Sampling and non-response bias on health-outcomes in surveys of the oldest old

Susanne Kelfve · Mats Thorslund · Carin Lennartsson

Published online: 26 March 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract Surveys of the oldest old population are associated with several design issues. Place of residence and possible physical or cognitive impairments make it difficult to maintain a representative study population. Based on a Swedish nationally representative survey among individuals 77+, the present study analyze the potential bias of not using proxy interviews and excluding the institutionalized part of the population in surveys of the oldest old. The results show that compared to directly interviewed people living at home, institutionalized and proxy interviewed individuals were older, less educated and more likely to be female. They had more problems with health, mobility and ADL, and a significantly increased mortality risk. If the study had excluded the institutionalized part of the population and/or failed to use proxy interviews, the result would have been severely biased and resulted in underestimated prevalence rates for ADL, physical mobility and psychologic problems. This could not be compensated for weighting the data by age and sex. The results from this study imply that accurate population estimates require a representative study population, in which all individuals are included regardless of their living conditions, health status, and cognitive ability.

Keywords Representative study population · Survey design · Proxy interviews · Institutionalized · Non-response · Oldest old

Introduction

During recent decades, researchers have dealt with inconsistent findings concerning health among the oldest sector in the population. There have been difficulties in generating comparable data, particularly data on the prevalence of old-age disability, both within and between countries (Freedman et al. 2004; Lafortune and Balestat 2007). Prevalence differences found between countries may indeed reflect real health differences, but studies based on the same population, within a country, are expected to generate similar results.

To explain the conflicting findings on health and disability among the oldest population, several problems related to study design have been discussed: exclusion or under-representation of institutionalized individuals, different time periods under study, different disability definitions or health indicators used, different interview modes, varying age distributions, different wording of questions, item non-response, differential loss to follow-up, and the use of proxy interviews (Freedman et al. 2002, 2004; Martin et al. 2010; Parker and Thorslund 2007; Wolf et al. 2005).

Here we will focus on the problems related to non-representative study populations, as a source of inconsistent findings in surveys of the oldest old. Our concern is that even with an appropriate choice of indicators and adequate measures, a non-representative study sample will provide bias in every outcome that correlates with the probability of certain individuals participating in the survey.

In surveys of the oldest old, the probability of both sample inclusion and response rate is correlated with health and functional ability (Chatfield et al. 2005; Hardy et al. 2009; Lundberg and Thorslund 1996; Manton and Suzman

Responsible Editor: D. J. H. Deeg.

S. Kelfve (✉)
Department of Sociology, Stockholm University, Stockholm,
Sweden
e-mail: susanne.kelfve@ki.se

S. Kelfve · M. Thorslund · C. Lennartsson
Aging Research Center, Karolinska Institutet and Stockholm
University, Stockholm, Sweden

1992). In very old age, institutional living and impaired health is more prevalent than in younger age groups. How to handle the institutionalized part of the population and individuals too frail or cognitively impaired to participate in a survey are issues every researcher has to consider when conducting a survey of the oldest old. Besides the ethical considerations when approaching a person with dementia, researchers also have to take into account country-specific laws and sometimes limited population registers for sampling.

Consequently, many surveys exclude institutionalized individuals. This is not controversial when the research question and the inferences made only concern the community-based population. However, in studies of the health status among the oldest old, institutionalized individuals comprise an important sector of the oldest population (Lafortune and Balestat 2007). The probability of living in an institution during old age is shown to be associated with several factors: age, sex, education, marital status, socioeconomic factors such as income and home ownership, smoking and alcohol consumption and health status, including chronic conditions, dementia, depressive symptoms, limitation in Activities of Daily Living (ADL), mobility problems, and hearing and vision impairment, (Asakawa et al. 2009; Larsson et al. 2006; Martikainen et al. 2008; Noël-Miller 2010; Rodgers and Herzog 1992; Wallace et al. 1992). Thus, excluding institutionalized individuals generates a sample-bias that might affect all factors that correlate with the probability of institutional living in old age. Furthermore, the fact that the probability of institutional living varies across countries and over time also presents problems for comparative studies. Policy changes entailing expansion of home-based care have, e.g., resulted in decreasing levels of institutional care in many countries (Huber and Hennessy 2005).

