

# Capabilities and quality of life in Dutch psycho-geriatric nursing homes: an exploratory study using a proxy version of the ICECAP-O

Peter Makai · Werner B. F. Brouwer ·  
Marc A. Koopmanschap · Anna P. Nieboer

Accepted: 15 August 2011 / Published online: 6 September 2011  
© The Author(s) 2011. This article is published with open access at Springerlink.com

## Abstract

**Purpose** To validate the ICECAP-O capability measure in psycho-geriatric elderly in nursing homes, we compared the capability scores of restrained and unrestrained clients. Both nursing staff and family were used as proxies for assessing clients' capabilities.

**Method** For 122 psycho-geriatric elderly, a total of 96 nursing professionals and 68 family members completed a proxy questionnaire. We investigated the convergent and discriminant validity of the ICECAP-O and measures of care dependency, health-related quality of life, and overall quality of life. We also directly compared ICECAP-O scores of the 56 clients for whom both nursing staff and family members had completed the questionnaire.

**Results** Convergent validity between ICECAP-O and care dependency, health-related, and overall quality of life measures could be established, as well as discriminant validity for the restrained and unrestrained groups. Nursing and family proxy ICECAP-O tariffs were not significantly correlated.

**Discussion** ICECAP-O measures a more general concept than health-related quality of life and can differentiate between restrained and non-restrained psycho-geriatric clients. Since nurses seem to be able to assess the current quality of life of clients using the ICECAP-O more precisely than the family proxies, for now the use of nursing proxies is recommended in a nursing home setting.

**Keywords** ICECAP-O · Capabilities · Psycho-geriatric elderly

## Abbreviations

ICECAP-O	ICEpop CAPability measure for older people
QoL	Quality of life
HrQoL	Health-related quality of life
CDS	Care dependency scale
EQ-5D	Euroqol
EQ-VAS	Visual analog scale
HADS	Hospital anxiety and depression scale

## Background

Services aimed at elderly living in psycho-geriatric (PG) wards, who mainly suffer from advanced dementia, are not often evaluated using cost-effectiveness analysis. Still, in general, trading off costs and benefits is as important in long-term care as it is in curative care. Especially in times of budget cuts or when care innovations find their way into the long-term care sector, considering the costs and benefits of interventions is important. In curative care, this is commonly done using cost-utility analysis, where the primary outcome is health-related quality of life (HrQoL). Preference-based health-related quality of life measures attach utility weights to specific health states in order to be able to compute utility gains from health changes. Such gains are then compared to the (incremental) costs of an intervention [1]. Using this approach to evaluate services for the PG elderly, however, is problematic.

A major problem is that health-related quality of life measures aim to detect and value changes in health and functioning, while services for the elderly may (be aimed

---

P. Makai (✉) · W. B. F. Brouwer · M. A. Koopmanschap ·  
A. P. Nieboer  
Institute of Health Policy and Management, Erasmus University  
Rotterdam, Rotterdam, The Netherlands  
e-mail: makai@bmg.eur.nl

to) affect quality of life more broadly [2–4]. For example, it is not uncommon to physically restrain the PG elderly to prevent them from falling [5], but doing so restricts freedom of movement, autonomy, and enjoyment of life. Removing such restraints would restore some control over their lives and allow more enjoyment through an increased capacity to fill their day with more varied activities. Whether the health of unrestrained patients would also improve, however, is questionable since freedom of movement may not directly affect existing health problems. Therefore, in determining the value for money of interventions aimed to reduce restraints, HrQoL is likely to be a too restrictive evaluative space, since it does not (directly) value self-control or enjoyment of life. HrQoL measures (such as the EQ-5D and SF-6D) may therefore not fully account for all benefits of such interventions, and using them in these contexts could misinform decision makers.

A promising approach to measure QoL more comprehensively in the PG elderly is to use the newly developed ICEpop (Investigating Choice Experiments for the Preferences of Older People) capability measure for older people (ICECAP-O). The ICECAP instruments can be seen as measuring capability QoL [6] achieved by the capacity to perform certain actions and achieve certain states [7]. The ICECAP-O measures five capability dimensions—attachment, security, role, enjoyment, and control—with one question per dimension. Each dimension can be scored on four levels. The ICECAP-O was developed using rigorous qualitative and quantitative approaches [8–11]. In order to obtain tariffs for the ICECAP-O, the attributes were valued using best-worst scaling, a special type of discrete choice analysis. The ICECAP-O has been used in the British general elderly population, demonstrating that it is related to, but not exclusively dependent on HrQoL [9]. The overwhelming majority of the included elderly lived at home and did not receive long-term health or social care. To date, ICECAP-O has not been used in populations receiving long-term care. This lack of validation is especially problematic for the vulnerable PG elderly populations, who consume substantial amounts of health and social services [12].

It needs noting that substantial effort has been put in recent years into developing dementia-specific QoL instruments for use in patients with mild to moderate dementia [13–20]) and in severe dementia [21]. However, not only do these instruments normally not have related utility weights, limiting their usefulness in cost-utility analysis, they also are, by definition, disease specific rather than generic, which limits their usefulness in decision making across diseases and sectors. Moreover, disease-specific measures can still focus on health-related, rather than general quality of life. Hence, here we focus on the generic ICECAP-O, with its preference-based tariffs.

The use of QoL instruments in a PG patient population is difficult, since due to their cognitive limitations, patients may not be able to assess their QoL accurately. It has been shown to be possible to develop user-friendly (disease specific) instruments for self-completion in this context, especially for mild to moderate dementia patients. However, with diminishing cognitive ability, this becomes increasingly difficult. Currently, to our knowledge, there are no generic QoL instruments with accompanying utility weights that are recommended for use in people with dementia. The lack of validation in this particular population, i.e., that PG elderly, is likely to be related to limited cognitive ability due to severe dementia [22], hampering self-completion of questionnaires. We therefore decided to use proxies, who complete the questionnaire on the patient's behalf. An important issue with proxies is that they may not complete the questionnaire as the client would have. A prerequisite in using proxies is that they can at least provide reasonable approximations of the patient's QoL [23]. Proxy measurement has been associated with a consistent negative bias in QoL measurement [13], although this may be more typical in case of informal carers of dementia patients [14]. It has been suggested that such proxy effects can be minimized using substituted judgment [14], asking the proxies to fill out the questionnaire *as if* they were the person with dementia.

