Age and Ageing 2019; **0:** 1–7
doi: 10.1093/ageing/afy191

© The Author(s) 2019. Published by Oxford University Press on behalf of the British Geriatrics Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Measuring health-related quality of life of care home residents: comparison of self-report with staff proxy responses

Adeela Usman[1], Sarah Lewis[2], Kathryn Hinsliff-Smith[1], Annabelle Long[1], Gemma Housley[3], Jake Jordan[4], Heather Gage[4], Tom Dening[5], John R F Gladman[6,7,8], Adam L Gordon[1,7,8,9]

[1]Division of Medical Sciences and Graduate Entry Medicine, School of Medicine, University of Nottingham, Derby, UK
[2]Division of Epidemiology and Public Health, School of Medicine, University of Nottingham, Nottingham, UK
[3]East Midlands Academic Health Science Network, Nottingham, UK
[4]Department of Clinical and Experimental Medicine, Surrey Health Economics Centre, University of Surrey, Guildford, UK
[5]Division of Psychiatry and Applied Psychology, School of Medicine, University of Nottingham, Nottingham, UK
[6]Division of Rehabilitation and Ageing, University of Nottingham, Nottingham, UK
[7]East Midlands Collaboration for Leadership in Applied Health Research and Care, Nottingham, UK
[8]Nottingham Biomedical Research Centre, University of Nottingham, Nottingham, UK
[9]School of Health Sciences, City, University of London, London, UK

Address correspondence to: Dr Adam Gordon, Derby Medical School, Royal Derby Hospital, Room 4113, Derby DE22 3NE, UK. Tel: +01332 724668; Fax: +01332 724697; Email: adam.gordon@nottingham.ac.uk

Abstract

Introduction: care home residents are often unable to complete health-related quality of life questionnaires for themselves because of prevalent cognitive impairment. This study compared care home resident and staff proxy responses for two measures, the EQ-5D-5L and HowRU.
Methods: a prospective cohort study recruited residents ≥60 years across 24 care homes who were not receiving short stay, respite or terminal care. Resident and staff proxy EQ-5D-5L and HowRu responses were collected monthly for 3 months. Weighted kappa statistics and intra-class correlation coefficients (ICCs) adjusted for clustering at the care home level were used to measure agreement between resident and proxies for each time point. The effect of staff and resident baseline variables on agreement was considered using a multilevel mixed effect regression model.
Results: 117, 109 and 104 matched pairs completed the questionnaires at 1, 2 and 3 months, respectively. When clustering was controlled for, agreement between resident and staff proxy EQ-5D-5L responses was fair for mobility (ICC: 0.29) and slight for all other domains (ICC ≤ 0.20). EQ-5D Index and Quality-Adjusted Life Year scores (proxy scores higher than residents) showed better agreement than EQ-5D-VAS (residents scores higher than proxy). HowRU showed only slight agreement (ICC ≤ 0.20) between residents and proxies. Staff and resident characteristics did not influence level of agreement for either index.
Discussion: the levels of agreement for EQ-5D-5L and HowRU raise questions about their validity in this population.

Keywords

Residential facilities, nursing homes, quality of life, proxy, older people.

Key points

- Dementia is prevalent in UK care homes, limiting the usefulness of self-reported quality of life measures.
- This study found agreement between resident and proxy responses for EQ-5D-5L and HowRU quality of life measures was inadequate.
- Further work is required to better describe health-related quality of life as an outcome measure in the care home sector.

Downloaded from https://academic.oup.com/ageing/advance-article-abstract/doi/10.1093/ageing/afy191/5274649 by University of Nottingham user on 09 January 2019

Introduction

Long-term care facilities in the UK are called care homes and classified as either care homes with or without nursing based upon the availability of registered nurses on-site. The types of residents cared for in both classifications of facility are similar and all UK care homes are included in the international consensus definition of a nursing home [1]. Around 425,000 people live in UK care homes [2] with most residents requiring care due to disability from long-term conditions. The majority of residents are aged over 85, 75–80% of residents live with dementia [3], and over half of the residents die within 12 months of arrival [4].

Improving the quality of care for older people in long-term care has become a focus of attention both within the UK and internationally [1], and an increasing number of evaluative research studies are testing the effectiveness and cost-effectiveness of interventions in this setting [5]. Residents' quality of life is frequently used as an outcome measure in these studies to maintain a patient-centred focus and facilitate health economic evaluation.