Another methodological issue faced by many researchers is the choice between non-response and obtaining proxy data. At high ages, the increased prevalence of poor health or cognitive impairment makes it impossible to conduct an interview with all older persons in the sample. One alternative to non-response, in these cases, is to conduct an indirect interview with a person close to the old individual, a person usually referred to as a proxy. To include or exclude the institutionalized individuals in the target population is a decision made by the researcher. The conduction of proxy interviews, on the other hand, is dependent both on whether this option is provided by the researcher and on decisions taken in the interaction between the interviewer and the respondent when planning the interview.

Extensive research on the use of proxy information is available. In general, the use of proxies shows fairly good validity for easily observed variables such as mobility and

physical function and poorer validity for more subjective health measures. The overall tendency is that proxies report more health problems than older persons and that this tendency is positively correlated with care-giving burden (e.g., Neumann et al. 2000).

Despite the widely acknowledged importance of a representative study population, there are considerable differences in studies of the oldest population regarding sample design, response rate, and the use of proxies (Gudex and Lafortune 2000). It is well-known that proxy interviewed and institutionalized individuals tend to be older, less healthy, and have a higher risk of mortality. The primary objective in population surveys is that the interviewed group does not differ substantially from the population that the sample is supposed to represent. In this study we estimate the magnitude of bias, caused by coverage problems and the use or non-use of proxy interviews, in surveys of the oldest old.

Aims

The aim is to analyze the effect of study design in a population-based survey of the oldest old on prevalence rates for various health outcomes. Our main focus is on the consequences of including or excluding institutionalized individuals and of using or not using proxy interviews.

First, we describe the characteristics, health and mortality risk of the proxy interviewed and the institutionalized and compare them to the directly interviewed individuals who were living at home. Second, we describe how exclusion of the subgroups (proxy interviewed and institutionalized) would have influenced differences in health and mortality risk between the interviewed and the non-interviewed groups in the sample. Third, we analyze how these exclusions would have affected estimations of health prevalences and mortality risks. Finally, we test whether weighting the data by sex and age can reduce the effect of excluding these subgroups, assuming that age and sex is commonly accessible for the total population and often accessible for non-responders in cross-sectional studies.

Methods

The study is based on the 2002 wave of a Swedish nationally representative survey of individuals aged 77+ (SWEOLD). The sample is originally obtained using the Swedish system of unique identification numbers for all Swedish residents. Hence, the sample of 735 people, aged 77–99, is representative of the Swedish older population, irrespective of residence, health and cognitive status, as living situation did not affect the probability of inclusion.

During the fieldwork, 618 individuals of the sample were interviewed (84.1 %) with an acceptable level of item response to the health questions. To avoid non-response due to impaired health or poor cognition we use several interview methods including proxy-, telephone- and face-to-face interviews. When approaching a respondent, our first aim was to perform a direct face-to-face interview. If the respondent was unable to participate or refused we offered to conduct the interview by telephone or with the support of a close relative.

In the SWEOLD study 13.3 percent of the interviews were performed as indirect interviews (Table 1), meaning that all questions were answered by a proxy. In 7.3 % of the interviews, the respondent needed help of a proxy to answer parts of the interview (mixed interview). Since the reason for both indirect and mixed interviews were the respondents inability or resistance to participate these groups were treated as one group in the analyses. Respondents interviewed without help from a proxy are referred to as directly interviewed.

Table 1 Descriptive statistics of the Swedish panel study of living conditions of the oldest old (SWEOLD) 2002

	Men	Women	Total
<i>Originally sampled</i>			
Age (mean)	82.6	83.5	83.1
Education ^a (mean number of years)	8.2	7.7	7.9
Non-response rate (%)	12.2	18.3	15.9
<i>n</i>	287	448	735
%	39.0	61.0	100
<i>Interviewed sample (directly or by proxy)</i>			
Age			
Range	77–98	77–99	77–99
Mean	82.7	83.7	83.3
Education (year)			
Mean	8.3	7.8	8.0
Median	7	7	7
Type of interview (%)			
Direct ^b	83.3	76.8	79.4
Proxy ^c	9.9	15.6	13.3
Mixed ^d	6.7	7.7	7.3
Living situation (%)			
Institution	19.4	24.3	22.3
Living at home	80.6	75.7	77.7
<i>n</i>	252	366	618
%	40.8	59.2	100

