

1 Measuring health related quality of life of care home residents,  
2 comparison of self-report with staff proxy responses for EQ-5D-  
3 5L and HowRu: Protocol for Assessing Proxy Reliability In Care  
4 home Outcome Testing

5  
6 Adeela Usman <sup>1</sup>

7 Sarah Lewis <sup>2</sup>

8 Kathryn Hinsliff-Smith <sup>1</sup>

9 Annabelle Long <sup>1</sup>

10 Gemma Housley <sup>3</sup>

11 Jake Jordan <sup>4</sup>

12 Heather Gage <sup>4</sup>

13 Tom Dening <sup>5</sup>

14 John R F Gladman <sup>6,7</sup>

15 Adam L Gordon <sup>1,8</sup>

16 1. Division of Medical Sciences and Graduate Entry Medicine, School of Medicine,  
17 University of Nottingham, UK.

18 2. Division of Epidemiology and Public Health, School of Medicine, University of  
19 Nottingham, UK.

20 3. East Midlands East Midlands Academic Health Science Network, Nottingham, UK.

21 4. School of Economics, University of Surrey, UK.

22 5. Institute of Mental Health, University of Nottingham, UK.

23 6. Division of Rehabilitation and Ageing, University of Nottingham, UK

24 7. East Midlands Collaboration for Leadership in Applied Health Research and Care, UK

25 8. School of Health Sciences, City, University of London, UK.

26

27 Corresponding Author: Adam Gordon, Division of Medical Sciences and Graduate  
28 Entry Medicine, University of Nottingham, Room 4113, Derby Medical School, Derby.  
29 DE22 3NE.

30 Email: [adam.gordon@nottingham.ac.uk](mailto:adam.gordon@nottingham.ac.uk); Tel: 01332 724668; Fax: 01332 724697

31

32

33 Word Count: 2793 words

34

35

36

37

38 **Abstract**

39 **Introduction**

40 Research into interventions to improve health and wellbeing for older people living in  
41 care homes is increasingly common. Health-Related Quality of Life (HRQoL) is  
42 frequently used as an outcome measure but collecting both self-reported and proxy  
43 HRQoL measures is challenging in this setting. This study will investigate the reliability  
44 of UK care home staff as proxy respondents for the EQ-5D-5L and HowRu measures.  
45

46 **Methods and Analysis**

47 This is a prospective cohort study of a sub-population of care home residents recruited  
48 to the larger **Proactive Healthcare for Older People in Care Homes (PEACH)** study. It  
49 will recruit residents  $\geq 60$  years across 24 care homes and not receiving short stay or  
50 respite care. The sample size is 160 participants. Resident and care home staff proxy  
51 EQ-5D-5L and HowRu responses will be collected monthly for three months. Weighted  
52 kappa statistics and intraclass correlation adjusted for clustering at the care home  
53 level will be used to measure agreement between resident and proxy responses. The  
54 extent to which staff variables (gender, age group, length of time caring, role, how well  
55 they know the resident, length of time working in care homes and in specialist  
56 gerontological practice) influence the level of agreement between self-reported and  
57 proxy responses will be considered using a multilevel mixed-effect regression model.  
58

59 **Ethics and Dissemination**

60 The PEACH study protocol was reviewed by the UK Health Research Authority and  
61 University of Nottingham Research Ethics Committee and was determined to be a  
62 service development project. We will publish this study in a peer-reviewed journal with  
63 international readership and disseminate it through relevant national stakeholder  
64 networks and specialist societies.  
65  
66  
67

68

69

70

71

72

73

74

## 75 Introduction

76 Long term care facilities in the UK are called care homes, and are classified as either  
77 care homes with or without nursing based upon the availability of registered nurses  
78 on-site (1). The types of residents cared for in both classifications of facility are similar  
79 and all UK care homes are included in the international consensus definition of a  
80 nursing home (2). Around 425,000 people live in care homes in the UK (3) with most  
81 residents requiring care due to disability from long-term conditions. The majority of  
82 residents are aged over 85, 75-80% of residents live with dementia (4), and over half  
83 of residents die within 12 months of admission to care home (5).

