern of change in perceived autonomy in a substantial sample of very old community-dwelling individuals.

Conducting a representative study of very old people also means giving people who, because of their health, can no longer be interviewed themselves, a chance to participate. In dementia and stroke research, proxies are seen as valuable sources of information (Burks et al., 2021; Leontjevas et al., 2012; Sneeuw et al., 1997). With respect of the QoL results considered here, proxy interviews are more common in well-being research (Martyr et al., 2018), but rarely employed to assess personal values or autonomy. There is critical debate with respect to the validity and potential measurement error in proxy assessment (Vuorisalmi et al., 2012). In a total-survey-error framework, however, such measurement errors need to be weighed against representation errors resulting from systematic non-response of particularly vulnerable individuals. For the current study, first analyses show that including vulnerable people via proxy interviews comes with substantive benefits regarding the representativity of the sample but only minor measurement bias in core QoL indicators such as positive affect (Kaspar et al., 2022).

Life Chances

Livable environment refers to social norms or values and the opportunities to achieve own goals (Wagner et al., 2018). Of the plethora of aspects of the natural, cultural, social, built or material human environment, the following environmental aspects of life chances were exemplarily included in this paper: social contact, financial resources, internet usage, and societal values. Research on *social contact* shows that larger social networks go along with higher social support (Cornwell & Schafer, 2016) and are positively associated with subjective well-being (SWB) as well as mental or physical health in older adults (Hawton et al., 2011; Huxhold et al., 2013; Smith & Victor, 2019). A meta-analysis showed that the quality of social contact correlates more strongly with SWB than the quantity (Pinquart & Sörensen, 2000). The frequency of contacts was, however, found to be associated with QoL (Netuveli et al., 2006) in older adults. Available *financial resources* characterize the socioeconomic position of an individual within society. Research has identified an indirect effect of financial resources on QoL via health inequalities as well as direct effects on different QoL results (Pratschke et al., 2017). For older people (65+) with reduced self-care capacity, poor financial resources are associated with lower life satisfaction (C. Borg et al., 2006). A literature review shows significant correlations of income with SWB operationalized as either life satisfaction, self-esteem, or happiness (Pinquart & Sörensen, 2000). Results of the Survey of Health and Retirement in Europe show that there are significant associations between income and QoL in the age group 65+ (Knesebeck et al., 2007). Information and communication technologies (ICT) have become indispensable in most areas of life and represent a major means of societal participation. *Internet usage* allows for knowledge acquisition, entertainment, or social interaction, and has the potential to preserve independency and autonomy (Czaja & Lee, 2007). Previous research indicated that internet use in old age increases self-efficacy (Erickson & Johnson, 2011), perceived control, social support, and life satisfaction, while reducing depressive symptoms and loneliness (Heo et al., 2015; Shapira et al., 2007; Szabo et al., 2019). According to the SIM study (Medienpädagogischer Forschungsverbund Südwest, 2022) 86% of the very old who use the internet do so at least once a week. *Societal values* co-determine opportunities of expression, pursuit, and attainment of individual goals (Bobowik et al., 2011; Sagiv et al., 2004). Due to societal and technological change, people in old age are regarded particularly likely to experience discrepancies between societal and their own values (Brandtstädter & Wentura, 1994), leading to feelings of alienation, disorientation, or exclusion (Sagiv et al., 2004). Such experiences – referred to here as *anomie* – provoke loneliness (Kaspar, 2004), decreased life satisfaction (Brandtstädter & Wentura, 1994), and suicidal tendencies (Schaller, 2008).

Life-ability of the person refers to resources within the individual, such as values, skills, or experience and knowledge (Wagner et al., 2018). Functional, cognitive and mental health (i.e., independence in activities of daily living, cognitive impairment and depressive symptoms) as well as personal values were considered in this study. *Functional health* has been identified as an indicator of individual (health) resources and as a key predictor of QoL results in old age. Henchoz et al. (2019) reported an association between an improvement in activities of daily living (ADL) and an increase of QoL. Schilling et al. (2013) found a strong linear effect of ADL deterioration on decline in overall life satisfaction among community-dwelling very old adults. This was also observed among institutionalized individuals (Liu et al. 2020). Furthermore, the ability to perform ADL was positively associated with perceived autonomy in both community-dwelling (Sánchez-

García et al., 2019) and institutionalized older adults (Liu et al., 2020). In old age, *dementia* and *depression* belong to the most prevalent psychiatric diseases (Lilford and Hughes 2020) and they were identified as the main cause of reduction in disability-adjusted life years (Wittchen et al. 2011). The prevalence of major depression and subclinical depressive symptoms continues to increase across older age groups (Charlson et al. 2019). Symptoms of depression such as negative mood and reduced interest for enjoyment (WHO, 1992) directly affect subjective well-being. As a trans-situational reference system, *personal values* serve as benchmarks for cognitive judgements of satisfaction or more implicit experiences of coherence or achievement. Different values might be adaptive in different stages of life. For instance, a prioritization of values focusing on loss prevention instead of growth is expected to help manage a shift towards a less positive balance between gains and losses in later life (Ebner et al., 2006). An age-related decrease of self-centeredness and an increase of values focusing on the welfare of others can be expected based on Erikson’s stage theory of human development (1998) and the model of gero-transcendence (Tornstam, 1997). This view is supported by previous studies on age differences in values (Robinson, 2013; Tulviste et al., 2017).

Research Aims

The aim of this paper is twofold. First, to offer a comprehensive and reliable account on the status quo of QoL chances and results in Germany’s very old population and subpopulations characterized by retained or limited abilities to communicate. Second, to test the validity of the CHAPO model by relating selected QoL chances for older adults to QoL results including successful life conduct.

Research Design and Methods

Participants and Procedures

Data are from a representative survey NRW80+ conducted in Germany's most populated state North-Rhine Westphalia. The random community sample of persons aged 80 and over was selected in a multi-stage sampling procedure assuring adequate coverage of very old persons living both in private households and institutional settings. Details about the study and sample design have been reported by Hansen et al. (2021). A total of 1,863 computer assisted personal interviews were conducted at participants' homes to assess a wide array of QoL chances and subjective QoL results. Study nonresponse was found to be unrelated to respondent characteristics such as age, gender and living privately or in long-term care facilities (Kaspar et al., 2022). Mean age of the realized sample at time of interview was 87.0±4.5 years (age range 80.1 to 102.9 years). A total of 211 interviews (11.3%) were conducted in nursing homes. The sample includes 176 interviews with proxy informants in the case of target persons that were willing to be included in the study but were not able to conduct the 90 min interview themselves due to health impairments. Proxy informants had a relationship of mutual trust with the target person. They included 42 (23.9%) spouses, 85 (48.3%) children, 9 (5.1%) professional caregivers, 9 (5.1%) children in law and 31 (17.6%) persons with other types of relationship. Most proxy informants were female (124, 70.4%). Age of proxy informants ranged from 19 to 99 years (mean age: 66.2±14.9 years).

Outcome Measures

Burdensomeness was used as an indicator of *appreciation by others* with the question "Do you have the feeling that society treats you like a burden (e.g., due to physical impairments)?" and a 4-point response scale from 1 (*not at all*) to 4 (*very much*).

The *positive affect* subscale of the short form of the Positive and Negative Affect Schedule (Kercher, 1992) was used as an indicator of a*ppreciation of own life*. Frequency of positive affective states (e.g., "enthusiastic", "excited") across the past 12 months was reported on a five-point response scale from 1 (*never*) to 5 (*very often*). Favorable psychometric properties of this very brief instrument were reported for age-diverse and old-old samples (Hilleras et al., 1998; Kercher, 1992; Mackinnon et al., 1999). Scale consistency in our sample was satisfying (MacDonald's ω = 0.88).

Autonomy was used as indicator of *successful life conduct*. Perceived autonomy was assessed by responses to the question "Do you lead your life according to your own preferences?" on a 4-point scale from 1 (*not at all*) to 4 (*very much*).

Covariates

Regarding *social contact*, respondents were asked how much time they spend with other people such as friends or family. Response options were 1 (*never*) to 5 (*very often*).

Net equivalence income was computed based on monthly all-source net income in private households. For respondents residing in long-term care (LTC), actual costs for accommodation and care per month were considered as a lower threshold estimate of *financial resources*.

Internet use across the last 12 months (yes, no) was considered as a lower bound estimate of digital resources available to the individual. Internet use comprised both the use of computer devices (e.g., laptop, desktop) and digital devices such as smartphones, tablets or wearables running web-based applications.

To measure respondents' reflection of *societal values*, two items (capturing difficulties with modern societal lifestyle, and perceived discrepancy between societal and personal values) from the Future Time Perspective Scale by Brandtstädter et al. (1997) and one item (capturing feelings of disorientation) from the anomia scale suggested by Gümüs et al. (2014) were used. Response options range from 1 (*not at all*) to 4 (*very much*). Scale consistency was found to be moderate in the current sample (MacDonald's ω = 0.69) and comparable to Cronbach's alpha of 0.72 reported for the five-item obsolescence subscale by Brandtstädter et al. (1997).

Independence in ADL (Fillenbaum, 1988; Lawton & Brody, 1969) is used as a measure of *functional health*. Items included seven basic ADL (e.g., eating, dressing, walking) and six instrumental ADL (IADL, e.g., preparing meals, handling finances) with response options 0 *(not possible without help*), 1 (*some help needed*), or 2 (*no help needed*). Reliability of the ADL and IADL scales in the current sample was high (MacDonald's ω = 0.92 and 0.93, respectively). Four domain-heterogeneous parcels of ADL and IADL items were used as observed indicators of the second-order latent construct of functional health (Cole et al., 2016).

The DemTect has been developed as a brief screening tool for *mild cognitive impairment* (MCI) and early stages of *dementia* (Kalbe et al., 2004). It comprises subtests targeting immediate/delayed word recall, digit span, number transcoding and verbal fluency. Favorable diagnostic properties in identifying MCI have been reported (Kalbe et al., 2013) and age-specific scoring instructions for persons 80 years or older have been reported by Kessler et al. (2014). In proxy interviews, cognitive status was reported with the Global Deterioration Scale (GDS) (Reisberg et al., 1982) in seven stages from 1 (*no cognitive impairment*) to 7 (*most severe*). Reisberg et al. (2011) aligned GDS stage 3 to correspond to a clinical presentation of MCI.

Depressive symptoms across the past two weeks were assessed with the short form of the Depression in Old Age Scale (DIA-S4; Heidenblut & Zank, 2014, 2020). The four items ask for the occurrence of lack of motivation, feelings of sadness, worrying, or inability to enjoy life. The DIA-S4 was shown to distinguish groups of depressed and non-depressed geriatric inpatients (Heidenblut & Zank, 2014). In this non-clinical sample, we use a continuous latent variable estimated from reported symptoms as an indicator of depressive state.

Personal values were assessed with the ten-item version of the Portrait Values Questionnaire (PVQ) (Datler et al., 2013; infas, 2006; Schwartz, 2003; Verkasalo et al., 2009). Based on results of a feasibility study, items were modified into direct questions about the importance of certain goals or areas of life with response options 1 (*not important at all*) to 4 (*very important*). In line with previous research on the dimensionality of the PVQ (Verkasalo et al., 2009), responses in the current sample could best be represented by a three-factor model comprising „self-transcendence“ (with the markers universalism, benevolence, self-direction, and tradition), „self-enhancement“ (containing conformity, security and stimulation), and „conservation vs. openness to change“ (containing power, achievement, hedonism and stimulation) (Reissmann et al., 2021). From proxy reports on the value-system of target persons with limited communicative abilities, only a single factor (termed “generic value orientation” here) could be extracted.

Plan of Analysis

In a first step, we describe levels and heterogeneity of person and environment chances as well as the prevalence of QoL results in very old age. All analyses include available proxy information in case the targeted individual was unable to respond (i.e. persons with limited ability to communicate, PLC) to avoid under-representation of this most vulnerable part of the population. However, because perspectives of persons able to communicate themselves (PC) and proxies of PLC may diverge based on item content (e.g., factual versus subjective), analyses were conducted in parallel for both subgroups. For multi-item scales, measurement equivalence was tested using a structural equation model multi-group (MG-SEM) approach and cross-group parameter constraints of factor loadings and item intercepts. If at least partial measurement invariance (MI) was established, information for the whole population with combined PC and PLC data was additionally reported and the expected bias when omitting PLC from the analysis was discussed.

In a second step, we use multiple linear regression to estimate the relative contribution of personal and environmental chances to predict differences in QoL results (i.e. positive affect, autonomy, perceived appreciation by others). Generalized logistic regressions were estimated for the ordinal outcome measures autonomy and perceived burdensomeness. In addition, we use correlation analysis to estimate the empirical overlap between these QoL results in very old age.

All analyses used weights to correct for the disproportional sampling design and survey as well as panel nonresponse (Valliant et al., 2013). Standard errors of parameter estimates were corrected for the clustering due to the two-stage sampling procedure. All analyses were performed using SAS 9.4 (SAS Institute, Cary, NC) and Mplus 8.4 (Muthén & Muthén 1998-2017) software.

Results

Descriptive Results

Estimated population means and prevalences for QoL chances and results are given in Table 1 for the PC and PLC subgroups separately.

The majority of older adults does not report feelings of being treated like a burden by society. However, one in eight adults aged 80 years or older reports such experiences quite a bit or very much. The share of older adults reporting such potentially ageist sentiments in PLC is approximately double the size of that in PC (i.e., 24.9% versus 12.9%).

With respect to positive affect, MI testing suggested that latent scores from proxy reports could be validly combined with those from self-reports to estimate levels of positive appraisal of own life in the very old population. Participants capable of self-report showed an overall mean of 3.3 on the five-point scale, corresponding to the “sometimes” and “often” categories. Average positive affect in PLC was significantly lower (2.6).

On the four-point autonomy scale, more than nine out of ten PC reported high levels of perceived autonomy (very much or quite a bit). In contrast, the majority of PLC (60.2%) was characterized as not living their lives according to their own preferences by proxy informants.

Frequency of social contact was high and similar for PLC (3.4) and PC (3.6). Average financial resources were lower for PC than for PLC (1,944€ vs. 2,581€, respectively). As Figure 2 illustrates, monthly costs for people receiving full inpatient care were higher than the net equivalence income reported for community-dwelling people or people in less institutionalized care settings (e.g., assisted living). In PC, more than one in five (21.2%) participants have been actively using the internet in the past 12 months. In contrast, this rate was considerably lower (1.6%) in PLC. Participants reported average scores of 2.5 (PC) and 2.4 (PLC) with regards to anomic feelings, representing values in the middle area of the scale. MI of the latent variable "anomie" has been established across PC and PLC groups, allowing for direct group comparison. Average anomie scores in both groups were not statistically different from one another.

Functional health was high in PC (1.6 on the 0–2-point scale) but well beyond scale midpoint in PLC (0.5). However, tests for MI of the functional health scale showed that scores on the latent variable cannot be validly compared across PC and PLC groups. With respect to cognitive status, the overall prevalence of beginning dementia was 9.8% in PC. In contrast, 73.5% of PLC showed symptoms of at least beginning dementia. The overall prevalence of MCI in PC was 16.4%. In PLC, the overall prevalence of MCI was 7.4%. Latent estimates of depressive symptoms were higher in PLC (1.6) than in PC (0.3). Tests for MI indicated that information from PC and PLC groups could validly be compared and also combined to estimate levels of depressive symptoms in the very old population. With respect to the system of values held by older adults, no factor score means are presented in Table 1 because these have been fixed at zero in the exploratory structural equations model. However, details of modelling value factors based on the Schwartz scale and value priorities in this sample have been described elsewhere (Reissmann et al., 2021).

Validity of the CHAPO Model

Associations between different facets of QoL results in very old age (Table 2) were found to be of medium to large effect size according to the classification suggested by Brydges (2019) for gerontological research (r = .2 and .3, respectively). The strongest link was observed between positive affect and perceived autonomy (PC: r = .41, PLC: r = .47). Older individuals who feel being treated as a burden to society, however, showed lower levels of autonomy and vice versa. While this moderate effect was apparent irrespective of study group, a negative association of feelings of burdensomeness with positive affect was only found in PC (r = -.18).

Approximately 32% of variance in self-reported *burdensomeness* could be explained by QoL chances, whereas individual background characteristics of PC such as gender or age had no predictive value (Table 3). The risk of negative appraisals by others was significantly reduced with higher functional health and more self-transcending values. In contrast, higher feelings of anomie strongly increased the risk of experiencing negative appraisal by others. Interestingly, higher financial resources were slightly but significantly related to higher risk of being treated like a burden. In PLC, the set of predictors was limited in its potential to systematically explain differences in feelings of being treated like a burden to society, as none of the estimated Odds Ratios reached statistical significance.

Model determination for *positive affect* was above .46 in both study groups, indicating a strong contingency of subjective positive experiences upon individual and environmental life chances. Self-enhancement values were strong positive predictors of SWB, while higher self-transcendence values were weakly and negatively associated with positive affect. Depressive symptoms were also weakly and negatively related to positive affect. Social contact as an environmental resource showed a positive and moderate effect on well-being. A small but significant contribution to well-being in very old age was also found for internet use. In the PLC subgroup, a similar pattern emerged. Global value orientation was strongly related to positive affect. Depressive symptoms exhibited a stronger effect on well-being in PLC. Moreover, functional health, cognitive impairment and societal values (i.e., anomie) played a more critical role in explaining variance in positive affect in this more vulnerable subpopulation.

The set of predictors was able to explain 33% of variance in *autonomy* in PC and 37% in the more vulnerable (and less autonomous) subgroup of PLC. In the latter, autonomy appeared to be solely and strongly dependent on functional health. The pattern of predictive relationships was more differentiated in the PC subpopulation. Here, in addition to strong effects for functional health and moderate effects for depressive symptoms, self-enhancement values and internet use moderately contributed to higher perceived autonomy. More financial resources (or financial turnaround in persons living in nursing homes) were associated with less perceived autonomy.