The EuroQoL EQ-5D questionnaires are widely used preference-based health-related quality of life measures suitable for use in economic evaluations. They were specifically designed to be quick and easy to complete. The first version of EQ-5D measured five domains of quality of life on three levels (EQ-5D-3L). EQ-5D-3L has been shown to have good construct validity for self-report [6] and has been used to measure quality of life of older people living in their own homes and in care homes [7] The five-level version, EQ-5D-5L, was developed subsequently to deal with identified issues with sensitivity and a ceiling effect on the EQ-5D-3L which limited its ability to discriminate between health states, particularly in those with higher quality of life [8]. The EQ-5D-5L version measures health-related quality of life across five domains (mobility, self-care, usual activities, pain, and anxiety/depression) with the scale for each domain ranging from level 1 (no problems) to level 5 (extreme problems). The responses from the five domains are converted to QoL index scores (utilities) generated from a given country's general population [9]. These index scores can be used to calculate quality-adjusted life years (QALYs), which are a measure of the person's state of health. One QALY equates to one year in perfect health. The cost per QALY gained from an intervention when compared to usual care is the chosen cost-utility measure for determining eligibility for funding support through the UK National Health Service [10]. EQ-5D-5L also includes a visual analogue scale which asks respondents to indicate on a thermometer how they feel that day, with anchor points of 100 (best possible health) and 0 (worst possible health).

The prevalence of frailty and cognitive impairment in the care home population means that collecting self-reported quality of life measures from residents is challenging. In response to this, proxy responses to quality of life items have sometimes been used [11]. For these, a consultee, drawn from care home staff, or a relative or friend who has regular ongoing visits, answers questions on the resident's behalf. Using proxy respondents can be unreliable in care home settings. There may be lack of continuity of care home staff contact with individual residents due to shift working and staff turnover, and family and friends may not be well placed to judge QoL domains if they visit residents for only short periods [11].

There is limited evidence comparing self-reported and proxy responses to the EQ-5D-5L in care home populations [12]. There is a particular paucity of data in UK care home populations.

HowRu ('How Are You') is a patient-rated outcome measure which has been specifically designed for use in long-term care settings to address quality of life in a way that is practical for older people [13, 14]. It records four variables (pain or discomfort, feeling low or worried, limitation in activities and dependency on others) related to QoL at a fixed point in time ('How are you doing today?') on a four-point scale (none, slight, quite a lot, extreme). The HowRu score is calculated by summing up the values for each domain to give a value on a 13-point scale ranging from 0 (worst) to 12 (best). HowRU may have greater cogency and immediacy than EQ-5D-5L. In a comparison with EQ-5D in patients attending a cardiovascular outpatient clinic, HowRu was reported to have better readability, higher completion rate and report a wider range of states [13]. HowRu has not been evaluated for older people living in care homes. It is not known whether proxy responses in this setting may be useful for HowRU.

This study was conducted as a part of a programme of work focussing on improving quality of care in UK care homes which used EQ-5D-5L and HowRu as outcome measures. To inform our use of proxy measures, we set out to establish the reliability of staff proxy responses for both indices.

Method

The full study protocol has been published [15]. Participants were a sub-population of care home residents recruited as a part of the Proactive Health Care in Care Homes (PEACH) study. The PEACH study includes an open cohort stepped wedge randomised trial to assess the impact of Comprehensive Geriatric Assessment implemented by Quality Improvement Collaborative. Comprehensive Geriatric Assessment is widely recognised as a gold-standard way to deliver care for older people with frailty [16]. PEACH uses EQ-5D-5L and HowRu as outcome measures and understanding their measurement properties represented important preparatory work. The measurements for this study coincided with months 3–6 of the PEACH study, with data collection for PEACH continuing for a further 6 months. The findings of this study were therefore able to inform the PEACH analysis.

All residents of care homes participating in PEACH who were aged ≥60 years were eligible for inclusion. Those

Downloaded from https://academic.oup.com/ageing/advance-article-abstract/doi/10.1093/ageing/afy191/5274649 by University of Nottingham user on 09 January 2019

who were admitted for short term respite or immediately approaching end of life were excluded. Informed consent was obtained from residents who had mental capacity and from an appropriate consultee when residents lacked mental capacity. Capacity assessments were based on the guidelines in the 2005 Mental Capacity Act for England and Wales.

