Item non-response when measuring elder abuse: influence of methodological choices

Liesbeth De Donder[1], Gert Lang[2], Bridget Penhale[3], José Ferreira-Alves[4], Ilona Tamutiene[5], Dominique Verté[1], Minna-Liisa Luoma[6]

1 Department of Educational Sciences, Faculty of Psychology and Educational Sciences, Vrije Universiteit Brussel, Brussels, Belgium
2 Research Institute of the Red Cross, Vienna, Austria
3 School of Nursing, Faculty of Medicine and Health Sciences, University of East Anglia, Norwich, UK
4 Human Cognition Lab, School of Psychology, University of Minho, Braga, Portugal
5 Department of Sociology, Vytautas Magnus University, Kaunas, Lithuania
6 Department of Living Conditions, Health and Wellbeing, National Institute for Health and Welfare (THL), Helsinki, Finland

Correspondence: Liesbeth De Donder, Vrije Universiteit Brussel, Faculty of Psychology and Educational Sciences, Pleinlaan 2 1050 Brussels, Belgium.

Background: Efficient actions to fight elder abuse are highly dependent on reliable dimensions of the phenomenon. Accurate measures are nevertheless difficult to achieve owing to the sensitivity of the topic. Different research endeavours indicate varying prevalence rates, which are explained by different research designs and definitions used, but little is known about measurement errors such as item non-responses and how outcomes are affected by modes of administration. **Methods:** A multi-national study was developed to measure domestic abuse against home-dwelling older women (aged >60 years) in Europe. The measurement instrument covered six forms of abuse, adapted from the Conflict Tactics Scale. 2880 individuals were interviewed by three different data collection methods (i.e. postal, face-to-face, telephone). **Results:** Principal component analysis of missing values of 34 indicators of abuse showed various patterns of item non-response. Moreover, principal component analysis indicated several response patterns across different types of data collection. A binary logistic regression explained that item non-response and abuse prevalence 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. **Conclusion:** The discussion revolves around how these findings could help improving measuring elder abuse. Advantages and disadvantages of the questionnaire and type of data collection methods are discussed in relation to three potential types of response errors: item positioning effect, acquiescence and social desirability.

Introduction

Abuse of older people is a significant global problem.[1] It is a violation of Article 25 of the EU Charter of Fundamental Rights[2] resulting in suffering, decreased long-term health and emotional well-being.[3] Accurate estimates of elder abuse are fundamental for developing effective prevention strategies and interventions by health service providers and policy makers,[4] and for evaluating the public health impact of preventive actions.[5] To obtain correct estimates of the prevalence of elder abuse, the measurement instrument needs to be reliable and valid. However, a European review on prevalence studies[6] demonstrates that publications mostly focus on results, whereas exceptionally few studies pay attention to methodological evaluation.[7,8] A widely applied definition on elder abuse originates from the UK's Action on Elder Abuse,[9,11] subsequently adopted by the International Network for Prevention of Elder Abuse and the WHO: Elder abuse is '*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.* Six forms of elder abuse can be differentiated: physical, psychological, sexual, financial abuse, neglect and violation of personal rights (Ibid.).

In terms of methodological evaluation, two important research gaps can be detected: First, more multi-national research is needed.[11] However, little information is available on study designs for prevalence studies on elder abuse, and the impact of different data collection methods on outcomes. The European Prevalence Study on Abuse and Violence against Older Women (AVOW) (For further information about the AVOW study: www.thl.fi/avow) collected data in Austria, Belgium, Finland, Lithuania and Portugal can be used to provide detailed insights in this matter. Second, literature names several kinds of response errors,[12] which can influence outcomes such as item effects (e.g. item non-response), item positioning effect (e.g. sections and order of forms of abuse), acquiescence effect (i.e. tendency to agree or to answer without a relation to the question), interviewer and attendance effects and social desirability (i.e. reluctance of people to disclose socially undesirable attitudes or behaviour or those perceived as such).[13,14]

To improve research development, further continued study on non-response patterns is clearly needed, its associations with individual characteristics and differences according to data collection methods.[7,15] Moreover, not only information about non-responders but also relating to item non-response requires debate from a methodological perspective.[16]

Given the above, the purpose of this article is threefold:

(1) Provide methodological background information on the research process of the AVOW study (e.g. study design, modes of administration, sampling procedures).
(2) Examine item non-response patterns and investigate associations with different forms of abuse (e.g. neglect, physical abuse), different modes of administration (e.g. type of data collection or receiving help from an interviewer) and individual characteristics

(e.g. socio-demographic, health and socio-economic features) as controls.

