Interviewing elderly in nursing homes – Respondent and survey characteristics as predictors of item nonresponse

Survey Methods: Insights from the Field, Special issue: Probability and Nonprobability Sampling: Sampling of hard-to-reach survey populations

Patrick Kutschar | Martin Weichbold

How to cite this article : Kutschar P. & Weichbold M. (2019). Interviewing elderly in nursing homes – Respondent and survey characteristics as predictors of item nonresponse. Survey Methods: Insights from the Field. Retrieved from https://surveyinsights.org/?p=11064

DOI : 10.13094/SMIF-2019-00015

Copyright : © the authors 2019. This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0)

Abstract : Survey methodology is applied regularly in medical, nursing or social science studies examining elderly populations. Research in nursing home residents, where age-related or pathological declines in cognitive function are highly prevalent, faces several methodological challenges. The quality of survey data may be subject to population-specific measurement errors. In this article, data of two studies about pain in nursing homes are used to examine which respondent-, survey- and item characteristics predict item nonresponse. Chances for non-substantial answers are higher for older residents, for females and for those with more cognitive impairment. If residents are in pain, valid answers are more likely. Chances for item nonresponse are less if interviewed by interviewers who are familiar to the respondents. Nonresponse increases with question length and order, and in case questions were preceded by a filter. Less nonresponse is observed for dichotomous answer formats and in case more words per answer were used. These effects are considerably influenced by the respondents' cognitive state and capacity. Results let us assume that respondent, interviewer and item characteristics affect the data quality in nursing home populations significantly. The findings illustrate the necessity of further methodological studies to improve survey data quality in elderly populations with cognitive decline.

Introduction

Information used in studies examining elderly populations often relies on self-report methods (Knäuper et al., 2016). However, the inherent logics of surveying old and oldest old populations may differ substantially from those taking place in younger and less vulnerable populations, where certain health problems are fewer and significantly less likely (Schwarz, 2006). With progressing age, people often experience a range of symptoms including fatigue, musculoscelatal disorders or sensory impairments in auditory and visual processing (Hall, Longhurst, & Higginson, 2009; Huh, 2018). Age-associated physiological degeneration also affects different areas of the human brain, moderating cognition and information processing (Belbase & Sanzenbacher, 2016). Cognitive changes regard all individuals in unique ways and trajectories vary significantly. Typically, basic functions of attention, short-term, long-term and working memory decline with age. Higher-level systems like speech and language abilities,

decision-making capacity as well as executive control and reasoning are likely to be affected (Glisky, 2007; Luo & Craik, 2008). While a certain amount of decline is part of normal aging, dementia is one major cause of more advanced cognitive impairment. Dementia is an umbrella term for symptoms of various types of progressive brain diseases, with Alzheimer's disease and vascular dementia being the most common neurodegenerative pathologies (Livingston et al., 2017). The syndrome is characterized by deterioration in cognitive function beyond of what might be accepted for normal aging accompanied by negative impacts on daily functioning and independent living (Palm et al., 2016; World Health Organization, 2016). Consequences are significant loss of memory, instabilities of emotional states, decreasing abilities to communicate, limited means of comprehension, and declines in semantic and working memory (Palm et al., 2016; Weatherhead & Courtney, 2012). More advanced dementia is a common cause for affected people to move into nursing homes. Study estimates for the number of people with documented dementia in nursing homes often exceed rates of 50% in developed countries (Hoffmann, Kaduszkiewicz, Glaeske, van den Bussche, & Koller, 2014). Around 80% of residents are assumed to be cognitively impaired (Bisla, Calem, Begum, & Stewart, 2011).

Given such high proportions, efforts to include nursing home residents (NHR) in large-scale population surveys or clinical trials are promoted throughout various disciplines. Topics of particular scientific interest are quality of life, care decision-making or management of pain and depression (Corbett et al., 2014; Menne & Whitlatch, 2007; Sivertsen, Bjorklof, Engedal, Selbaek, & Helvik, 2015). Such important phenomena have often been studied applying proxy interviewing or using information from process-generated data sources. However, third-party sources cannot substitute information reported by those affected (Mozley et al., 1999) and gerontological researchers have stated repeatedly that elderly with cognitive impairment can be included in survey research (e.g. Clark, Whitlatch, Tucke, & Rose, 2005; Tyler et al., 2010; Whitlatch & Menne, 2009).

Under reference to the 'total survey error (TSE)' framework, survey data of elderly populations may be subject to a variety of sources of error originating from both TSE components of representation (i.e. to infer from a sample to the target population) and measurement (i.e. to infer from an answer and the underlying construct to the measurement) (Groves & Lyberg, 2010). Attending briefly to the former, it has been stated that response rates, refusals or panel attrition are correlated with individuals' age and health status. Such observations indicate a higher risk of an age-related sample selectivity in surveys of elderly populations: Although detailed findings are somewhat inconsistent, the problem of nonresponse increases with age and constraints in cognitive and physical functions (e.g. Baltes, Schaie, & Nardi, 1971; Herzog & Rodgers, 1988; Kelfve, Thorslund, & Lennartsson, 2013; Oris et al., 2016). Impairment in cognitive function may have fundamental consequences for the measurement process. Answering to survey questions requires the performance of several cognitive tasks (Schwarz, 2007; Tourangeau, Rips, & Rasinski, 2012): Respondents have to understand a given question and its meaning, determine what is being asked for and which kind of information they should provide. Relevant information must be retrieved from memory or certain events have to be recalled. An answer has to be generated based on knowledge retrieved from memory or new judgements have to be computed 'on the spot'. These answers have to be formatted and edited given the presented response alternatives. Before a response is communicated, the answer may be altered to fit the situational context. According to Krosnick (1991), the provision of accurate and valid answers challenges the respondent to successfully search for the most relevant information in mind. If respondents do not run through the question-answer-process as thoroughly as possible, they provide satisfactory instead of optimal answers. Such 'satisficing' is a function of question difficulty, respondents' motivation and cognitive abilities. Whether respondents are less accurate in running through all stages (weak satisficing) or completely skip the tasks of retrieval or judgement (strong satisficing), satisficing theory has been frequently used to explain response effects and bias like context effects, acquiescence, straight-lining, or item nonresponse. If the question-answer-

process is flawed in any way, measurement is erroneous and may result in biased answers or no answers at all.