The aim of this study is to explore the validity of the ICECAP-O for the PG elderly. To that end, we first investigated the convergent validity of the ICECAP-O by comparing it to other care-related HrQoL and overall QoL instruments. We used a sample of elderly in Dutch psychogeriatric nursing homes to establish the discriminant validity of the ICECAP-O by (1) comparing a restrained group to a non-restrained group and (2) investigating whether the ICECAP-O was indeed measuring a concept broader than health. To complete the validation exercise, we compared questionnaires filled out by two appropriate proxies, namely nursing staff and family members. This is to our knowledge the first study of its kind.

## Method

### Design

The ICECAP-O questionnaire was forward–backward translated into Dutch by two independent translators. For our study, we used the baseline measurement from an economic evaluation study of a quality improvement intervention that aimed to reduce restraints in the Care for Better quality collaborative in Dutch long-term care [24, 25]. Four nursing homes and a total of 122 clients from different geographic regions in the Netherlands

participated in the study. All 72 clients in restraints participated and 50 randomly selected non-restrained clients in the same departments served as a control group. We distributed two copies of the questionnaire for each client, one to be filled out by nursing staff that personally cared for the client (nursing version) and one for family members (family version) asking proxies to use substituted judgement. Since data collection of the nursing version was carried out in the context of a national quality improvement program, no ethical committee approval was necessary under Dutch law [26, 27]. Informed consent was obtained for the family version. The researchers received no personal information about the clients during the study.

## Measures

Besides the ICECAP-O (as shown in [Appendix](#)), the questionnaire contained the following QoL measures: the EQ-5D, EQ-VAS instrument, Cantril's ladder, and overall life satisfaction. It also contained the Hospital Anxiety and Depression Scale (HADS). The nursing version contained the care dependency scale (CDS), which needs to be completed by care professionals. The EQ-5D [28] measures HrQoL along five dimensions (mobility; self-care; daily activities; pain and discomfort; and anxiety and depression) with three levels each (1 = no problems, 2 = moderate problems, and 3 = extreme problems). It has been used with proxies in a large number of studies, including clients with Alzheimer and severe dementia [29]. The EQ-VAS is a one-dimensional HrQoL measure frequently used alongside the EQ-5D in validation studies and has also been used with proxies [23]. The EQ-VAS comprises a single scale ranging from zero (worst imaginable health) to 100 (best imaginable health). Cantril's ladder is a classic one-dimensional overall quality of life scale [30], with the bottom rung representing no quality of life and the top representing full quality of life. It has been used with proxies [31]. We also used an overall life satisfaction scale, a one-dimensional index ranging from zero (completely dissatisfied) to 10 (completely satisfied) [32]. The HADS scale was originally developed for use in hospitals but has since been used in various populations [33] and with proxies [34] to assess anxiety and depression symptoms. HADS consists of two 7-item scales, one for depression and one for anxiety, which can be also used in a composite index (Cronbach's  $\alpha = 0.82$  nursing version, 0.87 family version in this current study, comparable to self-reported values in Dutch elderly [35]) with values ranging from 0 (no problems) to 42 (severe depression and anxiety). The care dependency scale (CDS) developed by Dijkstra [36] contains 15 dimensions measuring the amount of independence the patient has retained with regard to dimensions such as eating and drinking, body posture,

incontinence, learning ability, ability to structure the day, communication, and autonomy. The CDS has scores that range from 15 (completely care dependent) to 75 (completely care independent). The CDS has been used and validated extensively [37, 38] and is a useful instrument for assessing need for care. CDS scores have been shown to be associated with a number of problems in elderly care, such as fall-risk, pressure ulcers, and so on and are designed to be completed by nurses and professional caregivers [37, 38].

## Hypotheses

For convergent validity, we expect the ICECAP-O to correlate with overall measures (Cantril's ladder and overall life satisfaction) and HrQoL measures, as well as with CDS scores, since all measurement instruments differentiate between better and worse states. With respect to discriminant validity, we expect to find differences between the non-restrained and restrained groups in terms of ICECAP-O scores and other overall QoL measures, but not in HrQoL measures, since we expect the two groups to be in a similar health state, while their non-health circumstances differ. To test whether capabilities are indeed measuring a concept broader than health, we expect to observe a difference in ICECAP-O scores between the restrained and the non-restrained clients even when controlling for HrQoL, demographic variables, and care dependency. For proxy agreement, we expect the nursing and family proxies for each client to be correlated and the scores to be not significantly different from each other.

## Analysis

We performed an item-level analysis to determine non-response for all scales in the questionnaire. We used multiple imputations to treat item non-response for nursing and family questionnaires separately with the Markov chain Monte-Carlo method (MCMC) [39]. We also tested the assumption of multivariate normality underlying the MCMC method. Following multiple imputations, utility and sum-scores were computed where relevant (see [Appendix](#)). For the CDS, which was only included in the nursing version, the nursing scores for the patients for which a family version was also present were also used in the analysis pertaining to the family version. Remaining missing observations were imputed.

We used descriptive statistics to analyze demographic characteristics. Means and standard deviations were computed for continuous variables, medians for ordinal variables. All comparisons between demographic variables were performed using the Mann–Whitney-U test, except in the case of education, where a Chi-square was performed. Data were analyzed using STATA 11.

Concurrent validity was assessed using correlations in the nursing and family versions separately. To test discriminant validity, we employed chi-square tests and Mann–Whitney-U tests to compute mean differences between the restrained and non-restrained groups. We performed this comparison on the nursing and the family proxy separately. To further investigate whether the ICECAP-O could both discriminate between the groups and measure a concept broader than HrQol, we performed multivariate regressions. For this purpose, we controlled for demographic variables and care dependency. Two multivariate ordinary least squares regression (OLS) models were fitted on the ICECAP-O index, one using nursing proxy variables and one using family proxy variables. Regression assumptions were checked.

To test agreement between the two proxy groups, we used the Mann–Whitney-U tests and correlations for the questionnaires for which both proxies were available.