(3) Examine the influence of different modes of administration on the abuse prevalence rates obtained.

Methods

This article reports on the prevalence study of AVOW that took place between 2009 and 2011 and was funded by the EU's Daphne III programme concerning violence against women and children.

Measurement instrument

Based on an extensive literature review, a measurement instrument was developed with the aim to obtain prevalence figures of abuse against older women and data concerning its patterns.[6] Elder abuse was operationalized along six forms of abuse. Each form was measured by multiple items indicating a concrete abusive behaviour or incident.

These indicators of abuse were adapted from the Conflicts Tactics Scale.[8,17] In the measurement instrument, neglect and emotional abuse were operationalized by nine indicators. Financial, physical, sexual abuse and violation of personal rights comprised four indicators each. To avoid serious re-cognition and recall problems, we asked for abusive incidents committed by someone close (i.e. partner/spouse, daughter/son (in law), other relative, neighbour, friend or paid carer) to the respondent in the past 12 months. For each indicator a 4-point answer scale was offered to the respondent representing how often (the frequency) the incident happened (1 = never, 2 = 1–6 times, 3 = once a month, 4 = weekly). For items representing neglect, an answer category was added (0 = no, did not need help). The different items were summed and dichotomized, creating a new variable 'have been abused', indicating they had experienced at least one kind of abuse in the previous 12 months by someone who was close to them.

To identify individual characteristics, several indicators were used: Age was measured by asking the year of birth. The marital status was recoded in two categories (0 = married or living a civil partnership, 1 = not married nor living in a civil partnership). In terms of socio-economic situation, participants were asked how many years of formal education they had completed (in categories), and how the household managed with their available income (ranging from very badly to very easily). Next, to assess health status, we asked the respondents to rate their subjective health on a 5-point scale (ranging from very good to very poor). Respondents were asked about potential depressive feelings using two items—(i) In the past 2 weeks, have you felt sad, low or depressed most of the time? (ii) In the past 2 weeks have you been less interested in most things, or less able to enjoy the things you used to enjoy most of the time?)—which were combined creating an index. In addition, one question indicated whether respondents needed help filling in the questionnaire and by whom.

Data collection methods

Owing to the multi-national research design, different data collection procedures were applied, which were considered most suitable for the national context. In Lithuania, for example, it was not possible to collect information from a postal survey owing to a lack of widespread postal survey tradition. Therefore, data were collected using face-to-face interviews. In three countries, a postal survey (Belgium, Finland and Portugal) was undertaken, in two countries a face-to-face survey (Belgium, Lithuania) was used and in one country a telephone survey (Austria) was realized. In Belgium, the survey method was split: a postal survey for older women aged >60 years and a face-to-face survey of those aged >75 years were used [More information on this issue can be found in the national reports on the AVOW webpage (www.thl.fi/avow)].

Sampling and participants

The target population of the study was defined as older home-dwelling women aged ≥60 years. Despite different data collection methods being used across different countries, all were based on a (simple or multi-stage) random sampling procedure. In Austria, Belgium and Portugal, random probability or stratified random sampling methods—either by community or age groups—from different registers (telephone or post office registers) were used. In Finland, a simple random sampling was put into practice on the basis of the population register. In Lithuania, a multi-stage random sampling was applied.

Data were collected between April and July 2010. In total, $n = 2880$ individuals responded in the survey comprising of $n = 678$ respondents in Finland, $n = 649$ in Portugal, $n = 593$ in Austria, $n = 515$ in Lithuania and $n = 426$ in Belgium ($n = 318$ postal, $n = 108$ face-to-face). Response rates ranged from 49.1 in Austria (telephone survey), 41.7 in Lithuania (face-to-face), 40.9 in Portugal (postal), 39.9 in Finland (postal) to 26.1% in Belgium (21.2% postal, 43.2% face-to-face).