^a No information for 16 individuals

^b 9 % of the direct interviews were conducted by telephone

^c 88 % of the proxy interviews were conducted by telephone

^d 2 % of the mixed interviews were conducted by telephone

Most of the proxy interviews were conducted by telephone. Few studies have evaluated the effect of using telephone in proxy interviews. It is reported in a US-study (Segal et al. 1996) that the concordance between the proxy and the respondent was not affected by the use of the telephone and that all contradictory measures were attributed to the proxy interview itself. In SWEOLD, except for the proxy interviews, telephone interviews were used when the interview person refused a visit (9 %). It has been suggested that one contributory cause to conflicting measurements acquired from different interview modes is that the structure or the questions in the questionnaire are changed to suit the interview mode (Dillman and Christian 2005). In SWEOLD the same questionnaire was used by the interviewer, irrespective of the interview mode, thereby minimizing some of the effect of using a mixed mode.

In SWEOLD 2002, the non-response rate (including 2 individuals with high item non-response on the health questions) was 15.9 % (Table 1). The Swedish system of the unique personal identification number and obtainable information on current address and phone number from the register is one explanation for SWEOLD's relatively high response rate. All individuals in the sample were found and a proxy could be located when necessary. In 19 cases, when no relatives were found, healthcare personnel were used as proxies. Thus, all non-response was caused by active refusal to participate on the part of the older person or a close relative. More women than men chose not to participate (18.3 % compared with 12.2 %). Consequently, the non-response rate was more than 6 % points higher for women than for men.

Still, the interview group consisted of approximately 60 % women and 40 % men, which is close to the sex distribution in the Swedish population 77 and older. As compared with the women, the men were generally younger and had a higher education level, which is also in line with the structure of the Swedish oldest population (Statistics Sweden 2001).

Due to the lack of exclusion criteria in SWEOLD, a high percent of the interviewees, 22.3 %, were living in institutions. Institutionalized refers in this study to individuals living in nursing homes, group homes for people with dementia or different types of assisted living facilities. Due

Table 2 Type of interview by living situation

Type of interview (n)	%	Ordinary living	Living in institution	Total
Direct		90.8 (436)	39.9 (55)	79.4 (491)
Proxy		6.0 (29)	38.4 (53)	13.3 (82)
Mixed		3.1 (15)	21.7 (30)	7.3 (45)
Total		100 (480)	100 (138)	100 (618)

to the association between the probability of institutional living and of requiring an indirect interview, these groups of 138 and 127 individuals, respectively, overlapped by 83 individuals. Table 2 provides information on the distribution of proxy interviews by living situation. Telephone interviews were used in 88 % of the proxy interviews. The direct interviews were primarily conducted in person, but in those cases when the older person refused a visit, a telephone interview was considered acceptable, even though none of the tests could be carried out. Of the direct interviews 9 % were conducted by telephone. The corresponding number for mixed interviews were 2 %.

Measures

Age, sex, and education

Age and sex were registered at the time of sample construction. Education was measured using self-reported years of education from earlier survey waves. Education is a measure of socioeconomic status associated with disability (Crimmins and Saito 2001; Moe and Hagen 2011). Education is also positively associated with cognitive functioning later in life (Fors et al. 2009).

Diseases/symptoms

The magnitude of diseases and symptoms were measured using an index constructed from a battery of 22 indicators of common diseases and symptoms in the SWEOLD questionnaire. All items were based on the question: “Have you had any of the following diseases or disorders during the past 12 months?” The index was constructed to capture both the number and degree of diseases and symptoms, using the response alternatives: “no” (0 points), “yes, mild” (1 point), and “yes, severe” (3 points) and then summing the scores for each respondent. In order to disentangle the different types of diseases and symptoms, the index was also divided into 15 mainly physical diseases/symptoms (e.g., diabetes, weight loss, stroke, dizziness, high blood pressure, and chest pain), 2 measures of psychologic problems (anxiety and depression), 3 measures of musculoskeletal pain, and 2 measures of vision and hearing acuity.

Activities of daily living, ADL

Disability was measured using one summarized index of self-reported problems with activities of daily living (ADL), including the ability to eat, use the toilet, dress/undress, get up/go to bed, cut one’s toenails, and wash one’s hair. The index ranged from 0 to 12 points; the response alternatives were: “Yes, completely by myself”

(0 points), “Yes, with help” (1 point), and “No, not at all” (2 points).

Mobility

The index measuring mobility problems included the self-reported ability to, without difficulties, walk 100 m fairly briskly, walk up and down stairs, and get up from a chair (without using the armrests). The index ranged between 0 and 3 points, where positive answers gave 0 points and negative answers gave 1 point.