84

85 Improving the quality of care for older people in long-term care has become a focus of  
86 attention for health and social care providers, both within the UK and internationally  
87 (2), and an increasing number of evaluative research studies are being published that  
88 test the effectiveness and cost-effectiveness of interventions in this setting. Residents'  
89 quality of life (QoL) is frequently used as an outcome measure in these studies both  
90 to maintain a patient-centred focus and to facilitate health economic evaluation. The  
91 prevalent frailty and cognitive impairment in the care home population, however,  
92 means that collecting self-reported QoL measures from residents is challenging. As a  
93 response to this, proxy responses to QoL items have sometimes been used. For these  
94 a consultee, drawn from one of care home staff, or a relative or friend, answers  
95 questions on the resident's behalf. Using proxy respondents can be unreliable in care  
96 home settings. There may be lack of continuity of care home staff contact with  
97 individual residents due to shift working and staff turnover, and family and friends may  
98 not be well placed to judge QoL domains if they visit residents for only short periods  
99 (6).

100

101 The EurQoL questionnaire is a widely-used preference-based health-related quality of  
102 life (HRQoL) measure suitable for use in economic evaluations. The EQ-5D-5L version  
103 measures HRQoL across five domains (mobility, self-care, usual activities, pain,  
104 anxiety/depression) with the scale for each domain ranging from level 1 (no problems)  
105 to level 5 (extreme problems). The responses from the five domains are converted to  
106 QoL index scores (utilities) generated from a given country's general population (7).  
107 These index scores can be used to calculate quality adjusted life years (QALYs), which  
108 are a measure of the person's state of health – one QALY equates to one year in  
109 perfect health. QALYs are calculated using the area under the curve (8) defined by  
110 utility scores at the different assessment points over the study period. The cost per  
111 QALY gained from an intervention when compared to usual care is the chosen cost-  
112 utility measure for determining eligibility for public support of the intervention through  
113 the UK National Health Service (9).

114

115 The scale for the first version of EQ-5D had only three levels (EQ-5D-3L). EQ-5D-3L  
116 has been shown to have good construct validity for self-report (10), and has been used  
117 to measure QoL of older people living in their own homes and in care homes (11). The  
118 5L version was developed subsequently to deal with identified issues with sensitivity  
119 and a ceiling effect on the EQ-5D-3L which limited its ability to discriminate between  
120 health states, particularly in those with higher quality of life (12). There is also an EQ-  
121 5D visual analogue scale (VAS) used to assess overall health status, ranging from 0  
122 (worst imaginable) to 100 (best imaginable). VAS is recognised to have specific  
123 strengths and weaknesses(13) but is recommended to be used routinely alongside

124 the self-classification questionnaire by the EuroQoL group because of its usefulness  
125 in establishing global health status(14).

126  
127 It is recognised that the EQ-5D, in all its forms, is limited by consequence of being a  
128 generic measure that fails to take account of the difference in what constitutes “quality  
129 of life” in a long-term care setting. It doesn’t take account of shifts in emphasis about  
130 what constitutes wellbeing as residents enter long-term care, which means that social  
131 care related quality of life (SCRQoL) measures such as the Adult Social Care  
132 Outcomes Toolkit (ASCOT) may be preferable in this setting (15-16). A further critique  
133 has been that it fails to separate capability (what a resident is able to do) from  
134 preference (what a resident chooses to do under the circumstances), with the result  
135 that some authors have championed capability-based outcome measures, such as the  
136 ICEpop Capability Measure for Older People (ICECAP-O), in care homes (17-18).  
137 Best practice suggests that, if EQ-5D is used in this setting, it is used in combination  
138 with more specific instruments.