Table 1 gives an overview of the main characteristics of the sample, for each country and for the total sample. In the total sample, the age of the female respondents varied between 60 and 97 years (mean = 71.4, median = 70.0, SD = 8.2). About half of the older women (50.5%) were married, lived in a civil partnership or were co-habiting with their partner, and around one-third (31.8%) were widowed. 49.7% of participants lived with someone else in her household, and about two-fifths (38.2%) lived alone. One-third (32.0%) of respondents had completed between 5 and 9 years of schooling, and 13.4% has completed <5 years of education.

Analytic strategy

To analyse responsiveness across indicators of elder abuse, missing values were analysed in several steps: First, non-response indicators were calculated (0 = item completed, 1 = missing value), and item non-response was detailed on the basis of a simple frequency analysis. Second, principal component analyses (PCA with varimax rotation) were run to find item non-response patterns both in general and across the three different data collection methods. Third, binary logistic regression analyses were applied explaining overall non-response (non-response on all indicators of abuse; if someone answered one item, she was recoded as 'responded') and prevalence rates by aspects of data administration (i.e. data collection method, help needed in answering) and controlling for individual background variables (e.g. age, income, health). In a first analysis also, 'country' was incorporated as control variable in the logistic regression analysis. However, there were substantial correlations among country and data collection method, which could produce multi-collinearity and distort the model fit. Checks of variance inflation factor (VIF) values showed values >2.0. This indicated multi-collinearity problems and, consequently, country was excluded from further analysis. Statistical analysis was performed with SPSS in which statistical significance was set at $P < 0.05$.

Results

Item non-response of the 34 indicators of abuse varied between 1.8% (i.e. medication neglect) and 4.2% (i.e. excluding, ignoring and destroying possessions) (see table 2). In relation to the different forms of abuse, the lowest non-response rates were found in indicators of neglect (between 1.8 and 2.4%). The highest non-response was found in indicators of emotional abuse (between 3.5 and 4.2%).

PCA on missing values identified three different non-response patterns: The first pattern encompassed neglect; the second pattern covered emotional abuse; the third pattern financial, physical, sexual abuse and violation of personal rights. Moreover, response patterns

Table 1 Respondents and data collection characteristics per country (*n* = 2880)

	Austria	Belgium	Finland	Lithuania	Portugal	Total
Sample size *n*	593	436	687	515	649	2880
Age						
Mean	71.0	74.2	72.4	72.1	68.5	71.4
Median	70	76	71	72	66	70
Standard deviation	8.1	8.3	8.8	7.1	7.3	8.2
Marital status						
Married or cohabiting	49.4%	53.9%	50.2%	40.4%	57.7%	50.5%
Single	8.3%	6.9%	7.6%	4.1%	6.1%	6.7%
Divorced	13.5%	5.3%	13.8%	8.7%	11.4%	11.0%
Widowed	28.8%	33.8%	28.5%	46.8%	24.8%	31.8%
Household composition						
Living alone	40.6%	39.6%	44.6%	40.2%	25.4%	38.2%
Living with 1 person	48.2%	54.7%	53.5%	42.3%	49.8%	49.7%
Living with more persons	11.1%	5.7%	1.9%	17.5%	24.8%	12.1%
Education						
<5 years	0.8%	1.6%	5.7%	26%	30.8%	13.4%
5–9 years	43.5%	19.9%	43.3%	21.2%	26.2%	32.0%
>10 years	55.6%	78.5%	51.0%	52.8%	43.0%	54.7%
Mode of administration	Telephone	Postal + face-to-face	Postal	Face-to-face	Postal	
Receiving help by						
None	0.0%	72.5%	93.0%	24.1%	87.7%	57.2%
Interviewer	100%	3.7%	0.0%	74.4%	0.0%	34.3%
Someone else	0.0%	23.9%	7.0%	1.6%	12.3%	8.3%
Prevalence of abuse (% non-response)						
Total abuse	26.5% (1.9%)	33.0% (8.3%)	26.6% (10.9%)	22.8% (0.2%)	41.6% (8.2%)	30.1% (6.1%)
Neglect	6.4% (0.2%)	5.5% (4.6%)	2.5% (1.7%)	2.5% (0%)	9.9% (0.5%)	5.5% (1.3%)
Emotional	20.3% (1.3%)	28.9% (3.2%)	23.4% (5.4%)	18.6% (0%)	34.9% (2.8%)	25.3% (2.7%)
Financial	5.3% (1.7%)	5.9% (3.4%)	6.8% (4.4%)	9.7% (0%)	17.0% (4.2%)	9.2% (2.8%)
Physical	0.5% (1.7%)	2.2% (4.1%)	2.9% (5.1%)	4.5% (0%)	3.1% (4.8%)	2.6% (3.3%)
Sexual abuse	2.2% (1.9%)	2.4% (3.9%)	5.0% (4.4%)	2.3% (0.2%)	3.9% (4.6%)	3.3% (3.1%)
Violation of personal rights	3.8% (1.7%)	4.5% (3.9%)	5.7% (5.2%)	4.9% (0.2%)	12.8% (4.9%)	6.5% (3.3%)