In summary, the set of QoL chances considered here were able to explain up to nearly half of the variance observed in QoL results in very old adults. However, personal and environmental life chances contributed differently to distinct aspects of QoL results. Moreover, QoL results in the subpopulation of PLC appeared to be influenced by specific (additional) determinants. Individual background characteristics such as age group or gender did not predict QoL results in this sample of older adults when QoL chances within the person and the environment were simultaneously considered.

Discussion

This paper gives a comprehensive and multidimensional account of QoL in the population aged 80 years or older in Germany. By including persons with limited ability to communicate in this population-based assessment, the results decisively expand the state of research.

The descriptive results on the three outcome measures of *life results* showed that respondents did not fully negate *feelings of being treated as a burden to society*. This is all the more concerning against the background that lacking appreciation was also associated with lower well-being and less perceived autonomy in this study. An “acceptance problem” of very old age becomes evident even if maybe not as pronounced as described by Fürstenberg (2013). According to van Dyk (2009), the programmatic “discovery” of potentials of old age in response to lopsided deficit models of ageing has led to societal pressure on individuals to preserve and use those potentials for the benefits of society, while those who do not live up to such normative expectations are subject to devaluation and marginalization.

Descriptive results on autonomy and positive affect run contrary to the deficit-oriented model of ageing. Participants reported frequent positive affect as a marker of *appreciation of own life*. In PLC, positive affect is less frequent. A high degree of autonomy was observed in PC, signaling *successful life-conduct* in very old age. Reported autonomy was, however, lower in PLC. In all three examined life results, PLC scored lower than PC, indicating a special need for support in this group.

The main results showed that – in line with the CHAPO-Model – a multidimensional discussion of QoL in very old age needs to also consider QoL results beyond appreciation of own life and include both appreciation by others and successful life conduct as interrelated, yet distinct QoL facets. In addition, findings underscore that life chances are systematically associated with late-life QoL results. Unique contributions were observed for different indicators with respect to different life results and between PCs and PLCs.

Frequency of *social contact* was a significant predictor of positive affect but unrelated to autonomy and burdensomeness. The effect on positive affect is in line with existing evidence from other countries and age groups (Netuveli et al., 2006; Pinquart & Sörensen, 2000). So far, research on the association between social relations and autonomy has mainly focused on social support (Cimarolli et al., 2006; Warner et al., 2011). Social support may be more relevant for autonomy than the frequency of contact itself. It is also possible that a more differentiated assessment of social contact (e.g., regarding contact mode or relationship) is needed to better understand associations with autonomy or perceived burdensomeness.

Financial resources were not significantly associated with positive affect but were found to signal a risk factor for autonomy in PC, but also for burdensomeness in PLC. These findings might appear contradictory to existing evidence, but they can be traced back to the higher financial turnover in individuals residing in LTC compared to individuals in private households in this study.

With respect to *digital resources*, findings that only 21% of those 80 years or older are online shows a wide digital gap between generations (Initiative D21, 2020). While the growth of the online population is expected to continue (e.g., today's ageing generation being more experienced in the use of technology; Erickson & Johnson, 2011; van Eimeren & Frees, 2014), the gap in very old age is expected to persist, as age-related losses in (e.g. sensory, cognitive) abilities impede use of (new) technology (Seifert & Schelling, 2015). In fact, the small share of onliners (1.6%) in PLC is in line with previous findings that good physical health and cognitive functioning are predictors of internet use (Erickson & Johnson, 2011). Internet use contributed to the prediction of positive affect and autonomy in this sample. However, Huang (2010) reported small but negative associations between internet use and well-being in her meta-analysis. Detailed examination of sources for this contradictory evidence is beyond the scope of this paper and readers are referred to a more thorough discussion of antecedents and consequences of ICT use in this sample (Schlomann, Seifert, Zank, & Rietz, 2020; Schlomann, Seifert, Zank, Woopen, & Rietz, 2020).

This study drew on the concept of *anomie* as important marker of life chances. Our results indicate that anomic feelings may be not considerably higher in very old age than in a younger sample (mean age 73; Schmidt, 2017) and suggest that the subjective relationship with society as a whole may remain stable after a shift of personal priorities that occurs already before the age of 80. Anomie was a significant predictor of burdensomeness and positive affect (particularly in PLC). Hence, our findings do not support Schmidt's (2017) claim that anomie as perceived distance from the outside world may itself be adaptive in the sense of disengagement theory (Cumming & Henry, 1961). Structural gaps in the recognition of priorities of today's very old population or lack of opportunity for societal engagement do not create a "privileged" space for coping with developmental tasks in late life.

Tests for MI indicated that the measure of *functional health* was not comparable in PC and PLC and we observed differences in its predictive power across study groups. Functional health predicted positive affect only in the more vulnerable PLC, in which adaptation of well-being despite health impairments as supposed by theories of psychological coping and self-regulation (Schilling et al., 2013) might no longer be possible. Similarly, feelings of burdensomeness were found to be unrelated to functional health only in PLC. We speculate that in this subgroup of individuals with pronounced functional impairment and elevated levels of perceived burdensomeness, negative external appraisal may be more strongly related to specific support or care constellations (e.g., no caregiving relative) than to functional status itself. Not surprisingly, functional health was the strongest predictor of autonomy in all participants. Similar effects have been reported for community-dwelling (Henchoz et al., 2019; Sánchez-García et al., 2019) and institutionalized individuals (Liu et al., 2020). Hence, our findings are in line with expectations from self-determination theory (Hertz, 1996) that low ADL impairment results in high perceived ability to act autonomously in meeting needs and goals.

Although PLC showed lower levels of *cognitive functioning* than PC, relationships with respect to age and gender were comparable. Because PLC are often excluded from surveys, comparison of prevalence for MCI and dementia estimated in PLC in this study with existing studies is limited. Research with the Quality of Life-Alzheimer’s Disease scale has shown that cognition has a greater impact on QoL outcomes in proxy ratings than in self-ratings (Römhild et al., 2018). Our results for the prediction of positive affect are in line with these well-established results. While research emphasizes the importance of involving people with dementia in (supported) decision-making (Wied et al., 2019), we found no indication of an adverse effect of impaired cognition on perceived autonomy in our sample.

Depressive symptoms predicted positive affect and autonomy in both study groups. However, effect sizes were small in PC and moderate to large in PLC. Contrary to our expectation, depressive symptoms were not found to be related to the feeling of burdensomeness.

In respondents able to give detailed report on their *personal values*, a differentiated pattern of relationships with QoL results was observed. The finding that self-transcendence was negatively related to positive affect was unexpected, as palliative care patients with such “macro worries” have been found to show higher well-being (concerned with self-enhancement values; Fegg et al., 2005). In line with our finding of a positive predictive value of self-enhancement, we speculate that well-being for most individuals in their eighties and nineties is still contingent upon agency and priorities for engagement and self-enhancement. In contrast, self-transcendence values may signal concern for others (e.g. in caregiving relationships) that represent risks to one's own well-being. This interpretation is also in line with our additional finding that self-transcendence values are protective against negative external appraisal. Although a shift towards conservation with ageing has been discussed as an adaptive process (I. Borg et al., 2017) and has been found to be positively related to QoL results (Fegg et al., 2005; Fung et al., 2016; Robinson, 2013), we found no evidence of associations with the range of QoL results considered here. In the PLC subgroup, the single generic value orientation factor was positively related to positive affect, highlighting the importance of explicating one's priorities for successful life conduct with more pronounced health impairment.

Limitations and Future Directions

Smaller sample size effectively limited the potential to statistically substantiate small effects in PLC. We still consider including them in a population-based sample an important improvement as compared to excluding them completely. It allowed us to assess MI and to compare QoL results directly. Explained total variance was higher in PLC, suggesting higher relevance of the explanative model in more vulnerable individuals. Interestingly, whereas burdensomeness was negatively associated with positive affect in PC, no such effect was observed in the subpopulation of very old individuals represented in the survey by interviews with family members or close affiliates. We cautiously consider this a hint towards a protective effect of an engaged social network that buffers and compensates for negative external appraisal. However, this interpretation needs further substantiation as information bias in proxy interviews might have impacted the results.

Actual costs for respondents residing in LTC facilities were used in this study as a useful marker for financial asset in the institutionalized population; nevertheless, correspondence with net household income in private settings is limited. Second, although data represents adults 80+ in Germany's most populous Federal state, generalizability to other regions may be limited.

Finally, further insights into mechanisms linking QoL chances to results will be possible upon completion of the second wave of data collection. Using panel data instead of cross-sectional data, changes in QoL results can analyzed over time.

Statements and Declarations

The study was advised by the ethical board of the medical faculty at University of Cologne (17-169). The observational study NRW80+ was preregistered at the German Clinical Trials Register (DRKS00011924).

The authors have no relevant financial or non-financial interests to disclose.

References

Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 1–34).

Bobowik, M., Basabe, N., Páez, D., Jiménez, A., & Bilbao, M. Á. (2011). Personal values and well-being among Europeans, Spanish natives and immigrants to Spain: Does the culture matter? *Journal of Happiness Studies, 12*(3), 401–419. https://doi.org/10.1007/s10902-010-9202-1

Borg, C., Hallberg, I. R., & Blomqvist, K. (2006). Life satisfaction among older people (65+) with reduced self-care capacity: The relationship to social, health and financial aspects. *J Clin Nurs, 15*(5), 607–618. https://doi.org/10.1111/j.1365-2702.2006.01375.x

Borg, I., Hertel, G., & Hermann, D. (2017). Age and personal values: Similar value circles with shifting priorities. *Psychology and Aging, 32*(7), 636–641. https://doi.org/10.1037/pag0000196

Botes, R., Vermeulen, K. M., Gerber, A. M., Ranchor, A. V., & Buskens, E. (2019). Health-related quality of life and well-being health state values among Dutch oldest old. *Patient Preference and Adherence, 13*, 721–728. https://doi.org/10.2147/PPA.S193171

Brandtstädter, J., & Wentura, D. (1994). Veränderungen der Zeit- und Zukunftsperspektive im Übergang zum höheren Erwachsenenalter: Entwicklungspsychologische und differentielle Aspekte. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 26(1), 2–21.*

Brandtstädter, J., Wentura, D., & Schmitz, U. (1997). Veränderungen der Zeit- und Zukunftsperspektive im Übergang zum höheren Alter: Quer- und längsschnittliche Befunde. *Zeitschrift für Psychologie, 205*(4), 377–395.

Brydges, C. R. (2019). Effect size guidelines, sample size calculations, and statistical power in gerontology. *Innovation in Aging, 3*(4), igz036. https://doi.org/10.1093/geroni/igz036

Burks, H. B., Des Bordes, J. K. A., Chadha, R., Holmes, H. M., & Rianon, N. J. (2021). Quality of life assessment in older adults with dementia: A systematic review. *Dementia and Geriatric Cognitive Disorders, 50*(2), 103–110. https://doi.org/10.1159/000515317

Carstensen, L. L. (1991). Socioemotional selectivity theory: Social activity in life-span context. *Annual Review of Gerontology and Geriatrics, 11*, 195–217.

Cho, J., Martin, P., Margrett, J., Macdonald, M., & Poon, L. W. (2011). The Relationship between physical health and psychological well-being among oldest-old adults. *Journal of Aging Research, 2011*, 605041. https://doi.org/10.4061/2011/605041

Cimarolli, V. R., Reinhardt, J. P., & Horowitz, A. (2006). Perceived overprotection: Support gone bad? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 61*(1), 18-23. https://doi.org/10.1093/geronb/61.1.s18

Cole, D. A., Perkins, C. E., & Zelkowitz, R. L. (2016). Impact of homogeneous and heterogeneous parceling strategies when latent variables represent multidimensional constructs. *Psychological Methods, 21*(2), 164–174. https://doi.org/10.1037/met0000047

Cornwell, B., & Schafer, M. H. (2016). Social networks in later life. In L.K. George & K.F. Ferraro (Eds.) *Handbook of Aging and the Social Sciences* (pp. 181–201). Elsevier. https://doi.org/10.1016/B978-0-12-417235-7.00009-3

Cumming, E., & Henry, W. E. (1961). *Growing old. The process of disengagement.* Basic Books.

Czaja, S. J., & Lee, C. C. (2007). The impact of aging on access to technology. *Universal Access in the Information Society, 5*(4), 341–349. https://doi.org/10.1007/s10209-006-0060-x

Datler, G., Jagodzinski, W., & Schmidt, P. (2013). Two theories on the test bench: Internal and external validity of the theories of Ronald Inglehart and Shalom Schwartz. *Social Science Research, 42*(3), 906–925. https://doi.org/10.1016/j.ssresearch.2012.12.009

Dionigi, R. A. (2015). Stereotypes of aging: Their effects on the health of older adults. *Journal of Geriatrics, 2015*(1), 1–9. https://doi.org/10.1155/2015/954027

Ebner, N. C., Freund, A. M., & Baltes, P. B. (2006). Developmental changes in personal goal orientation from young to late adulthood: From striving for gains to maintenance and prevention of losses. *Psychology and Aging, 21*(4), 664–678. https://doi.org/10.1037/0882-7974.21.4.664

Erickson, J., & Johnson, G. M. (2011). Internet use and psychological wellness during late adulthood. *Canadian Journal on Aging, 30*(2), 197–209. https://doi.org/10.1017/S0714980811000109

Erikson, E. H. (1998). *Life cycle completed: A review.* W.W.Norton.

Fegg, M. J., Wasner, M., Neudert, C., & Borasio, G. D. (2005). Personal values and individual quality of life in palliative care patients. *Journal of Pain and Symptom Management, 30*(2), 154–159. https://doi.org/10.1016/j.jpainsymman.2005.02.012

Fillenbaum, G. G. (1988). *Multidimensional functional assessment of older Adults: The Duke Older Americans Resources and Services Procedures.* Duke University Center for the Study of Aging and Human Development.

Fung, H. H., Ho, Y. W., Zhang, R., Zhang, X., Noels, K. A., & Tam, K.-P. (2016). Age differences in personal values: Universal or cultural specific? *Psychology and Aging, 31*(3), 274–286. https://doi.org/10.1037/pag0000082

Fürstenberg, F. (2013). Very old age as a problem of acceptance: A position paper. *Zeitschrift für Gerontologie und Geriatrie, 46*(1), 16–20. https://doi.org/10.1007/s00391-012-0460-x

Gümüs, A., Gömleksiz, M., Glöckner-Rist, A., & Balke, D. (2014). *Anomie: Zusammenstellung sozialwissenschaftlicher Items und Skalen.* ZIS - GESIS Leibniz Institute for the Social Sciences. https://doi.org/10.6102/zis145

Hansen, S., Kaspar, R., Wagner, M., Woopen, C., & Zank, S. (2021). The NRW80+ study: Conceptual background and study groups. *Zeitschrift für Gerontologie und Geriatrie, 54*(2), 76–84. https://doi.org/10.1007/s00391-021-01970-z

Hawton, A., Green, C., Dickens, A. P., Richards, S. H., Taylor, R. S., Edwards, R., Greaves, C. J., & Campbell, J. L. (2011). The impact of social isolation on the health status and health-related quality of life of older people. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation, 20*(1), 57–67. https://doi.org/10.1007/s11136-010-9717-2

Heckhausen, J. (1997). Developmental regulation across adulthood: Primary and secondary control of age-related challenges. *Developmental Psychology, 33,* 176–187.

Heckhausen, J., Wrosch, C., & Schulz, R. (2013). A lines-of-defense model for managing health threats: A review. *Gerontology, 59*(5), 438–447. https://doi.org/10.1159/000351269

Heidenblut, S., & Zank, S. (2014). Screening for depression with the Depression in Old Age Scale (DIA-S) and the Geriatric Depression Scale (GDS15). *GeroPsych, 27*(1), 41–49. https://doi.org/10.1024/1662-9647/a000101

Heidenblut, S., & Zank, S. (2020). Screening for depression in old age with very short instruments: The DIA-S4 compared to the GDS5 and GDS4. *Gerontology & Geriatric Medicine, 6,* 2333721420981328. https://doi.org/10.1177/2333721420981328

Henchoz, Y., Abolhassani, N., Büla, C., Guessous, I., Goy, R., & Santos-Eggimann, B. (2019). Change in quality of life among community-dwelling older adults: Population-based longitudinal study. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation, 28*(5), 1305–1314. https://doi.org/10.1007/s11136-019-02108-w

Heo, J., Chun, S., Lee, S., Lee, K. H., & Kim, J. (2015). Internet use and well-being in older adults. *Cyberpsychology, Behavior and Social Networking, 18*(5), 268–272. https://doi.org/10.1089/cyber.2014.0549

Hertz, J. E. (1996). Conceptualization of perceived enactment of autonomy in the elderly. *Issues in Mental Health Nursing, 17*(3), 261–273. https://doi.org/10.3109/01612849609049919

Hilleras, P. K., Jorm, A. F., Herlitz, A., & Winblad, B. (1998). Negative and positive affect among the very old. *Research on Aging, 20*(5), 593–610. https://doi.org/10.1177/0164027598205003

Huang, C. (2010). Internet use and psychological well-being: A meta-analysis. *Cyberpsychology, Behavior and Social Networking, 13*(3), 241–249. https://doi.org/10.1089/cyber.2009.0217

Huxhold, O., Fiori, K. L., & Windsor, T. D. (2013). The dynamic interplay of social network characteristics, subjective well-being, and health: The costs and benefits of socio-emotional selectivity. *Psychology and Aging, 28*(1), 3–16. https://doi.org/10.1037/a0030170

infas. (2006). *Weltwertestudie: Fragebogen Version A*. infas.

Initiative D21. (2020). *D21-Digital-Index 2019/ 2020: Wie digital ist Deutschland? Jährliches Lagebild zur Digitalen Gesellschaft* (1. Auflage). D21-Digital-Index. Initiative D21.

