Evaluating a falls prevention intervention in older home care recipients: a comparison of SF-6D and EQ-5D

Maria Bjerk[1] · Therese Brovold[1] · Jennifer C. Davis[2] · Astrid Bergland[1]

Abstract

Purpose Health-related quality of life (HRQOL) is an important outcome in economic evaluations of health care interventions for older adults. The aim of this study was to compare two commonly used preference-based utility measures, SF-6D and EQ-5D, to provide knowledge on their applicability when evaluating falls prevention interventions in primary health care.

Methods The study is a secondary analysis of longitudinal data from a randomised controlled trial, which included 155 older home care recipients participating in a falls prevention intervention in Norway. HRQOL was measured by SF-6D and EQ-5D. Physical function was measured by Berg Balance Scale, 4-m walk test, 30-s sit-to-stand and Falls Efficacy Scale International. Assessments were performed at baseline, 3 months and 6 months. The agreement between SF-6D and EQ-5D was examined using Bland–Altman plots and Spearman correlations. Elasticities from regression analysis were employed to compare the instruments' responsiveness.

Results SF-6D and EQ-5D were strongly correlated (0.71), but there were differences in the instruments' agreement and domains of HRQOL covered. Participants with a higher mean HRQOL and/or better physical function scored generally higher on EQ-5D. Participants with a lower mean HRQOL and/or poorer physical function achieved a relatively higher score on SF-6D. EQ-5D was more responsive to changes in physical function compared to SF-6D.

Conclusions SF-6D and EQ-5D have both similarities and differences regarding sensitivity, domains covered and responsiveness to changes when evaluating a falls prevention intervention. Selecting the appropriate instrument depends on the characteristics of the participants and the intervention being evaluated.

Keywords Health-related quality of life · Home care · Falls prevention · Economic evaluation · SF-6D · SF-36 · EQ-5D

Introduction

Falls in older adults are a leading cause of disability, pain and reduced health-related quality of life (HRQOL), contributing considerably to the global burden of disease [1–3]. Due

✉ Maria Bjerk
maria.bjerk@oslomet.no

Therese Brovold
therese.brovold@oslomet.no

Jennifer C. Davis
jennifer.davis@ubc.ca

Astrid Bergland
astrid.bergland@oslomet.no

[1] Department of Physiotherapy, OsloMet – Oslo Metropolitan University, PO Box 4, St. Olavs Plass, 0130 Oslo, Norway

[2] Faculty of Management, University of British Columbia Okanagan, Kelowna, Canada

to the large consequences of falls, maintaining and improving HRQOL is one of the main goals of falls prevention [4]. Home care recipients have a high incidence of falls and low level of HRQOL compared to the general older population [5, 6]. Effective interventions to prevent falls and to improve HRQOL in older home care recipients are necessary to maintain functional abilities and well-being, contributing to healthy ageing [1, 3], as well as reducing costs [7–9].

HRQOL is an important outcome in evaluations assessing the effects of health care and policy interventions for older adults [10]. Given the rise in number of older adults and the associated increase in health care costs, evaluations are necessary to inform about the cost-effectiveness of interventions [4]. Preferences over health states in older adults can be measured by various instruments which in turn can influence the outcome of the evaluation [4, 10, 11].

The Short Form 36 Health Survey (SF-36) and the EuroQOL EQ-5D (EQ-5D) are two of the most widely used

generic measures of HRQOL, which have shown to be valid and reliable when assessing older adults [12]. In the population of community-dwelling older adults, EQ-5D has been recommended due to its shortness and high response rate, while SF-36 has been recommended for more detailed and broad assessments [12, 13]. Scores from SF-36 have been translated into the preference-based utility index SF-6D, comparable to EQ-5D. Both EQ-5D and SF-6D can be employed to calculate quality adjusted life years (QALYs) in economic evaluations [11]. Comparing EQ-5D and SF-6D in older adults can provide knowledge on how the instruments respond to changes in the underlying health state in this frail group. [13]. This can assist clinicians and researchers when choosing an instrument, and policymakers interpreting recommendations.

Differences and similarities in EQ-5D and SF-6D have been studied previously in various populations. Although both instruments measure HRQOL, differences have been demonstrated, in particular in the lower end of the utility scale [14]. Across various patient groups, floor effects in SF-6D and ceiling effects in EQ-5D have been shown [11, 13, 15]. In the domains pain and mood/mental health, similar utilities have been found, while EQ-5D puts more weight on physical functioning and SF-6D on social functioning [14, 15]. In the general population, EQ-5D seems to be more responsive to chronic physical conditions [16]. This was also evident when comparing the responsiveness of EQ-5D and the age-specific index of capability for older adults (ICE-CAPO) detecting changes over a 12-month period in fallers with mobility impairments [4]. However, in older people with severe chronic obstructive pulmonary disease, SF-6D was more efficient in detecting differences among subgroups in disease severity, but this observation was made at one time-point and not longitudinally [17, 18].

In economic evaluations, EQ-5D has demonstrated larger health gains and lower cost-utility ratios compared to SF-6D [19, 20]. Due to large differences in the incremental cost-utility ratios, the comparability of results from these instruments has been disputed [20, 21]. In patient groups with mild health conditions, the probability of acceptance of the incremental cost-utility ratio was larger when using EQ-5D, while in patient groups with worse health conditions the probability of acceptance was larger when using SF-6D [21]. Hence, when selecting a HRQOL instrument for evaluative purposes, the characteristics of the participants and the content of the intervention is relevant to consider. The group of frail older home care recipients has not been focused on previously.