varied across the three different kinds of data collection. Within the postal survey, five distinct patterns arose (with factor loadings >0.45 on the respective factor): neglect, emotional abuse, financial abuse, violation of personal rights and physical together with sexual abuse combined to build an individual response pattern. Within face-to-face interviews, only two patterns emerged: neglect items vs. all other abuse questions. Third, within telephone interviews, three patterns appeared: violation of personal rights, psychological, financial, physical and sexual abuse (first component), neglect (second component) and the indicators of psychological abuse also loaded high on the third component.

Table 3 presents the result of binary logistic regression analysis estimating the probability of overall non-response when taking individual characteristics and modes of administration into account.

In terms of individual characteristics, the results indicate that higher age was associated with greater odds of having missing data on all abuse indicators. In other words, the older the people were, the less often they filled in questions on elder abuse. In addition, older people living in situations of better incomes answered questions on elder abuse more often (0.87, $P = 0.031$), and those with higher levels of education showed a tendency to do so (0.90, $P = 0.056$). Although health was not significantly related (1.01, $P = 0.955$), a one-unit change in the index of poor mental health increased the odds of having missing data on all abuse items by a factor of 1.34 ($P = 0.001$), signifying that older people with lower/ poorer mental health were more likely not to answer than older people with better mental health. Marital status was not significantly related with non-response to the indicators.

Differences in the responses were also found in relation to modes of administration: First, in relation to postal surveys, face-to-face interviews produced significantly less missing values (0.33, $P = 0.002$), whereas there were no significant differences within telephone interviews (2.45, $P = 0.073$). Second, the presence of another person mattered (i.e. from whom the respondent received

helped during the data collection) because there were significantly more common valid responses if there was an interviewer (0.23, $P = 0.002$) or someone else (0.46, $P = 0.002$) present, e.g. the partner/ spouse, daughter/son, other relative, neighbour, paid care worker.

The overall prevalence rate of abuse was 30.1%, indicating that 3 of 10 older women participants experienced abuse at least once in the previous year (see table 1). As shown in table 3, older women who participated in the AVOW study by a face-to-face survey had significantly lower odds of experiencing abuse than in the postal survey, whereas using a telephone interview did not generate differences. In terms of those respondents who had been assisted when completing the questionnaire, the findings revealed a significant impact of receiving help from someone other than a professional interviewer, but not for help from a professional interviewer.

Discussion

This article reports on the study design and provides methodological information of the AVOW study as a rare example of a prevalence study of elder abuse with/using a multi-national study design (following the ABUEL survey[7]). More specifically, this article focuses on the effects of modes of administration on outcomes (i.e. response rates and prevalence rates). Looking at the cooperation rates, it becomes apparent that mode of administration played a significant role in the co-variation of missing data. Unit-response rate (i.e. participating in the survey) is the highest for telephone interviews, then face-to-face and lowest among postal surveys. However, the initial/preliminary literature review undertaken not only stressed the relevance of investigating the profiles of non-responders[7] but also the importance of examining the missing values of responders (item non-response) and those patterns and profiles.