A number of studies have shown that the response quality in older population studies is likely to be limited due to consequences of normal or pathological aging processes: Information about attitudes and behaviour provided by older respondents tends to be less precise than those of younger ones (Andrews & Herzog, 1986); Elderly seem to have more problems in assigning numeric values to answer categories (Schwarz & Hippler, 1995); The chances for question order effects are reduced but those for response order effects are increased in older persons with impaired short-term memory (Holbrook, Krosnick, Moore, & Tourangeau, 2007; Knäuper, Schwarz, Park, & Fritsch, 2007); Frequency reports about mundane events are less accurate in the elderly, although more accurate information about physical symptoms and the subjective health state is reported (Knäuper, Schwarz, & Park, 2004); Accuracy and consistency of answers is influenced by the severity of cognitive impairment (Whitlatch, Feinberg, & Tucke, 2005); Cognitively less impaired persons seem to be more consistent in answering to fact-based questions than are more severe impaired persons (Clark, Tucke, & Whitlatch, 2008); Elderly seem to favour unidirectional and dichotomous question-answer formats (Krestar, Looman, Powers, Dawson, & Judge, 2012); Age-associated cognitive declines trigger acquiescent response styles (Lechner & Rammstedt, 2015); Respondents with lower cognitive abilities are more likely to provide non-substantial answers which appears to be moderated by question type, format or difficulty (Colsher & Wallace, 1989; Fisher, Burgio, Thorn, & Hardin, 2006; Fuchs, 2009; Knäuper, Belli, Hill, & Herzog, 1997).

Systematic and original methodological studies on data quality from the elderly are still rather scarce. Many findings support the assumption that respondent and interviewer characteristics, formal survey features, or survey modes and administration procedures interact in complex ways. While some findings still await repeated testing, others seem to be contradictory or only observable under certain study conditions (see Knäuper et al., 2016 for a comprehensive review).

Aim and purpose

The quality of survey data is susceptible to age-associated decline and cognitive impairment. This is of special concern when it comes to conducting research in nursing home populations. In this article, we focus on item nonresponse (INR) as one specific measure of survey data quality. Item nonresponse occurs, when answers to particular questions are missing (e.g. item refusal, no answer), provided data is unavailable for analysis (e.g. technical error) or non-substantive answers are given (e.g. ‘don’t know’, ‘cannot say’) (De Leeuw & Hox, 2008).

The aims of the present study are to identify respondent characteristics predicting INR and analyse effects of survey aspects and item features on INR using survey data from two studies in nursing home residents (NHR).

Data, sample and method

We use data from two studies examining pain in German nursing homes (NH). Both were conducted by the Institute of Nursing Science and Practice, Paracelsus Medical University Salzburg. The studies aimed to improve the NHR’ pain situation after an implementation of nursing interventions and computer-assisted personal interviewing was applied.

Study I: ABSM (Action-alliance pain-free city of Muenster)

Design and setting. As one part of the health services research project ABSM (Osterbrink et al., 2010), a pre-post-observational study design was applied to assess the pain situation and the pain management of around 900 NHR in the city of Muenster, Germany. NHR were interviewed using self-report questionnaires or observed with proxy assessment tools, depending on the severity of NHR' state of cognitive impairment. After the intervention phase, same procedures as in the pre-test (2010-2011) were applied for the post-test (2012-2013).

Participants. Residents being at least 65 years of age and provided written informed consent by themselves (or legal guardians, respectively) were initially screened with Mini-Mental State Examination (MMSE) (Folstein, Folstein, & Fanjiang, 2001). MMSE measures cognitive function and results in a score ranging from 0-30 points with lower scores indicating more severe impairment. Residents with 'no or mild' (MMSE 18-30) and 'moderate' (MMSE 10-17) cognitive impairment were interviewed.

Data collection. 13 out of 32 NH participated in the study. Data collection was carried out over a period between four and six weeks per NH and executed by specifically trained interviewers. Interviewers were either health care professionals or students of nursing science or healthcare management. All interviewers were externs to the examined facilities. Interviewers read all questions from survey netbook aloud, while residents obtained print versions of the questionnaire as a visual aid.

Measures. Residents' age, gender, care level (classification of 'need for care' into three levels primarily based on physical abilities), diagnoses, and analgesics were collected directly from the residents' medical records. Participants were interviewed with standardised questionnaires. The questionnaire comprised 32 questions about residents' pain situation and perceived pain management, e.g. duration (*For how long have you been experiencing your pain?* 'less than or equal to three months', 'more than three months', 'more than six months', 'more than a year'), reporting behaviour (*Do you report to the nursing staff when in pain?* 'yes, every time', 'sometimes', 'no, never') or pharmacological treatment (*Have you received pain drugs during the last three days?* 'yes', 'no'). Pain intensity at rest and on movement were assessed using a verbal rating scale ('no, 'mild, 'moderate, 'severe, 'unbearable') (Melzack, 1975). Quality of life was assessed by EQ-5D-3L (The EuroQol Group, 1990), which consists of five dimensions (e.g. mobility, anxiety/depression) with three answer levels ('no problems', 'some problems', 'extreme problems') and a self-rating of health on a visual analogue scale ('0-worst imaginable health state' – '100-best imaginable health state').

Study II: PIASMA (Project to implement an adequate pain management in nursing homes)

Design and setting. In this cluster-randomized controlled trial (DRKS, 2016; ID: DRKS00011062), almost 1,000 residents of 15 randomly selected NH of one single nursing home provider in Bavaria, Germany, were examined to optimize the multiprofessional pain management. Analogous to the procedures of the ABSM study, NHR with up to moderate impairment were interviewed with self-report instruments. Between baseline (2016-2017) and follow up (2017-2018) measurement, an intervention phase took place.

Participants. Residents aged 60 or older were recruited if written informed consent was provided. Similar to ABSM, the nurses acted as gatekeepers and were responsible for recruiting and organization of written consent. NHR were stratified into the same MMSE groups ('no/mild', 'moderate', 'severe') and

those with severe impairment were excluded from self-report methodology.

Data collection. Data collection per NH lasted one week on average and was carried out by interviewers who were either students of health sciences, health care professionals or nurses of the examined nursing homes. Both external (i.e. recruited by the research institution) and internal (i.e. recruited by the nursing home managers) interviewers received the same training. Again, a standardised interviewing approach was pursued; questions were read aloud from survey tablets and presented visually as print versions.