Jopp, D., Rott, C., & Oswald, F. (2008). Valuation of life in old and very old age: The role of sociodemographic, social, and health resources for positive adoption. *The Gerontologist, 48*(5), 646–658. https://doi.org/10.1093/geront/48.5.646

Jyväkorpi, S. K., Urtamo, A., Pitkälä, K. H., & Strandberg, T. E. (2018). Nutrition, daily walking and resilience are associated with physical function in the oldest old men. *The Journal of Nutrition, Health & Aging, 22*(10), 1176–1182. https://doi.org/10.1007/s12603-018-1136-z

Kalbe, E., Calabrese, P., Fengler, S., & Kessler, J. (2013). Demtect, PANDA, EASY, and MUSIC: Cognitive screening tools with age correction and weighting of subtests according to their sensitivity and specificity. *Journal of Alzheimer's Disease: JAD, 34*(4), 813–834. https://doi.org/10.3233/JAD-122128

Kalbe, E., Kessler, J., Calabrese, P., Smith, R., Passmore, A. P., Brand, M., & Bullock, R. (2004). Demtect: A new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. *International Journal of Geriatric Psychiatry, 19*(2), 136–143. https://doi.org/10.1002/gps.1042

Kaspar, R. (2004). Technology and loneliness in old age. *Gerontechnology, 3*(1), 42–48. https://doi.org/10.4017/gt.2004.03.01.007.00

Kaspar, R., Brijoux, T., Albrecht, A., Zimmermann, J., Wenner, J., Fey, J., Reissmann, M., Wagner, M., & Zank, S. (2023). Challenges and benefits of including the institutionalized, non-communicative and cognitively impaired in a representative survey of the very old. *Survey Research Methods*

Kercher, K. (1992). Assessing subjective well-being in the old-old: The PANAS as a measure of orthogonal dimensions of positive and negative affect. *Research on Aging, 14*(2), 131–168. https://doi.org/10.1177/0164027592142001

Knesebeck, O. von dem, Wahrendorf, M., Hyde, M., & Siegrist, J. (2007). Socio-economic position and quality of life among older people in 10 European countries: Results of the SHARE study. *Ageing and Society, 27*(2), 269–284. https://doi.org/10.1017/S0144686X06005484

Kruse, A. (2021). *Vom Leben und Sterben im Alter: Wie wir das Lebensende gestalten können.* W. Kohlhammer Verlag.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist* (9), 179–186.

Leontjevas, R., Gerritsen, D. L., Vernooij-Dassen, M. J. F. J., Smalbrugge, M., & Koopmans, R. T. C. M. (2012). Comparative validation of proxy-based Montgomery-Åsberg depression rating scale and Cornell scale for depression in dementia in nursing home residents with dementia. *The American Journal of Geriatric Psychiatry, 20*(11), 985–993. https://doi.org/10.1097/JGP.0b013e318233152b

Levy, B. R. (2009). Stereotype embodiment. *Current Directions in Psychological Science, 18*(6), 332–336. https://doi.org/10.1111/j.1467-8721.2009.01662.x

Levy, B. R., Slade, M. D., Kunkel, S. R., & Kasl, S. V. (2002). Longevity increased by positive self-perceptions of aging. *Journal of Personality and Social Psychology, 83*(2), 261–270. https://doi.org/10.1037//0022-3514.83.2.261

Liu, L.-H., Kao, C.-C., & Ying, J. C. (2020). Functional capacity and life satisfaction in older adult residents living in long-term care facilities: The mediator of autonomy. *The Journal of Nursing Research: JNR, 28*(4), e102. https://doi.org/10.1097/JNR.0000000000000362

Mackinnon, A., Jorm, A. F., Christensen, H., Korten, A. E., Jacomb, P. A., & Rodgers, B. (1999). A short form of the Positive and Negative Affect Schedule: Evaluation of factorial validity and invariance across demographic variables in a community sample. *Personality and Individual Differences, 27*(3), 405–416. https://doi.org/10.1016/S0191-8869(98)00251-7

Martyr, A., Nelis, S. M., Quinn, C., Wu, Y.-T., Lamont, R. A., Henderson, C., Clarke, R., Hindle, J. V., Thom, J. M., Jones, I. R., Morris, R. G., Rusted, J. M., Victor, C. R., & Clare, L. (2018). Living well with dementia: A systematic review and correlational meta-analysis of factors associated with quality of life, well-being and life satisfaction in people with dementia. *Psychological Medicine, 48*(13), 2130–2139. https://doi.org/10.1017/S0033291718000405

Medienpädagogischer Forschungsverbund Südwest (Ed.). (2022). *SIM-Studie 2021: Senior*innen, Information, Medien. Basisuntersuchung zum Medienumgang von Personen ab 60 Jahren in Deutschland.* Stuttart. https://www.mpfs.de/fileadmin/files/Studien/SIM/2021/Web_SIM-Studie2021_final_barrierefrei.pdf

Nahemow, L., & Lawton, M. P. (1973). Toward an ecological theory of adaptation and aging. *Environmental Design Research, 1*(24), 24–32.

Neise, M., Janhsen, A., Geithner, L., Schmitz, W., & Kaspar, R. (2019). Lebensqualitäten in der Hochaltrigkeit. In K. Hank, F. Schulz-Nieswandt, M. Wagner, & S. Zank (Eds.), *Alternsforschung: Handbuch für Wissenschaft und Praxis* (1st ed., pp. 583–607). Nomos.

Netuveli, G., Wiggins, R. D., Hildon, Z., Montgomery, S. M., & Blane, D. (2006). Quality of life at older ages: Evidence from the English longitudinal study of aging (wave 1). *Journal of Epidemiology and Community Health, 60*(4), 357–363. https://doi.org/10.1136/jech.2005.040071

Pinquart, M., & Sörensen, S. (2000). Influences of socioeconomic status, social network, and competence on subjective well-being in later life: A meta-analysis. *Psychology and Aging, 15*, 187–224.

Pratschke, J., Haase, T., & McKeown, K. (2017). Direct and indirect influences of socio-economic position on the wellbeing of older adults: A structural equation model using data from the first wave of the Irish Longitudinal Study on Ageing. *Ageing and Society, 37*(9), 1770–1797. https://doi.org/10.1017/S0144686X1600060X

Reisberg, B., Ferris, S. H., Leon, M. J. de, & Crook, T. (1982). The Global Deterioration Scale for assessment of primary degenerative dementia. *The American Journal of Psychiatry, 139*(9), 1136–1139. https://doi.org/10.1176/ajp.139.9.1136

Reisberg, B., Jamil, I. A., Khan, S., Monteiro, I., Torossian, C., Ferris, S., Sabbagh, M., Gauthier, S., Auer, S., Shulman, M. B., Kluger, A., Franssen, E., & Wegiel, J. (2011). Staging Dementia. In A. Kumar, C. L. E. Katona, & M. T. Abou-Saleh (Eds.), *Principles and practice of geriatric psychiatry* (3rd ed., pp. 162–169). Wiley. https://doi.org/10.1002/9780470669600.ch31

Reissmann, M., Storms, A., & Woopen, C. (2021). Individual values and spirituality and their meaning for affective well-being and engagement with life in very old age. *Zeitschrift für Gerontologie und Geriatrie, 54*(2), 85–92. https://doi.org/10.1007/s00391-021-01974-9

Robinson, O. C. (2013). Values and adult age: Findings from two cohorts of the European Social Survey. *European Journal of Ageing, 10*(1), 11–23. https://doi.org/10.1007/s10433-012-0247-3

Römhild, J., Fleischer, S., Meyer, G., Stephan, A., Zwakhalen, S., Leino-Kilpi, H., Zabalegui, A., Saks, K., Soto-Martin, M., Sutcliffe, C., Rahm Hallberg, I., & Berg, A. (2018). Inter-rater agreement of the Quality of Life-Alzheimer's Disease (QoL-AD) self-rating and proxy rating scale: Secondary analysis of RightTimePlaceCare data. *Health and Quality of Life Outcomes, 16*(1), 131. https://doi.org/10.1186/s12955-018-0959-y

Sagiv, L., Roccas, S., & Hazan, O. (2004). Value pathways to well-being: Healthy values, valued goal attainment, and environmental congruence. In P. A. Linley & S. Joseph (Eds.), *Positive psychology in practice* (pp. 68–85). John Wiley & Sons, Inc. https://doi.org/10.1002/9780470939338.ch5

Sánchez-García, S., García-Peña, C., Ramírez-García, E., Moreno-Tamayo, K., & Cantú-Quintanilla, G. R. (2019). Decreased autonomy in community-dwelling older adults. *Clinical Interventions in Aging, 14*, 2041–2053. https://doi.org/10.2147/CIA.S225479

Schaller, S. (2008). Transactional model of suicidal behavior in the elderly. *Zeitschrift für Gerontologie und Geriatrie, 41*(1), 14–21. https://doi.org/10.1007/s00391-008-0518-y

Schilling, O. K. (2006). Development of life satisfaction in old age: Another view on the "paradox". *Social Indicators Research, 75*(2), 241–271. https://doi.org/10.1007/s11205-004-5297-2

Schilling, O. K., Wahl, H.-W., & Oswald, F. (2013). Change in life satisfaction under chronic physical multi-morbidity in advanced old age: Potential and limits of adaptation. *Journal of Happiness Studies, 14*(1), 19–36. https://doi.org/10.1007/s10902-011-9313-3

Schlomann, A., Seifert, A., Zank, S., & Rietz, C. (2020). Assistive technology and mobile ICT usage among oldest-old cohorts: Comparison of the oldest-old in private homes and in long-term care facilities. *Research on Aging, 42*(5-6), 163–173. https://doi.org/10.1177/0164027520911286

Schlomann, A., Seifert, A., Zank, S., Woopen, C., & Rietz, C. (2020). Use of information and communication technology (ICT) devices among the oldest-old: Loneliness, anomie, and autonomy. *Innovation in Aging, 4*(2), igz050. https://doi.org/10.1093/geroni/igz050

Schmidt, L. (2017). *Technikhandhabung im höheren Alter: Zur Rolle von kognitiver Leistungsfähigkeit, Technikeinstellung und Technikerfahrung.* Springer Fachmedien Wiesbaden. http://gbv.eblib.com/patron/FullRecord.aspx?p=4741454

Schwartz, S. H. (2003). *A Proposal for Measuring Value Orientations across Nations.* The Hebrew University of Jerusalem.

Seifert, A., & Schelling, H. R. (2015). *Digitale Senioren: Nutzung von Informations- und Kommunikationstechnologien (IKT) durch Menschen ab 65 Jahren in der Schweiz im Jahr 2015*. Pro Senectute Schweiz. https://doi.org/10.13140/RG.2.1.4183.5600

Shapira, N., Barak, A., & Gal, I. (2007). Promoting older adults' well-being through internet training and use. *Aging & Mental Health, 11*(5), 477–484. https://doi.org/10.1080/13607860601086546

Smith, K. J., & Victor, C. R. (2019). Typologies of loneliness, living alone and social isolation, and their associations with physical and mental health. *Ageing and Society, 39*(8), 1709–1730. https://doi.org/10.1017/S0144686X18000132

Sneeuw, K. C., Aaronson, N. K., Haan, R. J. de, & Limburg, M. (1997). Assessing quality of life after stroke: The value and limitations of proxy ratings. *Stroke, 28*(8), 1541–1549. https://doi.org/10.1161/01.str.28.8.

Szabo, A., Allen, J., Stephens, C., & Alpass, F. (2019). Longitudinal analysis of the relationship between purposes of internet use and well-being among older adults. *The Gerontologist, 59*(1), 58–68. https://doi.org/10.1093/geront/gny036

Tesch-Römer, C., & Wahl, H.-W. (2017). Toward a more comprehensive concept of successful aging: Disability and care needs. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 72*(2), 310–318. https://doi.org/10.1093/geronb/gbw162

Tornstam, L. (1997). Gerotranscendence: The contemplative dimension of aging. *Journal of Aging Studies, 11*(2), 143–154. https://doi.org/10.1016/S0890-4065(97)90018-9

Tulviste, T., Kall, K., & Rämmer, A. (2017). Value priorities of younger and older adults in seven European countries. *Social Indicators Research, 133*(3), 931–942. https://doi.org/10.1007/s11205-016-1392-4

Valliant, R., Dever, J. A., & Kreuter, F. (2013). *Practical tools for designing and weighting survey samples.* Springer New York. https://doi.org/10.1007/978-1-4614-6449-5

van Dyk, S. (2009). Das Alter: adressiert, aktiviert, diskriminiert. *Berliner Journal für Soziologie, 19*(4), 601–625. https://doi.org/10.1007/s11609-009-0114-z

van Eimeren, B., & Frees, B. (2014). 79 Prozent der Deutschen online – Zuwachs bei mobiler Internetnutzung und Bewegtbild: Ergebnisse der ARD/ZDF-Onlinestudie 2014. *MEDIA PERSPEKTIVEN(7-8)*, 378–396. https://www.ard-werbung.de/fileadmin/user_upload/media-perspektiven/pdf/2014/0708-2014_Eimeren_Frees_.pdf

Veenhoven, R. (2000). The Four Qualities of Life: Ordering concepts and measures of the good life. *Journal of Happiness Studies, 1*(1), 1–39. https://doi.org/10.1023/A:1010072010360

Verkasalo, M., Lönnqvist, J.-E., Lipsanen, J., & Helkama, K. (2009). European norms and equations for a two dimensional presentation of values as measured with Schwartz's 21-item portrait values questionnaire. *European Journal of Social Psychology, 39*(5), 780–792. https://doi.org/10.1002/ejsp.580

Vuorisalmi, M., Sarkeala, T., Hervonen, A., & Jylhä, M. (2012). Among nonagenarians, congruence between self-rated and proxy-rated health was low but both predicted mortality. *Journal of Clinical Epidemiology, 65*(5), 553–559. https://doi.org/10.1016/j.jclinepi.2011.11.001

Wagner, M., Rietz, C., Kaspar, R., Janhsen, A., Geithner, L., Neise, M., Kinne-Wall, C., Woopen, C., & Zank, S. (2018). Quality of life of the very old: Survey on quality of life and subjective well-being of the very old in North Rhine-Westphalia (NRW80+). *Zeitschrift für Gerontologie und Geriatrie, 51*(2), 193–199. https://doi.org/10.1007/s00391-017-1217-3

Wahl, H.-W., Iwarsson, S., & Oswald, F. (2012). Aging well and the environment: Toward an integrative model and research agenda for the future. *The Gerontologist, 52*(3), 306–316. https://doi.org/10.1093/geront/gnr154

Warner, L. M., Ziegelmann, J. P., Schüz, B., Wurm, S., Tesch-Römer, C., & Schwarzer, R. (2011). Maintaining autonomy despite multimorbidity: Self-efficacy and the two faces of social support. *European Journal of Ageing, 8*(1), 3–12. https://doi.org/10.1007/s10433-011-0176-6

Wettstein, M., Schilling, O. K., Reidick, O., & Wahl, H.-W. (2015). Four-year stability, change, and multidirectionality of well-being in very-old age. *Psychology and Aging*, *30*(3), 500–516. https://doi.org/10.1037/pag0000037

Wettstein, M., Schilling, O. K., & Wahl, H.-W. (2016). "Still feeling healthy after all these years": The paradox of subjective stability versus objective decline in very old adults' health and functioning across five years. *Psychology and Aging, 31*(8), 815–830. https://doi.org/10.1037/pag0000137

WHO. (1992). T*he ICD-10 Classification of Mental and Behavioural Disorders: Clinical descriptions and diagnostic guidelines*. https://icd.who.int/en

WHO. (2002). *Active Ageing: A Policy Framework*. Madrid. World Health Organisation.

Wied, T. S., Knebel, M., Tesky, V. A., & Haberstroh, J. (2019). The human right to make one's own choices – Implications for supported decision-making in persons with dementia: A systematic review. *European Psychologist (24)*, Article 2, 146–158. https://doi.org/10.1027/1016-9040/a000372

World Health Organisation (WHO). (1997). Whoqol Measuring Quality of Life.

Table 1. QoL chances and results in very old adults with and without limited capacity to answer interview questions themselves.

M (SD) or %	PC (self-report)	PLC (no self-report)
QoL chances of the individual		
Functional health (0-2)[a]	1.6 (0.4)	0.5 (0.5)
Cognitive status		
- MCI	16.4%	7.4%
- Beginning (suspected) AD	9.8%	73.5%
Depressive symptoms (0-4)	0.28 (1.32)	1.55 (1.36)
QoL chances of the environment		
Social contact (1-5)	3.6 (0.9)	3.4 (1.1)
Financial resources (in Euro)	1,944 (1,282)	2,581 (1,821)
Digital resources (internet use)	21.1%	1.6%
Societal values (anomie, 1-4)[a,b]	2.5 (0.5)	2.4 (0.9)
QoL results		
Appraisal of own life (positive affect, 1-5)[a,b]	3.3 (0.8)	2.6 (0.8)
Appraisal by others (burden to society)		
- Not at all/rather not	87.1%	75.1%
- Very much/quite a bit	12.9%	24.9%
Successful life conduct (autonomy)		
- Not at all/rather not	8.5%	60.2%
- Very much/quite a bit	91.5%	39.8%

Note. Weighted data. PC = persons able to conduct the interview themselves (self-report), PLC = persons with limited ability to conduct the interview (proxy report). [a]Estimated latent means given in original response scale. [b]Valid comparison of figures for PC and PLC possible because measurement equivalence was established.

Table 2. Associations between facets of QoL results in very old age.

Correlation	Group	Appreciation of own life (positive affect)	Successful life conduct (autonomy)
Successful life conduct (autonomy)[a]	PC	.41***	1
	PLC	.47***	1
Appreciation by others (burden to society)[a]	PC	-.18***	-.33***
	PLC	.01 n.s.	-.25**

Note. Weighted data. *** p<.001, ** p<.01, * p<.05, n.s. p≥.05. [a] Bivariate correlations for categorical observed variables are based on underlying continuous latent response variables.

Table 3. Prediction of QoL results by QoL chances.