In responding to the research question on the item non-response patterns, the findings illustrate that in general, older women are

Table 2 Item non-response frequencies and missing data patterns (n = 2880)

Indicators of abuse by forms	% Non-response	PCA general			PCA postal					PCA face-to-face		PCA telephone		
		1	2	3	1	2	3	4	5	1	2	1	2	3
Neglect														
Shopping	1.9			0.85	0.83						0.96	0.64		
Preparing meals or eating	1.9			0.89	0.88						0.97	0.75		
Doing routine housework	1.9			0.86	0.86						0.95	0.62		
Travel or transport	2.4			0.83	0.89						0.96	0.33		
Getting in and out of bed	1.9			0.93	0.93						0.92	0.96		
Washing or bathing	1.8			0.92	0.91						0.97	0.96		
Dressing or undressing	1.9			0.91	0.90						0.97	0.96		
Using the toilet	2.0			0.92	0.92						0.97	0.75		
Taking care medication	1.8			0.93	0.92						0.97	0.96		
Psychological abuse														
Threats on the phone, by mail, email or text message	3.5		0.82		0.86					0.92		0.67		0.57
Threats face to face	3.5		0.82		0.85					0.92		0.71		0.58
Insulting	3.9		0.79		0.84					0.83		0.48		0.64
Undermining or belittling	3.7		0.85		0.89					0.93		0.52		0.73
Excluding or repeatedly ignoring	4.2		0.83		0.86					0.97		0.47		0.76
Preventing from seeing others that one cares about	4.0		0.84		0.88					0.87		0.62		0.54
Shouting or yelling	3.8		0.85		0.88					0.88		0.66		0.55
Doing something to spite	3.9		0.82		0.87					0.88		0.64		0.53
Destroying possessions	4.2		0.80		0.84					0.92		0.80		
Financial abuse														
Taking financial advantage	3.3	0.71							0.75	0.97		0.85		
Blackmailing	3.4	0.77							0.78	0.97		0.95		
Prohibiting to make financial decisions	3.4	0.76							0.75	0.97		0.92		
Steal money or possessions	3.6	0.76							0.73	0.92		0.84		
Physical abuse														
Restraining	3.4	0.84						0.86		0.97		0.92		
Hitting	3.5	0.83						0.86		0.97		0.89		
Thrown a hard object	3.5	0.85						0.87		0.97		0.89		
Giving too much medicine to control	3.5	0.84						0.86		0.97		0.91		
Sexual abuse														
Talking in a sexual, uncomfortable way	3.3	0.85					0.67			0.95		0.85		
Watching pornography	3.3	0.87					0.68			0.95		0.91		
Touched in a sexual way	3.3	0.86					0.68			0.90		0.91		
Sexual intercourse /relations	3.5	0.85					0.67			0.95		0.88		
Violation of personal rights														
Hindering in personal decisions	3.6	0.83				0.81				0.95		0.85		
Hindering from reading mail	3.7	0.84				0.81				0.91		0.87		
Hindering to take part in activities	3.6	0.85				0.80				0.90		0.91		
Hindering from meeting friends	3.6	0.84				0.80				0.90		0.91		

PCA with Kaiser Normalization (Eigenvalues > 1.0) and varimax rotation technique. Factor loadings < 0.45 were omitted from the table.

willing to respond sensitive questions about elder abuse. However, item response rates differed by sensitivity with neglect showing the lowest item non-response. Next, item response rate differs by mode of administration. Postal surveys demonstrate the highest item non-response, as also found where older people did not receive help when filling in the questionnaire. Furthermore, in line with other work,[18–20] older people who are most at risk of being abused have higher age, lower socio-economic status and more depressive feelings. Moreover, these 'at-risk' older people are also more inclined to leave questions about elder abuse unanswered. This suggests that individuals with lower socio-economic status, as well as mentally vulnerable persons (e.g. with depressive moods) must be approached in different ways or with different measures to make the probability of missing answers as low as possible. There are clear implications for the way that research is conducted about sensitive issues such as abuse, with a need to promote maximum participation and responsiveness, with assistance provided to individuals through professional interviewers and attention paid to the ordering of questions and a range of questions about a topic funnelling down to specific questions about abuse. However, there are also implications for health and social care professionals who wish to approach older people who may be at risk of abuse concerning, e.g. the ordering of questions during interview and trying to reduce the extent of effects such as social desirability or acquiescence. Further work could usefully be undertaken in this area to draw out any additional recommendations concerning asking questions about such sensitive topics in a range of different contexts and settings.