Measures. Gender, age, care grade ('care level' was superseded in 2017 by five 'care grades' based on physical, mental and psychological impediments to independent living), diagnoses, and anaesthetic prescription were extracted from the centralized electronic record system. In contrast to ABSM, the construction of a questionnaire was waived but three validated instruments comprising 43 items were used. Pain was measured by the German version of BPI–Brief Pain Inventory (Budnick et al., 2016), depressive symptoms by the German GDS–Geriatric Depression Scale-SF (Gauggel & Birkner, 1999), and quality of life by EQ-5D-3L (see above). BPI measures pain along two scales: the pain intensity scale assesses least, worst, average and actual pain intensity (e.g. *What number describes your pain on average during the last 24 hours best?* '0-no pain at all'–'10-worst imaginable pain'). The pain interference scale comprises seven items to assess the impact of pain (e.g. *What number describes best how, during the last 24 hours, pain has interfered with your general activity?* '0-does not interfere'–'10-completely interferes'). Further questions about pain localization, duration and quality were added. GDS is a brief questionnaire with 15 dichotomous questions ('yes', 'no') about how residents felt over the past week (e.g. *Do you often feel helpless?*).

Computing primary outcome INR

Respondent-level INR. In both studies, answers were defined as INR if respondents either stated 'I don't know' (DK) or interviewers rated 'cannot be answered' (CA). CA was recorded if NHR couldn't give a substantial answer, couldn't distinguish between answer categories or stated 'cannot say'. Refusals were not counted as INR and breakoffs were excluded from analyses. Although also the causes for DK and CA may differ, we focus on analysing the volume of non-substantial answers. Consequently, INR was computed as each data cases' absolute number of non-substantial answers in relation to the individual number of administrated questions. Hence, the respondent-level INR represents the proportions of non-substantial answers per NHR within each study. INR cannot be prevented completely (De Leeuw & Hox, 2008). Considering extremely skewed distributions, INR was dichotomised for explanatory analyses using a threshold of 'less than 5%' and '5% or more' for considerable INR (Czaja & Blair, 2005; Fowler, 2013; Hardy, Allore, & Studenski, 2009).

Item-level INR. Within each study, every item was transformed in a dichotomous variable representing 'valid answers' or 'INR'. INR was calculated as the share of non-substantial answers in all answers per item—for total sample and both MMSE groups. Data from pre- and post-test were combined. Hence, item-level INR represents the proportions of non-substantial answers for each item within study I and II. In light of a more balanced distribution, INR was applied as a continuous variable in the regression analyses.

Coding of independent variables

Respondent-level. Age in years and survey duration in minutes are included as continuous variables. Presence of diagnoses and medication, gender of respondent and interviewer and, if applicable, affiliation to the nursing home are available as dichotomous variables. Survey daytime and residents' care

level/grade are categorical variables with higher values representing interviews later in the day and more severe health conditions, respectively. The cognitive status is introduced as continuous MMSE score (10-30, higher points indicate less cognitive decline) for descriptives, but as dichotomised explanatory variable MMSE 18-30 (no/mild) and 10-17 (moderate). Items of EQ-5D were computed into a quality of life index (0-100, higher points indicate a better quality of life). Presence of pain ('no' vs. 'yes') was derived from reporting pain at rest or on movement (ABSM), or reporting acute or chronic pain within BPI (PIASMA).

Item-level. Several formal item characteristics which influence the difficulty of a question (Clark et al., 2008; Schaeffer & Dykema, 2011) were coded for each item: Number of words per question were counted (continuous); Question type was categorised into 'factual' (i.e. simple and straightforward questions referring to knowledge), 'attitude' (i.e. questions referring to attitudes or state-dependent expressions, opinions or preferences), and 'behavioural/frequency'-questions (i.e. questions referring to reports about frequency of behaviour, events and other quantities); Presence of introductory phrases ('no' vs. 'yes'); Presence of a preceding filter ('no' vs. 'yes'); Ascending question order within each survey (continuous); Number of words of answer categories (continuous); Number of answer categories (excl. DK/CA, continuous); Answer format ('dichotomous' vs.' polytomous'). We combined factual and behavioural question types into one category due to an infrequent use of factual questions. As an approximate measure for the complexity to map 'mental' answers to presented categories, the ratio of the number of words to the number of answer categories was computed. Thus, a higher ratio indicates more differentiation in answer options (i.e. more words per category).

Statistical analyses

Respondent and item characteristics are presented by common descriptive statistics. Multivariable logistic models were used to regress respondent-level characteristics on considerable INR (>5%). A pre- vs. post-test variable was introduced as further control. Adjusted odds-ratios (aOR), 95%-confidence intervals, and type-I error levels are reported for each independent variable. A multiple linear regression model was applied to measure the adjusted influence of instrument and item features on continuous item-level INR in the combined data set. This model was repeated for each MMSE group to determine different effects at varying levels of cognitive impairment. Unstandardised and standardised coefficients and type-I error levels are presented within each model. To adjust for possible differences between the two studies, the binary variable 'study I vs. II' was integrated. All statistical analyses were conducted using IBM Statistics 24.0.