Standardized regression parameter estimate or odds ratio (95% confidence limits)	Appreciation by others (burden to society[a])		Appreciation of own life (positive affect[b])		Successful life conduct (autonomy[b])	
	PC	PLC	PC	PLC	PC	PLC
			Life-ability of the person			
Functional health	**0.44 (0.23 - 0.70)**	0.84 (0.09 – 5.91)	-0.06	0.29**	**3.98 (2.31 – 7.40)**	**9.29 (2.67 – 58.1)**
Cognitive status (ref: age-adequate)						
- MCI	0.82 (0.48 – 1.48)	3.16 (0.0 – 14.2)	-0.02	-0.11	1.20 (0.63 – 1.79)	0.89 (0.06 – 4.39)
- Beginning (suspected) AD	0.90 (0.56 – 1.35)	3.11 (0.2 – 40.8)	-0.06	-0.24(*)	1.04 (0.48 – 2.09)	0.66 (0.16 – 1.79)
Depressive symptoms	1.18 (0.96 – 1.43)	1.35 (0.56 – 2.75)	-0.17*	-0.38**	**0.73 (0.61 – 0.84)**	0.75 (0.52 – 1.08)
Personal values						
- Generic value orientation	-	1.28 (0.42 – 3.94)	-	0.36**		1.34 (0.61 – 2.66)
- Self-transcendence	**0.59 (0.07 – 0.80)**	-	-0.19*	-	1.13 (0.86 – 1.54)	-
- Conservation	1.00 (0.62 – 1.70)	-	-0.01	-	0.94 (0.40 – 1.29)	-
- Self-enhancement	0.70 (0.48 – 1.16)	-	0.51***	-	**1.74 (1.28 – 2.38)**	-
			Livable environment			
Social contact	0.89 (0.72 – 1.21)	1.17 (0.44 – 2.71)	0.26***	0.26**	1.19 (0.98 – 1.41)	1.34 (0.75 – 2.51)
Financial resources[c]	**1.01 (1.00 – 1.02)**	1.03 (0.97 – 1.06)	-0.01	0.15	**0.99 (0.97 - <1.0)**	0.98 (0.95 – 1.03)
Digital resources (ref: no internet use)	0.90 (0.56 – 1.35)	-	0.09**	-	**1.98 (1.28 – 3.32)**	-

Societal values (anomie)	**3.26 (1.69 – 6.16)**	3.86 (0.71 – 18.8)	-0.06	0.25*	0.98 (0.64 – 1.50)	1.00 (0.21 – 3.07)
			Socio-demographic characteristics			
Gender (ref: male)	1.02 (0.58 – 1.70)	0.44 (0.02 – 5.60)	0.01	-0.12	0.67 (0.34 – 1.18)	0.56 (0.20 – 2.36)
Age group (ref: 80-84 yrs)						
- 85-89 yrs	1.43 (0.96 – 2.29)	0.22 (0.0 – 2.40)	-0.06	-0.17	0.69 (0.35 – 1.06)	0.48 (0.05 – 7.86)
- 90 yrs or older	0.79 (0.47 – 1.39)	0.24 (0.01 – 13.5)	-0.05	-0.14	0.89 (0.46 – 1.58)	1.65 (0.18 – 10.5)
Interaction						
- Female 85-89 yrs	0.81 (0.39 – 1.43)	2.07 (0.03 – 166.9)	0.05	0.34	2.13 (0.94 – 4.22)	1.74 (0.08 – 51.8)
- Female 90+ yrs	1.50 (0.46 – 3.43)	1.18 (0.0 – 154.7)	0.07	0.31	1.36 (0.62 – 3.39)	0.62 (0.09 – 7.64)
Model R^2	0.318	0.448	0.461	0.486	0.328	0.366

Notes. PC = persons able to conduct the interview themselves (self-report), PLC = persons with limited ability to conduct the interview (proxy report). [a] Generalized logistic regression of categorical dependent variable. [b] MI has been established across PC and PLC groups. [c] Square root transformation was used to normalize net household equivalence income. *** p<.001, ** p<.01, * p<.05, (*) p<.10 (power restriction in PLC).

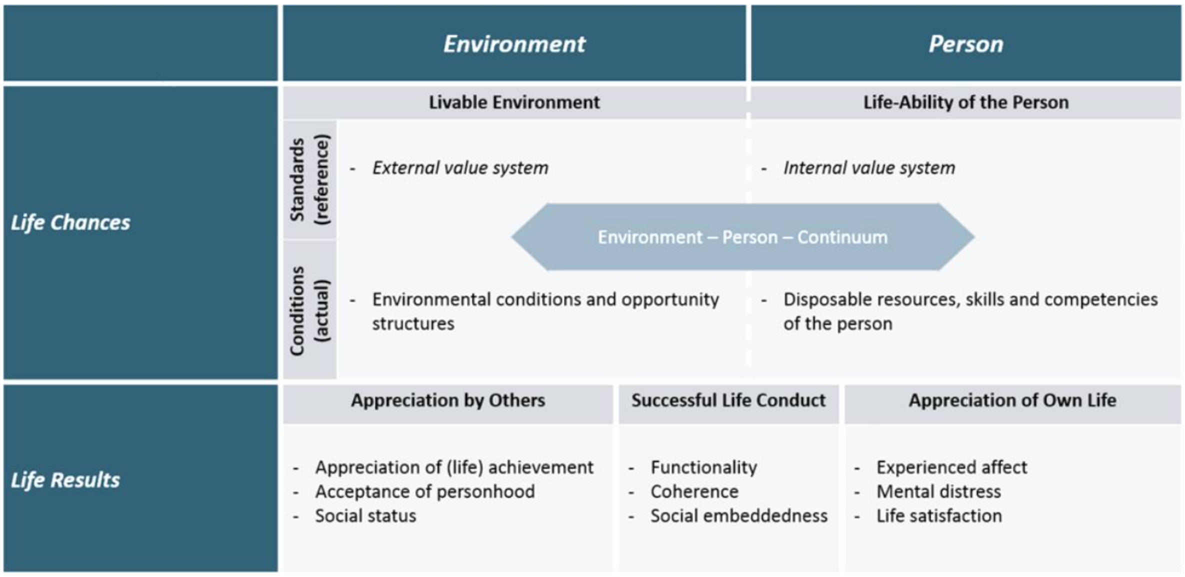


Fig. 1 Challenges and potentials model of Quality of Life in very old age (CHAPO, Wagner et al. 2018)

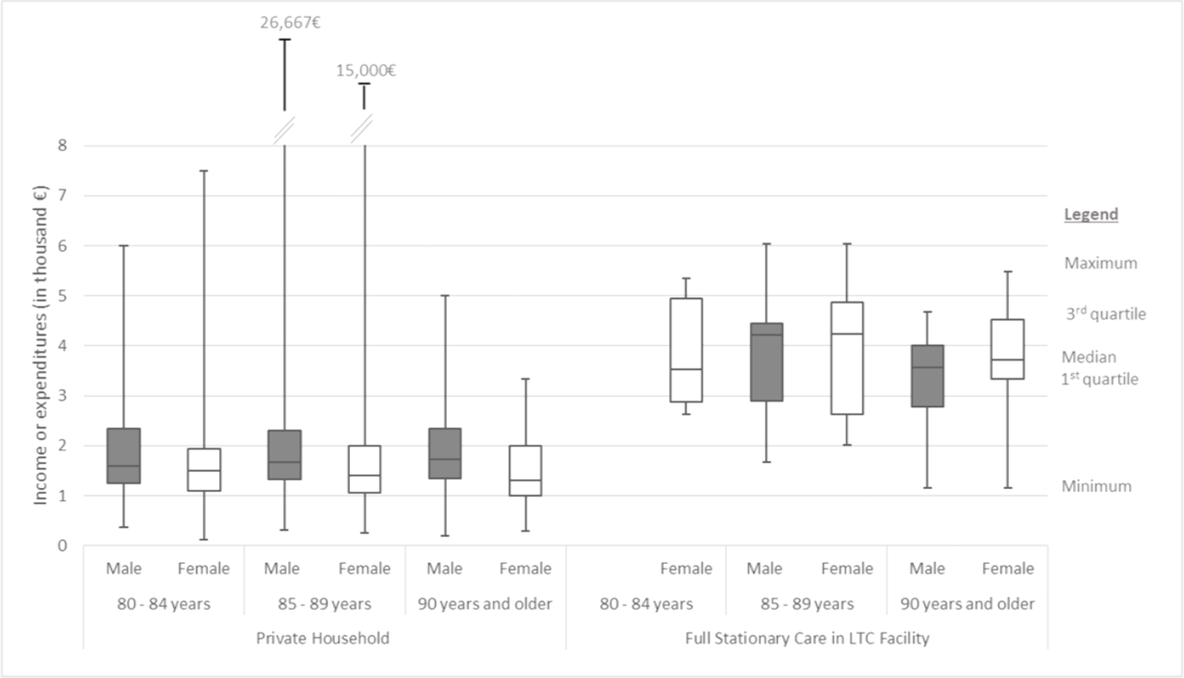


Fig. 2 Monthly cost for full stationary care in LTC facilities and net equivalence household income in private dwellings or care settings other than full stationary care. In the group LTC - 80 to 84 years - Male, there was only a single individual (4,842€)

4.5 Multimorbidity in old age and its impact on life results

The fifth paper examines the frequency of multimorbidity and common diseases in old age and the impact of multimorbidity on autonomy, life satisfaction, ADL, and IADL in different age and sex groups of the old age. Multimorbidity is measured via self-reports of currently treated diseases.

Zeitschrift für
Gerontologie+Geriatrie

Z Gerontol Geriat 2021 · 54 (Suppl 2):S108–S113
https://doi.org/10.1007/s00391-021-01920-9
Received: 17 March 2021
Accepted: 7 May 2021

Thomas Brijoux[1] · **Cristiane Woopen**[1] · **Susanne Zank**[2]

[1] cologne center for ethics, rights, economics, and social sciences of health, Universität zu Köln, Cologne, Germany
[2] Lehrstuhl für Rehabilitationswissenschaftliche Gerontologie, Humanwissenschaftliche Fakultät, Universität zu Köln, Cologne, Germany

Multimorbidity in old age and its impact on life results

Aging is often described as a process of gains and losses. While in social and emotional contexts, aging is also associated with gains and plasticity, the emphasis on loss dominates with respect to physical health [30]. Multiple studies like the Berlin aging study (BASE) [ibid.] or the German aging survey [36] have examined the physical health of the older population in Germany and analyzed the most common diagnoses. In persons above the age of 70 years, cardiovascular issues (e.g., hyperlipidemia and cerebral arteriosclerosis) and varicosis are the most common diagnoses according to BASE (ibid.). For people above the age of 75, the Federal Health Monitoring System names high blood pressure and arthrosis as the diseases with the highest prevalence in women and high blood pressure and cardiovascular disease in men [25]. For most diseases, the prevalence is higher in older age groups [25].

Concerning psychiatric disorders, depression and dementia have the highest prevalence in old age [5, 26]. For people aged above 75 years, the prevalence of major depression is 7.2% and 17.1% for subclinical depressive symptoms [19]. In older age groups, prevalence rises for subclinical depressive symptoms (ibid.). Dementia prevalence in old age ranges between 10.7% for men aged 80–84 and 44.8% for women aged above 90 [1]. Prevalence is in general higher for women than for men and higher in older age groups (ibid.). Both diseases, dementia and depression, have a huge impact on people's lives [25] and influence their quality of life (QoL), their participation in society, their success in geriatric treatments [9] and the management of other diseases [26]. Different aspects of QoL are impacted differently by dementia. While functioning or autonomy are severely reduced [20] the consequences for hedonic measures elements of QoL, especially in self-ratings at the beginning of the disease, are still substantial but of a lesser extent [22].

With higher prevalence for each disease in old age, the number of diseases per person also increases [5, 13]. The coexistence of multiple health conditions is called multimorbidity [14], different approaches exist concerning the number and kind of diseases for the operationalization of multimorbidity [34]. The first large-scale examination of multimorbidity in Germany was done in the Berlin Aging Study [30] and found that 28.0% of the examined persons above the age of 70 were treated for at least 5 diagnoses [30]. The BASE study not only asked the participants for which diagnoses they received medical treatment, participants also underwent detailed diagnostic procedures which allowed statements on undertreatment and underdiagnoses. While 27.6% of men and 25.6% of women aged older than 85 reported 5 or more diseases diagnosed by their general practitioner, 40.9% of men and 54.3% of old age women had 5 or more diseases diagnosed by the study team. These discrepancies were much smaller in persons below the age of 85. In contrast to the cut-off score of five used in BASE, more recent analyses define multimorbidity as the coexistence of two or more diagnoses as reported by participants [14]. In the German aging survey, 57% of the age group of 55–69 reported to be affected by more than 2 self-reported diagnoses [36]. In the German telephone health interview (GEDA) study, 75.4% of women aged between 65 and 74 were affected by 2 health conditions [5]. In women aged 75+, this increased to 81.7%. Fewer multiple health conditions were reported for men (65–74: 68%; 75+: 74.2%) and younger persons [5]. These influences of age and gender have also been shown in a recent international review [32] and are the most frequently studied determinants of multimorbidity [32]. Although multimorbidity serves as a marker for care, dependency, functional impairment, and frailty [7, 37], which occur more frequently among the oldest old, most of these studies analyzed multimorbidity only up to the age of 80. There is, thus, a specific need to look at multimorbidity among older adults.

Multimorbidity among the people aged above 65 years entails higher disability risks as well as poorer health-related QoL and healthcare utilization [21]. In addition, researchers and practitioners have stressed the need to consider multimorbidity explicitly in the healthcare setting [23]. Treating patients with multimorbidity is challenging and specific guidelines have been established to avoid polypharmacy and single disease treatment that cumulatively lead to adverse effects or interactions [2, 31].

The higher portion of incurable chronic diseases also results in a paradigm shift in the therapeutic treatment of geriatric patients. While in younger age, the physical cure of a disease is often the primary therapeutic goal, this moves towards a preservation of functioning,

Table 1 Percentages of the number of treated diseases in different demographic groups

Number of treated diseases	Men (%)			Women (%)		
	80–84 years	85–89 years	90+ years	80–84 years	85–89 years	90+ years
0	4.2	6.1	3	4.9	5.4	4.8
1	17.1	17.3	12	11.8	16.2	10.5
2	17.3	17.5	18	19.4	14.3	22.5
3	19.3	18.6	20.1	17.1	18.2	19
4	15.6	12	16.8	12.8	12.4	10.5
5	10.1	10.5	9	11.6	9	9.9
6	7.3	8.6	9	8.5	9.5	11.1
7	3.3	5.3	4.5	6.5	5.1	6.9
8	2.2	2.5	3	3.1	3.6	2.5
9	1.9	1.6	3	2.3	3.2	0.7
10	0.5	0.1	1.5	1.2	1.9	0.7
11	0.5	0	0	0.8	0.6	0.5
12	0.7	0	0	0	0.8	0.6
13	0	0	0	0	0	0

QoL, and autonomy in older age groups [13]. Examinations on the association between QoL and multimorbidity have focused on health-related QoL [18]; however, some evidence shows that multimorbidity is also associated with a decrease in functional health [21] and autonomy [29]. As functioning, QoL, and autonomy are important therapeutic goals in geriatric settings, influences of multimorbidity on these concepts need to be examined for old age.

The theoretical background for this examination is the challenges and potential (CHAPO) model [33] which describes interactions of different facets of QoL, like personal and environmental life chances and life results. Within this model, multimorbidity is understood as a lack of disposable skills and competences, life satisfaction as a result of appreciation of one's own life, and functioning as well as autonomy as life results in successful life conduct.

The goals of this study were to describe the occurrence of multimorbidity and the most common physical and psychiatric diseases in old age, and to evaluate the influence of multimorbidity on functioning, life satisfaction, and autonomy.

Methods

Multimorbidity was analyzed with the NRW80+ dataset. The NRW80+ study is a representative study on old people living in North Rhine-Westphalia, Germany's most populous state, and was conducted from 2016 to 2019. A general description of the NRW80+ study design is given elsewhere [8].

Variables

Multimorbidity was measured by asking participants if they were medically treated ("Yes" or "No") for 20 different health issues. The "Yes" answers were counted and interpreted as the number of treated diseases. The list of health issues is derived from the self-administered comorbidity questionnaire [28], the multimorbidity index in age [3], and the list of diseases used by the German aging survey [35].

Mental health conditions assessed in this study include cognition and depression. The DemTect was used to detect cases of dementia and mild cognitive impairment. It is a screening tool for cognitive impairment and shows good classification among the oldest old [15, 17]. We used the cut-off values of 9–12 for mild cognitive impairment and 8 or lower for dementia. The sum scores were derived from age-specific normation of the subtest on people aged 80 and older [17]. In the proxy interviews, the DemTect could not be applied. Proxies were asked to rate the cognitive status of the person they represented on the global deterioration scale [24]. Depressiveness was analyzed by means of the DIA-S4 [12]. The DIA-S4 is a four-item short form of the depression in age scale [11] that is interpreted as the number of depressive symptoms.

Analyzed life results are autonomy, life satisfaction, and functioning. A single item, which could be answered in four points, was used to measure autonomy ("do you arrange your life according to your own ideas?"). Life satisfaction was measured with one item ("all in all, how satisfied are you currently with your life?") that could be answered in 11 categories between 0 and 10 and had been used similarly in the socioeconomic panel [16]. Functional status was measured with the activities of daily living (ADL) and the instrumental activities of daily living (IADL) subscales of the older Americans resources and services questionnaire [4].

Age and gender were the analyzed sociodemographic variables. Age was analyzed in the 3 age groups 80–84 years, 85–89 years, and above 90 years.

Sample and bias

We used the full sample of the NRW80+ data, which includes 1687 self-report interviews and 176 proxy interviews. Survey weights are used to balance the sample for age, gender, household size, nursing home status, family status, and region, and to minimize bias in these variables. In the weighted sample 1012 persons are 80–84 years old, 573 respondents are 85–89 years old and 279 are 90 years or older, 676 persons are male 1187 are female.