All indexes were constructed so that a higher value indicates poorer health, greater disability, or poorer mobility. Individuals with missing values on only one or two of the variables in one index were considered to not have any problems in relation to that variable.

Mortality

Date of death was obtained from the Swedish National Cause of Death Registry using the personal identification number. Mortality by date of death was followed from the interview day until April 18, 2005, comprising an average follow-up time of 903 days (about 2.5 years). During the follow-up period, 156 individuals (21.2 %) died. Regarding the non-response group there was no given starting date for the follow-up period. The interviews were conducted during a time period of 6 months (84 % within the first 3 month). The recruitment process went on continuously during this time period and the usual approach was to re-contact the hesitant respondents again after a while. The non-response was thereby randomly distributed during the whole time period. As a starting date for the non-response group, the median value for interview date was chosen, due to the moderately skewed distribution of the interview date variable.

Analyses

Initially, we compared how the respondents interviewed by proxy and the institutionalized group differed from the directly interviewed respondents who lived at home (reference group) concerning age, sex, education, diseases and symptoms, mobility, disability, and mortality risk. Next, by adding different subgroups to the non-response group, we compared how the non-interviewed group of the sample would differ from the interviewed group concerning age, sex, education, and mortality risk, depending on study design. Mortality was analyzed using Cox proportional hazard regression. The assumption of proportionality was fulfilled in all Cox regression models.

Subsequently, we investigated how the estimated prevalence of three of the measures—ADL, mobility, and

Table 3 Proxy interviewed and institutionalized individuals characteristics, health and physical function problems and mortality risk compared to directly interviewed individuals who live at home

	Directly interviewed living at home	Proxy interviewed	Living in institution
Characteristics			
Sex ^a (% women)	57.1	66.9**	64.5
Age ^b (mean)	82.1	86.2***	86.4***
Education ^b (mean number of years)	8.2	7.6**	7.6**
Mean value of health index^b (number of items in the index)			
Diseases/symptoms (all 22)	8.7	10.9***	11.1***
Physical problems (15)	3.9	4.9**	5.1**
Psychologic problems (2)	0.5	1.4***	1.3***
Musculoskeletal pain (3)	2.7	2.6	2.8
Vision and hearing (2)	1.5	2.0**	1.9**
Mobility problems (3)	0.9	2.2***	2.3***
ADL limitations (6)	0.8	5.9***	5.5***
Mortality risk^c			
Number of deaths	47	65	70
Crude HR (95 %CI)	1.00 (ref)	6.57 (4.51–9.58)	6.45 (4.44–9.33)
Adjusted ^d HR (95 %CI)	1.00 (ref)	4.88 (3.22–7.39)	4.34 (2.87–6.56)
<i>n</i>	436	127	138

Note There is an overlap of 83 individuals in the proxy interviewed group and the group that are living in institutions

Level of significance * 10 % ** 5 % *** 1 %

^a χ^2 -test

^b *t* test, independent sample, equal variance not assumed

^c 2.5 years follow-up time

^d Adjusted for sex and age

psychologic problems—would differ depending on whether or not proxy interviewed and institutionalized individuals were included in the study.

Finally, these estimated prevalence measures were adjusted for age and sex distribution. All the samples, containing various subgroups of the originally sampled group, were weighted, to represent the original sample concerning age and sex. By doing this, we were able to examine whether or not variations in age and sex distribution in different subgroups affected the estimates.

Results

Characteristics by interview mode and living situation

Because two-thirds of the persons interviewed by proxy also lived in institutions (Table 2), these two groups showed similar patterns (Table 3) compared to the reference group, that is, the directly interviewed individuals who were living at home. In both groups, women were overrepresented and, compared with the reference group,

approximately 4 years older (86.2 and 86.4 years, compared to 82.1 years). The mean education level of the proxy interviewed and institutionalized individuals were also lower than that of the reference group, 7.6 years compared to 8.2 years. However, analysis using logistic regression revealed that the effect of lower education on the probability of institutionalized living or proxy interview almost disappeared when controlling for age and gender (not shown).

The proxy interviewed persons and the institutionalized group also reported significantly more diseases and symptoms than did the reference group. The mean value of the general diseases/symptoms index was 10.9 and 11.1 points, respectively, for these groups compared to 8.7 points for the reference group. When looking at the index by type of diseases and symptoms, these differences remained with only one exception: There were no significant differences between the interview groups regarding reported musculoskeletal pain.