For this study, in addition to the routine collection of EQ-5D-5L and HowRu from residents undertaken as part of PEACH, proxy responses to EQ-5D-5L and HowRu were gathered from staff. We included staff such as care assistants, care home managers and registered nurses, who were identified by the care home manager as most familiar with the resident. This placed emphasis on staff providing personal care to the resident on the day of data collection because both EQ-5D-5L and HowRu ask about the resident's health today. We excluded staff employed in a supportive role, such as activity coordinators, since their orientation to supporting residents is more variable and they are less likely to be involved in personal care.

Data were collected from proxies in three consecutive months during 2017 and matched with resident data for those months. The EQ-5D-5L questionnaire, including the EQ-5D visual analogue score, was used. Responses from the five domains were transformed into utilities (index scores) derived from the UK general population. This was done using the crosswalk value set [17]. For residents with proxy and self-reported responses at all three time points, QALYs were calculated using the area under the curve.

HowRu has four domains scaled from 0 to 3. Values for each domain were none = 3, slight = 2, quite a lot = 1 and extreme = 0. These were summed to give a 13-point scale ranging from 0 (worst health) to 12 (best health).

To standardise responses, taking account of residents who were unable to read or write, the researchers read the questionnaire for both EQ-5D-5L and HowRu to participants and then recorded their responses. For staff responses, we asked them to consider the proxy–resident's perspective when completing the questionnaire using the following statement: 'Please rate how you (staff) think the resident will rate his/her own health-related quality of life, if the resident was to communicate'. A researcher sat with staff and directly addressed any questions whilst they completed their responses but did not otherwise direct them. Responses for both self-reported and proxy questionnaires were completed on the same day.

Analysis was based on cross-sectional analysis of agreement at each time point. For the EQ-5D-5L and HowRu domain levels, the levels of agreement between self-reported and staff responses were calculated using percent agreement and weighted kappa statistics at 1, 2 and 3 months. Weighted kappa helps to distinguish between small and large differences in agreement ratings assigned to the different levels of each domain but with equal importance given to disagreement [18, 19]. We used linear weights for the weighted kappa: this assigns the same importance to the difference between any two categories within the response scale [20]. The 95% confidence interval for the weighted kappa was calculated at each time by bootstrapping using Stata 15 (Statacorp, LLC, 2015) with 1000 replications.

The kappa statistic ranges from −1 to 1, and the strength of the agreement was interpreted with regards to published guidelines [21] with agreement being:

- Poor, if kappa ≤ 0.00
- Slight, if kappa = 0.01–0.20
- Fair, if kappa = 0.21–0.40
- Moderate, if kappa = 0.41–0.60
- Substantial, if kappa = 0.61–0.80
- Almost perfect, if kappa ≥ 0.81

For the EQ-5D visual analogue scale, EQ-5D-5L index scores, QALYs and HowRu scores, the levels of agreement between the self-reported and proxy responses were assessed by calculating the intra-class correlation coefficient (ICC) at each time point using a two-way mixed effect analysis of variance (ANOVA) model [22]. Although the ANOVA model has been reported to be robust to deviation in normality, bootstrapping was run to assess if it made any difference to the estimated ICCs. The same benchmarks used for kappa were used for the intra-cluster correlation coefficients.

ICCs were calculated for EQ-5D-Self, EQ-5D-5L-Proxy, HowRu-Self and HowRu-Proxy. Since the calculation of kappa and ICC assumes independence of observations, we adjusted for clustering. In our study, clustering may have occurred at three levels. First, at the care home level where residents within the same care home have similar characteristics and are different from those in other care homes. Second, at staff level where staff members within a care home respond on behalf of multiple residents, and third at the individual level where responses are clustered within each resident.

For the ICCs, clustering was adjusted for using a multilevel mixed effect model by fitting a two-level random effect model with a random effect for care home and individuals. For the kappa statistics, clustering was adjusted for using a variance formula [15].

The study sample size was 160 residents, based upon a kappa of 0.145 and a confidence level width of 0.153 taken from a previous study [23].

Results

117, 109 and 104 matched pairs completed the questionnaires at 1, 2 and 3 months, respectively. The mean (SD) age of the residents was 86.8 (7.6) years and 68% were female. Forty-four percentage of participants had a documented diagnosis of dementia or cognitive impairment in their care home record. The characteristics of staff who provided proxy responses are reported in Table 1.