## Results

### Response

For a total of 122 clients, 96 nurses and 68 family members completed the questionnaires, implying response rates of 78 and 56%, respectively. For the 96 nursing questionnaires, 62 clients (64%) were in restraints; for the 68 family questionnaires, 47 clients (69%) were in restraints. For 56 clients, we received both types of proxy questionnaires. Item non-response was not systematic and averaged around 2% across all items in the nursing questionnaires, and 4% in the family questionnaires. In the nursing version, multiple imputations allowed for using 96 cases instead of 88–91 in bivariate analysis and 87 in multivariate analyses. In the family version, multiple imputations allowed using 68 cases instead of 58–61 in bivariate analysis and 47 in multivariate analysis.

### Descriptive characteristics and relationship between different proxies

Client's demographic and care-related characteristics can be seen in Table 1, split according to the two versions of the questionnaire.

### Convergent validity

As can be seen in Tables 2 and 3, there was a significant correlation between capabilities and HrQol, as shown by the significant correlation between the ICECAP-O tariffs and the EQ5D and the EQ-VAS health measures. The correlation, however, was not particularly strong. The

**Table 1** Demographic and care-related characteristics

Demographics	All cases	
	Nursing version <i>n</i> = 96	Family version <i>N</i> = 68
Age <sup>a</sup>	82 (9.1)	82 (7.3)
Sex <sup>b</sup>	68% female	67% female
Education		
Primary	45.8%	51.5%
Secondary (general)	27.1%	14.7%
Secondary (vocational)	14.6%	16.2%
Secondary (scientific)	5.2%	7.4%
Tertiary (university, college)	7.3%	10.4%
CDS <sup>a,c</sup>	31.14 (15.09)	33.53 (15.16)
HADS <sup>a</sup>	23.68 (3.91)	24.30 (3.17)
<i>ICECAP-O</i>		
Attachment <sup>a,b</sup>	2.34 (0.78) 2	2.79 (0.89) 3
Security <sup>a,b</sup>	3.38 (0.79) 4	3.15 (1.10) 4
Role <sup>a,b</sup>	1.78 (0.81) 2	1.43 (0.68) 1
Enjoyment <sup>a,b</sup>	2.21 (0.81) 2	1.88 (0.82) 2
Control <sup>a,b</sup>	1.40 (0.71) 1	1.15 (0.50) 1
ICECAP-O tariffs <sup>a</sup>	0.50 (0.20)	0.43 (0.17)
<i>EQ-5D</i>		
Mobility <sup>a,b</sup>	2.09 (0.62) 2	2.25 (0.53) 2
Self-care <sup>a,b</sup>	2.71 (0.50) 3	2.82 (0.46) 3
Usual activities <sup>a,b</sup>	2.75 (0.52) 3	2.88 (0.37) 3
Pain, discomfort <sup>a,b</sup>	1.81 (0.57) 2	1.69 (0.58) 3
Anxiety, depression <sup>a,b</sup>	1.70 (0.63) 2	1.76 (0.65) 2
EQ-5D tariffs <sup>a</sup>	0.49 (0.21)	0.46 (0.20)
EQ-VAS <sup>a</sup>	55.33 (17.24)	45.87 (16.56)
Cantrill's ladder <sup>a</sup>	4.62 (2.01)	4.11 (4.02)
Overall life satisfaction <sup>a</sup>	4.72 (2.34)	4.76 (3.22)

<sup>a</sup> Mean, (SD)

<sup>b</sup> Median

<sup>c</sup> Only included in nursing version

ICECAP-O tariffs were also correlated with Cantrill's ladder and the overall life satisfaction measures. There was also a significant relationship between ICECAP-O tariffs in both versions of the questionnaire and CDS, though the correlation was stronger in the nursing version. The HADS was not correlated with the ICECAP-O tariffs in either the nursing or family questionnaires.

### Discriminative validity

The demographic and care-related characteristics for the respondents of the restrained and unrestrained client groups can be seen in Table 4. Age and gender were not significantly different for the two groups. There was no significant

**Table 2** Convergent and discriminant validity nursing version

Nursing version/ nursing version n = 96	ICECAP-O tariffs nursing version	EQ-5D nursing version	EQ-VAS nursing version	Cantril's ladder nursing version	Overall life satisfaction nursing version	Care dependency scale nursing version	HADS nursing version
ICECAP-O tariffs nursing version	1.00						
EQ-5D nursing version	0.48**	1.00					
EQ-VAS nursing version	0.55**	0.49**	1.00				
Cantril's ladder nursing version	0.60**	0.51**	0.70**	1.00			
Overall life satisfaction nursing version	0.52**	0.34**	0.65**	0.70**	1.00		
Care dependency scale nursing version	0.56**	0.50**	0.34**	0.47**	0.23*	1.00	
HADS nursing version	-0.18	0.08	-0.10	-0.19	-0.08	-0.16	1.00

\* Significance on the 5% level

\*\* Significance on the 1% level

**Table 3** Convergent and discriminant validity family version

Family version/ family version n = 68	ICECAP-O tariffs family version	EQ-5D family version	EQ-VAS family version	Cantril's ladder family version	Overall life satisfaction family version	Care dependency scale nursing version	HADS family version
ICECAP-O tariffs family version	1.00						
EQ-5D family version	0.57**	1.00					
EQ-VAS family version	0.43**	0.36**	1.00				
Cantril's ladder family version	0.33**	0.20	0.32**	1.00			
Overall life satisfaction family version	0.48**	0.37**	0.28*	0.80**	1.00		
Care dependency scale nursing version	0.32**	0.10	0.45**	0.20	0.16	1.00	
HADS family version	-0.01	0.32**	0.01	0.07	0.11	-0.21	1.00

\* Significance on the 5% level

\*\* Significance on the 1% level

association between education and being in restraints. The mean CDS score was significantly lower in the group in restraints, indicating higher dependency. HADS scores differed significantly for clients in restraints in the nursing version; they were more depressed and anxious. In the nursing version, there was a significant difference between the groups in all ICECAP-O dimensions except for security.

In the family version, two dimensions—role and enjoyment—were significantly different. A difference was also observed in the ICECAP-O tariffs. Clients without restraints score somewhat higher on HrQol as measured by EQ5D and EQ-VAS, but the difference was not significant at the 5 percent confidence level. A Mann–Whitney–U test indicated that there was a significant difference in terms of capabilities.