The measures of mobility problems and ADL limitations showed significant differences between the interview groups. A mean value of 0.9 compared to mean values above 2 indicates that the reference group generally had

Table 4 Characteristics and mortality risk for the non-interviewed group compared to the interview group, depending on use or non-use of proxy interviews and the inclusion or exclusion of institutionalized individuals

	Interview group	Non-response group	Subgroups added to the non-response group		
			Non-response group + proxy interviewed	Non-response group + institutionalized	Non-response group + proxy
Number	618	117	244	255	299
Percent of sample	84.1	15.9	33.2	34.7	40.7
Characteristics					
Women %	59.2	70.1	68.4	67.1	66.6
Age (mean)	83.3	82.1	84.3	84.5	84.6
Education ^a (mean)	8.0	7.4	7.5	7.5	7.5
Mortality risk ^b					
Number of deaths	164	18	84	89	109
Crude HR (95 % CI)	1.0	0.67 ^c (0.41–1.09)	2.75 ^d (2.01–3.77)	2.90 ^e (2.11–3.98)	4.05 ^f (2.88–5.71)
Adjusted HR ^g (95 % CI)	1.0	0.86 (0.52–1.41)	2.33 (1.69–3.22)	2.30 (1.66–3.20)	3.15 (2.20–4.50)

^a No information for 16 individuals^b 2.5 years follow-up time^c Reference group is all participating respondents^d Reference group is all directly interviewed respondents^e Reference group is all respondents living at home^f Reference group is all directly interviewed respondents living at home^g Adjusted for sex and age

fewer than one mobility problem, while the proxy interviewed and institutionalized individuals generally had more than 2 mobility limitations. The same pattern is shown for ADL problems. The reference group generally had fewer than one ADL limitation, while the proxy interviewed and institutionalized people suffered from limitations in several of the ADL measures in the index.

Finally, Table 3 shows that proxy interviewed and institutionalized individuals had a significantly increased mortality risk compared to the reference group. Their risk of death—adjusted for gender, age, and education—was more than 4 times higher during the 2.5 years follow-up time.

Differences between interviewed and non-interviewed

Obtainable information on the actual non-response group in SWEOLD is given in the second column of Table 4. More women than men refused to participate; thus 70.1 % of the non-response group was women, compared to 59.2 % in the interview group, shown in the first column. This can be compared with the 61.0 % in the original

sample (Table 1), which is the expected percent of women in the interview and the non-response groups, if women and men had participated to the same extent in the survey. The non-response group was also generally slightly younger and less educated than the interview group. When comparing the mortality risk between the non-response group and the interview group, no significant differences were found.

The remaining columns in Table 4 show how the non-interviewed group would differ from the interviewed group depending on the study design. By excluding either the proxy group or the institutionalized group, and adding them to the non-response group, the percent of the sample not interviewed would increase from 15.9 to 33.2 and 34.7 %, respectively. These created non-interviewed groups which would still differ from the sample by having more women, and a slightly lower mean age and level of education, although these differences would decrease. However, the main change would be the increased difference in mortality risk between the non-interviewed group and the interviewed group. The relative risk of death during the 2.5 years follow-up time would be more than double for

Table 5 Estimated prevalences for individuals with ADL limitations, mobility problems, and psychologic problems, depending on use or non-use of proxy interviews and the inclusion or exclusion of institutionalized individuals (%)

	Proxy interviewed	Living in institution	Directly interviewed living at home <i>n</i> = 436 = 59 %	+ Proxy interviewed living at home <i>n</i> = 480 = 65 %	+ All institutionalized individuals <i>n</i> = 618 = 84 %
Indicator					
ADL limitations (3+)	59.8	52.9	2.8 (2.7)	5.6 (6.0)	16.2 (16.0)
Mobility problems (2+)	73.2	75.4	29.4 (30.3)	31.9 (33.3)	41.6 (41.5)
Psychologic problems (1+)	54.3	52.2	27.8 (28.4)	29.8 (30.3)	34.8 (35.1)

Prevalences weighted to correspond to sex and age distribution in the original sample are presented in *brackets*

the non-interviewed group compared to the interviewed group, if the institutionalized or the proxy interviewed individuals were added to the non-response group, even after adjusting for age and gender.