139  
140 The R-outcome tool howRu has been specifically designed for use in long term care  
141 settings in order to address quality of life in a straightforward way that is practical with  
142 older people. HowRu is a Patient Recorded Outcome Measure (PROM) that records  
143 four variables (pain or discomfort, feeling low or worried, limitation in activities, and  
144 dependency on others) related to quality of life at a fixed point in time (“How are you  
145 doing today?”) on a four-point scale (none, slight, quite a lot, extreme) (19, 20). The  
146 HowRu score is calculated by summing up the values for each domain to give a value  
147 on a 13-point scale ranging from 0 (worst) to 12 (best). The HowRu PROM was  
148 designed with older adults in mind (19,20), and may have a cogency and immediacy  
149 that improves upon some of the measurement uncertainty introduced by the relative  
150 abstraction of the questions included in highly validated general population indices  
151 such as EQ-5D-5L.

152  
153 In a comparison with EQ-5D in patients attending a cardiovascular outpatient clinic,  
154 HowRu was reported to have better readability, higher completion rate, wider range of  
155 states used and smaller ceiling effect(17). No national tariffs exist to enable calculation  
156 of HowRu indices that would facilitate its use as a preference-based measure in cost-  
157 utility analysis. Understanding how and whether R outcomes correlate with EQ-5D  
158 scores in the care home setting is, however, helpful when considering additional  
159 information that can help triangulate our understanding of how interventions affect  
160 quality of life in this context. Knowing that HowRu correlates with EQ-5D may provide  
161 further justification for using it in clinical settings where broad judgements about impact  
162 on quality of life have to be made without the need for detailed cost-utility analysis.

163  
164 Proxies have been used to capture EQ-5D-3L responses from people with dementia,  
165 although poor agreement between patient and proxy ratings has raised concerns  
166 (15,16) and differences between professional and family carer ratings have led to  
167 questions about the appropriate choice of proxy (16). In a comparison of clinicians and  
168 family carers as proxies, clinician responses had better construct validity in the more  
169 observable domains of mobility and self-care, and family carer responses had better  
170 construct validity in the less observable domains of usual activities and  
171 anxiety/depression (16). There is limited evidence, however, comparing self-reported  
172 and proxy responses to the EQ-5D-5L in care home populations (17). There is, in  
173 particular, a paucity of data as to how it performs in UK care home populations. This

174 is important because institutional care is structured differently between nations, with  
175 differing professional carer competencies, patterns of working and job roles. This  
176 means that carers in different countries will have differential exposure to residents and  
177 different competencies in terms of their ability to interpret residents' experiences, and  
178 a tool that works for professional proxy response in the US may not, therefore, work  
179 as well in the UK.

180  
181 The HowRu, as a recently developed PROM, is yet to be fully evaluated for older  
182 people living in care homes (21). It is not known whether proxy responses in this  
183 setting may be useful in completing HowRU and there are no data on how HowRu  
184 correlates with EQ-5D in the care home populations.

185 This study seeks to fill these evidence gaps.

186

## 187 Aim

188 The Assessing Proxy Reliability In Care home Outcome Testing (APRICOT) study is  
189 a preparatory piece of work for the Proactive Healthcare for Older People Living in  
190 Care Homes (PEACH) study. It aims to examine the level of agreement between the  
191 responses to EQ-5D-5L and HowRu by care home staff and residents, and between  
192 EQ-5D-5L and HowRu as quality of life measures. Findings will assist in the  
193 interpretation of quality of life data gathered for the larger PEACH study.

194

## 195 Objectives

196 To determine the level of agreement between:

- 197 • Resident EQ-5D-5L and staff-proxy EQ-5D-5L responses
- 198 • Resident HowRu and staff-proxy HowRu responses
- 199 • Resident EQ-5D-5L and HowRu responses
- 200 • Proxy EQ-5D-5L and HowRu responses

## 201 Methods

### 202 Setting

203 24 care homes in the East Midlands area of England. These are long-term care  
204 institutions, housing predominantly older people with frailty who can no longer be  
205 cared for at home. Detailed descriptions of the UK care home sector and the residents  
206 living within it have been published elsewhere (4).