These initial findings can refer to three potential types of response errors: First, item positioning effect (e.g. order of forms of abuse presented) may influence the item non-response rate. For example, the fact that the questions of neglect were placed in the beginning of the questionnaire could have influenced the lower item non-response rate. Second, acquiescence is the tendency of the respondent to agree or to answer the item without a specific relation to the question (i.e. choosing the category 'never have been abused'). This can be an individual strategy to reduce unwanted consequences or to be accepted, as found for example in research with some underprivileged persons.[21,22] Third, social desirability refers to the reluctance of people to disclose socially undesirable attitudes or behaviour, or perceptions about these aspects. Such answers might be given because of the perceived expectations owing to social (e.g. gender roles) or personal characteristics (e.g. low self-esteem). Social desirability is a particular

Table 3 Logistic regression coefficients of individual characteristics and modes of administration on overall non-response and prevalence rate

	Overall non-response		Have been abused	
	Exp(B)	P-value	Exp(B)	P-value
Age (in years)	1.04	0.000	0.97	0.000
Marital status (not married)	0.98	0.863	0.75	0.003
Income (ease of managing budget)	0.87	0.031	0.84	0.000
Education (completed years of education)	0.90	0.056	1.09	0.039
Health status (poor)	1.01	0.955	1.40	0.009
Depressive feelings (higher)	1.34	0.001	1.46	0.000
Data collection method				
Postal (ref.)		0.000		0.000
Face-to-face	0.33	0.000	0.47	0.000
Telephone	2.45	0.073	0.81	0.379
Receiving help?				
No help (ref.)		0.000		0.000
Interviewer	0.23	0.002	1.02	0.936
Someone else	0.46	0.002	2.29	0.000

Non-response: ($n = 2586$); Cox and Snell (R^2) = 4.8%; Nagelkerke (R^2) = 9.6%; 7 iterations.
Have been abused: ($n = 2460$); Cox and Snell (R^2) = 7.3%; Nagelkerke (R^2) = 10.3%; 4 iterations.

problem within research on taboo topics and when sensitive questions (such as abuse) are involved.[23,24] For example, a lower item non-response rate for neglect could be caused by the fact there was no exact wording concerning abuse in the phrasing of the question, perhaps therefore making the question less sensitive to respond to. Social desirability can also be related to the situational effect owing to the mode of administration (e.g. form of data collection, provision of interviewer or other help).[22] Further, one possible interpretation is that within postal surveys, respondents could be more inclined to leave the question unanswered, whereas in face-to-face and telephone interviews, social desirability and acquiescence factors could induce people to answer 'abuse never happened'.[13] Alternatively, individuals participating in the two latter forms of research might perceive the task to be difficult and thus be more willing to respond if receiving some type of help to complete the survey (from the telephone interviewer noting answers to questions, for example).

Following this finding, the next research question could be even more prominent. If face-to-face and telephone interviews are more affected by social desirability and acquiescence, then this should be reflected in a lower prevalence rate. As hypothesised, in relation to the data collection methods used, face-to-face interviews generated the lowest prevalence rate. However, this finding is not replicated for telephone interviews, indicating that social desirability may only influence the prevalence rate in face-to-face interviews. A similar hypothesis could be developed for receiving help in filling in the questionnaire. But contrary to our expectations, the presence of a professional interviewer did not impact on prevalence rate. Our findings lend support to the idea that is not merely the issue of provision of an interviewer that is of importance, but rather the location of the interviewer, i.e. whether he/she is directly physically present. Finally, receiving help from someone else (e.g. family, friends, paid carer) is associated with higher prevalence rates. However, this finding is somewhat counter-intuitive, as it is apparent that, as older people may be at most risk of abuse from family, friends and paid carers, it might have been more likely for respondents to report or disclose abuse to professional interviewers assisting them to complete the survey rather than to those who might also be implicated in abusive situations, and this was not evident in this study.

One possible explanation could be that the most vulnerable older people both need more help in participating in the study, and that

this group of individuals are more vulnerable to elder abuse, and that they also report this within the research context. It is also possible that those most at risk exercise a choice to report/disclose to someone about a type of abuse not associated with the person assisting them to complete the survey, e.g. to report abuse by a family member to a paid carer or financial abuse by a paid carer to a friend or family member. The analyses that have been undertaken to date have not allowed us to disaggregate the data to this extent; future research in this area may help to clarify this apparent anomaly.