The agreement between proxies and residents for individual domains of the EQ-5D-5L and HowRU, respectively, are summarised in Tables 2 and 3, respectively. The strength of agreement found between staff and residents for HowRu measure was weaker than for EQ-5D-5L. The intra-cluster correlation showed clustering of measures within care homes. When kappa values were adjusted for

Downloaded from https://academic.oup.com/ageing/advance-article-abstract/doi/10.1093/ageing/afy191/5274649 by University of Nottingham user on 09 January 2019

Table 1. Characteristics of care home staff

Characteristic	*N* (%) at baseline
Age group (*n* = 117)	
18–35 years	39 (33.3%)
36–55 years	50 (42.7%)
Aged 56 or older	28 (23.9%)
Sex (*n* = 117): numbers of female staff	103 (88.0%)
Role/rank (*n* = 117)	
Care worker or health care assistant	89 (76.1%)
Registered nurse	6 (5.1%)
Other	22 (18.8%)
Care home assistant practitioner	2 (1.7%)
Nursing assistant	6 (5.1%)
Senior care assistant	13 (11.1%)
Deputy Manager	1 (0.9%)
Length of time working in study care home (*n* = 117)	
Less than 6 months	15 (12.8%)
6–11 months	5 (4.3%)
1–5 years	42 (35.9%)
More than 5 years	55 (47.0%)
Length of time working in care of older people (*n* = 117)	
Less than 6 months	7 (6.0%)
6–11 months	3 (2.6%)
1–5 years	31 (26.5%)
More than 5 years	76 (65.0%)
Frequency of delivering care to resident (*n* = 107)	
Most/all of the time	86 (73.5%)
Sometimes	19 (16.2%)
Rarely	2 (1.7%)

clustering, agreement was fair for the mobility domain of EQ-5D-5L and slight for all other domains. Agreement was slight for all domains of HowRU when clustering was accounted for.

Mean total resident EQ-5D-5L (0.57, 0.50, 0.58) and HowRu (9.4, 9.2, 9.6) scores were higher than proxy EQ-5D-5L (0.43, 0.42, 0.42) and HowRu (8.4, 8.3, 9.0) scores at all three time points. By contrast, the mean EQ-5D-VAS was higher in proxies (68, 74, 72) compared to residents (65, 63, 69) across all time points.

The strength and magnitude of agreement between residents and proxies for EQ-5D visual analogue scale (cluster adjusted ICC: 0.24) was less than for index EQ-5D-5L scores (cluster adjusted ICC: 0.55), and agreement for the latter was less than for QALYs (cluster adjusted ICC: 0.70). At all-time points EQ-5D visual analogue scale showed a slight to fair agreement, EQ-5D-5L index scores showed moderate–substantial agreement, and QALYs showed substantial agreement between residents and proxy responses.

When regression analysis was conducted on resident and staff characteristics to consider their impact on the difference between EQ-5D-5L and HowRu scores, no statistically significant associations were identified.

Discussion

This study compared UK care home resident and staff proxy responses to the EQ-5D-5L and HowRu. Agreement for the domains of both the EQ-5D-5L and HowRu were slight when clustering was accounted for, with the exception of the EQ-5D-5L mobility domain where agreement was still only fair. EQ-5D-5L total and HowRu scores reported by residents were higher than when reported by proxies, yet the mean EQ-5D visual analogue scale was higher in proxies than residents, indicating further concern about the measurement of health-related quality of life in care home residents.

A strength of this study was that analyses controlled for clustering, important due to the heterogeneity of UK care homes, where the resident case-mix and staff skill-mix vary substantially between institutions [3]. We were further able to understand how resident and staff attributes influenced agreement by using regression analysis to consider their impact and found that they had no influence by doing so.

The main limitation of the study was in the low prevalence of dementia reported in the cohort. Other studies have shown the prevalence of dementia in UK care homes to be close to 80% [3, 24]. We found that a formal diagnosis of cognitive impairment did not impact on the level of agreement between residents and proxies and therefore, even though the sample here was unrepresentative of the care home population as a whole, we do not think this influences the validity of our findings about EQ-5D-5L and HowRU proxy measurements in this setting.