207

### 208 Brief description of the Proactive Healthcare in Older Peoples' Care Homes (PEACH) 209 study

210 The PEACH intervention involves using Quality Improvement Collaboratives as a  
211 mechanism to encourage implementation of Comprehensive Geriatric Assessment  
212 (CGA) as a unifying framework for assessment and delivery of healthcare in UK care  
213 homes. CGA is widely recognised as a gold-standard way to deliver care for older  
214 people with frailty (22). The aim of PEACH is to improve quality of care and quality of  
215 life for care home residents. Outcome data quantifying healthcare resource use and  
216 resident level quality of life will be collected on a monthly basis to assess the impact  
217 of the intervention.

218

219 Two instruments are being used in PEACH to assess residents' quality of life, the EQ-  
220 5D-5L and HowRu. The rationale is that these reflect measurable differences in the

221 patient experience that may translate, with some interpretation, into an understanding  
222 of how CGA influences quality of care and general wellbeing. APRICOT has been  
223 designed as a preparatory sub-study within PEACH to better enable interpretation of  
224 proxy EQ-5D-5L and HowRU responses collected as part of outcome measurement.

225

## 226 Participants

227 Care home recruitment for PEACH took place between October 2016 and January  
228 2017, with individual resident recruitment from January 2017. A prospective cohort of  
229 a sub-population of residents will be included in the comparison of proxy and self-  
230 report measurement of EQ-5D-5L and HowRu in APRICOT. Residents included in the  
231 study will be those  $\geq 60$  years across 24 care homes and not receiving short stay or  
232 respite care. To have a better reflection of self-reported and proxy agreement in a care  
233 home setting we will include residents with and without mental capacity. Care home  
234 managers will provide guidance with regards to residents with and without capacity to  
235 participate. When managers are unsure, researchers will make judgements based on  
236 the framework for mental capacity outlined in the 2005 Mental Capacity Act for  
237 England and Wales (23) and in keeping with the recommendations of that Act for  
238 inclusion research, for residents that lack capacity to provide consent to participation  
239 an appropriate person will be consulted to make a decision with regards to  
240 participation in the study.

241

242 This study will be conducted in parallel to the main PEACH study. In addition to the  
243 routine collection of EQ-5D-5L and HowRu from residents recruited to PEACH, proxy  
244 responses to EQ-5D-5L and HowRu will be gathered from staff. We will include staff  
245 such as care home assistants, care home manager and registered nurses, who  
246 know the resident well. We will exclude staff employed in a supportive role, such as  
247 activity coordinators, since their orientation to supporting residents is more variable.

248

## 249 Data collection

250 Data from proxies will be collected on three consecutive months. Due to the high staff  
251 turnover amongst care home staff, and to enable the influence of carer characteristics  
252 on the level of agreement to be estimated, data on the carer will be gathered at each  
253 assessment. Repeated measures are required for the final analysis in the PEACH  
254 study for calculating costs per QALY gained (comparing the intervention with usual  
255 care condition) and understanding how the agreement changes at different time points  
256 is therefore of interest.

257

258 Staff proxies will be asked to consider the proxy-resident's perspective when  
259 completing the questionnaire using the following statement: "*Please rate how you  
260 (staff) think the resident will rate his/her own health-related quality of life, if the resident  
261 was to communicate*" (24). Both self-reported and proxy EQ-5D will be completed on  
262 the same day to minimize any variations in responses.

263

264 The EQ-5D VAS will be used in the study in keeping with the recommendations of the  
265 EuroQoL group.

266

## 267 Primary analysis

268 An overall agreement between the self-reported and staff proxy responses on the  
269 domain levels of the EQ-5D-5L and HowRu will be computed. Weighted kappa statistic  
270 and Intra-class correlation (ICC) will be used to calculate the level of agreement for  
271 categorical and continuous outcomes respectively. All reliability indices will be  
272 calculated at the domain levels and overall index scores/QALYs for the EQ-5D-5L. To  
273 calculate the EQ-5D-5L index scores, responses from the descriptive system will be  
274 transformed into index scores derived from the UK general population. This will be  
275 done using the crosswalk value set (25), which maps the 5L descriptive system data  
276 onto the 3L valuation.