Health care interventions for older fallers receiving home care are important to maintain their physical function and quality of life and further enable them to remain at home. Evaluations are necessary to develop efficient and high-quality health care services for this group of frail older adults.

By comparing SF-6D and EQ-5D, further knowledge on the instruments' impact and their effect on evaluations can be obtained. This can be important information for health care managers and policy makers conducting evaluations to prioritise between different health care interventions for this population. We therefore conducted a study to examine the agreement between EQ-5D and SF-6D using longitudinal data on older home care recipients. We also assessed differences and similarities in HRQOL domains covered by the instruments and the responsiveness to changes in physical function over time.

Methods

Design

The study is a secondary analysis of data from a randomised controlled trial (RCT). Longitudinal data from three time points, baseline, 3 months and 6 months, was employed. The randomised controlled trial has been described elsewhere [5, 22, 23]. Reporting follows the STROBE guidelines [24].

Participants and setting

This study was conducted in primary health care service in six municipalities in Norway. Participants were recruited from lists of people receiving home care by health professionals. All assessments were performed in the participants' homes by trained research assistants. Participants provided informed written consent before baseline testing.

Recruitment was conducted from February 2016 to February 2017 and follow-up assessments were carried out until September 2017. The sample size was determined based on power calculations for the RCT [23]. The inclusion criteria for the participants were that they received professional home care, either home nursing, practical assistance or safety alarm service. They were 67 years or older and had experienced at least one fall during the previous 12 months. They were able to walk with or without a walking aid and could understand Norwegian. Exclusion criteria were that they had any medical contraindications to exercise or a life expectancy below 1 year. They were also excluded if they had a score below 23 on the Mini Mental State Examination (MMSE), indicating cognitive decline, or if they already were participating in another falls prevention programme or trial.

Outcome measures

HRQOL was measured by two common self-report assessment tools, Short Form 36 Health Survey (SF-36) and EuroQOL EQ-5D (EQ-5D). SF-36 is a generic and validated

questionnaire, which has been translated into Norwegian [25]. It consists of 36 questions on physical functioning, role limitations due to physical functions, role limitations due to emotional problems, bodily pain, general health perception, vitality, social functioning and mental health. Scores for the different items range from 1 to 6, where 1 is good and 6 is worse. Based on SF-36, SF-6D utility indexes can be calculated, with scores ranging from 0.29 to 1 [26]. SF-6D derives from the responses on 11 items of the SF-36 which are combined into six dimensions of health with four to six levels each [18]. The six dimensions are: physical functioning, role participation, social functioning, pain, mental health and vitality. EQ-5D is a generic and validated tool, but briefer, and comprises five domains: mobility, usual activities, pain/discomfort and emotions [27]. Scores for the different domains range from 1 to 5, where 1 is good and 5 is worse. The corresponding utilities in EQ-5D are ranging from -0.59 to 1 [19]. The utility scores for SF-6D and EQ-5D were calculated. The published algorithm with parametric preference weights for United Kingdom was employed as there is no Norwegian value set [27, 28]. Measurements were taken as interviews to increase the response and completion rate [12].

Physical function consisted of measures of balance, usual walking speed and leg muscle strength. Static and dynamic balance was assessed by the Berg Balance Scale (BBS), a 14-item scale measuring performance on a scale from 0 (cannot perform) to 4 (normal performance) [29]. The sum score ranges from 0 to 56, and a score below 45 indicates a high risk of falling. Usual walking speed was based on the time that was required to walk 4 meters in their usual speed (4MWT) and was expressed in meters per second [30]. The participants could use a walking aid if needed. Leg muscle strength was assessed using the functional 30 s sit-to-stand test (STS), which reports the number of rises from a chair within 30 s [31].

Falls self-efficacy was assessed by the Falls Efficacy Scale International (FES-I) [32]. This questionnaire measures concern about falling in 16 different activities in older adults and has been recommended for clinical practice as well as research [33]. Each activity has a four-point scale from 1 (not concerned) to 4 (very concerned), and the sum score ranges from 16 to 64.

Background variables like sex, age, falls history and health care services were collected at baseline. Mini Mental State Examination (MMSE) was performed at baseline to exclude participants with cognitive decline [34].

Statistical analyses

The statistical analyses were conducted using STATA/SE 14. Descriptive statistics on the study population are reported. Mean and standard deviations (SD) were calculated for continuous data and percentages for categorical data. The distribution of SF-6D and EQ-5D scores over domains and levels is presented. Furthermore, for different baseline characteristics a threshold (median) value has been calculated and mean utility scores are reported for the groups below and above this threshold. Complete case analyses were conducted.

Bland–Altman plots illustrate the agreement between EQ-5D and SF-6D. Plots were drawn for absolute values and changes at different time points. In the Bland–Altman plots, a fitted regression line and boundaries of agreement were drawn. Outliers were investigated to check characteristics, but were not excluded as they did not substantially affect the results. To study the associations between the domains of SF-6D and EQ-5D, and physical measures, Spearman rank correlations were calculated as most of the data was ordinal. The strength of correlations was interpreted according to Cohen, where 0.10 to 0.29 is weak, 0.3 to 0.49 is moderate, and 0.5 to 1.0 is strong [35].