This was also true for the overall QoL as measured by Cantril's ladder and overall life satisfaction.

Table 5 shows how the ICECAP-O tariffs discriminated between clients with and without restraints using a multivariate analysis. Being in restraints independently discriminated between capability QoL in the nursing version, but not in the family version, when controlling for HrQoL measures, demographic measures, and care dependency. The individual influences of the EQ5D and CDS on the ICECAP-O tariffs are pronounced in both versions.

#### Relationship between the two proxies

Table 6 shows the agreement between the nursing and family assessment of the variables using Mann–Whitney-U and correlations for the 56 clients for whom both proxy versions were available. The results of the Mann–Whitney-U show that three out of five ICECAP-O dimensions had a significantly different distribution between the proxy groups, while the average tariffs were not significantly different. The distributions of the EQ-5D were not

**Table 4** Comparison of restrained and non-restrained clients

	Nursing version N = 96			Family version N = 68		
	Restrained	Not restrained	P value	Restrained	Not restrained	P value
Age <sup>a</sup>	83 (7.7)	80 (10)	0.232	82 (8.1)	83 (5.3)	0.456
Sex <sup>b</sup>	67% Female	68% Female	0.407	64% Female	72% Female	0.975
Education <sup>c</sup>			0.154			0.495
Primary	37.1%	61.8%		55.3%	42.9%	
Secondary (general)	30.6%	20.6%		10.6%	23.8%	
Secondary (vocational)	16.1%	11.8%		12.8%	23.8%	
Secondary (scientific)	8.1%	6.7%		8.5%	4.8%	
Tertiary (university, college)	8.0%	5.8%		12.8%	4.8%	
CDS <sup>b,d</sup>	27.58 (11.62)	40.47 (15.09)	0.000**	28.70 (12.17)	44.33 (15.89)	0.000**
HADS <sup>a</sup>	24.16 (4.18)	22.82 (3.25)	0.048*	24.38 (3.34)	24.14 (2.83)	0.826
<i>ICECAP-O</i>						
Attachment <sup>a,b</sup>	2.17 (0.78) 2	2.64 (0.69) 3	0.001**	2.74 (0.92) 3	2.90 (0.83) 3	0.533
Security <sup>a,b</sup>	3.43 (0.80) 4	3.29 (0.76) 3	0.266	3.13 (1.10) 4	3.19 (1.17) 4	0.744
Role <sup>a,b</sup>	1.53 (0.78) 1	2.20 (0.69) 2	0.000**	1.32 (0.63) 1	1.67 (0.73) 2	0.028*
Enjoyment <sup>a,b</sup>	2.05 (0.76) 2	2.53 (0.75) 2	0.005**	1.74 (0.70) 2	2.19 (0.98) 2	0.034*
Control <sup>a,b</sup>	1.26 (0.65) 1	1.65 (0.73) 2	0.000**	1.11 (0.48) 1	1.24 (0.54) 1	0.121
ICECAP-O tariffs <sup>a</sup>	0.43 (0.19)	0.63 (0.16)	0.000**	0.40 (0.16)	0.51 (0.18)	0.033*
<i>EQ-5D</i>						
Mobility <sup>a,b</sup>	2.21 (0.55) 2	1.88 (0.69) 2	0.016*	2.30 (0.59) 2	2.14 (0.36) 2	0.194
Self-care <sup>a,b</sup>	2.84 (0.41) 3	2.47 (0.56) 2	0.000**	2.87 (0.40) 3	2.71 (0.56) 3	0.160
Usual activities <sup>a,b</sup>	2.82 (0.46) 3	2.61 (0.60) 3	0.043*	2.89 (0.37) 3	2.86 (0.36) 3	0.496
Pain, discomfort <sup>a,b</sup>	1.76 (0.62) 2	1.91 (0.45) 2	0.156	1.72 (0.62) 2	1.62 (0.50) 2	0.589
Anxiety, depression <sup>a,b</sup>	1.77 (0.62) 2	1.61 (0.65) 2	0.260	1.83 (0.64) 2	1.62 (0.67) 2	0.193
EQ-5D tariffs <sup>a</sup>	0.46 (0.22)	0.53 (0.18)	0.200	0.43 (0.21)	0.53 (0.18)	0.073
EQ-VAS <sup>a</sup>	53.18 (15.52)	59.26 (19.66)	0.072	43.96 (16.50)	50.14 (16.28)	0.230
Cantrill's ladder <sup>a</sup>	4.15 (2.00)	5.49 (1.96)	0.000**	3.96 (4.05)	4.45 (4.04)	0.931
Overall life satisfaction <sup>a</sup>	4.48 (2.16)	4.16 (2.63)	0.118	4.46 (2.23)	5.42 (3.14)	0.319

\* Significance on the 5% level

\*\* Significance on the 1% level

<sup>a</sup> Mean (SD)

<sup>b</sup> Median

<sup>c</sup> Chi-square test

<sup>d</sup> Only included in nursing version

**Table 5** Regression results

Independent variables	Dependent variable: ICECAP-O	
	Nursing version (n = 96)	Family version (n = 68)
Constant	3.61 (0.326)	-5.68 (0.242)
Restrained	-0.141 (0.001)**	-0.043 (0.403)
Age	0.002 (0.349)	0.003 (0.225)
Sex	-0.008 (0.804)	0.029 (0.526)
Education level	0.012 (0.289)	0.002 (0.871)
Organization dummy 1	0.055 (0.357)	-0.067 (0.293)
Organization dummy 2	0.062 (0.125)	0.030 (0.642)
Organization dummy 3	0.059 (0.274)	-0.037 (0.518)
Pilot or control department	0.041 (0.437)	-0.073 (0.172)
CDS <sup>a</sup>	0.003 (0.007)**	0.003 (0.033)*
EQ5D	0.325 (0.000)**	0.411 (0.000)**
R square (adj.)	0.500 (0.441)	0.263 (0.373)

\* Significance on the 5% level

\*\* Significance on the 1% level

<sup>a</sup> Only included in nursing version

significantly different except for the mobility dimension. The EQ-VAS was significantly different. Overall Qol measures were the same in both proxy groups.