277  
278 Weighted kappa helps to distinguish between small and large difference in agreement  
279 ratings assigned to the different levels of each domain but equal importance is given  
280 to disagreement (26, 27). The weighting for kappa will be done using linear weight -  
281 this assigns the same importance to the difference between any two categories within  
282 the response scale (28). The confidence interval for the weighted kappa will be  
283 calculated by bootstrapping in Stata 15 (Statcorp, LLC, 2015) with 1000 replications.  
284 This will be done for each time point.

285  
286 The kappa statistic ranges from -1 to 1, and the strength of the agreement will be  
287 interpreted with regards to published guidelines (29) with agreement being:

- 288 • Poor, if  $\text{kappa} \leq 0.00$
- 289 • Slight, if  $\text{kappa} = 0.00$  to  $0.20$
- 290 • Fair, if  $\text{kappa} = 0.21$  to  $0.40$
- 291 • Moderate, if  $\text{kappa} = 0.41$  to  $0.60$
- 292 • Substantial, if  $\text{kappa} = 0.61$  to  $0.80$
- 293 • Almost perfect, if  $\text{kappa} \geq 0.80$

294 Unadjusted ICC will be calculated using two-way mixed effect ANOVA model to  
295 examine the level of agreement between the self-reported and proxy responses for  
296 the EQ-5D-VAS, EQ-5D index scores and total QALYs. ANOVA models are reported  
297 to be robust to deviations in normality and have been used in other quality of life  
298 agreement studies (16,30).The Bland-Altman graph (plotting the mean difference  
299 between the EQ-5D-5L-S (self-report) and EQ-5D-5L-P (proxy) against the mean of  
300 the two measures) will be constructed to supplement the ICC .

301  
302 Analysis will be done at each time point for Kappa and ICC. However, a single ICC  
303 value for QALYs will be calculated for individuals with data on all three consecutive  
304 months as this will be used in practice in the PEACH study, where analysis will be  
305 done on consecutive measures made over time.

306  
307 To allow for comparability of the level of agreement at the domain and index score  
308 level; the same benchmarks used for kappa will be used for the ICC.

309  
310 **Clustering**

311 Clustering will be adjusted for because the calculation of kappa and ICC assumes  
312 independence of observations. In our study, clustering could occur at three levels at  
313 each time point. Firstly, at the care home level, where residents within the same care  
314 home have similar characteristics and are different from those in other care homes.  
315 Secondly, at the individual level, where responses are clustered within each resident

316 and lastly, at staff level, where staff members within a care home responds on behalf  
317 of multiple residents.

318  
319 The ICC value will be calculated allowing for clustering using a nested two-way  
320 mixed effect model calculated by fitting a 2 level random effect model with a random  
321 effect for care home and individuals.

322  
323 A cluster adjusted kappa will be calculated using a variance formula. This will include  
324 calculating kappa and its variance for each care home, then estimating the within  
325 cluster variance  $\sigma_{\omega}^2$  using the average of the individual variances, and between cluster  
326 variance  $\sigma_b^2$  using the variance of the individual kappa. Using these estimates the  
327 clustered Kappa and its variance will be calculated using the formulas below:  
328

$$\omega_i = \frac{n_i}{(1+(n_i-1)\rho_k)}$$

329

330 Where

331  $n_i$  = Size of cluster  $i$

332  $\rho$  = Intra-cluster correlation coefficient for Kappa

$$\rho_k = \frac{\sigma_b^2}{(\sigma_{\omega}^2 + \sigma_b^2)}$$

333

334

335 Variance of the cluster adjusted kappa will be obtained using the equation below  
336 (24):

$$\sigma_{clustered\ k}^2 = \frac{\sum_{i=1}^k \omega_i^2 \left( \sigma_b^2 + \frac{\sigma_{\omega}^2}{n_i} \right)}{\left( \sum_{i=1}^k \omega_i \right)^2}$$

337

338

339 The clustered kappa and its variance will then be divided by the square root of the  
340 number of individuals to get the standard error. The 95% confidence interval will be  
341 calculated using this standard error.