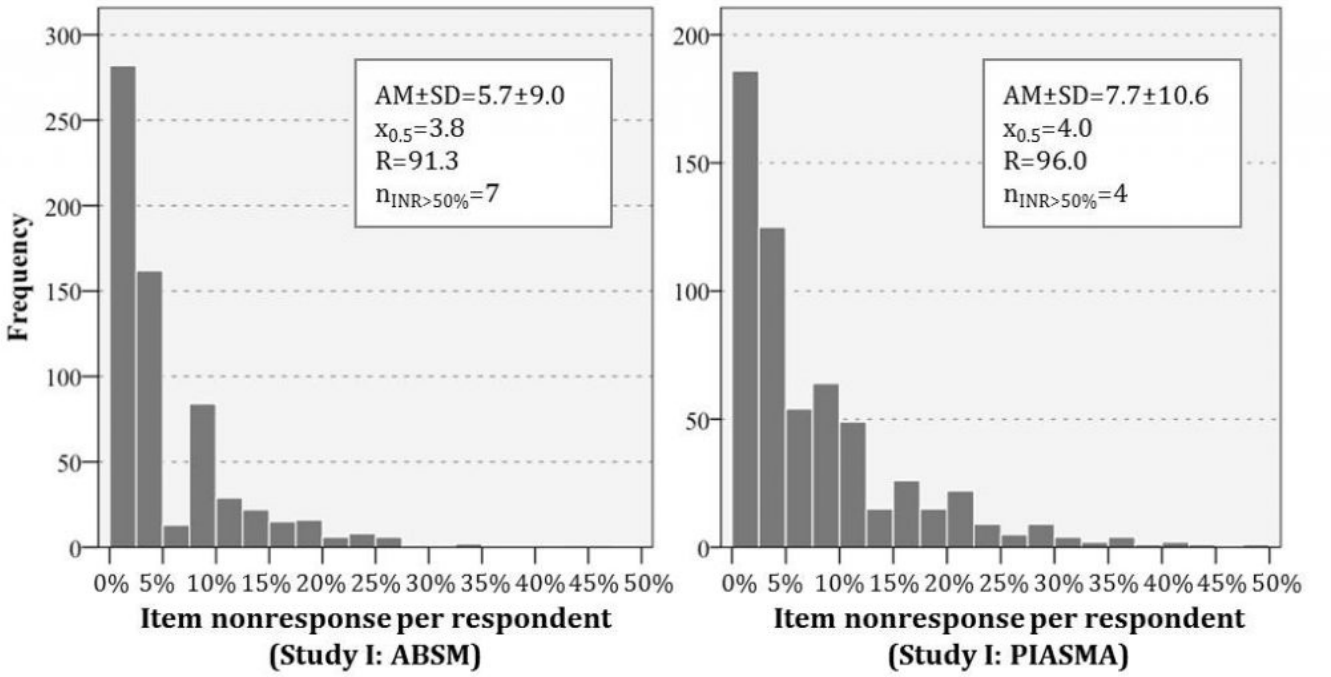
Results

Respondent and item characteristics

Details of respondent and survey characteristics are displayed in *Table 1*. Among the ABSM sample (n=659), three quarters were female, the mean age was 84.5 years, and the respondents averaged 22 points on MMSE. In comparison of the descriptive sample characteristics, the PIASMA sample (n= 598 NHR) comprises fewer women (62.6%), the respondents were younger (mean age 81.0 years) and scored slightly worse on MMSE (20.8 points on average). While every fifth resident showed moderate cognitive impairment in ABSM, this holds true for almost every third resident in PIASMA. Care level (median level I) and grade (median grade III) distribution varied most likely due to the different classification system. Depression (approx. 18%) and dementia (approx. 32% vs. 35%) were documented equally frequent. The studies' distributions for thematical outcome was alike for pain presence (approx. 65%) but varied slightly

for quality of life (mean score 61.2 vs. 67.6). ABSM respondents exhibited an average INR of 5.7%, which was slightly more distinct (mean 7.7%) in PIASMA (*Figure 1*). Interviews lasted 21 minutes on average in ABSM and 19 minutes in PIASMA. Time of the day differed in so far as more PIASMA interviews were conducted after noon. About 70% (ABSM) and 76% (PIASMA) were female interviewers; almost every third was a member of the nursing staff in PIASMA (i.e. internal interviewer).

Figure 1: Distribution of item nonresponse on the respondent-level for study I: ABSM and study II: PIASMA



The question length of 32 ABSM and 43 PIASMA items averages 12.1 and 15.3 words, respectively. In study I, two questions assessed factual knowledge, 13 questions attitudes and situation-dependent information, and 17 questions asked for behaviour reports or frequencies. Two questions contained introductory phrases and eleven items were preceded by filter questions. Rather differently, PIASMA instruments comprised 39 attitude questions, one factual question and three behaviour/frequency questions, 14 items included general instructions, and seven items succeeded filter questions. The mean number of words per answer format are comparable (8.1 vs. 8.2), but more categories were presented in PIASMA (mean 3.0) than in ABSM (mean 5.1). The latter study contained less dichotomous (13 vs. 19) answer formats than the former. On average, ABSM items exhibited 6.4% non-substantial answers, mean item-level INR was 7.3% in PIASMA (*Figure 2*).

Figure 2: Distribution of item nonresponse on the item-level for total sample, MMSE 18-30, and MMSE 10-17