Missing values

In this study 2.1% of the data analyzed were missing but missing values exceeded 18.6% in cognition and 7.3% in depression. Complete case reports lead to biased estimators; therefore, missing values were treated via multiple imputation. A total of 20 datasets were imputed, which is sufficient for this missing structure [6]. The presented frequencies and means are the average results of these 20 datasets. Standard deviations for multiple imputed datasets were calculated using Rubin's formula [27].

Statistical methods

We present the number of treated diseases in different age and sex groups and use ANOVA and t-test to explore significant differences between these sociodemographic groups. Linear models with main and interaction terms were used to estimate the relationship between multimorbidity and autonomy, ADL, IADL, and life satisfaction for different age groups and genders. Tests for multicollinearity showed no problems. Heteroscedasticity was addressed by use of robust standard errors HC3 [10]. The analyses were done with SPSS Version 27 (IBM Corp. Released 2020. IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY, USA).

Results

Multimorbidity

The mean number of treated diseases per person was 3.62 (Confidence Interval (CI)-95%: 3.51–3.73). Women reported more treated diseases then men (3.73 vs. 3.43, T = 2.74, $p < 0.01$). No treated diagnoses were reported by 4.9% of the sample (women: 5.1%, men 4.6%), 31.4% of the old-aged population reported 5 or more treated diagnoses, while women reported 5 or more treated diseases more frequently than men (33.6% vs. 27.5%) (see **Table 1**). No uniform influence of age on the number of treated diseases can be shown. In men, the number of treated diseases is highest in the oldest age group (80–84: 3.42; 85–89: 3.36; 90+: 3.66), while in women, the oldest age group reports the least number of treated diseases (80–84: 3.73; 85–89: 3.77; 90+: 3.64). In none of the imputed datasets, did the age group have a significant influence on the number of reported diseases (F = 0.07–0.185, $p = 0.831–0.993$).

High blood pressure was the most common treated disease, reported by 59.4% of the sample (60.5% in women and 57.7% in men). Treatment of joint and bone diseases, which were the second most common type of treated diseases,

Abstract · Zusammenfassung

Z Gerontol Geriat 2021 · 54 (Suppl 2):S108–S113 https://doi.org/10.1007/s00391-021-01920-9

T. Brijoux · C. Woopen · S. Zank

Multimorbidity in old age and its impact on life results

Abstract

Background. High prevalence diseases, such as high blood pressure, dementia and depression in old age can lead to multimorbidity, which is often defined as the presence of more than one health condition in an individual. Multimorbidity has negative consequences on health-related quality of life and healthcare utilization. As many age-associated diseases are not curable, therapeutic goals like preservation of autonomy, functioning, and life satisfaction become more important in old age patients.
Objective. The prevalence of multimorbidity dementia and depressive symptoms and the consequences of multimorbidity on autonomy, functioning, and life satisfaction among the oldest old were examined.
Material and methods. In personal computer-assisted interviews, participants of the representative study NRW80+ were asked for which health issues they received medical treatment.
Results. On average, people above the age of 80 years were treated for 3.62 diseases and 31.4% of older people received medical treatment for 5 or more diseases. A connection between multimorbidity and age group could not be shown. Autonomy, functioning, and life satisfaction are reduced in association with multimorbidity.
Conclusion. Multimorbidity is a frequent phenomenon among old people. A lack of diagnostic procedures and medical treatment can be a reason for the missing age trends. The results illustrate the importance of multimorbidity for patient-relevant outcomes and reveal the need to identify patients with multimorbidity.

Keywords
Multimorbidity · Oldest old · Depression · Dementia · Quality of life

---

Multimorbidität in der Hochaltrigkeit und ihre Auswirkung auf Lebensergebnisse

Zusammenfassung

Hintergrund. Die hohe Prävalenz von Erkrankungen wie Bluthochdruck, Demenz oder depressiven Störungen in der Hochaltrigkeit führt vielfach zu Multimorbidität, dem gleichzeitigen Vorliegen mehrerer Erkrankungen. Multimorbidität wirkt sich negativ auf die gesundheitsbezogene Lebensqualität und Inanspruchnahme von Hilfe aus. Da viele der alterskorrelierten Erkrankungen nicht heilbar sind, werden therapeutische Ziele wie der Erhalt von Funktionalität, Autonomie und Lebensqualität in der Hochaltrigkeit bedeutender.
Fragestellung. Die Häufigkeit von Multimorbidität, demenzieller und depressiver Symptomatik sowie die Auswirkungen von Multimorbidität auf Funktionalität, Autonomie und Lebenszufriedenheit bei über 80-Jährigen wurden untersucht.
Material und Methode. Teilnehmende der repräsentativen Studie NRW80+ wurden in computergestützten persönlichen Interviews befragt, zu welchen gesundheitlichen Problemen sie aktuell ärztlich behandelt werden.
Ergebnisse. Durchschnittlich waren die über 80-Jährigen von 3,62 Erkrankungen betroffen. 31,4 % der Hochaltrigen wurden wegen 5 oder mehr gesundheitlichen Problemen ärztlich behandelt. Zusammenhänge zwischen Alter und Anzahl der behandelten Erkrankungen konnten nicht nachgewiesen werden. Autonomie, Funktionalität und Lebenszufriedenheit sind bei Multimorbidität reduziert.
Schlussfolgerung. Multimorbidität ist in der Hochaltrigkeit ein häufiges Phänomen. Das Fehlen klarer Alterstrends kann auch mit Unterdiagnostik und Unterbehandlung begründet werden. Die Ergebnisse verdeutlichen die Signifikanz von Multimorbidität für patientenrelevante Outcomes. Der Identifikation multimorbider Patienten im Gesundheitssystem kommt eine zentrale Rolle zu.

Schlüsselwörter
Multimorbidität · Hochaltrigkeit · Depression · Demenz · Lebensqualität

Table 2 Treated diseases by gender

Reported treatment	Overall (%)		Men (%)		Women (%)	
	Yes	No	Yes	No	Yes	No
Myocardial infection	7.4	92.6	12.5	87.5	4.5	95.5
Heart disease (e.g., heart failure)	34.9	65.2	32.1	67.9	36.5	63.5
Hypertension	59.4	40.5	57.7	42.3	60.4	39.6
Stroke	7.9	92.2	7.1	92.9	8.4	91.6
Mental disorder (e.g., phobia, depression)	6.7	93.3	4.4	95.6	8.1	91.9
Cancer	8	92.0	11.9	88.1	5.8	94.2
Diabetes	16.5	83.6	17.2	82.8	16.1	83.9
Respiratory or lung disease	12.6	87.4	12.8	87.2	12.6	87.4
Back pain	32.6	67.3	27.5	72.5	35.6	64.4
Gastrointestinal disease	13.3	86.8	10.8	89.2	14.7	85.3
Kidney disease	7.3	92.7	8	92	7	93
Liver disease	2	98.3	1.6	98.4	2.2	97.8
Blood disorders (e.g., anemia)	2.5	97.5	2.2	97.8	2.7	97.3
Bone or joint disorder	45.8	54.3	35.8	64.2	51.5	48.5
Urinary disorder	17	83.1	20.2	79.8	15.2	84.8
Insomnia	14.2	85.9	10.8	89.2	16.1	83.9
Vision impairment	30.5	69.7	28.5	71.5	31.5	68.5
Hearing impairment	20.9	79.3	23.7	76.3	19.2	80.8
Neurologic disease	9.9	90.3	8.6	91.4	10.6	89.4
Other chronic disease	12.5	87.3	9.1	90.9	14.5	85.5

was reported by 45.2%. In joint and bone diseases, a huge gender gap can be observed, with higher prevalence for women (51.5% vs. 35.8%). 34.9% of the subjects reported treatment of heart diseases, which occured more frequently in women than in men (36.5% vs. 32.1%). A full account of the reported treatments is given in ◘ **Table 2**.

Mental health

Concerning psychiatric diseases, prevalence for problems in cognition as analyzed by DemTect and GDS is 16.5% for dementia and 16% for MCI. About half of the sample (47.5%) shows no depressive symptoms, 24% showed 1 symptom and 28.4% of the people showed 2 or more out of 4 symptoms on the DIA-S4.

Associations between morbidity and QOL

Multimorbidity influences life satisfaction, autonomy, ADL and IADL. While in each QoL dimension the main effects for age group are significant, most of the interaction terms for multimorbidity and age group are not. All regression coefficients and their corresponding confidence intervals and *p*-values are presented in ◘ **Table 3**.

Discussion

In this study 31% of old age people report treatment for at least 5 different diagnoses, which is comparable to the rate found in BASE. This study allows a differentiation within the oldest age groups; however, no age-related effects can be observed in this sample. This surprising result can be better understood on the basis of results from the BASE study, where underdiagnoses were most common in the oldest age group. In the NRW80+ study, our operationalization of multimorbidity relies on self-report, which is strongly affected by underdiagnoses [30]. As the NRW80+ study did not involve medical examinations, evidence-based statements about undertreatment and underdiagnoses cannot be made from this sample, indicating the need for studies of the oldest old that include detailed medical diagnostic procedures.

The most frequently reported disease in old age is high blood pressure, which is in line with the results of the Federal Health Monitoring System [25]. Although in depth diagnostic procedures were not feasible in this study the use of the screening tools DemTect and GDS shows that problems with cognition and depressive symptoms are common in old age.

Multimorbidity has a negative impact on the investigated life results, indicating the importance of the construct and the subsequent need to identify patients with multimorbidity in the healthcare setting by use of screening tools. Subsequently, policies that are already a part of guidelines [2], like checks for multimedication and a constant comparison between medical goals and priorities of the patients, need to be put in practice. In all analyzed life results (life satisfaction, autonomy, ADL, and IADL), age effects can be observed. Furthermore, we observe gender differences in IADL which can be explained by higher prevalence rates for dementia among women. The effects of multimorbidity on life satisfaction are reduced in the older age groups which underlines the necessity to differentiate within the oldest age groups in analysis of life satisfaction. Most of the interaction terms remain insignificant, hence it can be assumed that the influence of multimorbidity on autonomy, ADL and IADL does not change in the different age groups.

The usual limitations occur in interpreting cross-sectional data: causal interpretations cannot be made. Multimedication which happens alongside with multimorbidity could not be addressed in this article. In the second wave of the NRW80+ study multimedication is a part of the questionnaire and can be included in future analysis. To examine desired research questions regarding the interactions between different health-related variables, coping behavior, and life results, new study types need to be developed. The questionable circular interactions need to be examined in longitudinal data collected in short time intervals.

Practical implications

- Multimorbidity affects about one third of the old-aged people in North Rhine-Westphalia.

Table 3 Effects of multimorbidity, age group, and gender on functioning, autonomy, and life satisfaction

–	**Autonomy**				**Life satisfaction**			–
Variable	*B*	*B-CI-95%*	*T-value*	*P-value*	*Beta*	*B CI*	*T-value*	*P-value*
Age group: 80–84 (Ref.)								
Age group: 85–89	–0.175	–0.324––0.025	–2.288	0.022*	–0.22	–0.547—0.107	–1.321	0.107
Age group: 90+	–0.509	–0.770––0.247	–3.813	<0.001***	–1.242	–1.870—0.613	–3.872	<0.001***
Gender: men (Ref.)								
Gender: women	–0.074	–0.221––0.072	–1.000	0.317	0.126	–0.191–0.443	0.781	0.435
#Treated diseases	–0.044	–0.078––0.011	–2.577	0.01*	–0.165	–0.237—0.093	–4.502	<0.001***
#Treated diseases * Age (85–89)	0.002	–0.036–0.041	0.107	0.915	0.011	–0.073–0.096	0.266	0.79
#Treated diseases * Age (90+)	0.014	–0.053–0.081	0.417	0.677	0.162	0.018–0.306	2.201	0.028*
#Treated diseases * Gender	0.007	–0.046–0.032	0.349	0.727	–0.04	–0.124–0.043	0.781	0.435
Constant	3.771	3.654–3.888	63.026	<0.001***	8.55	8.283–8.818	62.655	<0.001***
R^2	–	–	–	0.05	–	–	–	0.05
–	**ADL**				**IADL**			–
Variable	*Beta*	*B CI*	*T-value*	*P-value*	*Beta*	*B CI*	*T-value*	*P-value*
Age group: 80–84 (Ref.)								
Age group: 85–89	–0.207	–0.292––0.122	–4.785	<0.001***	–0.352	–0.461—0.243	–6.318	<0.001***
Age group: 90+	–0.559	–0.720––0.398	–6.818	<0.001***	–0.817	–0.995—0.638	–8.948	<0.001***
Gender: men (Ref.)								
Gender: women	–0.078	–0.159–0.003	–1.892	0.059	–0.13	–0.234—0.025	–2.432	0.015*
#Treated diseases	–0.06	–0.079––0.040	–5.912	<0.001***	–0.075	–0.099—0.051	–6.088	<0.001***
#Treated diseases * Age (85–89)	0.017	–0.004–0.039	1.581	0.114	0.018	–0.008–0.044	1.343	0.179
#Treated diseases * Age (90+)	0.030	–0.008–0.069	1.533	0.125	0.033	–0.007–0.073	1.611	0.107
#Treated diseases * Gender	0.004	–0.018–0.026	0.318	0.751	0.00	–0.027–0.028	–0.018	0.986
Constant	2.002	1.932–2.071	56.150	<0.001***	1.942	1.856–2.027	44.465	<0.001***
R^2	–	–	–	0.13	–	–	–	0.20

Number of; *B* regression coefficient; *B CI-95%* 95% confidence intervall of the regression coefficient. Positive values in B indicate higher values in the corresponding outcome when the independent variable increase

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

- Life results like life satisfaction, autonomy, and functioning are associated with multimorbidity. Therefore, diagnostic procedures are needed to identify these patients.
- Further studies are needed to examine undertreatment and underdiagnosis in the oldest old.

Corresponding address

Thomas Brijoux
cologne center for ethics, rights, economics, and social sciences of health, Universität zu Köln
Cologne, Germany
thomas.brijoux@uni-koeln.de

Acknowledgements. The project NRW80+ belongs to the key research area "Aging and Demographic Change" (Speaker: Susanne Zank) of the Cologne Center for Ethics, Rights, Economics, and Social Sciences of Health (ceres), which is directed by Christiane Woopen at the University of Cologne. Members of the project board are Michael Wagner, Christiane Woopen and Susanne Zank. NRW80+ is funded by the Ministry of Innovation, Science and Research of North Rhine-Westphalia.

Funding. Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest. T. Brijoux, C. Woopen and S. Zank declare that they have no competing interests.

Ethical standards. All procedures performed in studies involving human participants or on human tissue were in accordance with the ethical standards of the institutional and/or national research committee and with the 1975 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

1. Alzheimer Europe (2019) Alzheimer Europe 2019 dementia in Europe yearbook 2019. Estimating the prevalence of dementia in Europe
2. DEGAM Multimorbiditaet S3-Leitlinie. https://www.degam.de/files/Inhalte/Leitlinien-Inhalte/Dokumente/DEGAM-S3-Leitlinien/053-047_Multimorbiditaet/053-047l_%20Multimorbiditaet_redakt_24-1-18.pdf. Accessed 20 Jan 2021
3. Diederichs C-P (2011) Entwicklung eines Multimorbiditätsindex zur standardisierten Erfassung von chronischen Erkrankungen in der älteren Bevölkerung. Dissertation, Hannover.
4. Fillenbaum GG (1988) Multidimensional functional assessment of older adults. The Duke Older