342

343 Cluster adjusted kappa (EQ-5D-5L and HowRu domain levels) and ICC (EQ-VAS and  
344 EQ-5D index scores) will be reported for each time point. However, QALYs will be  
345 presented over time, three months, similar to how it will be calculated in the PEACH  
346 study.

347

348 Sample size calculation

349 We need a sample size of 160 residents assuming a kappa of 0.145 and a confidence  
350 level width of 0.153 derived from a previous study (16), given that 50% of residents  
351 will have any problems.

352

353 Secondary analyses

354 The effect of age, sex and length of stay in care home (for residents), length time  
355 working in care of older people/care homes and role/rank (for staff) at baseline on the  
356 difference between staff and proxy EQ-5D-5L-S scores will be analysed using a  
357 multilevel mixed effect regression model.



358  
 359 To investigate the reliability of using HowRu as a quality of life measure in the care  
 360 home population compared with EQ-5D-5L, we will assess agreement between these  
 361 indices using weighted kappa statistics. This will involve testing the level of agreement  
 362 between domains with similar construct on both scales (31) as shown in table 1.

363  
 364 *Table 1 - Planned pair-wise alignment of HowRU and EQ-5D-5L domains for agreement analysis using kappa*

HowRu domains	EQ-5D-5L domains
Pain or discomfort	Pain/discomfort
Feeling low or worried	Anxiety/depression
Limited in what you can do	Mobility Self-care Usual activities
Dependent on others	Mobility Self-care Usual activities

365  
 366 **Patient and Public Involvement**  
 367 The APRICOT and PEACH studies were developed and designed in discussion with  
 368 both Care Home Sector and Patient and Public Involvement (PPI) representatives.  
 369 The initial research proposal and protocol was presented, prior to submission for  
 370 funding, to the Dementia and Frail Older Person’s PPI group hosted in the Division of  
 371 Rehabilitation and Ageing, University of Nottingham. Amendments were made to the  
 372 proposal and protocol in light of their feedback. The PEACH study team includes  
 373 one PPI and two care home sector representatives who are present at all study  
 374 management meetings, with oversight for the APRICOT sub-study. We keep all  
 375 participating care homes working with PEACH updated through quarterly newsletters  
 376 which will include dissemination of our findings in lay terms as these become  
 377 available.

378  
 379 **Strengths and Limitations of this study**

- 380 • To our knowledge, this is the first study to investigate the reliability of staff  
 381 proxy using the EQ-5D-5L and HowRu, and the reliability of HowRu as a  
 382 quality of life measure in a UK care home setting.
- 383 • The scales will be administered to residents and staff on the same day, thus  
 384 reducing any variations that will affect the level of agreement.
- 385 • Analysing the characteristics of staff respondents will help generate insights  
 386 into how proxies might better be selected or prepared to maximise the validity  
 387 of proxy responses.
- 388 • The use of multiple time points will help increase the sample size and account  
 389 for the influence of staff turnover on the quality of proxy-response.
- 390 • Working with residents with mental capacity will optimise their responses as a  
 391 gold standard against which proxies can be compared. The way in which staff  
 392 interpret quality of life in this more communicative and able group may,  
 393 however, be different to the way in which the interpret quality of life in  
 394 advanced dementia.

396 Ethics and dissemination

397 This study is part of preparatory work for the larger PEACH study. PEACH was  
398 reviewed by both the UK Health Research Authority and the University of  
399 Nottingham Research Ethics Committee and determined by both to be a service  
400 development and evaluation project. We will obtain informed consent from residents  
401 who have mental capacity and use a consultee when residents lack mental capacity.  
402 The PEACH study protocol has been reviewed as part of good governance by the  
403 Nottinghamshire Healthcare Foundation Trust. We aim to publish this study in a  
404 peer-reviewed journal with international readership and disseminate it further using  
405 relevant national stakeholder networks and specialist societies.