If we were to exclude both proxy interviewed and institutionalized people, 40.7 % of the original sample would not have been interviewed. This hypothetical non-interviewed group (now including the actual non-response group, the institutionalized and the proxy interviewed) would have a mortality risk more than three times as high as the individuals left in the interview group. This indicates a strong bias in all results correlated with mortality.

Estimated prevalence rates depending on design and inclusion criteria

Table 5 shows the estimated prevalence of ADL limitation, mobility problems and psychologic problems depending on study design. The first two columns give the prevalences for the proxy interviewed and the institutionalized group. The third column gives the prevalences estimated from the most restricted sample, the directly interviewed individuals who were living at home. If the researchers involved in the SWEOLD study had excluded institutionalized individuals and failed to conduct proxy interviews, the prevalence of ADL limitations, based on these 436 individuals, would have been estimated at 2.8 %. The second column shows how the decision to conduct proxy interviews would increase this prevalence to 5.6 %, now based on 480 individual from the original sample of 735. Finally, by adding the institutionalized population in the last column, the estimated prevalence of ADL limitations increased to 16.2 %.

The picture is the same for the prevalence of mobility and psychologic problems. Inclusion of the proxy interviewed and institutionalized individuals would increase the prevalence of mobility problems from 29.4 to 41.6 %. For the measure of psychologic problems, the estimated prevalence would increase from 27.8 to 34.8 %.

Weighting

Finally, we weighted all the restricted samples, to represent the originally sampled group as regards age and sex. This gave the estimated prevalence that each sample would provide if they contained the same proportion of men and women, for every one-year-age group, as the originally sampled group. However, weighting the data resulted in only a negligible change in the estimated prevalences, presented in brackets in Table 4.

Magnitude of underestimation

One solution when dealing with this type of incomplete data would be to estimate the magnitude of the underestimation. Our results suggest that the size of the error is dependent on the type of variable and the specific characteristics of the excluded subgroups. Given that the best possible estimates are achieved when all 618 individuals are included, a study excluding institutionalized individuals and proxies would require the prevalence of ADL limitations to be multiplied almost six times to reach the value of 16.2 %. The corresponding factors for mobility and psychologic problems are, however, only 1.4 and 1.2, respectively.

Discussion

In the present study, we have shown that the exclusion of the institutionalized population and the non-use of proxy interviews, in representative studies of the oldest old, lead to reduced prevalence rates for all studied health outcomes. In line with previous research, the results showed that old people living in an institution as well as the proxy interviewed were significantly older, less educated and more likely to be female, compared to the directly interviewed people who were living at home. They also had significantly more diseases and symptoms, mobility problems,

and ADL limitations, and their mortality risk was 4 times higher than the other group.

If the researchers involved in the SWEOLD study had either excluded institutionalized individuals or chosen not to use proxy interviews, the non-interviewed part of the sample would have been more than doubled. Besides the “regular” non-respondents, almost 20 % of the oldest population would have been excluded, including more disabled individuals with poor health status. The proportion of old people in the population with problems regarding ADL, physical mobility, and mental health would have been underestimated.

Our conclusion is that a high non-response rate and/or a low level of proxy interviews in a survey of the oldest old most probably indicate a non-response bias in the sampled group. This is in addition to the sample-bias many surveys have due to the exclusion of institutionalized individuals. The magnitude of the discrepancies that resulted from the non-representative study population leads us to conclude that the issue of representativeness should be addressed before other methodological issues when estimating prevalence of old age disability.

A non-representative study population leads to uncertainty about the real health status of the population. Problems are compounded in comparative studies. Do prevalence differences reflect health differences or different response rates? In cross-country comparisons, country-specific differences in rates of institutionalization can lead to prevalence differences in health when institutionalized persons are excluded.

Studies of health trends over time (Lafortune and Balestat 2007) face similar challenges. It is difficult to maintain consistency in survey design over long periods of time. Survey waves may differ over time in response rates, use of proxies or inclusion criteria. Changes in institutionalization, due to changes in policy or health needs, will affect prevalence rates if this group is excluded.

The probability for when an older person is institutionalized varies across countries, communities, and over time. In Sweden, for example, during recent years the group of institutionalized older people has become smaller, but more dependent and fragile. This is an effect of raising the threshold for admission to institutions (NBHW 2009). The exclusion of institutionalized individuals from a population-based survey will thus not only affect the result by not including the most ailing and frail group, but there is also a risk that the health and function trends over time will be affected, as well as the comparability of data across countries or regions. Inconsistencies in prevalence across countries or a significant trend toward, e.g., changed physical ability among the oldest old, could merely be the effects of a change in social policies regarding the availability of institutional beds.