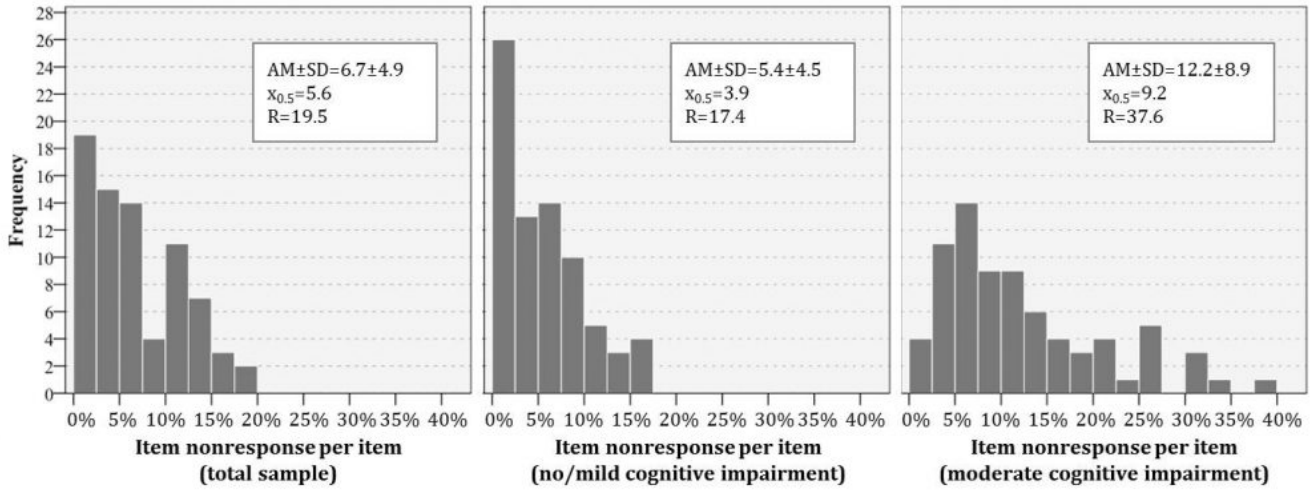


Table 1: Characteristics at respondent- and item-level

	Study I: ABSM	Study II: PIASMA
Respondent-level		
Female %	75.9 %	62.6%
Age in years AM (SD), R	84.5 (7.4), 65-102	81.0 (9.8), 60-101
MMSE score AM (SD)	22.0 (5.3)	20.8 (5.7)
MMSE 10-17 %	20.9 %	30.8%
Care level/grade %		
L0/G0,I	11.5%	8.1%
LI/GII	57.3%	30.7%
LII/GIII	28.0%	38.3%
LIII/GIV,V	3.3%	22.9%
Depression %	18.0%	17.8%
Dementia %	31.7%	34.5%
Morbus Parkinson %	6.6%	-
Stroke %	17.1%	-
Scheduled analgesics %	46.8%	-
Analgesics as needed %	46.5%	-
Anaesthetics %	-	9.1%
Pain presence %	64.5%	64.2%
Quality of life Score AM (SD)	61.2 (22.3)	67.6 (21.9)
Survey daytime %		
<11:00	39.0%	36.0%
11:00 - <13:00	26.5%	14.5%
13:00 - <15:00	18.0%	30.6%
15:00 >	16.5%	18.9%
Survey duration in minutes AM (SD)	20.9 (41.5)	19.2 (49.4)
Interviewer		
Female %	70.4%	75.5%
NH-intern %	-	28.2%
Item nonresponse rate AM (SD)	5.7 (8.9)	7.7 (10.6)
>5% item nonresponse rate %	32.2%	47.5%
	n(total)=656	n(total)=598
Item-level		
Number of words (question) AM (SD)	12.1 (9.2)	15.3 (10.2)
Question type n		
Factual/knowledge	2	1
Attitude/situation-dependent	13	39
Behavioural/frequency	17	3
Questions containing instructions n	2	14
Filtered questions n	11	7
Number of words (answer) AM (SD)	8.1 (8.1)	8.2 (10.2)
Number of categories (answer) AM (SD)	3.0 (1.5)	5.1 (4.4)
Dichotomous answer format	13	19
Item nonresponse rate AM (SD)	6.4 (5.5)	7.3 (4.6)
Categories of item nonresponse rate n		
No INR	-	2
>0 - 5%	18	14
>5 - 10%	5	13
>10%	9	14
	n(total)=32	n(total)=43

Notes: MMSE Mini-Mental Status Examination, AM arithmetic mean, SD standard deviation, R range (min-max), L level, G grade, NH nursing home, INR Item nonresponse, n valid frequency, n(total) sample size, - not applicable

Predictors of INR at the respondent-level

Table 2 shows the results for potential respondent-level predictors of INR in terms of adjusted odds ratios (aOR) within each study. Within the ABSM data, INR was independently predicted by respondents' age, cognitive impairment, care level, dementia diagnosis and pain. Adjusted odds for exhibiting INR of more than 5% increase with age (aOR=1.03, 95% CI=1.01-1.06). The odds to feature INR>5% in NHR with moderate cognitive impairment are more than twice of that in NHR with no or mild impairment (aOR=2.28, 95% CI=1.39-3.75); those for considerable INR increases by 59% (aOR=1.59, 1.04-2.45) if dementia is present. Compared to residents with the least need for care (level 0), care level II (aOR=2.55, 95% CI=1.14-5.70) and III (aOR=5.81, 95% CI=1.61-21.01) were associated with higher odds of INR. NHR being in pain are less likely to provide INR than NHR without pain (aOR=0.59, 95% CI= 0.38-0.93). In study II, INR-proportions of more than 5% were predicted significantly by respondents' gender and cognitive impairment, and interviewers' affiliation to the NH. Odds for males were 41% less of those for females (aOR=0.59, 95% CI=0.36-0.95); those for NHR with moderate cognitive impairments were 1.81-times higher compared to NHR with no or mild (95% CI=1.05-3.13). Compared to being interviewed by external interviewers, being examined by internal interviewers induces a 58% decrease in odds for INR (aOR=0.42, 95% CI=0.24-0.72). Variability ($R^2_{Nagelkerke}$) of INR outcome is explained to 20.3% in ABSM and 12.9% in PIASMA.

Table 2: Multivariable logistic regression models predicting respondent-level INR>5%

	Study I: ABSM aOR [CI95%]	Study II: PIASMA aOR [CI95%]
Socio- and biometrics		
Gender (ref: female)	0.64 [0.39-1.04]ᵀ	**0.59 [0.36-0.95]***
Age in years	**1.03 [1.01-1.06]***	0.99 [0.97-1.02]
MMSE-group (ref: MMSE 18-30)	**2.28 [1.39-3.75]*****	**1.81 [1.05-3.13]***
Care level/grade (ref: L0/G0,I)		
LI/GII	1.86 [0.90-3.88]ᵀ	1.83 [0.77-4.35]
LII/GIII	**2.55 [1.14-5.70]***	1.01 [0.53-1.95]
LIII/GIV,V	**5.81 [1.61-21.01]****	1.12 [0.59-2.07]
Diagnoses		
Depression	0.82 [0.49-1.36]	0.67 [0.39-1.14]
Dementia	**1.59 [1.04-2.45]***	1.21 [0.73-1.99]
Morbus Parkinson	1.48 [0.70-3.10]	-
Stroke	1.10 [0.65-1.86]	-
Medication		
Scheduled analgesics	0.77 [0.52-1.16]	-
Analgesics as needed	0.92 [0.59-1.44]	-
Anaesthetics	-	1.56 [0.76-3.20]
Survey outcome		
Pain presence	**0.59 [0.38-0.93]***	1.68 [1.00-2.83]ᵀ
Quality of life Score	1.00 [0.99-1.01]	1.01 [1.00-1.02]
Survey features and design		
Survey daytime (ref : <11:00)		
11:00 - <13:00	1.32 [0.82-2.12]	0.82 [0.43-1.58]
13:00 - <15:00	0.83 [0.48-1.42]	1.26 [0.75-2.12]
15:00 >	0.87 [0.48-1.42]	0.57 [0.29-1.10]
Survey duration in minutes	1.00 [0.99-1.01]	1.00 [0.99-1.01]
Interviewer		
Sex (ref: female)	1.40 [0.86-2.29]	0.90 [0.53-1.53]
NH affiliation (ref : extern)	-	**0.42 [0.24-0.72]*****
Sample (ref: pre-test)	1.12 [0.66-1.88]	0.65 [0.41-1.03]ᵀ
	n(total)=622 X^2=97.9***, R^2=20.3%	n(total)=598 X^2=42.2***, R^2=12.9%