To examine the responsiveness of SF-6D and EQ-5D related to measures of physical outcomes, elasticities were calculated from a linear mixed regressions model with individual-specific effects [36]. An elasticity of an outcome variable y with respect to a predictor variable x is calculated such as it equals approximately the proportional change in y for a proportional change in x. The elasticities were calculated at the mean level of x and can conveniently be interpreted as the percentage change in y in response to a one percentage change in x at this level. This removes the unit of measurement and makes responses in different regression models more illustrative and directly comparable.

Results

Descriptive statistics

Eight-hundred and sixty-five older adults receiving home care from the primary health care services in Norway were contacted and checked for eligibility. One hundred and sixty-seven were assessed at baseline and 12 were excluded, ten due to a score lower than 23 on MMSE and two due to ill-health. One hundred and fifty-five participants agreed to participate, gave informed consent and were tested at baseline (T0). At intervention end at 3 months (T1), 138 completed the assessments. At follow-up at 6 months (T2), 136 participants completed the assessments. Missing data at the different time points were due to death, ill-health, declining to participate or other reasons. For EQ-5D complete data was obtained for 155 at T0, 135 at T1 and 135 at T2. For SF-6D complete data was obtained for 155 at T0, 136 at T1 and 133 at T2.

Table 1 shows baseline characteristics of the sample. The participants had a mean age of 82.7 (SD 6.7) and 73.3% were females. They all received home care services. Practical assistance (69.7%) and safety alarm service (75.5%) were most common. The mean number of falls in the previous 12 months was 2.7 (SD 3.7). 36.1% had higher education with more than 12 years of education and the mean score on MMSE was 27.4. The measures of physical function were generally low, with a mean of 39.1 points (SD 11.3) on BBS, a mean of 5.1 raises (SD 4.1) on STS, a mean of 0.6 m/s (SD 0.2) on 4MWT and a mean of 30.7 points (SD 9.8) on FES-I. The mean index score of SF-6D was 0.64 (SD 0.12) and the mean index score of EQ-5D was 0.62 (SD 0.23).

Tables 2 and 3 present the distribution of SF-6D and EQ-5D results at baseline. EQ-5D has a larger proportion of responses in the top category of each dimension compared to SF-6D. Very few responses are within level five of EQ-5D.

Table 1 Baseline characteristics of the sample $N = 155$

Characteristics	
Age, mean (SD)	82.7 (6.7)
Sex, males, %	20.7
Higher education (> 12 years), %	36.1
Falls in the last 12 months, mean (SD)	2.7 (3.7)
Mini Mental State Examination (MMSE), mean (SD)	27.4 (2.2)
Primary health care services	
Practical assistance, %	69.7
Nursing, %	30.3
Safety alarm service, %	75.5
EQ-5D, mean (SD)	
EQ-5D Index score	0.62 (0.23)
SF-6D, mean (SD)	
SF-6D Index score	0.64 (0.12)
Physical measures, mean (SD)	
Berg Balance Scale (BBS)	39.1 (11.3)
30 s sit-to-stand (STS)	5.1 (4.1)
4-m walk test m/s (4MWT)	0.6 (0.2)
Falls Efficacy Scale (FES-I)	30.7 (9.8)

SD Standard deviation

Table 3 Distribution of EQ-5D results at baseline (%) $N = 155$

Level	Mobility	Self-care	Usual activities	Pain/discomfort	Anxiety/depression
1	27.7	60.0	24.5	26.5	51.6
2	29.7	25.2	37.4	23.9	34.8
3	21.9	12.3	27.1	30.3	9.7
4	20.0	2.6	10.3	15.5	3.2
5	0.7	–	0.7	3.9	0.7

Both distributions of SF-6D and EQ-5D scores appear to be reasonable spread across the scales.

Table 4 presents a comparison of SF-6D and EQ-5D utility scores over selected groups of the sample. The table shows that there is a larger spread of values within EQ-5D utility scores compared to SF-6D between those with higher/lower age and better/worse physical function. The exception is on number of falls, where both instruments are more similar. Higher age and better scores on physical measures are associated with higher utility scores of both SF-6D and EQ-5D. Having lower scores on physical measures is associated with relatively higher scores on SF-6D compared to EQ-5D, and contrary.

More information on the flow of participants, on their characteristics at baseline, as well as the results of the RCT are presented elsewhere [5, 23].

Similarities and differences in SF-6D and EQ-5D

The index scores of SF-6D and EQ-5D are strongly correlated (0.71). This is also confirmed by the Bland–Altman plots in Figs. 1 and 2. Figure 1 shows the agreement of these two preference-based measures at the time points T0, T1 and T2. Many observations cluster around 0 difference between the instruments. Furthermore, the difference is proportional to the mean HRQOL scores, illustrated by the fitted regression line. A positive difference is associated with a higher mean score, and a negative difference is associated with a lower mean score. Participants with higher mean

Table 2 Distribution of SF-6D results at baseline (%) $N = 155$

Level	SF physical functioning (six levels)	SF role participation (four levels)	SF social functioning (five levels)	SF bodily pain (six levels)	SF mental health (five levels)	SF vitality (five levels)
1	–	18.7	37.4	24.5	36.1	1.9
2	7.7	32.9	19.4	16.1	30.3	7.7
3	27.7	5.2	16.1	13.6	19.4	15.5
4	26.5	43.2	21.9	14.2	11.6	33.6
5	32.9	–	5.2	20.7	2.6	41.3
6	5.2	–	–	11.0	–	–