As with the decision to include the institutionalized, the decision to use proxy interviews in a survey is made by the researcher. However, which respondents are interviewed indirectly using a proxy is dependent on additional factors and the recruiting process. It cannot be assumed that all proxy interviewed persons would have been non-respondents if indirect interviews were not offered. Nevertheless, in SWEOLD, proxy interviews were only conducted when it was considered impossible to conduct the interview without a proxy. A direct face-to-face interview was always the preferred option.

Furthermore, there is no clear evidence whether a proxy gives the same answer as the respondent would have given. Some of the differences between the directly and indirectly interviewed group in SWEOLD are most certainly explained by response patterns found in previous research—that proxies are more prone to report health problems than the older persons themselves. However, the research varies depending on the type of health variable considered, and has demonstrated the validity of proxy interviews on easily observed variables such as mobility and physical function. This supports our results regarding the measures of ADL and mobility; it is unlikely that proxy bias could explain the differences between the studied subgroups. In contrast, the measures of psychologic health should be interpreted with caution. The literature is more mixed concerning subjective measures that are difficult to observe, such as pain, self-rated health, and psychologic problems (Magaziner et al. 1996; Neumann et al. 2000; Smith and Goldman 2011).

Few studies can claim to be truly representative of an older population. The possibility to obtain a representative sample of the population and maintain a high non-biased response rate varies across countries, studies and over time. One way to deal with a non-representative study population is to weight the data to correspond to the population it is supposed to represent. In longitudinal surveys, an array of possible variables may be available from previous waves that could be used in a weighting procedure. However, in cross sectional surveys, researchers most often are restricted to the variables used when creating the sample, such as sex and age.

According to our results, it is unlikely that researchers who exclude institutionalized individuals, or do not use proxy interviews, would be able to compensate for this by giving less represented groups of women and men in different ages more weight in surveys of the oldest old.

Our data also suggests that the impact on the results, due to the exclusion of different subgroups of the oldest population, depends on the type of health outcome being studied. For instance, in SWEOLD, the relative differences become greater for ADL limitations than for mobility problems. On the other hand, mobility problems were more prevalent among all subgroups and showed a higher

absolute increase when adding proxy interviewed and institutionalized persons.

On several measures, SWEOLD has painted a more negative picture of health and function among the oldest old than many other studies (Parker and Thorslund 2007). In SWEOLD, one major goal has been used to obtain a representative picture of the living situation, including health and function, of the entire oldest Swedish population (Lundberg and Thorslund 1996). All possible efforts have been made to recruit even ailing and frail individuals. There is no clear answer as to how the mixed interview methods used in SWEOLD affect the results. What we do know is that by excluding institutionalized individuals and those too frail to participate in a survey we will underestimate the health problems in the oldest population.

Attention to survey methodology and country-specific knowledge about excluded subgroups—the non-response, the use of proxy interviews and the representation of institutionalized individuals—are necessary when interpreting results and drawing conclusions about health and disability among oldest populations. Accurate estimates regarding prevalence rates as well as trend analyses require non-biased study samples, where the oldest persons are represented regardless of their living conditions, health status, and cognitive ability.

Acknowledgments This work was supported by the Swedish Council for Working Life and Social Research, grants 2006-1622 and 2010-1684. The data collection for the 2002 SWEOLD was funded by the Swedish Research Council, grant 2001-6651.