Agreement between the nursing and family proxies was low for the ICECAP-O dimensions. A significant correlation existed only between the two versions for the

control dimension. Neither the ICECAP-O tariffs for both proxy groups nor the EQ5D scores were significantly correlated. Measures of overall quality of life were also uncorrelated between the two proxy groups; this is true for both Cantril's ladder and overall life satisfaction. On the other hand, there was a slightly significant correlation between the EQ-VAS scores in both proxy groups. The HADS score was significantly correlated in the two proxy groups.

## Discussion

### Summary of main results

Our study is the first attempt to measure Qol and capabilities of physically restrained psycho-geriatric nursing home clients. It was performed in the context of a validation exercise of the ICECAP-O. The ICECAP-O seems a promising, generic, preference-based instrument in the context of evaluating interventions in the psycho-geriatric context. Our study showed reasonable convergent and discriminant validity. Although related to HrQol, the relationship did not turn out to be very strong in our study. Given that and the multivariate regression results, the ICECAP-O appears to encompass a broader evaluative space than health alone. As expected, when the two groups were compared, clients in restraints had a lower Qol than

**Table 6** Analysis of selection of respondents for whom both versions were available

	Both responded N = 56			
	Nursing version	Family version	Mann-Whitney-U	Correlation
CDS <sup>b,c</sup>	32.00 (15.47)	32.00 (15.47)	n.a.	n.a.
HADS <sup>a</sup>	24.39 (3.22)	24.33 (3.23)	0.96	0.33*
<i>ICECAP-O</i>				
Attachment <sup>a,b</sup>	2.34 (0.78) 2	2.79 (0.89) 3	0.06	0.14
Security <sup>a,b</sup>	3.38 (0.79) 4	3.15 (1.10) 4	0.75	0.04
Role <sup>a,b</sup>	1.78 (0.81) 2	1.43 (0.68) 1	0.01*	0.23
Enjoyment <sup>a,b</sup>	2.21 (0.81) 2	1.88 (0.82) 2	0.03*	0.19
Control <sup>a,b</sup>	1.40 (0.71) 1	1.15 (0.50) 1	0.04*	0.30*
ICECAP-O tariffs <sup>a</sup>	0.49 (0.19)	0.44 (0.17)	0.24	0.15
<i>EQ-5D</i>				
Mobility <sup>a,b</sup>	2.09 (0.62) 2	2.25 (0.53) 2	0.02*	0.38**
Self-care <sup>a,b</sup>	2.71 (0.50) 3	2.82 (0.46) 3	0.18	0.23
Usual activities <sup>a,b</sup>	2.75 (0.52) 3	2.88 (0.37) 3	0.05	0.18
Pain, discomfort <sup>a,b</sup>	1.81 (0.57) 2	1.69 (0.58) 2	0.29	0.18
Anxiety, Depression <sup>a,b</sup>	1.70 (0.63) 2	1.76 (0.65) 2	0.85	-0.34
EQ-5D tariffs <sup>a</sup>	0.46 (0.22)	0.47 (0.20)	0.63	0.01
EQ-VAS <sup>a</sup>	54.48 (14.64)	46.56 (1.57)	0.00**	0.33*
Cantrill's ladder <sup>a</sup>	4.46 (2.24)	4.11 (4.14)	0.41	0.01
Overall life satisfaction <sup>a</sup>	4.44 (2.45)	4.76 (3.29)	0.62	0.03

\* Significance on the 5% level

\*\* Significance on the 1% level

<sup>a</sup> Mean (SD)<sup>b</sup> Median<sup>c</sup> Only included in nursing version



clients without restraints. Being in restraints discriminated in capability Qol in the nursing version, even when correcting for the influence of other variables. This was not the case in the family version. In general, little agreement between the family and nursing versions was found for the different variables, raising important questions about which proxy version to consider superior or most reliable.

#### Methodological limitations

There are some noteworthy methodological limitations to our study. Ours has been a relatively small-scale study in a particular setting, limiting the generalization of results. This is especially true since collection of additional data on diagnosis and disease severity was not feasible, given limited space in the questionnaire. Only the functional consequences of the disease were measured through the CDS. Unfortunately, using dementia-specific Qol measures also was not possible. This study was performed alongside a real-world economic evaluation, and adding additional instruments to the questionnaire would have decreased the number of participating organizations even further due to the increased burden caused by the study. Therefore, we necessarily restricted the scope of this study, focusing on generic quality of life instruments that are particularly useful in economic evaluation. Another limitation concerns the use and interpretation of the HADS. This instrument to our knowledge has not been validated in people with severe dementia. We nonetheless opted for inclusion of the HADS because symptoms of depression and anxiety could be particularly important in the context of restraints. Also, HADS is widely used making comparisons to other populations straightforward. Moreover, it is a relatively short instrument compared to instruments like the Cornell Scale for Depression and Dementia [40], while showing similar reliability [40].

Moreover, given the limited number of respondents, we used multiple imputations to retain the full sample in the analyses. Multiple imputations allow for a more valid statistical inference [39] than full-case analysis, as long as only a small percentage of the data are imputed even if the assumption of multivariate normality is not met, as in this case. In the current study, imputed results are comparable with a full-case analysis (not shown here). OLS estimates in the nursing version had non-normally distributed error terms; in our analysis of it, we thus used robust estimation techniques. Clearly, therefore, repeating studies like this, using larger samples is encouraged.

#### Security dimension

The nursing version of the ICECAP-O discriminated between restrained and non-restrained clients on all

dimension levels except for security. This may be related to the fact that the scores on this item were relatively high. This was somewhat surprising since the average score on the security dimension was low in the study among general British elderly [8]. This difference may have to do with the study setting or item phrasing. Regarding the former, it is quite possible that nursing home clients suffering from dementia really did not seem worried. This may imply that nursing homes provide a safe environment. It is also possible, however, that these patients may not have been (seen as being) able to worry about or have a grasp of their future. In future (proxy) studies, it may be worthwhile to further investigate this by, for instance, using alternative wording. Then, the underlying reasons for indicating being able to think about the future without worry (i.e., because there is nothing to worry about or one has lost the ability to worry) can be distinguished. Additionally, the proximity to death of some clients may influence the security dimension for the proxies completing the questionnaire. We have received anecdotal information from some respondents that the proximity to death makes questions regarding the future difficult to answer. It is also noteworthy in this context that in another version of the ICECAP, the ICECAP-A, the wording of the security dimension reads “feeling settled and secure” rather than “thinking about the future” [41].