Notes: aOR adjusted Odds-Ratio, CI95% 95% confidence interval, MMSE Mini-Mental Status Examination, ref reference category, L level, G grade, NH nursing home, n(total) sample size X^2 Chi-Square statistic, R^2 Nagelkerke R-Square statistic, - not applicable, * p<0.05, ** p<0.01, *** p<0.001, ᵀ p<0.10

Predictors of INR at the item-level

Multiple linear regression models were used to test which features of the instruments are associated independently to item nonresponse (*Table 3*). Based on the data of NHR with MMSE scores from 10 to 30 points, higher INR is predicted significantly the more words questions contain (β=0.44), the less words are presented per answer category (β=-0.68), the later a question is positioned in the survey (β=0.51), and if items are in the course of prior filter questions (β=0.35). If the model is repeated for each group of cognitive impairment, the effect pattern differs. The influence of words-to-answers ratio is considerably higher for NHR with MMSE scores between 10 and 17 points (β=0.71). While the effect of preceding filters disappears, an influence of the answer format (dichotomous vs. polytomous) is only present with more severe cognitive impairments (β=0.37). Behaviour or factual question types seem to yield higher INR in residents with moderate impairment—however, this association is not significant. The predictor variables account for 37.0% (R^2: no/mild), 44.7% (R^2: total) and 54.3% (R^2: moderate) of the item-level INR variability, while collinearity is assumed to be non-problematic based on only low to moderate collinearity

indices (CI<15, VIF<5).

Table 3: Multiple linear regression model predicting item-level INR

	Total		No/mild		Moderate	
	B	β	B	β	B	β
Question comprehension						
Number of words	0.23	**0.44****	0.16	**0.34***	0.36	**0.39****
Instructions included	-0.99	-0.08	-0.88	-0.08	2.33	0.11
Question type						
Attitude vs. behaviour/fact	0.83	0.08	0.10	-0.01	2.17	0.11
Answer format						
Ratio words : categories	-1.23	**-0.68*****	-0.92	**-0.57*****	-2.29	**-0.71*****
Dichotomous vs. polytomous	1.00	0.10	-0.72	-0.08	6.65	**0.37****
Layout features						
Question order	0.23	**0.51*****	0.17	**0.42****	0.41	**0.51*****
Preceding filter	4.04	**0.35*****	4.53	**0.43*****	3.43	0.17
Study	0.72	0.07	0.53	0.06	-1.15	-0.06
	n=69, F=7.88***, R^2=44.7%		n=69, F=5.99***, R^2=37.0%		n=69, F=11.11***, R^2=54.3%	

Notes: B unstandardised coefficient, β standardised coefficient, n sample size, F F-Statistic, R^2 Adjusted R-Square Statistic, CI Condition index<15, VIF Variance inflation factor <5, * p<0.05, ** p<0.01, *** p<0.001

Discussion and conclusion

In this article effects of respondent and item characteristics on volume of item nonresponse were analysed in two samples of old and oldest nursing home residents. While most of our results are in line with the relevant studies of data quality in the elderly, our study adds new evidence for some infrequently examined effects.

For respondent-level characteristics it was found that chances for item nonresponse increase if residents are female, older, exhibit higher cognitive impairment and need more care (i.e. worse physical shape)—effects which have been observed in a range of previous studies (e.g. Andrews & Herzog, 1986; Fuchs, 2009; Koyama et al., 2014; Slymen, Drew, Wright, Elder, & Williams, 1994; Whitlatch et al., 2005). Findings suggest that pain presence and being surveyed by NH-intern interviewers decreases item nonresponse. Respondents in an institutionalised setting seem to be more confined and motivated. They may put more effort into the question-answer-process if intimate and salient topics are asked by interviewers they are familiar to (see Weinreb, Sana, & Stecklov, 2018). Pain, depression and quality of life are subject to high self-relevance. Our findings regarding the effects of pain presence are somewhat controversial. In study I, residents with pain show less item nonresponse. Similar to Knäuper et al. (2004) or Borchelt (1999), we assume that people with pain are more eager to report their pain as accurate as possible due to the high subjective relevance, salience and motivation. A non-significant contradictory trend was observed in study II, probably attributable to different routing procedures: If respondents reported no pain, the four items on pain intensity and the seven items on pain interference were skipped. These questions were longer, more complex to understand and comprised an 11-point rating scale answer format. Therefore, NHR with pain are being asked considerably more difficult questions than those without pain, and possible saliency effects may be disguised.

Consistent with previous studies (e.g. Fisher et al., 2006; A. Holbrook, Cho, & Johnson, 2006; Knäuper et

al., 1997; Krestar et al., 2012; Schwarz & Knäuper, 2005), we found that item nonresponse increases with question length, ascending question order and in case questions were preceded by filters. Interestingly, item nonresponse decreases when more words are used to describe the answer categories. These effects interact with the respondents' progress of cognitive decline. The influence of question length and question order is slightly higher in elderly with more severe cognitive impairment. Longer questions are more complex and require greater working memory capacities—abilities which are frequently reduced in residents with moderate impairment. However, they are likely to be even less well performed in the course of the interview. We propose a similar explanation for the effect that less item nonresponse is observed the more words are used per answer category. It seems to be easier to map ones answer into given categories if answer options are more differentiated. With more advanced cognitive impairment, the question purpose and context information may be less available due to declines in short-term memory. A higher ratio of words to answer categories may serve as anchoring source of information if the question is difficult to comprehend or not totally clear. Declines in short-term and working memory may explain that preceding filters increase item nonresponse propensity only in those with no or mild impairments in cognitive functioning. Information and context originating from the initial filter questions may be recalled worse due to cognitive impairment. Hence, less information has to be searched mentally, complex cognitive matching processes are lacking and retrieval and judgment is rendered less erroneous.

Some limitations have to be acknowledged. Referring to the item-level regression models comprising only moderate sample sizes, even low to moderate multicollinearity may lead to upscaled standardised coefficients in the adjusted models. Although the coefficients predict the same nature of effects, they are higher in multivariate modelling. The sensitivity of effects may not only be a statistical issue but represent the complex interactions taking place between different item characteristics and cognitive response processes.