The main reason for wanting to use staff proxy responses for EQ-5D-5L in the care home population is the high prevalence of cognitive impairment. However, the findings here match those in studies using EQ-5D-3L, and done in long-term care sectors in other countries, which suggest that staff proxy ratings consistently differ from those of residents, for residents with or without cognitive impairment [11, 25]. The reason for these differences is not clear. It may be that staff and residents understand the domains included in health-related quality of life measures differently. It may, alternatively, be that indices developed in non-care home settings, do not include the sort of domains upon which care home staff feel they can reliably comment. Further work is required to understand the ways in which proxy respondents understand existing measures, whether these measures can be adapted to take account of current difficulties, or whether new care home-specific measures are required.

Several other health-related quality of life indices have been developed to take specific account of dementia. The most notable amongst these are QUALIDEM [26], DeMQoL [27] and QoL-AD [28]. DeMQoL was developed in non-institutional community settings and the proxy responses within it have not been validated for care homes, where carer relationships are different to those for patients cared for in their own homes. QUALIDEM was developed in the long-term care setting and is an observation-based measure, with good test–retest reliability but some issues with inter-observer agreement, with four out of the nine subscales showing poor–moderate agreement only [29]. Work comparing QoL-AD with EQ-5D in people with

Downloaded from https://academic.oup.com/ageing/advance-article-abstract/doi/10.1093/ageing/afy191/5274649 by University of Nottingham user on 09 January 2019

Table 2. Resident–proxy agreement using the exact percent agreement and Kappa values for the EQ-5D-5L at three points in time

Domain	Time point (month)	Kappa coefficient (95% CI)	Kappa adjusted for clustering (95% CI)	Number of clusters (range in cluster size)
Mobility	1	0.48	0.22	16
	(n = 117)	(0.36,0.59)	(0.01, 0.46)	(2–15)
	2	0.56	0.33	11
	(n = 109)	(0.44–0.65)	(0.16, 0.50)	(3–14)
	3	0.48	0.33	14
	(n = 104)	(0.35–0.60)	(0.16, 0.50)	(2–14)
Self-care	1	0.36	0.19	14
	(n = 117)	(0.26–0.46)	(0.04, 0.35)	(2–15)
	2	0.25	0.10	14
	(n = 109)	(0.15–0.38)	(0.00, 0.21)	(2–14)
	3	0.33	0.23	12
	(n = 104)	(0.21–0.44)	(0.12,0.35)	(2–14)
Usual activities	1	0.15	0.02	14
	(n = 117)	(0.02–0.28)	(−0.17, 0.21)	(2–15)
	2	0.26	0.13	13
	(n = 109)	(0.12–0.39)	(0.00, 0.25)	(2–14)
	3	0.17	0.09	11
	(n = 104)	(0.02–0.30)	(−0.02, 0.20)	(2–14)
Pain/discomfort	1	0.22	0.14	14
	(n = 117)	(0.11–0.34)	(−0.02,0.30)	(2–15)
	2	0.20	0.16	11
	(n = 109)	(0.08–0.31)	(0.05,0.28)	(5–14)
	3	0.14	0.11	10
	(n = 104)	(0.03–0.26)	(−0.00,0.23)	(5–14)
Anxiety/depression	1	0.08	0.05	10
	(n = 117)	(−0.03–0.21)	(0.09,0.19)	(2–15)
	2	0.10	0.08	9
	(n = 109)	(−0.03–0.23)	(− 0.02–0.18)	(2–14)
	3	0.24	0.14	12
	(n = 104)	(0.09–0.42)	(−0.07–0.35)	(2–14)

Table 3. Resident–proxy agreement using the exact percent agreement and Kappa values for HowRu at three points in time