Americans Resources and Services procedures. Lawrence Erlbaum, Hillsdale

5. Fuchs J, Busch M, Lange C et al (2012) Prevalence and patterns of morbidity among adults in Germany. Results of the German telephone health interview survey German Health Update (GEDA) 2009. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 55:576–586. https://doi.org/10.1007/s00103-012-1464-9
6. Graham JW, Olchowski AE, Gilreath TD (2007) How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci 8:206–213. https://doi.org/10.1007/s11121-007-0070-9
7. Hanlon P, Nicholl BI, Jani BD et al (2018) Frailty and pre-frailty in middle-aged and older adults and its association with multimorbidity and mortality: a prospective analysis of 493 737 UK Biobank participants. Lancet Public Health 3:e323–e332. https://doi.org/10.1016/S2468-2667(18)30091-4
8. Hansen S, Kaspar R, Wagner M, Woopen C, Zank S (2021) The NRW80+ study: Conceptual background and design decisions (in press). Z Gerontol Geriat 54. (In press)
9. Haupt M, Jänner M, Richert F (2020) Die Bedeutung psychischer Störungen für den stationär akutgeriatrischen Behandlungsverlauf: eine explorative Untersuchung. Z Gerontol Geriat 53:749–755. https://doi.org/10.1007/s00391-019-01671-8
10. Hayes A, Cai L (2007) Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. Behav Res Methods 39:709–722
11. Heidenblut S, Zank S (2010) Entwicklung eines neuen Depressionsscreenings für den Einsatz in der Geriatrie. Die „Depression-im-Alter-Skala“ (DIA-S). Z Gerontol Geriat 43:170–176. https://doi.org/10.1007/s00391-009-0067-z
12. Heidenblut S, Zank S (2020) Screening for depression in old age with very short instruments: the DIA-S4 compared to the GDS5 and GDS4. Gerontol Geriatr Med 6:2333721420981328. https://doi.org/10.1177/2333721420981328
13. Holzhausen M, Scheidt-Nave C (2012) Multimorbidität als Interventionsherausforderung. In: Wahl H-W, Tesch-Römer C, Ziegelmann JP (eds) Angewandte Gerontologie. Interventionen für ein gutes Altern in 100 Schlüsselbegriffen. Kohlhammer, Stuttgart, pp 48–53
14. Johnston MC, Crilly M, Black C et al (2019) Defining and measuring multimorbidity: a systematic review of systematic reviews. Eur J Public Health 29:182–189. https://doi.org/10.1093/eurpub/cky098
15. Kalbe E, Kessler J, Calabrese P et al (2004) DemTect: a new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. Int J Geriat Psychiatry 19:136–143. https://doi.org/10.1002/gps.1042
16. Kantar Public (2019) SOEP-Core – 2019: Personenfragebogen, Stichproben A-L3, M1-M2 + N-O. https://www.diw.de/documents/publikationen/73/diw_01.c.807938.de/diw_ssp0909.pdf. Accessed 19 Jan 2021
17. Kessler J, Fengler S, Kaesberg S et al (2014) DemTect 40- und DemTect 80+: Neue Auswertungsroutinen für diese Altersgruppen. Fortschr Neurol Psychiatr 82:640–645. https://doi.org/10.1055/s-0034-1385278
18. Lehnert T, Heider D, Leicht H et al (2011) Review: health care utilization and costs of elderly persons with multiple chronic conditions. Med Care Res Rev 68:387–420. https://doi.org/10.1177/1077558711399580
19. Luppa M, Sikorski C, Luck T et al (2012) Age- and gender-specific prevalence of depression in latest-life—Systematic review and meta-analysis. J Affect Disord 136:212–221. https://doi.org/10.1016/j.jad.2010.11.033
20. Mahlberg G, Gutzmann H (2009) Einleitung: Krankheitsbilder. In: Mahlberg G, Gutzmann H (eds) Demenzerkrankungen. erkennen, behandeln und versorgen. Deutscher Ärzte Verlag, Köln, pp 3–11
21. Marengoni A, Angleman S, Melis R et al (2011) Aging with multimorbidity: a systematic review of the literature. Ageing Res Rev 10:430–439. https://doi.org/10.1016/j.arr.2011.03.003
22. Martyr A, Nelis SM, Quinn C et al (2018) Living well with dementia: a systematic review and correlational meta-analysis of factors associated with quality of life, well-being and life satisfaction in people with dementia. Psychol Med 48:2130–2139. https://doi.org/10.1017/S0033291718000405
23. Navickas R, Petric V-K, Feigl AB et al (2016) Multimorbidity: What do we know? What should we do? J Comorbidity 6:4–11. https://doi.org/10.15256/joc.2016.6.72
24. Reisberg B, Ferris SH, de Leon MJ et al (1988) Global Deterioration Scale (GDS). Psychopharmacol Bull 24:661–663
25. Robert Koch-Institut (2015) Bericht Gesundheit in Deutschland 2015. Gesundheitsberichterstattung des Bundes gemeinsam getragen von RKI und DESTASIS
26. Robert Koch-Institut (RKI) (2015) Wie gesund sind die älteren Menschen? Gesundheit in Deutschland (Einzelkapitel)
27. Rubin DB (2009) Multiple imputation for nonresponse in surveys. John Wiley & Sons Inc., Hoboken
28. Sangha O, Stucki G, Liang MH et al (2003) The self-administered comorbidity questionnaire: a new method to assess comorbidity for clinical and health services research. Arthritis Rheum 49:156–163. https://doi.org/10.1002/art.10993
29. Scheidt-Nave C, Richter S, Fuchs J et al (2010) Herausforderungen an die Gesundheitsforschung für eine alternde Gesellschaft am Beispiel „Multimorbidität“. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 53:441–450. https://doi.org/10.1007/s00103-010-1052-9
30. Steinhagen-Thiessen E, Borchelt M (2010) Morbidität, Medikation und Funktionalität im Alter. In: Lindenberger U, Smith J, Mayer KU, Baltes PB (eds) Die Berliner Altersstudie. Akad.-Verl., Berlin, pp 175–207
31. The Academy of Medical Sciences (2018) Multimorbity: a priority for global health researcg. https://acmedsci.ac.uk/file-download/82222577. Accessed: 7 May 2021
32. Violan C, Foguet-Boreu Q, Flores-Mateo G et al (2014) Prevalence, determinants and patterns of multimorbidity in primary care: a systematic review of observational studies. PLoS One 9:e102149. https://doi.org/10.1371/journal.pone.0102149
33. Wagner M, Rietz C, Kaspar R et al (2018) Lebensqualität von Hochaltrigen: Studie zu Lebensqualität und subjektivem Wohlbefinden hochaltriger Menschen in Nordrhein-Westfalen (NRW80+). Z Gerontol Geriat 51:193–199. https://doi.org/10.1007/s00391-017-1217-3
34. Wetterling T (2020) Pathogenese der Multimorbidität – was ist bekannt? Z Gerontol Geriat. https://doi.org/10.1007/s00391-020-01752-z
35. Wiest M, Richter M, Krauel F et al (2014) German Ageing Survey, Deutscher Alterssurvey (DEAS): Documentation of instruments and variables 1996–2011
36. Wurm S, Schöllgen I, Tesch-Römer C (2010) Gesundheit. In: Motel-Klingebiel A, Wurm S, Tesch-Römer C (eds) Altern im Wandel. Befunde des Deutschen Alterssurveys (DEAS). Kohlhammer, Stuttgart, pp 90–117
37. Yarnall AJ, Sayer AA, Clegg A et al (2017) New horizons in multimorbidity in older adults. Age Ageing 46:882–888. https://doi.org/10.1093/ageing/afx150

5 Discussion

All five presented papers look at difficulties in achieving representativity. The first paper researches the exclusion of groups that are difficult to survey and finds that the exclusion of the nursing home population, people with dementia, and people who are represented by proxies leads to substantial bias and consequently to a reduction of representativity as estimated by R-indicators. Furthermore, it discovers that an exclusion of each of the three groups results in some bias in social economic status, functioning, and well-being.

The second paper implicitly explores measurement errors that derive from two interview settings that are common and sometimes necessary when surveying the oldest old: present third persons and proxy interviews. An analysis of measurement invariance shows that answers given by proxy informants as well as answers given when third persons are present differ in their factorial structure from answers given in interviews conducted with target persons only. While this result does not allow for an assessment of the validity of each measurement, it shows that the construct is understood differently between self-reported interviews and proxy interviews as well as interviews with present third persons. The paper concludes that self-reports are more purposeful for the examined research question because at a later stage, elder abuse is related to constructs that were also captured via the mode of self-reports.

The third paper addresses the exclusion of interviews with present third persons and interviews conducted with proxies. In Supplement 1, possible consequences for some variable estimates of the sample resulting from the exclusion of these interviews are displayed in Table 1. To reduce this bias, a recalculation of the survey weights was performed. Apparently, the resulting bias in some variables can be reduced, but does not completely disappear. In particular, people with dementia who are mainly represented by proxy interviews are significantly better represented after reweighting. A shortcoming of this approach is that the resulting sample only includes people within early stages of dementia who are meant represent the whole dementia population, including later stages. However, this frame error was wilfully accepted over the alternative of a measurement error resulting from a joint analysis of proxy interviews, interviews with present persons, and self-reports. It should be pointed out that the quality of the excluded interviews might be reduced because present third persons as well as proxy informants could be perpetrators of abuse. Therefore, it can be argued that in this case, the reduction of measurement error through the exclusion of certain interviews outweighs the reduction of frame error resulting from their inclusion.

The multidimensional measurement of different facets of quality of life is the focus of the fourth paper. With regard to representativity, a different approach was chosen and analyses of self-reports and proxy reports were presented separately. While this results in the disadvantage of abandoning the idea of having one estimator for the whole population, it has the advantage of reducing measurement error when allowing different factor loadings between the two presented groups within the chosen latent variable approach. Furthermore, it allows for different measuring instruments in both groups, which can further reduce measurement error. In this case, the DemTect was used for the measurement of cognition in self-reports (Kalbe et al., 2004), and the Global Deterioration Scale was used for proxy reports (Reisberg et al., 1982). The approach of reweighting used in the third paper had to balance measurement error against frame error and, therefore, had to accept one of the shortcomings. This is not necessary in this approach, which enables researchers to reduce measurement error while keeping frame error constant. However, the idea of having one estimator for the whole population cannot be maintained here. Representativity is achieved in the sense of a mosaic by a combination of different analyses and results. It should explicitly be pointed out that the definitions of strong and weak representativity by Schouten, Cobben, and Bethlehem (2009) do not require one estimator for the whole population but refer only to response probabilities, with R-indicators referring to the variance of response probabilities which are not influenced by this approach. However, this approach might not be appropriate if the research question requires one estimator for the population, which is often the case in political reports as well as in research on elder abuse..

The fifth paper states that multimorbidity is common in old age: 31% of the sample reports to be treated for five or more diseases, and only 5% of the sample is currently not treated for any disease. All data is based on reports of treated conditions. The Berlin Ageing Study (Steinhagen-Thiessen & Borchelt, 2010) finds a difference between diagnoses made by general practitioners and by project physicians and concludes that underdiagnosis is common in the very old. This raises the question whether the information provided in the paper is also affected by underdiagnosis and specification error. To rule this out, professional diagnostics would be required, which would mean an increase in costs. It would also make the survey more strenuous and therefore less accessible, which can create nonresponse bias towards healthy subpopulations. In addition, it should be pointed out that the interpretation of the Berlin Ageing Study that diagnoses made by project physicians are to be preferred over those made by general practitioners, some of whom have known the very old patients for decades, is not unambiguous.

Limitations and Future Directions

The constructs of dementia, multimorbidity, and elder abuse investigated in this thesis are not sufficient to describe old age in its entirety. However, the problems in surveying these constructs among the oldest old described herein are typical for gerontological research and can be applied to further constructs. Nevertheless, the advantages and disadvantages of including hard-to-survey groups must always be weighed and final decisions depend on the research context and are construct-specific.

In all examined constructs, analytical approaches for the quantification and measurement of different kinds of errors are scarce, which might stem from the common "publish or perish" culture that fosters surprisingly significant results.

> Der Publikationsdruck, der aufwändige Publikationsprozess und der wirtschaftliche Druck mögen jedoch in manchen Fällen auch nach der Replikationskrise dazu führen, dass mit Ergebnissen unkritisch umgegangen und aus Zeitmangel keine tiefgehende Reflexion stattfindet, insbesondere, was die Möglichkeit verschiedener Arten von Fehlern anbelangt. (Spitzer & Spitzer, 2020, p. 413)
>
> However, the pressure to publish, the elaborate publication process and the economic pressure may in some cases, even after the replication crisis, lead to an uncritical approach to results and a lack of deep reflection due to a lack of time, especially with regard to the possibility of different types of errors.

Therefore, gerontology needs to promote critical research that explores and quantifies different kinds of errors. In the topics presented here, there seems to be a better understanding of measurement errors that occur between self-reports, interviews with present third persons, and proxy-interviews in dementia research, and a more transparent handling of specification errors in elder abuse and multimorbidity research, where it seems that different constructs are explored under the same name. Gerontology also needs to enable researchers to weigh different kinds of survey errors. In a first step, this requires data generation within the hard-to-survey groups, such as the nursing home population, which can trigger subsequent phases of data collection, even if they result in increased costs for social research Institutes.

Conclusion

Old age is characterised by several overlapping subpopulations that are hard to survey. This includes people with dementia, the nursing home population, people who cannot participate in surveys by self-report and are represented by proxies, or people who can only participate when third persons are present who assist the target persons. The inclusion of these groups reduces frame error and increases representativity but can lead to measurement error that needs to be understood and can sometimes be handled. One method for the reduction of measurement error is the exclusion of some subpopulations and the subsequent recalibration of survey weights to improve the representation of the excluded populations. This approach is useful when one estimator for the whole population is needed or common in the respective research field, as is the case with elder abuse prevalence estimations. Secondly, the mosaic approach can control frame error while possibly reducing measurement error. This approach is suitable whenever it is not required to calculate an estimator for the whole sample.

6 References

Abbott, A. (2011). Dementia: A problem for our age. *Nature*, *475*(7355), S2-4. https://doi.org/10.1038/475S2a

Aboud, K. S., Huo, Y., Kang, H., Ealey, A., Resnick, S. M., Landman, B. A., & Cutting, L. E. (2019). Structural covariance across the lifespan: Brain development and aging through the lens of inter-network relationships. *Human Brain Mapping*, *40*(1), 125–136. https://doi.org/10.1002/hbm.24359

Alexander, C. M., Martyr, A., Savage, S. A., & Clare, L. (2019). Measuring awareness in people with dementia: Protocol for a scoping review. *Systematic Reviews*, *8*(1), 160. https://doi.org/10.1186/s13643-019-1078-5

Algar, K., Woods, R. T., & Windle, G. (2016). Measuring the quality of life and well-being of people with dementia: A review of observational measures. *Dementia (London, England)*, *15*(4), 832–857. https://doi.org/10.1177/1471301214540163

Alzheimer Europe. (2019). *Dementia in Europe Yearbook 2019: Estimating the prevalence of dementia in Europe*. https://www.alzheimer-europe.org/sites/default/files/alzheimer_europe_dementia_in_europe_yearbook_2019.pdf

Alzheimer's Disease International. (n.d.). *Symptoms of dementia*. Retrieved December 29, 2022, from https://www.alzint.org/about/symptoms-of-dementia/

Anstey, K. J., Ee, N., Eramudugolla, R., Jagger, C., & Peters, R. (2019). A Systematic Review of Meta-Analyses that Evaluate Risk Factors for Dementia to Evaluate the Quantity, Quality, and Global Representativeness of Evidence. *Journal of Alzheimer's Disease : JAD*, *70*(s1), S165-S186. https://doi.org/10.3233/JAD-190181

Balachandran, A., & James, K. S. (2019). A multi-dimensional measure of population ageing accounting for Quantum and Quality in life years: An application of selected countries in Europe and Asia. *SSM - Population Health*, *7*, 100330. https://doi.org/10.1016/j.ssmph.2018.100330

Baltes, M. M., & Carstensen, L. L. (1996). The Process of Successful Ageing. *Ageing and Society*, *16*(4), 397–422. https://doi.org/10.1017/S0144686X00003603

Biemer, P. P. (2010). Total Survey Error: Design, Implementation, and Evaluation. *Public Opinion Quarterly*, *74*(5), 817–848. https://doi.org/10.1093/poq/nfq058

Biemer, P. P., & Lyberg, L. E. (2003). *Introduction to Survey Quality*. John Wiley & Sons, Inc. https://doi.org/10.1002/0471458740

Bittmann, F. (2022). Improving Data Quality with Sociodemographic Matching? About the Effects and Implications of Age and Gender Matching in Face-to-Face Interviews. *Journal of Applied Social Science*, *16*(1), 242–260. https://doi.org/10.1177/1936724421998257

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2021). *Introduction to meta-analysis* (Second edition). John Wiley & Sons, Inc.

Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler: Limitierte Sonderausgabe* (7. Aufl. 2010). *Springer-Lehrbuch*. Springer Berlin Heidelberg. http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1554850

Brick, J. M. (2013). Unit Nonresponse and Weighting Adjustments: A Critical Review. *Journal of Official Statistics*, *29*(3), 329–353. https://doi.org/10.2478/jos-2013-0026

Brix, J., Steinacker, G., Stadler, M., & Schneekloth, U. (2018). *Methodenbericht: NRW80.* Kantar Public.

Bronfenbrenner, U. (1986). Ecology of the family as a context for human development: Research perspectives. *Developmental Psychology*, *22*(6), 723–742. https://doi.org/10.1037/0012-1649.22.6.723

Brydges, C. R. (2019). Effect Size Guidelines, Sample Size Calculations, and Statistical Power in Gerontology. *Innovation in Aging*, *3*(4), igz036. https://doi.org/10.1093/geroni/igz036

Bullock, L., Bedson, J., Jordan, J. L., Bartlam, B., Chew-Graham, C. A., & Campbell, P. (2019). Pain assessment and pain treatment for community-dwelling people with dementia: A systematic review and narrative synthesis. *International Journal of Geriatric Psychiatry*, *34*(6), 807–821. https://doi.org/10.1002/gps.5078

Burge, R., Dawson-Hughes, B., Solomon, D. H., Wong, J. B., King, A., & Tosteson, A. (2007). Incidence and economic burden of osteoporosis-related fractures in the United States, 2005-2025. *Journal of Bone and Mineral Research : The Official Journal of the American Society for Bone and Mineral Research*, *22*(3), 465–475. https://doi.org/10.1359/jbmr.061113

Burks, H. B., Des Bordes, J. K.A., Chadha, R., Holmes, H. M., & Rianon, N. J. (2021). Quality of Life Assessment in Older Adults with Dementia: A Systematic Review. *Dementia and Geriatric Cognitive Disorders*, *50*(2), 103–110. https://doi.org/10.1159/000515317

Cai, T., Parast, L., & Ryan, L. (2010). Meta-analysis for rare events. *Statistics in Medicine*, *29*(20), 2078–2089. https://doi.org/10.1002/sim.3964

Calderón-Larrañaga, A., Vetrano, D. L., Ferrucci, L., Mercer, S. W., Marengoni, A., Onder, G., Eriksdotter, M., & Fratiglioni, L. (2019). Multimorbidity and functional impairment-bidirectional interplay, synergistic effects and common pathways. *Journal of Internal Medicine*, *285*(3), 255–271. https://doi.org/10.1111/joim.12843.