The results suggest a number of implications for future research: Vulnerable older people (i.e. oldest, individuals with low levels of education, poorer mental health and high levels of frailty) are often under-represented within research, and therefore prevalence rates are possibly under-reported, as the rates do not include representative proportions of those most at risk. Responsiveness to studies may be enhanced by interviewers, and their skills and qualities, but the underlying results clearly point to acquiescence effects and social desirability in sensitive questions. Both effects contribute to a systematic under-representation within general prevalence rates. This can perhaps be minimized if respondents have the possibility/option/choice to respond to them in a self-administered way[24,25] or if interviewers are recruited from the same age, gender or peer-groups,[26] e.g. because the issue of trust may be of less concern to respondents from association with interviewers.

Furthermore, choosing the best type of data collection methods is a crucial point when researching older people (with their associated special population characteristics) because each has advantages and disadvantages. For instance, telephone and face-to-face interviews support interviewing those with good hearing and communication skills, postal surveys benefit those with good levels of reading and writing abilities. Nevertheless, this article points out that the type of data collection may bias the results, e.g. if these abilities are not present any more. Hence, final conclusions are difficult, but future research activities should pay more attention to the application of multi-trait-multi-method (MTMM) studies, which would allow for control of the effects of different data collection methods.

A final remark concerns the limitations of this article. First, this article investigated the differences between data collection methods, but did not fully address the differences that were found between the countries. Cultural factors, such as tolerance and attitudes to violence[27] might also influence both the levels of item-response and the resulting prevalence rate(s). Cross-national comparisons are rare, and further research could usefully examine the between-country variations and identify possible country-specific explanations.[28,29] Second, this study only targeted older women. Insights on gender specificity or how older men would respond to elder abuse items (e.g. possible higher impact of social desirability owing to masculine social roles or norms) is therefore lacking. Third, additional controls relating to social desirability could be performed when examining correlations with 'social desirability' scales (e.g. Marlowe–Crowne Scale).[30] Such correlations might helpfully indicate whether prevalence rates reflect correct estimations or whether they are distorted owing to the lack of willingness to report socially undesirable behaviour.

To conclude, this study demonstrates that mode of administration plays a significant role in terms of outcome of a prevalence study. Given such impact of methodological choices on the responsiveness (item and overall non-response rate) and study results (prevalence rates of abuse), it is crucial that those researchers who want to investigate abuse against older people are aware of these challenges and pay sufficient attention to them, particularly when designing and implementing/undertaking studies of such sensitive, but important issues.

Acknowledgements

Key points

- Accurate estimates of elder abuse are fundamental for developing effective prevention strategies and interventions by health service providers and policy makers, and for evaluating the public health impact of preventive actions.
- Earlier studies on elder abuse assessed unit non-response patterns, but did not examine item non-response patterns.
- Item non-response differs by forms of abuse (e.g. neglect, physical abuse), modes of administration (e.g. type of data collection and help) and individual characteristics (e.g. socio-demographic, health and socio-economic features).
- Type of data collection influenced the prevalence rate of abuse.
- Receiving help from a professional interviewer did not influence the prevalence rate, but receiving help from someone else in filling in the questionnaire was associated with higher prevalence rates and reports of abuse.

References

1   Bennett G, Levin SP, Straka S. Missing Voices. *Views of Older Persons On Elder Abuse*. Geneva: World Health Organization, 2000. Available at: http://www.who.int/ageing/projects/elder_abuse/missing_voices/en/ (30 April 2012, date last accessed).

2   European Community (EU). *EU Charter On Human Rights*. European Community, 2000, Available at: (http://www.europarl.europa.eu/charter/pdf/text_en.pdf (30 April 2012, date last accessed).

3   McGarry J, Simpson C, Hinchliff-Smith K. The impact of domestic abuse for older women: a review of the literature. *Health Soc Care Community* 2011;19:3–14.

4   Cooper C, Selwood A, Livingston G. The prevalence of elder abuse and neglect: a systematic review. *Age Ageing* 2008;37:151–60.

5   Maindal HT, Sokolowski I, Vedsted P. Adaptation, data quality and confirmatory factor analysis of the Danish version of the PACIC questionnaire. *Eur J Pub Health* 2012;22:31–6.