Domain	Time point (month)	Kappa coefficient (95% CI)	Kappa adjusted for clustering (95% CI)	Number of clusters (range in cluster size)
Pain/discomfort	1	0.25	0.14	14
	(n = 117)	(0.14–0.35)	(−0.02,0.30)	(2–15)
	2	0.18	0.16	11
	(n = 109)	(0.06–0.31)	(0.05,0.28)	(5–14)
	3	0.16	0.11	10
	(n = 104)	(0.05–0.29)	(−0.00,0.23)	(5–14)
Feeling low or worried	1	0.22	0.16	13
	(n = 117)	(0.09–0.37)	(0.02,0.30)	(3–15)
	2	0.20	0.16	11
	(n = 109)	(0.07–0.35)	(0.00,0.35)	(3–14)
	3	0.14	0.09	11
	(n = 104)	(0.01–0.28)	(−0.18,0.36)	(2–14)
Limited in what you can do	1	0.15	0.03	13
	(n = 117)	(0.02–0.27)	(−0.10,0.17)	(3–15)
	2	0.09	0.00	14
	(n = 109)	(−0.02–0.21)	(−0.13,0.14)	(2–14)
	3	0.15	0.18	12
	(n = 104)	(0.03–0.29)	(−0.04, 0.40)	(2–14)
Dependent on others	1	0.17	0.11	13
	(n = 117)	(0.07–0.27)	(−0.01,0.24)	(3–15)
	2	0.20	0.13	12
	(n = 109)	(0.10–0.30)	(0.03,0.23)	(3–14)
	3	0.21	0.10	11
	(n = 104)	(0.09–0.33)	(−0.05,0.26)	(2–14)

Downloaded from https://academic.oup.com/ageing/advance-article-abstract/doi/10.1093/ageing/afy191/5274649 by University of Nottingham user on 09 January 2019

dementia suggested that patients and their carer proxies applied different constructs and were influenced by different baseline variables, when providing quality of life ratings [30]. Our work here, which questions the utility of EQ-5D-5L and HowRU in care home residents with more advanced cognitive impairment, underlines the inability of current health-related quality of life indices to accurately inform research and practice in this group. Further work is needed.

In conclusion, we recommend that staff proxy responses for EQ-5D-5L are treated with caution in care home studies. Staff responses for HowRU are not a good proxy for resident responses and it is difficult to envisage a scenario in which they would be useful.

Acknowledgements: The authors would like to acknowledge the broader PEACH study team: Mr Zimran Alam, Ms Anita Astle, Professor Tony Avery, Dr Jaydip Banerjee, Professor Clive Bowman, Dr Neil H Chadborn, Mr Michael Crossley, Dr Reena Devi, Professor Claire Goodman, Professor Pip Logan, Professor Finbarr Martin, Professor Julienne Meyer, Professor Dominick Shaw, Professor David Stott and Dr Maria Zubair.

Declaration of Conflicts of Interest: None.

Declaration of Sources of Funding: This work has been conducted as a part of the Proactive Healthcare of Older People in Care Homes (PEACH) study, supported by the Dunhill Medical Trust, award number FOP1/0115.

Ethics

This study is a part of preparatory work for the larger Proactive Healthcare of Older People in Care Homes (PEACH) study. PEACH was reviewed by both the UK Health Research Authority and the University of Nottingham Research Ethics Committee and determined by both to be a service development and evaluation project (London Bromley REC ref: 205840; University of Nottingham ref: LT07092016). The PEACH study protocol has been reviewed as a part of good governance by the Nottinghamshire Healthcare Foundation Trust.