Chen, H., Zhang, N., Lu, X., & Chen, S. (2013). Caution regarding the choice of standard deviations to guide sample size calculations in clinical trials. *Clinical Trials (London, England)*, *10*(4), 522–529. https://doi.org/10.1177/1740774513490250

Clare, L. (2010). Awareness in people with severe dementia: Review and integration. *Aging & Mental Health*, *14*(1), 20–32. https://doi.org/10.1080/13607860903421029

Dasch, B., & Lenz, P. (2022). Der Sterbeort älterer Menschen mit einer Demenz : Epidemiologische Daten einer Beobachtungsstudie zu Sterbeorten in Deutschland (2001, 2011, 2017) [Place of death of older people with dementia : Epidemiological data from an observational study of places of death in Germany (2001, 2011, 2017)]. *Zeitschrift für Gerontologie und Geriatrie*, *55*(8), 673–679. https://doi.org/10.1007/s00391-021-01976-7

De Donder, L., Lang, G., Penhale, B., Ferreira-Alves, J., Tamutiene, I., Verté, D., & Luoma, M.-L. (2013). Item non-response when measuring elder abuse: Influence of methodological choices. *European Journal of Public Health*, *23*(6), 1021–1026. https://doi.org/10.1093/eurpub/cks172

Degnen, C. (2007). Minding the gap: The construction of old age and oldness amongst peers. *Journal of Aging Studies*, *21*(1), 69–80. https://doi.org/10.1016/j.jaging.2006.02.001

Deutsche Alzheimer Gesellschaft. (2019). *Demenz.Das Wichtigste.: Ein kompakter Ratgeber*. https://www.deutsche-alzheimer.de/fileadmin/Alz/pdf/Broschueren/Demenz-das_wichtigste_.pdf

Deutsche Alzheimer Gesellschaft. (2022). *Die Alzheimer-Krankheit*. https://www.deutsche-alzheimer.de/demenz-wissen/die-alzheimer-krankheit

Doblhammer-Reiter, G. (2012). *Demografie der Demenz* (1. Aufl.). *Verlag Hans Huber*. Huber. http://elibrary.hogrefe.de/9783456950501/C1

Dong, X. (2017). Prologue. In X. Dong (Ed.), *Elder Abuse* (pp. vii–xiv). Springer International Publishing.

Dong, X., Chen, R., Chang, E.-S., & Simon, M. (2013). Elder abuse and psychological well-being: A systematic review and implications for research and policy--a mini review. *Gerontology*, *59*(2), 132–142. https://doi.org/10.1159/000341652

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. vollständig überarbeitete, aktualisierte und erweiterte Auflage 2016). *SpringerLink Bücher*. Springer. http://link.springer.com/book/10.1007/978-3-642-41089-5 https://doi.org/10.1007/978-3-642-41089-5

Dumicic, K. (2011). Representative Samples. In M. Lovric (Ed.), *International Encyclopedia of Statistical Science* (pp. 1222–1224). Springer Berlin Heidelberg.

Engel, S. (2006). *Alzheimer und Demenzen: Unterstützung für Angehörige ; die Beziehung erhalten mit dem neuen Konzept der einfühlsamen Kommunikation*. Trias.

Fakhouri, T. H. I., Martin, C. B., Chen, T.-C., Akinbami, L. J., Ogden, C. L., Paulose-Ram, R., Riddles, M. K., van de Kerckhove, W., Roth, S. B., Clark, J., Mohadjer, L. K., & Fay, R. E. (2020). An Investigation of Nonresponse Bias and Survey Location Variability in the 2017-2018 National Health and Nutrition Examination Survey. *Vital and Health Statistics. Series 2, Data Evaluation and Methods Research*(185), 1–36.

Fang, B., Li, D., Yan, E., Zhou, Y., Yu, Z., & Hu, J. (2022). Associated factors of discrepancy between older adults and their family caregivers in reporting elder abuse. *Journal of Clinical Nursing.* Advance online publication. https://doi.org/10.1111/jocn.16283

Faulbaum, F. (2022). Total Survey Error. In N. Baur & J. Blasius (Eds.), *Handbuch Methoden der empirischen Sozialforschung* (pp. 567–584). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-37985-8_36

Ferrucci, L., & Kuchel, G. A. (2021). Heterogeneity of Aging: Individual Risk Factors, Mechanisms, Patient Priorities, and Outcomes. *Journal of the American Geriatrics Society*, *69*(3), 610–612. https://doi.org/10.1111/jgs.17011

Field, A., Miles, J., & Field, Z. (2014). *Discovering statistics using R* (Repr). Sage.

Foo, H., Mather, K. A., Thalamuthu, A., & Sachdev, P. S. (2019). The many ages of man: Diverse approaches to assessing ageing-related biological and psychological measures and their relationship to chronological age. *Current Opinion in Psychiatry*, *32*(2), 130–137. https://doi.org/10.1097/YCO.0000000000000473

Fortin, M., Stewart, M., Poitras, M.-E., Almirall, J., & Maddocks, H. (2012). A systematic review of prevalence studies on multimorbidity: Toward a more uniform methodology. *Annals of Family Medicine*, *10*(2), 142–151. https://doi.org/10.1370/afm.1337

France, E. F., Wyke, S., Gunn, J. M., Mair, F. S., McLean, G., & Mercer, S. W. (2012). Multimorbidity in primary care: A systematic review of prospective cohort studies. *The British Journal of General Practice : The Journal of the Royal College of General Practitioners*, *62*(597), e297-307. https://doi.org/10.3399/bjgp12X636146

Gaebel, W., Jessen, F., & Kanba, S. (2018). Neurocognitive disorders in ICD-11: The debate and its outcome. *World Psychiatry*, *17*(2), 229–230. https://doi.org/10.1002/wps.20534

Görgen, T. (2009). Sicherheitsgefährdungen im Alter - Schärfe- und Unschärfebereiche wissenschaftlicher Erkenntnis. In T. Görgen (Ed.), *"Sicherer Hafen" oder "gefahrvolle Zone"? Kriminalitäts- und Gewalterfahrungen im Leben alter Menschen* (pp. 511–520). https://www.bmfsfj.de/resource/blob/94186/727885d90cc7178a738226f533bce61d/kriminalitaets-gewalterfahrungen-leben-alter-menschen-langfassung-data.pdf

Görgen, T., Herbst, S., & Rabold, S. (2009). Kriminalitäts- und Gewalterfahrungen alter Menschen. In T. Görgen (Ed.), *"Sicherer Hafen" oder "gefahrvolle Zone"? Kriminalitäts- und Gewalterfahrungen im Leben alter Menschen* (pp. 66–174). https://www.bmfsfj.de/resource/blob/94186/727885d90cc7178a738226f533bce61d/kriminalitaets-gewalterfahrungen-leben-alter-menschen-langfassung-data.pdf

Griffiths, A. W., Smith, S. J., Martin, A., Meads, D., Kelley, R., & Surr, C. A. (2020). Exploring self-report and proxy-report quality-of-life measures for people living with dementia in care homes. *Quality of Life Research : An International Journal of Quality of Life*

Aspects of Treatment, Care and Rehabilitation, *29*(2), 463–472. https://doi.org/10.1007/s11136-019-02333-3

Groves, R. M., & Lyberg, L. (2010). Total Survey Error: Past, Present, and Future. *Public Opinion Quarterly*, *74*(5), 849–879. https://doi.org/10.1093/poq/nfq065

Gutzmann, H., & Zank, S. (2005). *Demenzielle Erkrankungen: Medizinische und psychosoziale Interventionen* (1. Aufl.). *Kohlhammer Urban-Taschenbücher: Vol. 767*. Kohlhammer. http://www.socialnet.de/rezensionen/isbn.php?isbn=978-3-17-017658-4

Halbesleben, J. R. B., & Whitman, M. V. (2013). Evaluating survey quality in health services research: A decision framework for assessing nonresponse bias. *Health Services Research*, *48*(3), 913–930. https://doi.org/10.1111/1475-6773.12002

Hank, K., Schulz-Nieswandt, F., Wagner, M., & Zank, S. (2019). *Alternsforschung: Handbuch für Wissenschaft und Praxis* (1. Auflage). Nomos. https://doi.org/10.5771/9783845276687

Hansen, S., Kaspar, R., Wagner, M., Woopen, C., & Zank, S. (2021). Die NRW80+ Hochaltrigenstudie: Konzeptueller Hintergrund und Untersuchungsgruppen [The NRW80+ study: Conceptual background and study groups]. *Zeitschrift Für Gerontologie Und Geriatrie*, *54*(Suppl 2), 76–84. https://doi.org/10.1007/s00391-021-01970-z

Ho, C. S., Wong, S. Y., Chiu, M. M., & Ho, R. C. (2017). Global Prevalence of Elder Abuse: A Meta-analysis and Meta-regression. *East Asian Archives of Psychiatry : Official Journal of the Hong Kong College of Psychiatrists = Dong Ya Jing Shen Ke Xue Zhi : Xianggang Jing Shen Ke Yi Xue Yuan Qi Kan*, *27*(2), 43–55.

Ho, I. S.-S., Azcoaga-Lorenzo, A., Akbari, A., Black, C., Davies, J., Hodgins, P., Khunti, K., Kadam, U., Lyons, R. A., McCowan, C., Mercer, S., Nirantharakumar, K., & Guthrie, B. (2021). Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies. *The Lancet Public Health*, *6*(8), e587-e597. https://doi.org/10.1016/S2468-2667(21)00107-9

Holden, T. R., Keller, S., Kim, A., Gehring, M., Schmitz, E., Hermann, C., Gilmore-Bykovskyi, A., & Kind, A. J. H. (2018). Procedural Framework to Facilitate Hospital-Based Informed Consent for Dementia Research. *Journal of the American Geriatrics Society*, *66*(12), 2243–2248. https://doi.org/10.1111/jgs.15525

Holzhausen, M., & Scheidt-Nave, C. (2012). Multimorbidität als Interventionsherausforderung. In H.-W. Wahl, C. Tesch-Römer, & J. P. Ziegelmann (Eds.), *Angewandte Gerontologie: Interventionen für ein gutes Altern in 100 Schlüsselbegriffen* (2nd ed., pp. 48–53). Kohlhammer.

Hörnsten, C., Littbrand, H., Boström, G., Rosendahl, E., Lundin-Olsson, L., Nordström, P., Gustafson, Y., & Lövheim, H. (2021). Measurement error of the Mini-Mental State Examination among individuals with dementia that reside in nursing homes. *European Journal of Ageing*, *18*(1), 109–115. https://doi.org/10.1007/s10433-020-00572-9

Hounsome, N., Orrell, M., & Edwards, R. T. (2011). Eq-5D as a quality of life measure in people with dementia and their carers: Evidence and key issues. *Value in Health : The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, *14*(2), 390–399. https://doi.org/10.1016/j.jval.2010.08.002

Jia, P., Lin, L., Kwong, J. S. W., & Xu, C. (2021). Many meta-analyses of rare events in the Cochrane Database of Systematic Reviews were underpowered. *Journal of Clinical Epidemiology*, *131*, 113–122. https://doi.org/10.1016/j.jclinepi.2020.11.017

Johannesen, M., & LoGiudice, D. (2013). Elder abuse: A systematic review of risk factors in community-dwelling elders. *Age and Ageing*, *42*(3), 292–298. https://doi.org/10.1093/ageing/afs195

Johnston, M. C., Crilly, M., Black, C., Prescott, G. J., & Mercer, S. W. (2019). Defining and measuring multimorbidity: A systematic review of systematic reviews. *European Journal of Public Health*, *29*(1), 182–189. https://doi.org/10.1093/eurpub/cky098

Kalbe, E., Kessler, J., Calabrese, P., Smith, R., Passmore, A. P., Brand, M., & Bullock, R. (2004). Demtect: A new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. *International Journal of Geriatric Psychiatry*, *19*(2), 136–143. https://doi.org/10.1002/gps.1042

Kieser, M. (2018). *Fallzahlberechnung in der Medizinischen Forschung: Eine Einführung Für Mediziner und Biostatistiker*. *Essentials*. Springer Fachmedien Wiesbaden GmbH. https://livivo.idm.oclc.org/login?url=https://ebookcentral.proquest.com/lib/zbmed-ebooks/detail.action?docID=5252877

Kruskal, W., & Mosteller, F. (1979a). Representative Sampling, I: Non-Scientific Literature. *International Statistical Review*, *47*(1), 13–24.

Kruskal, W., & Mosteller, F. (1979b). Representative Sampling, II: Scientific Literature, Excluding Statistics. *International Statistical Review*, *47*(2), 111–127.

Kruskal, W., & Mosteller, F. (1979c). Representative Sampling, III: The Current Statistical Literature. *International Statistical Review*, *47*(3), 245–265.

Kruskal, W., & Mosteller, F. (1980). Representative Sampling, IV: The History of the Concept in Statistics, 1895-1939. *International Statistical Review*, *48*(2), 169–195.

Kutzleben, M. von, Köhler, K., Dreyer, J., Holle, B., & Roes, M. (2017). Stabilität von häuslichen Versorgungsarrangements für Menschen mit Demenz : Entwicklung und Konsentierung einer Definition von Stabilität durch Expertenfokusgruppen [Stability of home based care arrangements for people with dementia : Development of a consensus definition of stability using expert focus groups]. *Zeitschrift für Gerontologie und Geriatrie*, *50*(3), 210–218. https://doi.org/10.1007/s00391-015-0990-0

Kydd, A., Fleming, A., Paoletti, I., & Hvalič Touzery, S. (2020). Exploring Terms Used for the Oldest Old in the Gerontological Literature. *The Journal of Aging and Social Change*, *10*(2), 53–73. https://doi.org/10.18848/2576-5310/CGP/v10i02/53-73

Lachs, M. S., & Pillemer, K. (2004). Elder abuse. *Lancet (London, England)*, *364*(9441), 1263–1272. https://doi.org/10.1016/S0140-6736(04)17144-4

Lampert, T., & Hoebel, J. (2019). Sozioökonomische Unterschiede in der Gesundheit und Pflegebedürftigkeit älterer Menschen [Socioeconomic differences in health and need for care among the elderly]. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, *62*(3), 238–246. https://doi.org/10.1007/s00103-019-02917-x

Lang, F. R., Rupprecht, R., Schulz, M., & Rohleder, N. (2019). *Die munteren Neunziger.: Leben in der zehnten Dekade in Nürnberg.* Friedrich-Alexander Universität Erlangen-Nürnberg. www.geronto.fau.de/lebendig

Lautenbacher, S., & Kunz, M. (2019). Schmerzerfassung bei Patienten mit Demenz [Assessing pain in patients with dementia]. *Schmerz (Berlin, Germany)*, *33*(6), 563–575. https://doi.org/10.1007/s00482-019-00411-6

Lavrakas, P. J. (Ed.). (2008). *Encyclopedia of survey research methods*. SAGE Publications. https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=996578

Lee, S., Brown, E. R., Grant, D., Belin, T. R., & Brick, J. M. (2009). Exploring nonresponse bias in a health survey using neighborhood characteristics. *American Journal of Public Health*, *99*(10), 1811–1817. https://doi.org/10.2105/AJPH.2008.154161

Lee, T., & Shi, D. (2021). A comparison of full information maximum likelihood and multiple imputation in structural equation modeling with missing data. *Psychological Methods*, *26*(4), 466–485. https://doi.org/10.1037/met0000381

Leonhart, R. (2009). *Lehrbuch Statistik: Einstieg und Vertiefung* (2., überarb. und erw. Aufl.). *Psychologie-Lehrbuch*. Huber. http://sub-hh.ciando.com/book/?bok_id=25456

Lindenberger, U. (2000). Intellektuelle Entwicklung über die Lebensspanne: Überblick und ausgewählte Forschungsbrennpunkte. *Psychologische Rundschau*, *51*(3), 135–145. https://doi.org/10.1026//0033-3042.51.3.135

Locker, D. (2000). Response and nonresponse bias in oral health surveys. *Journal of Public Health Dentistry*, *60*(2), 72–81. https://doi.org/10.1111/j.1752-7325.2000.tb03298.x

Lord, F. M. (1968). Some TEST THEORY FOR TAILORED TESTING *ETS Research Bulletin Series*, *1968*(2), i-62. https://doi.org/10.1002/j.2333-8504.1968.tb00562.x

Mahlberg, R., & Gutzmann, H. (2009). Einleitung: Krankheitsbilder. In R. Mahlberg, H. Gutzmann, & G. Adler (Eds.), *Demenzerkrankungen: Erkennen, behandeln und versorgen* (pp. 3–11). Deutscher Ärzte-Verl.