#### Clients in restraints

According to the nursing version of the ICECAP-O, clients in restraints are indeed worse off than non-restrained clients in terms of capabilities, indicating that physical freedom seems to be an empirically important element of Qol. This finding is in line with earlier studies that indicated that being restrained is not beneficial to the elderly [5, 42]. We do note that, since the current study did not measure cognition directly, only indirectly through the CDS, it is possible that unobserved differences in cognition between two groups may have influenced our results. Cognition, however, is not consistently identified as a predictor of using physical restraints [5]. Our study, in that sense, gives further rationale for efforts toward reducing the use of physical restraints in psycho-geriatric nursing homes [42].

#### Differences between proxies

The differences between the nursing and family versions of the questionnaire raise important, yet difficult to answer questions regarding suitable (and valid) proxies. The observed differences may well relate to a difference in reference points. Nursing staff might answer the questions



with similar clients in mind, while family members may assess the client's current capability Qol in relation to former capability Qol, i.e., before psycho-geriatric services were necessary. While both viewpoints can be relevant in their own right, for evaluating interventions aimed at improving the situation of clients in the care context, the nursing proxy seems the most logical choice.

An important limitation of the ICECAP-O to date is that its sensitivity to change has not been explored. Indirectly, our study provides some indication of it in that nursing proxies distinguish between restrained and non-restrained clients. The fact that family proxies apparently did not may strengthen the choice for using nursing staff as proxies. Still, it is necessary to test and explain the discrepancy between proxies further, also in relation to sensitivity to change.

### Decision making and transferability of tariffs

Concerning the use of ICECAP-O in cost-utility studies, it should be noted that on a theoretical level, the ICECAP instruments are rooted in capability theory rather than utility theory. In capability theory, developed by Sen [7], people's wellbeing is measured in terms of their capacity to perform certain actions and achieve certain states [7]. Its prescription for societal redistribution may be seen as maximizing capabilities or as guaranteeing basic capabilities for everyone [43]. Until recently, the approach did not have an empirically tested, well-defined list of capabilities [44], which can be a weakness if societal redistribution is an issue. It would also be possible to use capability-based Qol instruments in cost-effectiveness studies [45]. In such an evaluation, the final outcome would be based on capability attributes instead of HrQol attributes, allowing the computation of "capability QALYs" [6]. Such an approach could be considered to be consistent with the extra-welfarist framework [46] underlying cost-effectiveness analysis, which allows the broadening of the evaluative space to include (also) non-utility information [46]. On the other hand, there is considerable theoretical and empirical uncertainty about how such an approach might work [8, 45] with respect to the valuation of health and capabilities.

Our study used the British tariffs to compute capability valuation since Dutch tariffs are not (yet) available. Using Dutch tariffs would probably not have led to vastly different results, since, in the nursing proxy questionnaire, already four of the five dimensions of the ICECAP-O had significantly different scores for the restrained versus non-restrained group. Still, the weights attached to different capabilities may vary between countries.

Besides the problem of tariffs, the transferability of the capability dimensions themselves can also be a point of

discussion. According to Sen [7], who does not list specific capabilities, relevant capabilities should be tailored to the local population and hence generating a list should be performed on a more local level. On the contrary, Nussbaum [43] proposed that basic capabilities exist and can be used globally. Since the capability measure is a possible outcome used in optimization and redistributive policies, using a standardized descriptive system across health systems (and countries) to evaluate similar interventions aimed at basic capabilities is clearly advantageous. On the other hand, specific (non-basic) capabilities may be valuable for the relevant target group of a particular intervention. The issue here is whether the ICECAP-O measures basic capabilities, or at least transferable capabilities, or more specifically capabilities important to British elderly. The fact that the dimensions of the ICECAP-O resemble frequently reported universal subjective well-being measures [47] is indicative of the former, although the physical dimension is not measured directly. It seems, therefore, that the ICECAP-O is suitable as a more generic outcome measure in elderly care. As such, it may assist decision makers to make choices based on ensuring and enhancing basic capabilities for this group.

### Conclusion

The ICECAP-O instrument appears to be a promising tool for use in evaluations of interventions in psycho-geriatric care that do not necessarily or primarily improve health. The nursing proxy version of the questionnaire particularly demonstrated convergent and discriminant validity. Future research will have to confirm these findings in other settings, with particular attention paid to dementia severity, diagnosis, and validation alongside dementia-specific Qol measures. Additional research is also required on (1) the ICECAP-O's sensitivity to change, especially in evaluating interventions, (2) the relationship between overall quality of life, utilities, and capabilities for different settings, and (3) eliciting valid proxy information. With respect to the clients involved in this study, the ICECAP-O makes it clear that interventions aimed at removing restraints may well be worthwhile if capabilities are deemed important.

**Acknowledgments** We are grateful to Hareth Al-Janabi and Mathilde Strating for their useful suggestions on earlier versions of this paper, as well as the anonymous reviewers for their suggestions on the current one. This study was financed by ZonMw in the context of the Care for Better evaluation project.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

**Appendix: The ICECAP-O instrument proxy version**

General instructions for the proxy questionnaire

This appendix highlights the ICECAP-O instrument proxy version used in this study. It was based on the original ICECAP-O version, as developed by Joanna Coast and Terry Flynn, and introduced and validated in [9]. The original version can be found at <http://www.icecap.bham.ac.uk/documents/icecapquest.doc>.

We would like to ask you to fill out the questions below for the client/family member. Please try to answer the questions in manner as the client/family member would if he/she would be able to answer the questions. With every question, please tick the answer that the client/family member would give.

**ABOUT THE CLIENT'S QUALITY OF LIFE**

By placing a tick (✓) in ONE box in EACH group below, please indicate which statement best describes the clients quality of life at the moment.