To conclude, our study findings suggest that higher risks for item nonresponse originate from respondents' cognitive decline and age-related impairment. Besides respondent characteristics the manifestation of the respondent-interviewer familiarity and item characteristics influence the complexity of the question-answer process and affect the propensity of item nonresponse.

References

1. Andrews, F. M., & Herzog, A. R. (1986). The Quality of Survey Data as Related to Age of Respondent. Journal of the American Statistical Association, 81(394), 403-410.
2. Baltes, P. B., Schaie, K. W., & Nardi, A. H. (1971). Age and experimental mortality in a seven year longitudinal study of cognitive behavior. Developmental Psychology, 5, 18-26.
3. Belbase, A., & Sanzenbacher, G. T. (2016). Cognitive Aging: A Primer (Vol. 16-17). Boston: Boston College.
4. Bisla, J., Calem, M., Begum, A., & Stewart, R. (2011). Have we forgotten about dementia in care homes? The importance of maintaining survey research in this sector. Age and ageing, 40, 5-6.
5. Borchelt, M., Gilbert, R., Horgas, A. L., & Geiselmann, B. (1999). On the significance of morbidity and disability in old age. In B. P.B. & K. U. Mayer (Eds.), The Berlin aging study: aging from 70 to 100. (pp. 403-429). New York: Cambridge University Press.
6. Budnick, A., Kuhnert, R., Könner, F., Kalinowski, S., Kreutz, R., & Dräger, D. (2016). Validation of a Modified German Version of the Brief Pain Inventory for Use in Nursing Home Residents with Chronic Pain. J Pain, 17(2), 248-256. doi:10.1016/j.jpain.2015.10.016
7. Clark, P., Tucke, S. S., & Whitlatch, C. J. (2008). Consistency of information from persons with dementia: An analysis of differences by question type. Dementia, 7(3), 341-358. doi:10.1177/1471301208093288

8. Clark, P., Whitlatch, C., Tucke, S., & Rose, B. (2005). Reliability of information from persons with dementia. The Gerontologist, 45, 132-132.
9. Colsher, P. L., & Wallace, R. B. (1989). Data quality and age: health and psychobehavioral correlates of item nonresponse and inconsistent responses. Journal of Gerontology, 44(2), P45-52.
10. Corbett, A., Husebo, B. S., Achterberg, W. P., Aarsland, D., Erdal, A., & Flo, E. (2014). The importance of pain management in older people with dementia. British Medical Bulletin, 111(1), 139-148. doi:10.1093/bmb/ldu023
11. Czaja, R., & Blair, J. (2005). Designing Surveys. A Guide to Decicions and Procedures (2. ed.). Thousand Oaks, London, New Delhi: Pine Forge Press.
12. De Leeuw, E. D., & Hox, J. (2008). Missing Data. Retrieved 08.05.2018 http://sk.sagepub.com/reference/survey/n298.xml?fromsearch=true
13. DRKS (2016). German Clinical Trial Register. Retrieved 20.09.2018 https://www.drks.de/drks_web/setLocale_EN.do
14. Fisher, S. E., Burgio, L. D., Thorn, B. E., & Hardin, J. M. (2006). Obtaining Self-Report Data From Cognitively Impaired Elders: Methodological Issues and Clinical Implications for Nursing Home Pain Assessment. The Gerontologist, 46(1), 81-88.
15. Folstein, M. F., Folstein, S. E., & Fanjiang, G. (2001). MMSE Mini-Mental State Examination clinical guide. Odessa: Psychological Assessment Resources.
16. Fowler, F. J. (2013). Survey Research Methods. Applied Social Research Methods. Thousand Oaks et al.: Sage.
17. Fuchs, M. (2009). Item-Nonresponse in a Survey among the Elderly. The Impact of Age and Cognitive Resources. In M. Weichbold, J. Bacher, & C. Wolf (Eds.), Umfrageforschung: Herausforderungen und Grenzen (pp. 333-349). Wiesbaden: VS Verlag für Sozialwissenschaften. (Original work published in German)
18. Gauggel, S., & Birkner, B. (1999). Validity and reliability of a German version of the Geriatric Depression Scale (GDS). Zeitschrift für klinische Psychologie-Forschung und Praxis, 28(1), 18-27.
19. Glisky, E. L. (2007). Changes in Cognitive Function in Human Aging. In D. R. Riddle (Ed.), Brain Aging: Models, Methods, and Mechanisms. Boca Raton (FL): CRC Press/Taylor & Francis.
20. Groves, R. M., & Lyberg, L. (2010). Total Survey Error: Past, Present, and Future. Public Opinion Quarterly, 74(5), 849-879.
21. Hall, S., Longhurst, S., & Higginson, I. J. (2009). Challenges to conducting research with older people living in nursing homes. BMC geriatrics, 9, 38-46.
22. Hardy, S. E., Allore, H., & Studenski, S. A. (2009). Missing Data: A Special Challenge in Aging Research. Journal of the American Geriatrics Society, 57(4), 722-729.
23. Herzog, A. R., & Rodgers, W. L. (1988). Age and Response Rates to Interview Sample-Surveys. Journal of Gerontology, 43(6), S200-S205.
24. Hoffmann, F., Kaduszkiewicz, H., Glaeske, G., van den Bussche, H., & Koller, D. (2014). Prevalence of dementia in nursing home and community-dwelling older adults in Germany. Aging Clinical and Experimental Research, 26(5), 555-559. doi:10.1007/s40520-014-0210-6
25. Holbrook, A. L., Cho, Y. I., & Johnson, T. (2006). The impact of question and respondent characteristics on comprehension and mapping difficulties. Public Opinion Quarterly, 70(4), 565-595.
26. Holbrook, A. L., Krosnick, J. A., Moore, D., & Tourangeau, R. (2007). Response order effects in dichotomous categorical questions presented orally – The impact of question and respondent attributes. Public Opinion Quarterly, 71(3), 325-348. doi:10.1093/poq/nfm024
27. Huh, M. (2018). The relationships between cognitive function and hearing loss among the elderly. J Phys Ther Sci, 30(1), 174-176. doi:10.1589/jpts.30.174
28. 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.
29. Knäuper, B., Belli, R. F., Hill, D. H., & Herzog, R. A. (1997). Question Difficulty and Respondents' Cognitive Ability: The Effect on Data Quality. Journal of Official Statistics, 13(2), 181-199.