Table 4 Comparison of SF6D and EQ 5D utility scores over selected groups of baseline characteristics

Characteristics	SF-6D	EQ-5D
Age. Median 83 years		
Age < 83 years	0.61	0.56
Age ≥ 83 years	0.66	0.67
Falls baseline		
Falls = 1	0.65	0.61
Falls ≥ 2	0.63	0.62
Berg Balance Scale. Median 41 points		
BBS < 41	0.62	0.55
BBS ≥ 41	0.65	0.68
30-s sit-to-stand. Median 6 raises		
STS < 6	0.60	0.53
STS ≥ 6	0.67	0.70
4-m walk test. Median 0.603 m/s		
4MWT < 0.603	0.62	0.55
4MWT ≥ 0.603	0.66	0.69
Falls Efficacy Scale. Median 29 points		
FES-I < 29	0.68	0.70
FES-I ≥ 29	0.59	0.54

HRQOL have higher scores on EQ-5D compared to SF-6D, and analogously participants with lower mean HRQOL have generally lower scores on EQ-5D compared to SF-6D. This relationship becomes more evident from T0 to T2. Figure 2 illustrates the agreement on changes in HRQOL from T0 to T1, T1 to T2 and T0 to T2. A similar pattern can be observed as in Fig. 1. Participants with a positive change in mean HRQOL have a generally stronger positive change in EQ-5D than in SF-6D, while for the participants with negative changes in mean HRQOL, the negative change is generally stronger in EQ-5D than in SF-6D.

Associations between subdomains of SF-6D and EQ-5D and physical measures

Table 5 presents correlations between the different subdomains of SF-6D and EQ-5D and physical measures. First, three items from EQ-5D are strongly correlated with four items from SF-6D; EQ-5D Self-care and SF Physical Function (0.65), EQ-5D Pain/Discomfort and SF Bodily Pain (0.71) and EQ-5D Anxiety/Depression and SF Role Participation (0.61) and SF Mental Health (0.71). Weak to moderate correlations were shown between SF Social Functioning and SF Vitality and all domains of EQ-5D. Similarly, only weak to moderate correlations were shown between EQ-5D Mobility and all domains of SF-6D.

BBS is the only physical measure which is strongly correlated with one of the HRQOL domains, EQ-5D Mobility (0.54). On EQ-5D, moderate correlations were shown between all physical measures and Mobility, Self-care and Usual Activities. On SF-6D, moderate correlations were shown between BBS, STS and FES-I and Physical Functioning and Role Participation. In addition, BBS was moderately

Fig. 1 Bland–Altman plot on the agreement of EQ-5D and SF-6D at time points T0, T1 and T2
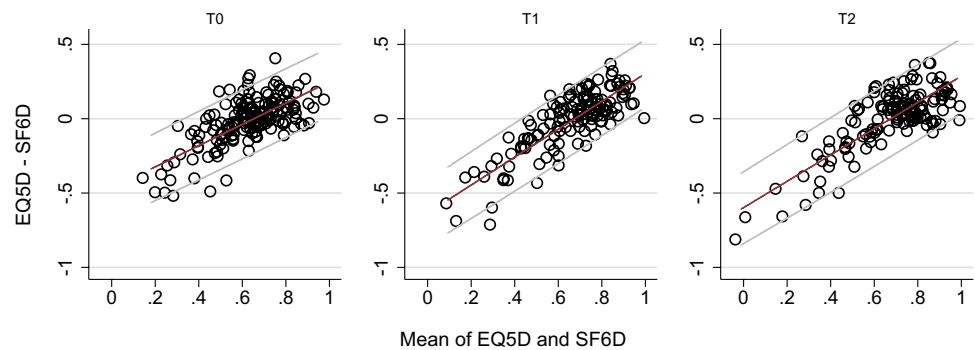

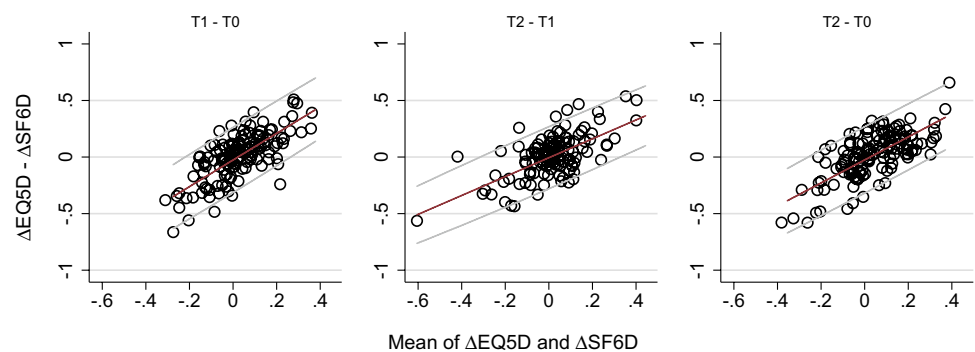**Fig. 2** Bland–Altman plot agreement on changes in EQ-5D and SF-6D from T0 to T1, T1 to T2 and T0 to T2