<p><b>1. Love and Friendship</b></p> <p>The client can have all of the love and friendship that he/she wants <input type="checkbox"/> 4</p> <p>The client can have a lot of the love and friendship that he/she want <input type="checkbox"/> 3</p> <p>The client can have a little of the love and friendship that he/she wants <input type="checkbox"/> 2</p> <p>The client cannot have any of the love and friendship that he/she wants <input type="checkbox"/> 1</p>	<p><b>Tick one box only in each section</b></p>
<p><b>2. Thinking about the future</b></p> <p>The client can think about the future without any concern <input type="checkbox"/> 4</p> <p>The client can think about the future with only a little concern <input type="checkbox"/> 3</p> <p>The client can only think about the future with some concern <input type="checkbox"/> 2</p> <p>The client can only think about the future with a lot of concern <input type="checkbox"/> 1</p>	
<p><b>3. Doing things that make you feel valued</b></p> <p>The client is able to do all of the things that make him/her feel valued <input type="checkbox"/> 4</p> <p>The client is able to do many of the things that make him/her feel valued <input type="checkbox"/> 3</p> <p>The client is able to do a few of the things that make him/her feel valued <input type="checkbox"/> 2</p> <p>The client is unable to do any of the things that make him/her feel valued <input type="checkbox"/> 1</p>	
<p><b>4. Enjoyment and pleasure</b></p> <p>The client can have all of the enjoyment and pleasure that he/she wants <input type="checkbox"/> 4</p> <p>The client can have a lot of the enjoyment and pleasure that he/she wants <input type="checkbox"/> 3</p> <p>The client can have a little of the enjoyment and pleasure that he/she wants <input type="checkbox"/> 2</p> <p>The client cannot have any of the enjoyment and pleasure that he/she wants <input type="checkbox"/> 1</p>	
<p><b>5. Independence</b></p> <p>The client is able to be completely independent <input type="checkbox"/> 4</p> <p>The client is able to be independent in many things <input type="checkbox"/> 3</p> <p>The client is able to be independent in a few things <input type="checkbox"/> 2</p> <p>The client is unable to be at all independent <input type="checkbox"/> 1</p>	

© Original: Joanna Coast & Terry Flynn

## Tariffs and scoring algorithm in STATA

The ICECAP-O has 5 attribute dimensions each having 4 levels, 44444 representing full capability and 11111 representing no capability. The tariffs for the ICECAP-O, based on the preferences of the 65+ population in the United Kingdom, were presented by Coast et al. [8]. The tariffs basically provide preference weights for the different ICECAP-O states, which are normalized in such a way that 0 represents the worst situation described on the ICECAP-O (11111) and 1 represents the best situation described on the ICECAP-O. Lower scores thus represent fewer, preference-based capabilities.

The complete explanation as to how to calculate them is fully described on the ICECAP website: <http://www.icecap.bham.ac.uk/tariffs.shtml>.

## References

- Drummond, M., Schulpher, M., Torrance, G., O'Brien, B., & Stoddart, G. (2005). *Methods for the economic evaluation of healthcare programmes* (p. 379). Oxford: Oxford University Press.
- Ryan, M., Netten, A., Skatun, D., & Smith, P. (2006). Using discrete choice experiments to estimate a preference-based measure of outcome—an application to social care for older people. *Journal of Health Economics*, *25*, 927–944.
- Makai, P., Nieboer, A. B., Koopmanschap, M. A., Stolk, E., & Brouwer, W. B. F. (2011). *Quality of life instruments for economic evaluations of elderly care*. Draft.
- Nieboer, A. P., Koolman, X., & Stolk, E. A. (2010). Preferences for long-term care services: Willingness to pay estimates derived from a discrete choice experiment. *Social Science and Medicine*, *70*, 1317–1325.
- Hamers, J. (2005). Why do we use physical restraints in the elderly? *Zeitschrift für Gerontologie und Geriatrie*, *38*, 19.
- Cookson, R. (2005). QALYs and the capability approach. *Health Economics*, *14*, 817–829.
- Sen, A. (1982). *Choice, welfare and measurement* (p. 440). Cambridge, MA: Harvard University Press.
- Coast, J., Flynn, T. N., Natarajan, L., Sproston, K., Lewis, J., Louviere, J. J., et al. (2008). Valuing the ICECAP capability index for older people. *Social Science and Medicine*, *67*, 874–882.
- Coast, J., Peters, T. J., Natarajan, L., Sproston, K., & Flynn, T. (2008). An assessment of the construct validity of the descriptive system for the ICECAP capability measure for older people. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *17*, 967–976.
- Flynn, T. N., Louviere, J. J., Marley, A. A. J., Coast, J., & Peters, T. J. (2008). Rescaling quality of life values from discrete choice experiments for use as QALYs: A cautionary tale. *Population Health Metrics*, *6*. doi:10.1186/1478-7954-6-6.
- Grewal, I., Lewis, J., Flynn, T., Brown, J., Bond, J., & Coast, J. (2006). Developing attributes for a generic quality of life measure for older people: Preferences or capabilities? *Social Science and Medicine*, *62*, 1891–1901.
- Macdonald, A., & Cooper, B. (2007). Long-term care and dementia services: An impending crisis. *Age and Ageing*, *36*, 16–22.
- Kwasky, A. N., Harrison, B. E., & Whall, A. L. (2010). Quality of life and dementia: An integrated review of literature. *Alzheimer's Care Today*, *11*, 186–195.
- Banerjee, S., Samsi, K., Petrie, C. D., Alvir, J., Treglia, M., Schwam, E. M., et al. (2009). What do we know about quality of life in dementia? A review of the emerging evidence on the predictive and explanatory value of disease specific measures of health related quality of life in people with dementia. *International Journal of Geriatric Psychiatry*, *24*, 15–24.
- Brod, M., Stewart, A. L., Sands, L., & Walton, P. (1999). Conceptualization and measurement of quality of life in dementia: The dementia quality of life instrument (DQoL). *The Gerontologist*, *39*, 25–36.
- Ready, R. E. (2003). Quality of life measures for dementia. *Health and Quality of Life Outcomes*, *1*(11). doi:10.1186/1477-7525-1-11.
- Logsdon, R. G., Gibbons, L. E., McCurry, S. M., & Teri, L. (1999). Quality of life in Alzheimer's disease: Patient and caregiver reports. *Journal of Mental Health and Aging*, *5*, 21–32.
- Selai, C. C. E. (2001). Assessing quality of life in dementia: Preliminary psychometric testing of the quality of life assessment schedule (QOLAS). *Neuropsychological Rehabilitation*, *11*, 219–243.
- Smith, S. C., Lamping, D. L., Banerjee, S., Harwood, R. H., Foley, B., Smith, P., et al. (2007). Development of a new measure of health-related quality of life for people with dementia: DEMQOL. *Psychological Medicine*, *37*, 737.
- Ettema, T. P., Dröes, R., de Lange, J., Mellenbergh, G. J., & Ribbe, M. W. (2007). QUALIDEM: Development and evaluation of a dementia specific quality of life instrument. scalability, reliability and internal structure. *International Journal of Geriatric Psychiatry*, *22*, 549–556.
- Hoe, J., Katona, C., Roch, B., & Livingston, G. (2005). Use of the QOL-AD for measuring quality of life in people with severe dementia—the LASER-AD study. *Age and Ageing*, *34*, 130–135.
- Neyens, J. C. L., Dijcks, B. P. J., Twisk, J., Schols, J. M. G. A., van Haastregt, J. C. M., van den Heuvel, W. J. A., et al. (2009). A multifactorial intervention for the prevention of falls in psychogeriatric nursing home patients, a randomised controlled trial (RCT). *Age and Ageing*, *38*, 194–199.
- Pickard, A. S., & Knight, S. J. (2005). Proxy evaluation of health-related quality of life: A conceptual framework for understanding multiple proxy perspectives. *Medical Care*, *43*, 493–499.
- Strating, M. M. H., Zuiderent, T., Nieboer, A. P., & Bal, R. A. (2008). Evaluating the care for better collaborative. *Results of the first year of evaluation*, pp. 1–132.
- Strating, M. M. H., Nieboer, A. P., Zuiderent-Jerak, T., & Bal, R. A. (2011). Creating effective quality-improvement collaboratives: A multiple case study. *BMJ Quality and Safety*, *20*, 344–350.
- Burgerlijke wetboek [civil code of the Netherlands]. (1992). Article 446–Article 468.
- Miller, F. G., & Emanuel, E. J. (2008). Quality-improvement research and informed consent. *New England Journal of Medicine*, *358*, 765–767.
- EuroQol. (1990). A new facility for the measurement of health-related quality of life. The EuroQol group. *Health Policy*, *16*, 199–208.
- Jonsson, L., Andreasen, N., Kilander, L., Soininen, H., Walde- mar, G., Nygaard, H., et al. (2006). Patient- and proxy-reported utility in Alzheimer disease using the EuroQoL. *Alzheimer Disease and Associated Disorders*, *20*, 49–55.
- Cantril, H. (1966). *The pattern of human concerns*. New Brunswick: Rutgers University Press.
- Bengtsson-Tops, A., Hansson, L., Sandlund, M., Bjarnason, O., Korkeila, J., Merinder, L., et al. (2005). Subjective versus interviewer assessment of global quality of life among persons with schizophrenia living in the community: A Nordic multicentre study. *Quality of Life Research*, *14*, 221–229.