References

- Asakawa K, Feeny D, Senthilselvan A, Johnson JA, Rolfson D (2009) Do the determinants of health differ between people living in the community and in institutions? *Soc Sci Med* 69:345–353
- Chatfield MD, Brayne CE, Matthews FE (2005) A systematic literature review of attrition between waves in longitudinal studies in the oldest shows a consistent pattern of dropout between differing studies. *J Clin Epidemiol* 58:13–19
- Crimmins E, Saito Y (2001) Trends in healthy life expectancy in the United States, 1970–1990: gender, racial, and educational differences. *Soc Sci Med* 52:1629–1641
- Dillman DA, Christian LM (2005) Survey mode as a source of instability in responses across surveys. *Field Methods* 17(1):30–52
- Fors S, Lennartsson C, Lundberg O (2009) Childhood living conditions, socioeconomic position in adulthood, and cognition in later life: exploring the associations. *J Gerontol: Soc Sci* 64B:S750–S757
- Freedman V, Aykan H, Martin L (2002) Another look at aggregate changes in severe cognitive impairment: further investigation into the cumulative effects of three survey design issues. *J Gerontol: Soc Sci* 57B:S126–S131
- Freedman V, Crimmins E, Schoeni R, Spillman B, Aykan H, Kramarow E, Land K, Lubitz J, Manton K, Martin L et al (2004) Resolving inconsistencies in trends in old-age disability: report from a technical working group. *Demography* 41:417–441
- Gudex C, Lafortune G (2000) An inventory of health and disability-related surveys in OECD countries. OECD Labour Market and Social Policy Occasional Papers, No. 44, OECD Publisher
- Hardy S, Allore H, Studenski S (2009) Missing data: a special challenge in aging research. *J Am Geriatr Soc* 57:722–729
- Huber M, Hennessy P (2005) Long-term care for older people. OECD Publishing
- Lafortune G, Balestat G (2007) Trends in severe disability among oldest people: assessing the evidence in 12 OECD countries and the future implications. OECD Health Working Papers, No. 26, OECD Publishing
- Larsson K, Thorslund M, Kåreholt I (2006) Are public care and services for older people targeted according to need? Applying the behavioural model on longitudinal data of a Swedish urban older population. *Eur J Ageing* 3:22–33
- Lundberg O, Thorslund M (1996) Fieldwork and measurement considerations in surveys of the oldest old. *Soc Indic Res* 37:165–187
- Magaziner J, Bassett S, Hebel J, Gruber-Baldini A (1996) Use of proxies to measure health and functional status in epidemiologic studies of community-dwelling women aged 65 years and older. *Am J Epidemiol* 143:283–292
- Manton K, Suzman R (1992) Conceptual issues in the design and analysis of longitudinal surveys of the health and functioning of the oldest old. In: Suzman R, Willis D, Manton K (eds) *The Oldest Old*. Oxford University Press, New York, pp 89–122
- Martikainen P, Nihtilä E, Moustgaard H (2008) The effects of socioeconomic status and health on transitions in living arrangements and mortality: a longitudinal analysis of oldest Finnish men and women from 1997 to 2002. *J Gerontol: Soc Sci* 63B:S99–S109
- Martin LG, Schoeni RF, Andreski PM (2010) Challenges in estimating trends in late-life disability from the American community survey. *J Gerontol: Med Sci* 65A:M517–M518
- Moe JO, Hagen TP (2011) Trends and variation in mild disability and functional limitations among older adults in Norway, 1986–2008. *Eur J Ageing* 8:49–61
- NBHW (2009) *Folkhälsorapport 2009* (Public health report 2009). National Board of Health and Welfare, Västerås
- Neumann P, Araki S, Guterman E (2000) The use of proxy respondents in studies of older adults: lessons, challenges and opportunities. *J Am Geriatr Soc* 48:1646–1654
- Noël-Miller C (2010) Spousal loss, children, and the risk of nursing home admission. *J Gerontol: Soc Sci* 65B:S370–S380
- Parker MG, Thorslund M (2007) Health trends in the oldest population: getting better and getting worse. *Gerontologist* 47:150–158
- Rodgers W, Herzog A (1992) Collecting data about the oldest old: problems and procedures. In: Suzman R, Willis D, Manton K (eds) *The oldest old*. Oxford University Press, New York, pp 135–156
- Segal ME, Gillard M, Schall RR (1996) Telephone and in-person proxy agreement between stroke patients and caregivers for the functional independence measure. *Am J Phys Med Rehabil* 75(3):208–212
- Smith KV, Goldman N (2011) Measuring health status: self-, interviewer, and physician reports of overall health. *J Aging Health* 23:242–266
- Statistics Sweden (2001) *Statistical yearbook of Sweden 2002*. Statistics Sweden, Örebro
- Wallace R, Kohout F, Colsher P (1992) Observations on interview surveys of the oldest old. In: Suzman R, Willis D, Manton K (eds) *The oldest old*. Oxford University Press, New York, pp 123–134
- Wolf D, Hunt K, Knickman J (2005) Perspectives on the recent decline in disability at older ages. *Milbank Q* 83:365–395