Table 5 Correlations between subdomains of SF-6D, subdomains of EQ-5D and physical outcomes

	SF physical functioning	SF role participation	SF social functioning	SF bodily pain	SF mental health	SF Vitality	Berg balance scale	30-s sit-to-stand	4-min walk test	Falls Efficacy Scale
EQ-5D mobility	0.38	0.35	0.25	0.25	0.14	0.21	−0.54	−0.40	−0.42	0.40
EQ-5D self-care	0.65	0.32	0.32	0.25	0.25	0.34	−0.36	−0.40	−0.27	0.40
EQ-5D usual activities	0.44	0.32	0.39	0.28	0.18	0.34	−0.38	−0.28	−0.36	0.37
EQ-5D pain/discomfort	0.23	0.27	0.39	0.71	0.28	0.25	−0.11	−0.17	−0.21	0.17
EQ-5D anxiety/depression	0.26	0.61	0.36	0.31	0.71	0.26	−0.10	−0.24	−0.18	0.34
Berg Balance Scale (BBS)	−0.36	−0.31	−0.20	−0.03	−0.07	−0.13				
30-s sit-to-stand (STS)	−0.34	−0.39	−0.24	−0.11	−0.21	−0.12				
4-min walk test (4MWT)	−0.28	−0.25	−0.15	−0.12	−0.06	−0.08				
Falls Efficacy Scale (FES-I)	0.35	0.35	0.27	0.27	0.30	0.25				

Table 6 Elasticities from linear mixed regressions with individual-specific effects

Physical measures	HRQOL	Elasticity	SE	p value
Berg Balance Scale	EQ-5D	0.54	0.11	0.000
	SF-6D	0.18	0.06	0.002
30-s sit-to-stand	EQ-5D	0.09	0.04	0.010
	SF-6D	0.02	0.02	0.194
4-min walk test	EQ-5D	0.20	0.08	0.013
	SF-6D	0.06	0.04	0.118
Falls Efficacy Scale International	EQ-5D	−0.24	0.07	0.001
	SF-6D	−0.13	0.03	0.000

SE standard error

correlated on SF-6D with Bodily Pain and FES-I with Mental Health.

Responsiveness assessed by elasticities

Table 6 presents elasticities from linear mixed regressions. Each regression includes as covariate a physical measure and individual-specific effects. Changes in physical outcomes have larger impacts on changes of HRQOL measured by EQ-5D compared to SF-6D. Elasticities can be interpreted for small changes in the dependent variable. For instance, a 10% increase in BBS score is expected to increase EQ-5D by 5.4% ($p < 0.001$) and SF-6D by 1.8% ($p < 0.001$). A 10% improvement of FES-I is expected to give an improvement of 2.4% ($p < 0.001$) on EQ-5D and 1.3% ($p < 0.001$) on SF-6D. On STS and 4MWT an increased score is expected only leading to significant changes ($p < 0.05$) of EQ-5D.

Discussion

This study compares two widely used metrics of HRQOL in this growing population of older fallers living at home. The results show that SF-6D and EQ-5D are strongly correlated in index scores and some subdomains. There are, however, some differences in the utilities' agreement and in their responsiveness to changes in physical function. EQ-5D gave generally higher scores than SF-6D at a higher mean HRQOL and/or when physical function was better. In contrary, SF-6D provided relatively higher scores than EQ-5D when mean HRQOL was lower and/or when physical function was poorer. EQ-5D showed to be more responsive to changes in physical function compared to SF-6D in older home care recipients who have experienced falls.

There seems to be a high level of agreement between SF-6D and EQ-5D, both in absolute values and in changes over time. There are also some differences. EQ-5D appears to be more sensitive to changes than SF-6D. Similar results

have been presented previously. A study comparing the two instruments in mental health patients showed that EQ-5D resulted in larger health gains, in particular in the subgroup with higher severity of mental health problems [19]. Another study showed that, even though SF-6D had better distributional properties, it did not improve the sensitivity to change [15]. EQ-5D has fewer health states, but it seems not limiting its ability to measure utilities [13]. Hence, the range and variability of the two outcomes, where SF-6D has a smaller range and lower variability, could contribute to its lower sensitivity [11]. The sensitivity of EQ-5D is even more evident at T2 (6 months) than at T0 (baseline). This result could be due to the change in HRQOL over the 6 months' time period, possibly caused by the intervention.

Another reason for limitations in the agreement could be differences in domains of HRQOL covered by the instruments [15]. HRQOL is a complex concept and includes both physical, mental and social domains of health [37]. SF Physical Function was strongly correlated with EQ-5D Self-Care, and SF Role Participation was strongly correlated with EQ-5D Anxiety/Depression. SF Bodily Pain was strongly correlated with EQ-5D Pain/Discomfort and SF Mental Health with EQ-5D Anxiety/Depression. Interestingly, SF Social Functioning and SF Vitality were only weakly to moderately correlated with EQ-5D subscales. This could be due to the discrepancy in domains included. Other elements than physical health, as mental health, might be under-represented in EQ-5D compared to SF-6D [38]. A previous study showed that SF-6D was more efficient at detecting external indicators of health status, for instance longstanding illness or disability, compared to EQ-5D [14]. SF-6D can tap into broader aspects of HRQOL, as role participation and social functioning, which could be the reason why it is more sensitive in complex health conditions. Moreover, a higher number of response items on each of the domains can result in a larger descriptive system with possibly greater sensitivity to the external health indicators. Enhancing several dimensions of health can be important when evaluating interventions for older home care recipients. This should be considered when selecting HRQOL measurement tools for this population.