406

407 Funding

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

411

412 Contributors

413 All authors meet the ICJME criteria for authorship. AU, SL, KHS, AL, GH, JJ, HG,  
414 TD, JRFG and ALG conceived the study at a PEACH study management meeting  
415 and specified the aims and objectives. AU, SL and ALG produced the initial draft of  
416 the protocol, with KHS, AL, GH, JJ, HG, TD and JRFG contributing to subsequent  
417 redrafts. AU and SL led on aspects of statistical design. JJ and HG provided  
418 specialist health economics input. AU, SL, KHS, AL, GH, JJ, HG, TD, JRFG and  
419 ALG all reviewed the final manuscript and approved it prior to submission.

420 Competing interests

421 None declared.

422

423 Acknowledgements

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

429 References

- 430 1. Aspden T, Bradshaw SA, Playford ED, Riazi A. Quality-of-life measures for use  
431 within care homes: a systematic review of their measurement properties. *Age and*  
432 *Ageing*. 2014; 43(5):596-603.
- 433 2. Sanford AM, Orrell M, Tolson D, Abbatecola AM, Arai H, Bauer JM, et al. An  
434 International Definition for "Nursing Home". *Journal of the American Medical*  
435 *Directors Association*. 2015;16(3):181-184.
- 436 3. Laing & Buisson. *Care of the Elderly People Market Survey*. London: Laing and  
437 Buisson 2010; 2009.
- 438 4. Gordon AL, Franklin M, Bradshaw L, Logan P, Elliott R, Gladman JRF. Health  
439 status of UK care home residents: a cohort study. *Age and Ageing*.  
440 2014;43(1):97-103.
- 441 5. Kinley J, Hockley J, Stone L, Dewey M, Hansford P, Stewart R, et al. The  
442 provision of care for residents dying in UK nursing care homes. *Age and Ageing*.  
443 2014;43(3):375-379.