6   De Donder L, Luoma ML, Penhale B, et al. European map of prevalence rates of elder abuse and its impact for future research. *Eur J Ageing* 2011;8:129–43.

7   Lindert J, Luna J, Torres-Gonzalez F, et al. Study design, sampling and assessment methods of the European study 'abuse of the elderly in the European region'. *Eur J Pub Health* 2012;22:662–6.

8   Straus MA. *Conflict Tactics Scales*. New York: Routledge, 2007.

9   AEA. *Action on Elder Abuse's Definition of Elder Abuse*. Action on Elder Abuse Bulletin, No. 11. London: Action on Elder Abuse, 1995.

10   WHO. *The Toronto Declaration on the Global Prevention of Elder Abuse*. Geneva: World Health Organization, 2002.

11   WHO. *European report on preventing elder maltreatment*. Copenhagen: World Heatlh Organization, 2011.

12   Lynn P. The problem of nonresponse. In: De Leeuw E, Hox JJ, Dillman DA, editors. *International Handbook of Survey Methodology*. New York: Psychology Press, 2008: 35–55.

13   Schnell R, Hill PB, Esser E. *Methods of Empirical Social Research,* 8th edn. [in German]. Munich: Oldenbourg, 2008.

14   Babbie ER. *The Practice of Social Research*, 9th edn. Belmont: Wadsworth, 2001.

15   Uusküla A, Kals M, McNutt LA. Assessing non-response to a mailed health survey including self-collection of biological material. *Eur J Pub Health* 2011;21: 538–42.

16   Tu SH, Liao PS. Social distance, respondent cooperation and item nonresponse in sex survey. *Qual Quant* 2007;41:177–99.

17   Straus MA. The revised Conflict Tactics Scales (CTS2). Development and preliminary psychometric data. *J Fam Issues* 1996;17:283–316.

18   Sev'er A. More than wife abuse that has gone old: a conceptual model for violence against the aged in Canada and the US. *J Comp Fam Stud* 2009;40:279–92.

19   Naughton C, Drennan J, Lyons I, et al. Elder abuse and neglect in Ireland: results from a national prevalence survey. *Age Ageing* 2012;41:98–103.

20   Acierno R, Hernandez MA, Amstadter AB, et al. Prevalence and correlates of emotional, physical, sexual, and financial abuse and potential neglect in the United States: the national elder mistreatment study. *Am J Pub Health* 2010;100: 292–7.

21   Couch A, Keniston K. Yeasayers and naysayers: agreeing response set as a personality variable. *J Abnorm Soc Psychol* 1960;60:151–74.

22   Ferrando PJ, Anguiano-Carrasco C. Acquiescence and social desirability as item response determinants: an IRT-based study with the Marlowe–Crowne and the EPQ Lie scales. *Pers Individ Diff* 2010;48:596–600.

23   Bradburn NM, Sudman S, Blair E, Stocking C. Question threat and response bias. *Pub Opin Q* 1978;42:221–34.

24   De Leeuw E, Hox JJ, Kef S. Computer-assisted self-interviewing tailored for special populations and topics. *Field Methods* 2003;15:223–51.

25   Görgen T, Herbst S, Kotlenga S, et al. Crime experiences and experiences of violence in the lives of older people. In: *Summary of Major Events of a Study on Risk of Older and Dependent People [in German]*. Berlin: Federal Ministry for Family Affairs, Senior Citizens, Women and Youth, 2009.

26   De Donder L, Verté D, Messelis E. Fear of crime and the elderly: key-factors that determine fear of crime among elderly people in West Flanders. *Ageing Int* 2005;4:363–376.

27   Schiamberg LB, Gans D. Elder Abuse by Adult Children: An Applied Ecological Framework for Understanding Contextual Risk Factors and the Intergenerational Character of Quality of Life. *Int J Aging Hum Dev* 2000;50:329–59.

28   Harkness JA, Van de Vijver FJR, Mohler PP, editors. *Cross-cultural Suvey Methods*. New Jersey: John Wiley & Sons, 2003.

29   De Leeuw E, Hox JJ, Dillman DA, editors. *International Handbook of Survey Methodology*. New York: Psychology Press, 2008.

30   Crowne D, Marlowe D. A new scale of social desirability independent of psychopathology. *J Consult Psychol* 1960;24:349–54.