The instruments' responsiveness to changes seems to be influenced by the type of intervention being evaluated [15]. In the present study, a relatively higher responsiveness to changes in physical function was shown in the scores of EQ-5D compared to SF-6D. Positive changes in HRQOL were associated with improvement of balance, measured by Berg Balance Scale, and falls self-efficacy, measured by Falls Efficacy Scale International. Results were illustrated using elasticities, another representation of regression coefficients enabling direct comparison of the estimates. The findings emphasise the importance of picking an instrument that is more sensitive to the elements that the intervention is aiming to change. When evaluating a falls prevention exercise intervention, responsiveness to underlying physical changes could therefore be of importance. However, if the falls prevention intervention had included actions also targeting other dimensions, as social functioning or vitality, SF-6D could be more responsive. The HRQOL outcome selected will influence the economic evaluations conducted and further the decision-making in public health policy.

Another element to consider when selecting HRQOL instruments is the older adults' health status and level of function at baseline. EQ-5D is thought to be sensitive in patient groups with more severe health states at baseline and less sensitive in patient groups with milder health states at baseline, and the opposite applies to SF-6D [21]. In comparison with normative data on SF-6D and EQ-5D, the participants in the present study had a lower level of self-perceived HRQOL [39, 40]. The older home care recipients had a mean index score of 0.62 on EQ-5D. In a study including Danish population norms, the index score was 0.85 in males and 0.82 in females aged 70–79 [39]. In SF-6D, the mean index score in the present study was 0.64. Norms from a sample of the British population demonstrated a mean index score ranging from 0.77 in the age-group 70–74, to 0.69 in the age-group above 85 [40]. While the Danish and British population of older adults are similar to the Norwegian, there might be some differences in socioeconomic status influencing the general health status [41]. Nevertheless, the low level of HRQOL in this population of home care receivers and fallers emphasise the need for effective health services in the primary care to maintain or improve their HRQOL contributing to healthy ageing.

In addition to health policy, this study has implications for patient management in primary care. Measuring HRQOL in older adults is increasingly seen as important in evaluations and there is a lack of tools that can be applied in clinical practise [4, 42, 43]. In the present study, both measurements were conducted as interviews as recommended for this population, achieving high completion rates, but are at the same time more time-consuming and costly [12]. Previous research including older adults has suggested that EQ-5D might be sufficient when brevity is required and the health changes are expected to be substantial, while SF-36 is more beneficial when several details are required and the health changes expected are less substantial [12, 44]. Although general measures of HRQOL provide relevant information in the group of older adults, there might be some important age-specific factors that are missing, as for instance sensory abilities and autonomy [42]. Instruments designed specifically for older adults could therefore be an important addition in clinical practise.

This study has both strengths and limitations. Due to a thorough follow-up a low number of missing for both SF-6D and EQ-5D was achieved. The participants were recruited from six municipalities to a falls prevention

exercise intervention, and the results might not representative for the general population of older home care recipients. A low percentage of males were included, but this is typical for the population of older home care recipients. The preference weights of SF-6D and EQ-5D have been developed specifically for the United Kingdom, but the Norwegian population could have different preferences. The two scales also differ in their range. To deal with this issue elasticities were calculated, where responsiveness is expressed as comparable changes in percentages. Few index scores are in the lowest end of the distribution, where EQ-5D has health states regarded worse than what SF-6D is able to generate. Finally, responsiveness was only explored related to important physical measures collected in the randomised controlled trial. Other outcomes might also be relevant for this population.

Conclusion

SF-6D and EQ-5D are strongly correlated, but there are some differences in their agreement, aspects of HRQOL covered and responsiveness to changes. Older adults with a generally higher level of HRQOL and/or better physical function achieved a relatively higher score on EQ-5D, and older adults with a generally lower level of HRQOL and/or poorer physical function achieved a relatively higher score on SF-6D. EQ-5D was more responsive to changes in physical function compared to SF-6D in older home care recipients who had experienced falls. This study shows that selecting a HRQOL instrument for evaluating an intervention may depend on the characteristics of the intervention and the studied population. The choice of instrument can affect the outcome of evaluations in the group of frail older adults and consequently health policy for this increasing population.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval The project proposal has been approved by The Regional Committee for Medical Research Ethics in South East Norway (Ref. 2014/2051).

Informed consent Written informed consent was obtained from all participants before baseline testing.