- 444 6. Kane RL, Kane RA, Bershadsky B, Degenholtz H, Kling K, Totten A, et al. Proxy  
445 sources for information on nursing home residents' quality of life. *The journals of*  
446 *gerontology Series B, Psychological sciences and social sciences.*  
447 2005;60(6):S318-s325.
- 448 7. Devlin N, Shah K, Feng Y, Mulhern B, van Hout B. Valuing health-related quality  
449 of Life: An EQ-5D-5L Value Set for England. *Health Economics.* 2017; 1-16.
- 450 8. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and  
451 utilities. *Br Med Bull.* 2010;96(1):5–21.
- 452 9. McCabe C, Claxton K, Culyer AJ. The NICE Cost-Effectiveness Threshold.  
453 *Pharmacoeconomics.* 2008;26(9):733-744.
- 454 10. Janssen MF, Simon Pickard A, Golicki D, Gudex C, Niewada M et al.  
455 Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across  
456 eight patient groups: a multi-country study. *Qual Life Res.* 2013; 22(7): 1717-27
- 457 11. Borowiak E, Kostka T. Predictors of quality of life in older people living at home  
458 and in institutions. *Aging Clinical and Experimental Research.* 2004;16(3):212-  
459 220.
- 460 12. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future.  
461 *Appl Health Econ Health Policy.* 2017;15(2):127–37.
- 462 13. Parkin D, Devlin N. Is there a case for using visual analogue scale valuations in  
463 cost-utility analysis? *Health Econ [Internet].* 2006 Jul 1 [cited 2018 Apr  
464 25];15(7):653–64. Available from: <http://doi.wiley.com/10.1002/hec.1086>
- 465 14. Feng Y, Parkin D, Devlin NJ. Assessing the performance of the EQ-VAS in the  
466 NHS PROMs programme. *Qual Life Res.* 2014; 23(3):977–89.
- 467 15. Hounsome N, Orrell M, Edwards RT. EQ-5D as a Quality of Life Measure in  
468 People with Dementia and Their Carers: Evidence and Key Issues. *Value in*  
469 *Health.* 2011;14(2):390-399.
- 470 16. Devine A, Taylor SJC, Spencer A, Diaz-Ordaz K, Eldridge S, Underwood M. The  
471 agreement between proxy and self-completed EQ-5D for care home residents  
472 was better for index scores than individual domains. *Journal of Clinical*  
473 *Epidemiology.* 2014;67(9):1035-1043.
- 474 17. Bulamu NB, Kaambwa B, Ratcliffe J. A systematic review of instruments for  
475 measuring outcomes in economic evaluation within aged care. *Health and Quality*  
476 *of Life Outcomes.* 2015;13(1).
- 477 18. van Leeuwen KM, Bosmans JE, Jansen AP, Hoogendijk EO, van Tulder MW, van  
478 der Horst HE, et al. Comparing Measurement Properties of the EQ-5D-3L,  
479 ICECAP-O, and ASCOT in Frail Older Adults. *Value in health : the Journal of the*  
480 *International Society for Pharmacoeconomics and Outcomes Research.*  
481 2015;18(1):35-43.
- 482 19. Benson T, Williams DH, Potts HWW. Performance of EQ-5D, howRu and Oxford  
483 hip & knee scores in assessing the outcome of hip and knee replacements. *BMC*  
484 *Health Services Research.* 2016;16(1).
- 485 20. Benson T, Potts HWW, Whatling JM, Patterson D. Comparison of howRU and  
486 EQ-5D measures of health-related quality of life in an outpatient clinic. *Inform*  
487 *Prim Care.* 2013;21(1):12-17.
- 488 21. Hendriks SH, Rutgers J, van Dijk PR, Groenier KH, Bilo HJG, Kleefstra N, et al.  
489 Validation of the howRu and howRwe questionnaires at the individual patient  
490 level. *BMC Health Services Research.* 2015;15(1).
- 491 22. Welsh TJ, Gordon AL, Gladman JR. Comprehensive geriatric assessment - a  
492 guide for the non-specialist. *Int J Clin Pract.* 2014. 68(3): 290–293
- 493 23. Dobson C. Conducting research with people not having the capacity to consent to  
494 their participation. A practical guide for researchers [Internet]. Leicester, UK: The  
495 British Psychological Society; 2017 [cited 1 December 2017]. Available from:  
496 [https://pdfs.semanticscholar.org/f97c/5a3d8605e24eaaf965b64e42c690d42a618](https://pdfs.semanticscholar.org/f97c/5a3d8605e24eaaf965b64e42c690d42a6182.pdf)  
497 [2.pdf](https://pdfs.semanticscholar.org/f97c/5a3d8605e24eaaf965b64e42c690d42a6182.pdf)

- 498 24. Selai C, Pickard S, Johnston JA. The collection of EuroQol EQ-5D data via  
499 proxies. EQ Plenary Meeting in York: September, 2002
- 500 25. van Hout B, Janssen MF, Feng Y-S, Kohlmann T, Busschbach J, Golicki D, et al.  
501 Interim Scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L Value  
502 Sets. *Value Heal.*;15(5):708–15.
- 503 26. Vanbelle S, Albert A. A note on the linearly weighted kappa coefficient for ordinal  
504 scales. *Statistical Methodology*. 2009; 6(2):157-163.
- 505 27. Brenner H, Kliebsch U. Dependence of Weighted Kappa Coefficients on the  
506 Number of Categories. *Epidemiology*. 1996;7(2):199-202.
- 507 28. Cicchetti DV, Allison T. A New Procedure for Assessing Reliability of Scoring  
508 EEG Sleep Recordings. *American Journal of EEG Technology*. 1971;11(3):101-  
509 110.
- 510 29. Landis JR, Koch GG. The measurement of observer agreement for categorical  
511 data. *Biometrics*. 1977;33(1):159.
- 512 30. Harwell MR, Rubinstein EN, Hayes WS, Olds CC. Summarizing Monte Carlo  
513 Results in Methodological Research: The One- and Two-Factor Fixed Effects  
514 ANOVA Cases. *J Educ Stat*. 1992 Dec 23;17(4):315–39.
- 515 31. Kerry SM, Bland JM. Unequal cluster sizes for trials in English and Welsh general  
516 practice: implications for sample size calculations. *Statistics in Medicine*.  
517 2001;20(3):377-