Makovski, T. T., Schmitz, S., Zeegers, M. P., Stranges, S., & van den Akker, M. (2019). Multimorbidity and quality of life: Systematic literature review and meta-analysis. *Ageing Research Reviews*, *53*, 100903. https://doi.org/10.1016/j.arr.2019.04.005

Manti, S., & Licari, A. (2018). How to obtain informed consent for research. *Breathe (Sheffield, England)*, *14*(2), 145–152. https://doi.org/10.1183/20734735.001918

Marengoni, A., Angleman, S., Melis, R., Mangialasche, F., Karp, A., Garmen, A., Meinow, B., & Fratiglioni, L. (2011). Aging with multimorbidity: A systematic review of the literature. *Ageing Research Reviews*, *10*(4), 430–439. https://doi.org/10.1016/j.arr.2011.03.003

Maresova, P., Javanmardi, E., Barakovic, S., Barakovic Husic, J., Tomsone, S., Krejcar, O., & Kuca, K. (2019). Consequences of chronic diseases and other limitations associated with old age - a scoping review. *BMC Public Health*, *19*(1), 1431. https://doi.org/10.1186/s12889-019-7762-5

Marmolejo, I. I. (2008). *Elder Abuse in the Family in Spain.* http://www.inpea.net/images/Spain_Report_2008_Elder.pdf

Martyr, A., Nelis, S. M., Quinn, C., Wu, Y.-T., Lamont, R. A., Henderson, C., Clarke, R., Hindle, J. V., Thom, J. M., Jones, I. R., Morris, R. G., Rusted, J. M., Victor, C. R., & Clare, L. (2018). Living well with dementia: A systematic review and correlational meta-analysis of factors associated with quality of life, well-being and life satisfaction in people with dementia. *Psychological Medicine*, *48*(13), 2130–2139. https://doi.org/10.1017/S0033291718000405

Michalowsky, B., Xie, F., Kohlmann, T., Gräske, J., Wübbeler, M., Thyrian, J. R., & Hoffmann, W. (2020). Acceptability and Validity of the EQ-5D in Patients Living With Dementia. *Value in Health : The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, *23*(6), 760–767. https://doi.org/10.1016/j.jval.2020.01.022

Mulford, C. F., & Mao, A. (2017). Department of Justice: Elder Maltreatment Initiatives. In X. Dong (Ed.), *Elder Abuse* (pp. 637–651). Springer International Publishing. https://doi.org/10.1007/978-3-319-47504-2_31

Neise. In

Neise, M., Janhsen, A., Geithner, L., Schmitz, W., & Kaspar, R. (2019). Lebensqualitäten in der Hochaltrigkeit. In K. Hank, F. Schulz-Nieswandt, M. Wagner, S. Zank, H. Baranzke, H. Brandenburg, & S. Brose (Eds.), *Alternsforschung: Handbuch für Wissenschaft und Praxis* (1st ed., pp. 581–604). Nomos. https://doi.org/10.5771/9783845276687-581

Nguyen, Q. D., Moodie, E. M., Forget, M.-F., Desmarais, P., Keezer, M. R., & Wolfson, C. (2021). Health Heterogeneity in Older Adults: Exploration in the Canadian Longitudinal Study on Aging. *Journal of the American Geriatrics Society*, *69*(3), 678–687. https://doi.org/10.1111/jgs.16919

Olsen, C. F., & Bergland, A. (2017). "Reliability of the Norwegian version of the short physical performance battery in older people with and without dementia". *BMC Geriatrics*, *17*(1), 124. https://doi.org/10.1186/s12877-017-0514-4

Orgeta, V., Edwards, R. T., Hounsome, B., Orrell, M., & Woods, B. (2015). The use of the EQ-5D as a measure of health-related quality of life in people with dementia and their carers. *Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *24*(2), 315–324. https://doi.org/10.1007/s11136-014-0770-0

O'Shea, E., Hopper, L., Marques, M., Gonçalves-Pereira, M., Woods, B., Jelley, H., Verhey, F., Kerpershoek, L., Wolfs, C., Vugt, M. de, Stephan, A., Bieber, A., Meyer, G., Wimo, A., Michelet, M., Selbaek, G., Portolani, E., Zanetti, O., & Irving, K. (2020). A comparison of self and proxy quality of life ratings for people with dementia and their carers: A European prospective cohort study. *Aging & Mental Health*, *24*(1), 162–170. https://doi.org/10.1080/13607863.2018.1517727

Patrician, P. A. (2002). Multiple imputation for missing data. *Research in Nursing & Health*, *25*(1), 76–84. https://doi.org/10.1002/nur.10015

Pillemer, K., Burnes, D., & MacNeil, A. (2021). Investigating the connection between ageism and elder mistreatment. *Nature Aging*, *1*(2), 159–164. https://doi.org/10.1038/s43587-021-00032-8

Plugge, M. (2021). Successful ageing in the oldest old: Objectively and subjectively measured evidence from a population-based survey in Germany. *European Journal of Ageing*, *18*(4), 537–547. https://doi.org/10.1007/s10433-021-00609-7

Poth, A., Penger, S., Knebel, M., Müller, T., Pantel, J., Oswald, F., & Haberstroh, J. (2022). Empowering patients with dementia to make legally effective decisions: A randomized controlled trial on enhancing capacity to consent to treatment. *Aging & Mental Health*, 1–9. https://doi.org/10.1080/13607863.2021.2024797

Putnick, D. L., & Bornstein, M. H. (2016). Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. *Developmental Review : DR*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Puvill, T., Lindenberg, J., Craen, A. J. M. de, Slaets, J. P. J., & Westendorp, R. G. J. (2016). Impact of physical and mental health on life satisfaction in old age: A population based observational study. *BMC Geriatrics*, *16*(1), 194. https://doi.org/10.1186/s12877-016-0365-4

Reisberg, B., Ferris, S. H., Leon, M. J. de, & Crook, T. (1982). The Global Deterioration Scale for assessment of primary degenerative dementia. *The American Journal of Psychiatry*, *139*(9), 1136–1139. https://doi.org/10.1176/ajp.139.9.1136

Riedel-Heller, S. G., Schork, A., Matschinger, H., & Angermeyer, M. C. (2000). Recruitment procedures and their impact on the prevalence of dementia. Results from the Leipzig Longitudinal Study of the Aged (LEILA75+). *Neuroepidemiology*, *19*(3), 130–140. https://doi.org/10.1159/000026248

Rivadeneira, M. F., Mendieta, M. J., Villavicencio, J., Caicedo-Gallardo, J., & Buendía, P. (2021). A multidimensional model of healthy ageing: Proposal and evaluation of determinants based on a population survey in Ecuador. *BMC Geriatrics*, *21*(1), 615. https://doi.org/10.1186/s12877-021-02548-5

Robert Koch-Institut. (2015). *Gesundheit in Deutschland. Gesundheitsberichterstattung des Bundes.* https://doi.org/10.17886/rkipubl-2015-003

Roberto, K. A., & Teaster, P. B. (2017). Theorizing Elder ABuse. In X. Dong (Ed.), *Elder Abuse* (pp. 21–41). Springer International Publishing.

Rodríguez-García, M. D. P., Ayala, A., Rodríguez-Blázquez, C., Martínez-Martín, P., Forjaz, M. J., & Damián, J. (2019). Features and impact of missing values in the association of self-rated health with mortality in care homes: A longitudinal study. *Health and Quality of Life Outcomes*, *17*(1), 111. https://doi.org/10.1186/s12955-019-1184-z

Rosenman, L., & Scott, W. (2009). Financing Old Age: Why is There Still Gender Inequality? *Australian Social Work*, *62*(2), 287–298. https://doi.org/10.1080/03124070902835998

Rowe, J. W., & Kahn, R. L. (1997). Successful aging. *The Gerontologist*, *37*(4), 433–440. https://doi.org/10.1093/geront/37.4.433

Rowe, J. W., & Kahn, R. L. (2015). Successful Aging 2.0: Conceptual Expansions for the 21st Century. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *70*(4), 593–596. https://doi.org/10.1093/geronb/gbv025

Sadana, R. (2000). Measuring reproductive health: review of community-based approaches to assessing morbidity. *Bulletin of the World Health Organization*(78), 640–654. https://www.scielosp.org/article/ssm/content/raw/?resource_ssm_path=/media/assets/bwho/v78n5/v78n5a09.pdf

Särndal, C.-E., & Lundström, S. (2008). Assessing auxiliary vectors for control of nonresponse bias in the calibration estimator. *Journal of Official Statistics*, *24*, 167–191.

Sathyanarayana Rao, T. S., Jacob, K. S., Shaji, K. S., Raju, M. S. V. K., Bhide, A. V., Rao, G. P., Saha, G., & Jagiwala, M. (2017). Dementia and the International Classification of Diseases-11 (Beta Version). *Indian Journal of Psychiatry*, *59*(1), 1–2. https://doi.org/10.4103/psychiatry.IndianJPsychiatry_66_17

Scherpenzeel, A., Axt, K., Bergmann, M., Douhou, S., Oepen, A., Sand, G., Schuller, K., Stuck, S., Wagner, M., & Börsch-Supan, A. (2020). Collecting survey data among the 50+ population during the COVID-19 outbreak: The Survey of Health, Ageing and Retirement in Europe (SHARE). Advance online publication. https://doi.org/10.18148/srm/2020.v14i2.7738 (217-221 Pages / Survey Research Methods, Vol 14 No 2 (2020): Survey Research Methods During the COVID-19 Crisis).

Schminkey, D. L., Oertzen, T. von, & Bullock, L. (2016). Handling Missing Data With Multilevel Structural Equation Modeling and Full Information Maximum Likelihood Techniques. *Research in Nursing & Health*, *39*(4), 286–297. https://doi.org/10.1002/nur.21724

Schnell, R., Hill, P. B., & Esser, E. (2018). *Methoden der empirischen Sozialforschung* (11., überarbeitete Auflage). *De Gruyter Studium*. De Gruyter Oldenbourg. http://www.blickinsbuch.de/item/650655bd5c7d06de54d85e500b16a14d

Schouten, B., Cobben, F., & Bethlehem, J. (2009). Indicators for the representativeness of survey response. *Survey Methodology*, *35*(1), 101–113. https://www150.statcan.gc.ca/n1/pub/12-001-x/2009001/article/10887-eng.pdf

Schouten, B., Morren, M., Bethlehem, J., Shlomo, N., & Skinner, C. (2009). *How to use R-indicators: Workpackage 4 Delivarable 3.* RISQ Project. https://hummedia.manchester.ac.uk/institutes/cmist/risq/RISQ-Deliverable-3.pdf

Sheehan, B. D., Lall, R., Stinton, C., Mitchell, K., Gage, H., Holland, C., & Katz, J. (2012). Patient and proxy measurement of quality of life among general hospital in-patients with dementia. *Aging & Mental Health*, *16*(5), 603–607. https://doi.org/10.1080/13607863.2011.653955

Shlomo, N., Skinner, C., Schouten, B., Bethlehem, J., & Zhang, L.-C. (2009). *Statistical Properties of R-indicator*. https://hummedia.manchester.ac.uk/institutes/cmist/risq/RISQ-Deliverable-2-1-V2.pdf

Silva-Fhon, J. R., Del Rio Suarez, A. D., Motta Herrera, S. N., Coelho Fabricio-Webhe, S. C., & Partezani Rodrigues, R. A. (2015). VIOLENCIA INTRAFAMILIAR EN EL ADULTO MAYOR QUE VIVE EN EL DISTRITO DE BREÑA, PERÚ. *Revista De La Facultad De Medicina*, *63*(3), 367–375. https://doi.org/10.15446/revfacmed.v63n3.44743

Skou, S. T., Mair, F. S., Fortin, M., Guthrie, B., Nunes, B. P., Miranda, J. J., Boyd, C. M., Pati, S., Mtenga, S., & Smith, S. M. (2022). Multimorbidity. *Nature Reviews. Disease Primers*, *8*(1), 48. https://doi.org/10.1038/s41572-022-00376-4

Slaughter, S., Cole, D., Jennings, E., & Reimer, M. A. (2007). Consent and assent to participate in research from people with dementia. *Nursing Ethics*, *14*(1), 27–40. https://doi.org/10.1177/0969733007071355

Smeding, H. M., & Koning, I. de (2000). Frontotemporal dementia and neuropsychology: The value of missing values. *Journal of Neurology, Neurosurgery, and Psychiatry*, *68*(6), 726–730. https://doi.org/10.1136/jnnp.68.6.726

Spitzer, M. W. H., & Spitzer, M. (2020). Die Replikationskrise in der Psychologie. *Nervenheilkunde*, *39*(06), 404–416. https://doi.org/10.1055/a-1095-0144

Starfield, B., & Kinder, K. (2011). Multimorbidity and its measurement. *Health Policy (Amsterdam, Netherlands)*, *103*(1), 3–8. https://doi.org/10.1016/j.healthpol.2011.09.004

Statistisches Bundesamt. (2022). *Pflegequote nach Altersgruppen*. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Pflege/_inhalt.html;jsessionid=DD9A6835403EADC919F0D55A8660A73C.live712#250648

Stechl, E., Lämmler, G., Steinhagen-Thiessen, E., & Flick, U. (2007). Subjektive Wahrnehmung und Bewältigung der Demenz im Frühstadium - SUWADEM: Eine qualitative Interviewstudie mit Betroffenen und Angehörigen [Subjective perception of and coping with dementia in its early stages. A qualitative interview study with people with dementia and their relatives - SUWADEM]. *Zeitschrift fur Gerontologie und Geriatrie*, *40*(2), 71–80. https://doi.org/10.1007/s00391-007-0434-6

Steinhagen-Thiessen, E., & Borchelt, M. (2010). Morbidität, Medikation und Funktionalität im Alter. In U. Lindenberger (Ed.), *Interdisziplinäre Arbeitsgruppen. Forschungsberichte: Vol. 3. Die Berliner Altersstudie* (3rd ed., pp. 175–207). Akad.-Verl.

Storey, J. E. (2020). Risk factors for elder abuse and neglect: A review of the literature. *Aggression and Violent Behavior*, *50*, 101339. https://doi.org/10.1016/j.avb.2019.101339

Teaster, P. B., Wangmo, T., & Anetzberger, G. J. (2010). A glass half full: The dubious history of elder abuse policy. *Journal of Elder Abuse & Neglect*, *22*(1-2), 6–15. https://doi.org/10.1080/08946560903436130

Tesch-Römer, C., & Wahl, H.-W. (2017). Toward a More Comprehensive Concept of Successful Aging: Disability and Care Needs. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *72*(2), 310–318. https://doi.org/10.1093/geronb/gbw162

United Nations. (2021). *World Population Ageing 2019*. United Nations.

van Deun, B., Hobbelen, J. S. M., Cagnie, B., van Eetvelde, B., van den Noortgate, N., & Cambier, D. (2018). Reproducible Measurements of Muscle Characteristics Using the MyotonPRO Device: Comparison Between Individuals With and Without Paratonia. *Journal of Geriatric Physical Therapy (2001)*, *41*(4), 194–203. https://doi.org/10.1519/JPT.0000000000000119

Vandenberg, R. J., & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods*, *3*(1), 4–70. https://doi.org/10.1177/109442810031002

Veenhoven, R. (2013). The Four Qualities of Life Ordering Concepts and Measures of the Good Life. In A. Delle Fave (Ed.), *Happiness Studies Book Ser. The Exploration of Happiness: Present and Future Perspectives* (1st ed., pp. 195–226). Springer Netherlands. https://doi.org/10.1007/978-94-007-5702-8_11

Vidovičová, L., & Doseděl, T. (2018). Who Should Interview Older People? The Effect of Interviewer and Interviewee Characteristics in Surveys of Older People and Aging Topics. *Sociológia - Slovak Sociological Review*, *50*(6). https://doi.org/10.31577/sociologia.2018.50.6.28

Violan, C., Foguet-Boreu, Q., Flores-Mateo, G., Salisbury, C., Blom, J., Freitag, M., Glynn, L., Muth, C., & Valderas, J. M. (2014). Prevalence, determinants and patterns of multimorbidity in primary care: A systematic review of observational studies. *PloS One*, *9*(7), e102149. https://doi.org/10.1371/journal.pone.0102149

Visschers, J., Jaspaert, E., & Vervaeke, G. (2017). Social Desirability in Intimate Partner Violence and Relationship Satisfaction Reports: An Exploratory Analysis. *Journal of Interpersonal Violence*, *32*(9), 1401–1420. https://doi.org/10.1177/0886260515588922

Wagner, M., Rietz, C., Kaspar, R., Janhsen, A., Geithner, L., Neise, M., Kinne-Wall, C., Woopen, C., & Zank, S. (2018). Lebensqualität von Hochaltrigen : Studie zu Lebensqualität und subjektivem Wohlbefinden hochaltriger Menschen in Nordrhein-Westfalen (NRW80+) [Quality of life of the very old : Survey on quality of life and subjective well-being of the very old in North Rhine-Westphalia (NRW80+)]. *Zeitschrift Für Gerontologie Und Geriatrie*, *51*(2), 193–199. https://doi.org/10.1007/s00391-017-1217-3

Weuve, J., Proust-Lima, C., Power, M. C., Gross, A. L., Hofer, S. M., Thiébaut, R., Chêne, G., Glymour, M. M., & Dufouil, C. (2015). Guidelines for reporting methodological challenges and evaluating potential bias in dementia research. *Alzheimer's & Dementia : The Journal of the Alzheimer's Association*, *11*(9), 1098–1109. https://doi.org/10.1016/j.jalz.2015.06.1885

Wied, T. S., Haberstroh, J., Gather, J., Karakaya, T., Oswald, F., Qubad, M., Scholten, M., Vollmann, J., & Pantel, J. (2021). Supported Decision-Making in Persons With Dementia: Development of an Enhanced Consent Procedure for Lumbar Puncture. *Frontiers in Psychiatry*, *12*, 780276. https://doi.org/10.3389/fpsyt.2021.780276

Wied, T. S., Poth, A., Pantel, J., Oswald, F., & Haberstroh, J. (2021). Wie bewerten Demenzforscher:Innen Entscheidungsassistenz für die informierte Einwilligung zur Forschungsteilnahme von Menschen mit Demenz? [How do dementia researchers view support tools for informed consent procedures of persons with dementia?]. *Zeitschrift Für Gerontologie Und Geriatrie*, *54*(7), 667–675. https://doi.org/10.1007/s00391-020-01779-2

World Health Organization. *Global report an ageism*. https://apps.who.int/iris/rest/bitstreams/1336324/retrieve

World Health Organization. (2019). *ICD-10 Version: 2019*. https://icd.who.int/browse10/2019/en#/F00-F09

World Health Organization. (2022). *ICD-11*. https://icd.who.int/en

World Health Organization; Université de Genève. (2008). *A global response to elder abuse and neglect: Building primary health care capacity to deal with the problem worldwide ; main report*. World Health Organization. http://site.ebrary.com/lib/academiccompletetitles/home.action

Yon, Y., Mikton, C. R., Gassoumis, Z. D., & Wilber, K. H. (2017). Elder abuse prevalence in community settings: a systematic review and meta-analysis. *The Lancet Global Health*, *5*(2), e147-e156. https://doi.org/10.1016/S2214-109X(17)30006-2

Yon, Y., Ramiro-Gonzalez, M., Mikton, C. R., Huber, M., & Sethi, D. (2019). The prevalence of elder abuse in institutional settings: A systematic review and meta-analysis. *European Journal of Public Health*, *29*(1), 58–67. https://doi.org/10.1093/eurpub/cky093

Zank, S., Schacke, C., & Leipold, B. (2007). Längsschnittstudie zur Belastung pflegender Angehöriger von demenziell Erkrankten (LEANDER). *Zeitschrift Für Gerontopsychologie & -Psychiatrie*, *20*(4), 239–255. https://doi.org/10.1024/1011-6877.20.4.239

Zhou, S., & Shen, C. (2022). Statistical concerns for meta-analysis of rare events and small sample sizes. *The Lancet. Infectious Diseases*, *22*(8), 1111–1112. https://doi.org/10.1016/S1473-3099(22)00364-4