References

1. World Health Organization. (2015). *World report on ageing and health*. Geneva, Switzerland: World Health Organization.
2. Kannus, P., Sievänen, H., Palvanen, M., Järvinen, T., & Parkkari, J. (2005). Prevention of falls and consequent injuries in elderly people. *The Lancet, 366*(9500), 1885–1893. https://doi.org/10.1016/S0140-6736(05)67604-0.
3. Stenhagen, M., Ekstrom, H., Nordell, E., & Elmstahl, S. (2014). Accidental falls, health-related quality of life and life satisfaction: A prospective study of the general elderly population. *Archives of Gerontology and Geriatrics, 58*(1), 95–100. https://doi.org/10.1016/j.archger.2013.07.006.
4. Davis, J. C., Best, J. R., Dian, L., Khan, K. M., Hsu, C. L., Chan, W., et al. (2017). Are the EQ-5D-3L and the ICECAP-O responsive among older adults with impaired mobility? Evidence from the Vancouver Falls Prevention Cohort Study. *Quality of Life Research, 26*(3), 737–747. https://doi.org/10.1007/s11136-016-1487-z.
5. Bjerk, M., Brovold, T., Skelton, D. A., & Bergland, A. (2018). Associations between health-related quality of life, physical function and fear of falling in older fallers receiving home care. *BMC Geriatrics, 18*(1), 253. https://doi.org/10.1186/s12877-018-0945-6.
6. Fletcher, P. C., & Hirdes, J. P. (2004). Restriction in activity associated with fear of falling among community-based seniors using home care services. *Age and Ageing, 33*(3), 273–279. https://doi.org/10.1093/ageing/afh077.
7. Hektoen, L. F., Aas, E., & Lurås, H. (2009). Cost-effectiveness in fall prevention for older women. *Scandinavian Journal of Public Health, 37*(6), 584–589. https://doi.org/10.1177/1403494809341093.
8. Gillespie, L. D., Robertson, M. C., Gillespie, W. J., Sherrington, C., Gates, S., Clemson, L. M., et al. (2012). Interventions for preventing falls in older people living in the community. *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.CD007146.pub3.
9. El-Khoury, F., Cassou, B., Charles, M.-A., & Dargent-Molina, P. (2013). The effect of fall prevention exercise programmes on fall induced injuries in community dwelling older adults: Systematic review and meta-analysis of randomised controlled trials. *BMJ, 347,* f6234. https://doi.org/10.1136/bmj.f6234.
10. Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes* (4th ed.). Oxford, UK: Oxford University Press.
11. Brazier, J., Roberts, J., Tsuchiya, A., & Busschbach, J. (2004). A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Economics, 13*(9), 873–884. https://doi.org/10.1002/hec.866.
12. Haywood, K., Garratt, A., & Fitzpatrick, R. (2005). Quality of life in older people: A structured review of generic self-assessed health instruments. *Quality of Life Research, 14*(7), 1651–1668. https://doi.org/10.1007/s11136-005-1743-0.
13. Gerard, K., Nicholson, T., Mullee, M., Mehta, R., & Roderick, P. (2004). EQ-5D versus SF-6D in an older, chronically ill patient group. *Applied Health Economics and Health Policy, 3*(2), 91–102. https://doi.org/10.2165/00148365-200403020-00005.
14. Petrou, S., & Hockley, C. (2005). An investigation into the empirical validity of the EQ-5D and SF-6D based on hypothetical preferences in a general population. *Health Economics, 14*(11), 1169–1189. https://doi.org/10.1002/hec.1006.
15. van Stel, H. F., & Buskens, E. (2006). Comparison of the SF-6D and the EQ-5D in patients with coronary heart disease.

Health and Quality of Life Outcomes, 4(1), 20. https://doi.org/10.1186/1477-7525-4-20.

16. Cunillera, O., Tresserras, R., Rajmil, L., Vilagut, G., Brugulat, P., Herdman, M., et al. (2010). Discriminative capacity of the EQ-5D, SF-6D, and SF-12 as measures of health status in population health survey. *Quality of Life Research, 19*(6), 853–864. https://doi.org/10.1007/s11136-010-9639-z.

17. Chen, J., Wong, C. K., McGhee, S. M., Pang, P. K., & Yu, W.-C. (2014). A comparison between the EQ-5D and the SF-6D in patients with chronic obstructive pulmonary disease (COPD). *PLoS ONE, 9*(11), e112389. https://doi.org/10.1371/journal.pone.0112389.

18. Kontodimopoulos, N., Pappa, E., Papadopoulos, A. A., Tountas, Y., & Niakas, D. (2009). Comparing SF-6D and EQ-5D utilities across groups differing in health status. *Quality of Life Research, 18*(1), 87–97. https://doi.org/10.1007/s11136-008-9420-8.

19. Lamers, L., Bouwmans, C., van Straten, A., Donker, M., & Hakkaart, L. (2006). Comparison of EQ-5D and SF-6D utilities in mental health patients. *Health Economics, 15*(11), 1229–1236. https://doi.org/10.1002/hec.1125.

20. McDonough, C. M., & Tosteson, A. N. (2007). Measuring preferences for cost-utility analysis: How choice of method may influence decision-making. *Pharmacoeconomics, 25*(2), 93–106. https://doi.org/10.2165/00019053-200725020-00003.

21. Joore, M., Brunenberg, D., Nelemans, P., Wouters, E., Kuijpers, P., Honig, A., et al. (2010). The impact of differences in EQ-5D and SF-6D utility scores on the acceptability of cost-utility ratios: Results across five trial-based cost–utility studies. *Value in Health, 13*(2), 222–229. https://doi.org/10.1111/j.1524-4733.2009.00669.x.

22. Bjerk, M., Brovold, T., Skelton, D. A., & Bergland, A. (2017). A falls prevention programme to improve quality of life, physical function and falls efficacy in older people receiving home help services: Study protocol for a randomised controlled trial. *BMC Health Services Research, 17*(1), 559. https://doi.org/10.1186/s12913-017-2516-5.