References

1. Sanford AM, Orrell M, Tolson D *et al.* An international definition for 'nursing home'. J Am Med Dir Assoc 2015; 16: 181–4.
2. Laing & Buisson. Care of the Elderly People Market Survey, 2009. London: Laing and Buisson, 2010.
3. Gordon AL, Franklin M, Bradshaw L, Logan P, Elliott R, Gladman JR. Health status of UK care home residents: a cohort study. Age Ageing 2014; 43: 97–103.
4. Kinley J, Hockley J, Stone L *et al.* The provision of care for residents dying in UK nursing care homes. Age Ageing 2014; 43: 375–9.
5. National Institute of Health Research Dissemination Centre. Advancing Care: Research within Care Homes. Available at https://www.dc.nihr.ac.uk/themed-reviews/advancing-care-themed-review.pdf (20 July 2018, date last accessed).
6. Janssen MF, Pickard AS, Golicki D *et al.* Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res 2013; 22: 1717–27.
7. Borowiak E, Kostka T. Predictors of quality of life in older people living at home and in institutions. Aging Clin Exp Res 2004; 16: 212–20.
8. Devlin NJ, Brooks R. EQ-5D and the EuroQol group: past, present and future. Appl Health Econ Health Policy 2017; 15: 127–37.
9. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. Health Econ 2018; 27: 7–22.
10. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. Pharmacoeconomics 2008; 26: 733–44.
11. Kane RL, Kane RA, Bershadsky B *et al.* Proxy sources for information on nursing home residents' quality of life. J Gerontol B Psychol Sci Soc Sci 2005; 60: S318–25.
12. Bulamu NB, Kaambwa B, Ratcliffe J. A systematic review of instruments for measuring outcomes in economic evaluation within aged care. Health Qual Life Outcomes 2015; 13: 179.
13. Benson T, Potts HWW, Whatling JM, Patterson D. Comparison of howRU and EQ-5D measures of health-related quality of life in an outpatient clinic. Inform Prim Care 2013; 21: 12–7.
14. Benson T, Williams DH, Potts HWW. Performance of EQ-5D, howRu and Oxford hip & knee scores in assessing the outcome of hip and knee replacements. BMC Health Serv Res 2016; 16: 512.
15. Usman A, Lewis S, Hinsliff-Smith K *et al.* Measuring health related quality of life of care home residents, comparison of self-report with staff proxy responses for EQ-5D-5L and HowRu: protocol for assessing proxy reliability in care home outcome testing. BMJ Open 2018; 8: e022127.
16. Welsh TJ, Gordon AL, Gladman JR. Comprehensive geriatric assessment—a guide for the non-specialist. Int J Clin Pract 2014; 68: 290–3.
17. van Hout B, Janssen MF, Feng Y-S *et al.* Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health 2012; 15: 708–15.
18. Vanbelle S, Albert A. A note on the linearly weighted kappa coefficient for ordinal scales. Stat Methodol 2009; 6: 157–63.
19. Brenner H, Kliebsch U. Dependence of weighted kappa coefficients on the number of categories. Epidemiology 1996; 7: 199–202.
20. Cicchetti DV, Allison T. A new procedure for assessing reliability of scoring EEG sleep recordings. Am J EEG Technol 1971; 11: 101–10.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159–74.
22. Harwell MR, Rubinstein EN, Hayes WS, Olds CC. Summarizing Monte Carlo results in methodological research: the one- and two-factor fixed effects ANOVA cases. J Educ Stat 1992; 17: 315–39.
23. Devine A, Taylor SJC, Spencer A, Diaz-Ordaz K, Eldridge S, Underwood M. The agreement between proxy and self-completed EQ-5D for care home residents was better for index scores than individual domains. J Clin Epidemiol 2014; 67: 1035–43.

Downloaded from https://academic.oup.com/ageing/advance-article-abstract/doi/10.1093/ageing/afy191/5274649 by University of Nottingham user on 09 January 2019

24. Livingston G, Barber J, Marston L *et al*. Prevalence of and associations with agitation in residents with dementia living in care homes: MARQUE cross-sectional study. BJPsych Open 2017; 3: 171–8. doi:10.1192/bjpo.bp.117.005181.
25. Selai C, Pickard S, Johnston JA. The collection of EuroQol EQ-5D data via proxies. EQ Plenary Meet York Sept 2002. Available at: http//www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Selai_Proxy_Guidance.pdf (June 27 2011, date last accessed) 2002.
26. Ettema TP, Dröes R-M, de Lange J, Mellenbergh GJ, Ribbe WM. QUALIDEM: development and evaluation of a dementia specific quality of life instrument—validation. Int J Geriatr Psychiatry 2007; 22: 424–30.
27. Chua K-C, Brown A, Little R *et al*. Quality-of-life assessment in dementia: the use of DEMQOL and DEMQOL-Proxy total scores. Qual Life Res 2016; 25: 3107–18.
28. Logsdon RG, Gibbons LE, McCurry SM, Teri L. Assessing quality of life in older adults with cognitive impairment. Psychosom Med 2002; 64: 510–9.
29. Aspden T, Bradshaw SA, Playford ED, Riazi A. Quality-of-life measures for use within care homes: a systematic review of their measurement properties. Age Ageing 2014; 43: 596–603.
30. Sheehan BD, Lall R, Stinton C *et al*. Patient and proxy measurement of quality of life among general hospital in-patients with dementia. Aging Ment Health 2012; 16: 603–7.

Received 27 July 2018; editorial decision 9 November 2018

Downloaded from https://academic.oup.com/ageing/advance-article-abstract/doi/10.1093/ageing/afy191/5274649 by University of Nottingham user on 09 January 2019