32. Veenhoven, R. (2010). World database of happiness. From <http://worlddatabaseofhappiness.eur.nl>. Assessed 3 Jan 2010.
33. Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the hospital anxiety and depression scale: An updated literature review. *Journal of Psychosomatic Research*, *52*, 69–77.
34. Carod-Artal, F. J., Coral, L. F., Trizotto, D. S., & Moreira, C. M. (2009). Self- and proxy-report agreement on the stroke impact scale. *Stroke*, *40*, 3308–3314.
35. Spinhoven, P., Ormel, J., Sloekers, P. P., Kempen, G. I., Spekens, A. E., & Van Hemert, A. M. (1997). A validation study of the hospital anxiety and depression scale (HADS) in different groups of dutch subjects. *Psychological Medicine*, *27*, 363–370.
36. Dijkstra, A., Buist, G., & Dassen, T. (1996). Nursing-care dependency. Development of an assessment scale for demented and mentally handicapped patients. *Scandinavian Journal of Caring Sciences*, *10*, 137–143.
37. Valimaki, M., Dijkstra, A., Soini, H., M-, Antila., Zanotti, R., & Dassen, T. (2004). Elderly patient's care dependency. *Hallym International Journal of Aging*, *6*, 79–93.
38. Dijkstra, A., Coleman, M., Tomas, C., Valimaki, M., & Dassen, T. (2003). Cross-cultural psychometric testing of the care dependency scale with data. *Journal of Advanced Nursing*, *43*, 181–187.
39. Lee, K. J., & Carlin, J. B. (2010). Multiple imputation for missing data: Fully conditional specification versus multivariate normal imputation. *American Journal of Epidemiology*, *171*, 624–632.
40. Alexopoulos, G. S., Abrams, R. C., Young, R. C., & Shamoian, C. A. (1988). Cornell scale for depression in dementia. *Biological Psychiatry*, *23*, 271–284.
41. Al-Janabi, H., Flynn, T., & Coast, J. (2011). Development of a self-report measure of capability wellbeing for adults: The ICE-CAP-A. *Quality of Life Research*.
42. Huizinga, A. R., Hamers, J. P. H., Gulpers, M. J. M., & Berger, M. P. F. (2009). A cluster-randomized trial of an educational intervention to reduce the use of physical restraints with psychogeriatric nursing home residents. *Journal of the American Geriatrics Society*, *57*, 1139–1148.
43. Nussbaum, M. (Ed.). (1993). *The quality of life* (p. 453). Oxford: Clarendon Press.
44. Robeyns, I. (2005). The capability approach: A theoretical survey. *Journal of Human Development*, *6*(1), 93–114.
45. Coast, J., Smith, R., & Lorgelly, P. (2008). Should the capability approach be applied in health economics? *Health Economics*, *17*, 667–670.
46. Brouwer, W. B. F., Culyer, A. J., van Exel, N. J. A., & Rutten, F. F. H. (2008). Welfarism vs. extra-welfarism. *Journal of Health Economics*, *27*, 325–338.
47. Gullone, E., & Cummins, R. A. (Eds.). (2002). *Universality of subjective wellbeing indicators* (p. 196). Dordrecht: Kluwer.