23. Bjerk, M., Brovold, T., Skelton, D. A., Liu-Ambrose, T., & Bergland, A. (2019). Effects of a falls prevention exercise programme on health-related quality of life in older home care recipients: A randomised controlled trial. *Age and Ageing, 48*(2), 213–219. https://doi.org/10.1093/ageing/afy192.

24. von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gotzsche, P. C., Vandenbroucke, J. P., et al. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *The Lancet, 370*(9596), 1453–1457. https://doi.org/10.1016/S0140-6736(07)61602-X.

25. Loge, J. H., & Kaasa, S. (1998). Short Form 36 (SF-36) health survey: Normative data from the general Norwegian population. *Scandinavian Journal of Public Health, 26*(4), 250–258. https://doi.org/10.1177/14034948980260040401.

26. Walters, S. J., & Brazier, J. E. (2005). Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Quality of Life Research, 14*(6), 1523–1532. https://doi.org/10.1007/s11136-004-7713-0.

27. The EuroQol Group. (1990). EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy, 16*(3), 199–208. https://doi.org/10.1016/0168-8510(90)90421-9.

28. Brazier, J. E., & Roberts, J. (2004). The estimation of a preference-based measure of health from the SF-12. *Medical Care, 42*, 851–859.

29. Berg, K. O., Maki, B. E., Williams, J. I., Holliday, P. J., & Wood-Dauphinee, S. L. (1992). Clinical and laboratory measures of postural balance in an elderly population. *Archives of Physical Medicine and Rehabilitation, 73*(11), 1073–1080.

30. Peters, D., Fritz, S., & Krotish, D. (2013). Assessing the reliability and validity of a shorter walk test compared with the 10-Meter Walk Test for measurements of gait speed in healthy, older adults. *Journal of Geriatric Physical Therapy, 36*(1), 24–30. https://doi.org/10.1519/JPT.0b013e318248e20d.

31. Jones, C. J., Rikli, R. E., & Beam, W. C. (1999). A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. *Research Quarterly for Exercise and Sport, 70*(2), 113–119. https://doi.org/10.1080/02701367.1999.10608028.

32. Helbostad, J. L., Taraldsen, K., Granbo, R., Yardley, L., Todd, C. J., & Sletvold, O. (2010). Validation of the falls efficacy scale-international in fall-prone older persons. *Age and Ageing, 39*(2), 259. https://doi.org/10.1093/ageing/afp224.

33. Marques-Vieira, C. M. A., Sousa, L. M. M., Severino, S., Sousa, L., & Caldeira, S. (2016). Cross-cultural validation of the falls efficacy scale international in elderly: Systematic literature review. *Journal of Clinical Gerontology and Geriatrics, 7*(3), 72–76. https://doi.org/10.1016/j.jcgg.2015.12.002.

34. Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: A comprehensive review. *Journal of the American Geriatrics Society, 40*(9), 922–935. https://doi.org/10.1111/j.1532-5415.1992.tb01992.x.

35. Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. London: Routledge.

36. Greene, W. H. (2012). *Econometric analysis*. Boston: Pearson.

37. Karimi, M., & Brazier, J. (2016). Health, health-related quality of life, and quality of life: What is the difference? *Pharmacoeconomics, 34*(7), 645–649. https://doi.org/10.1007/s40273-016-0389-9.

38. Grund, S., Breitinger, E., Fricke, S., Alpers, G. W., Hundsdörfer, W., & Schäfer, H.-G. (2017). Health-related quality of life measurement inpatient geriatric rehabilitation: A comparison of the feasibility and suitability between the SF-36 and EQ-5D-5L questionnaires. *Journal of Gerontology & Geriatric Research, 6*(6), 458. https://doi.org/10.4172/2167-7182.1000458.

39. Sørensen, J., Davidsen, M., Gudex, C., Pedersen, K. M., & Brønnum-Hansen, H. (2009). Danish EQ-5D population norms. *Scandinavian Journal of Public Health, 37*(5), 467–474. https://doi.org/10.1177/1403494809105286.

40. van den Berg, B. (2012). Sf-6d population norms. *Health Economics, 21*(12), 1508–1512. https://doi.org/10.1002/hec.1823.

41. Christiansen, T., Lauridsen, J. T., Kifmann, M., Lyttkens, C. H., Ólafsdóttir, T., & Valtonen, H. (2018). Healthcare, health and inequality in health in the Nordic countries. *Nordic Journal of Health Economics, 6*(2), 10. https://doi.org/10.5617/njhe.5955.

42. Leegaard, M., Utne, I., Halvorsrud, L., Valeberg, B. T., Torbjørnsen, A., Bjørnnes, A. K., et al. (2018). A review of self-rated generic quality of life instruments used among older patients receiving home care nursing. *Health and Social Care in the Community, 26*(3), e321–e328. https://doi.org/10.1111/hsc.12447.

43. Hickey, A., Barker, M., McGee, H., & O'boyle, C. (2005). Measuring health-related quality of life in older patient populations: A review of current approaches. *Pharmacoeconomics, 23*(10), 971–993. https://doi.org/10.2165/00019053-200523100-00002.

44. Brazier, J. E., Walters, S. J., Nicholl, J., & Kohler, B. (1996). Using the SF-36 and Euroqol on an elderly population. *Quality of Life Research, 5*(2), 195–204. https://doi.org/10.1007/BF00434741.