30. Knäuper, B., Carriere, K., Chamandy, M., Xu, Z., Schwarz, N., & Rosen, N. O. (2016). How aging affects self-reports. European Journal of Ageing, 13(2), 185-193.
31. Knäuper, B., Schwarz, N., & Park, D. (2004). Frequency Reports Across Age Groups. Journal of Official Statistics, 20(1), 91-96.
32. Knäuper, B., Schwarz, N., Park, D., & Fritsch, A. (2007). The Perils of Interpreting Age Differences in Attitude Reports: Question Order Effects Decrease with Age. Journal of Official Statistics, 23(4), 515-528.
33. Koyama, A., Fukunaga, R., Abe, Y., Nishi, Y., Fujise, N., & Ikeda, M. (2014). Item non-response on self-report depression screening questionnaire among community-dwelling elderly. Journal of Affective Disorders, 162, 30-33.
34. Krestar, M. L., Looman, W., Powers, S., Dawson, N., & Judge, K. S. (2012). Including Individuals with Memory Impairment in the Research Process: The Importance of Scales and Response Categories Used in Surveys. Journal of Empirical Research on Human Research Ethics: An International Journal, 7(2), 70-79.
35. Krosnick, J. A. (1991). Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. Applied Cognitive Psychology, 5(3), 213-236. doi:DOI 10.1002/acp.2350050305
36. Lechner, C., & Rammstedt, B. (2015). Cognitive Ability, Acquiescence, and the Structure of Personality in a Sample of Older Adults. Psychological Assessment, 27(4), 1301-1311.
37. Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., . . . Mukadam, N. (2017). Dementia prevention, intervention, and care. The Lancet, 390(10113), 2673-2734. doi:https://doi.org/10.1016/S0140-6736(17)31363-6
38. Luo, L., & Craik, F. I. M. (2008). Aging and memory: A cognitive approach. Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie, 53(6), 346-353.
39. Melzack, R. (1975). The McGill Pain Questionnaire: major properties and scoring methods. Pain, 1(3), 277-299.
40. Menne, H. L., & Whitlatch, C. J. (2007). Decision-making involvement of individuals with dementia. The Gerontologist, 47(6), 810-819.
41. Mozley, C. G., Huxley, P., Sutcliffe, C., Bagley, H., Burns, A., Challis, D., & Cordingley, L. (1999). 'Not knowing where I am doesn't mean I don't know what I like': Cognitive impairment and quality of life responses in elderly people. International journal of geriatric psychiatry(14), 776-783.
42. Oris, M., Guichard, E., Nicolet, M., Gabriel, R., Tholomier, A., Monnot, C., . . . Joye, D. (2016). Representation of Vulnerability and the Elderly. A Total Survey Error Perspective on the VLV Survey. In M. Oris, C. Roberts, D. Joye, & M. E. Stähli (Eds.), Surveying Human Vulnerabilities across the Life Course (Vol. 3, pp. 27-65): Springer.
43. Osterbrink, J., Ewers, A., Nestler, N., Pogatzki-Zahn, E., Bauer, Z., Gnass, I., . . . van Aken, H. (2010). Health services research project "Action Alliance Pain-free City Munster". Der Schmerz, 24(6), 613-620. doi:10.1007/s00482-010-0983-2. (Original work published in German)
44. Palm, R., Junger, S., Reuther, S., Schwab, C. G., Dichter, M. N., Holle, B., & Halek, M. (2016). People with dementia in nursing home research: a methodological review of the definition and identification of the study population. BMC geriatrics, 16, 78. doi:10.1186/s12877-016-0249-7
45. Schaeffer, N. C., & Dykema, J. (2011). Questions for Surveys. Current Trends and Future Directions. Public Opinion Quarterly, 75(5), 909-961.
46. Schwarz, N. (2006). Measurement: Aging and the Psychology of Self-Report. In N. Schwarz (Ed.), When I'm 64 (pp. 219-230).
47. Schwarz, N. (2007). Cognitive Aspects of Survey Methodology. Applied Cognitive Psychology, 21, 277-287.
48. Schwarz, N., & Hippler, H. J. (1995). The numeric values of rating scales: A comparison of their impact in mail surveys and telephone interviews. International Journal of Public Opinion Research, 7, 72-74.
49. Schwarz, N., & Knäuper, B. (2005). Cognition, Aging, and Self-Reports. In D. Park & N. Schwarz (Eds.), Cognitive Aging – A Primer (2nd edition) (pp. 233-252). Philadelphia: Psychology Press.

50. Sivertsen, H., Bjorklof, G. H., Engedal, K., Selbaek, G., & Helvik, A.-S. (2015). Depression and Quality of Life in Older Persons: A Review. Dement Geriatr Cogn Disord, 40, 311-339.
51. Slymen, D. J., Drew, J. A., Wright, B. L., Elder, J. P., & Williams, S. J. (1994). Item non-response to lifestyle assessment in an elderly cohort. International Journal of Epidemiology, 23(3), 583-591.
52. The EuroQol Group. (1990). EuroQol-a new facility for the measurement of health-related quality of life. Health Policy, 16(3), 199-208.
53. Tourangeau, R., Rips, L. J., & Rasinski, K. (2012). The Psychology of Survey Response (Vol. 13). Cambridge et al.: Cambridge University Press.
54. Tyler, D. A., Shield, R. R., Rosenthal, M., Miller, S. C., Wetle, T., & Clark, M. A. (2010). How Valid Are the Responses to Nursing Home Survey Questions? Some Issues and Concerns. The Gerontologist, 51(2), 201-211.
55. Weatherhead, I., & Courtney, C. (2012). Assessing the signs of dementia. Practice Nursing, 23(3), 114-118.
56. Weinreb, A., Sana, M., & Stecklov, G. (2018). Strangers in the Field: A Methodological Experiment on Interviewer-Respondent Familiarity. Bulletin de Methodologie Socioloqique(137-138), 94-119.
57. Whitlatch, C. J., Feinberg, L. F., & Tucke, S. (2005). Accuracy and consistency of responses from persons with cognitive impairment. Dementia, 4(2), 171-183. doi:10.1177/1471301205051091
58. Whitlatch, C. J., & Menne, H. L. (2009). Don’t Forget About Me! Decision Making by People with Dementia. Generations – Journal of the American Society on Aging, 33(1), 66-73.
59. World Health Organization. (2016). Fact sheet No 362 – Dementia. Retrieved from http://www.who.int/mediacentre/factsheets/fs362/en/