



Review

Quality of life instruments for economic evaluations in health and social care for older people: A systematic review



Peter Makai^{a,*}, Werner B.F. Brouwer^b, Marc A. Koopmanschap^b, Elly A. Stolk^b, Anna P. Nieboer^b

^a Radboud University Medical Center, Nijmegen, The Netherlands

^b Institute of Health Policy and Management, Erasmus University Rotterdam, The Netherlands

ARTICLE INFO

Article history:

Available online 4 December 2013

Keywords:

Quality of life
Cost-utility analysis
Older people
Long-term care
Review

ABSTRACT

Gaining health may not be the main goal of healthcare services aimed at older people, which may (also) seek to improve wellbeing. This emphasizes the need of finding appropriate outcome measures for economic evaluation of such services, particularly in long-term care, capturing more than only health-related quality of life (HrQoL). This review assesses the usefulness of HrQoL and wellbeing instruments for economic evaluations specifically aimed at older people, focusing on generic and preference-based questionnaires measuring wellbeing in particular.

We systematically searched six databases and extracted instruments used to assess HrQoL and wellbeing outcomes. Instruments were compared based on their usefulness for economic evaluation of services aimed at older people (dimensions measured, availability of utility scores, extent of validation).

We identified 487 articles using 34 generic instruments: 22 wellbeing (two of which were preference-based) and 11 HrQoL instruments. While standard HrQoL instruments measure physical, social and psychological dimensions, wellbeing instruments contain additional dimensions such as purpose in life and achievement, security, and freedom.

We found four promising wellbeing instruments for inclusion in economic evaluation: Ferrans and Powers QLI and the WHO-QoL OLD, ICECAP-O and the ASCOT. Ferrans and Powers QLI and the WHO-QoL OLD are widely validated but lack preference-weights while for ICECAP-O and the ASCOT preference-weights are available, but are less widely validated. Until preference-weights are available for the first two instruments, the ICECAP-O and the ASCOT currently appear to be the most useful instruments for economic evaluations in services aimed at older people. Their limitations are that (1) health dimensions may be captured only partially and (2) the instruments require further validation. Therefore, we currently recommend using the ICECAP-O or the ASCOT alongside the EQ-5D or SF-6D when evaluating interventions aimed at older people.

© 2013 Elsevier Ltd. All rights reserved.

Introduction

The growing number of older people worldwide and the associated higher demand for healthcare increasingly put pressure on public funds. Hence, there is growing need to make funding decisions about various health and social services aimed at older people. **Cost-utility analysis (CUA)** can support policy makers to optimally allocate health and social care resources within limited budgets by comparing two or more healthcare interventions to investigate their relative value for money (Drummond, Sculpher, Torrance, O'Brien, & Stoddart, 2005). CUA is increasingly used in

the curative sector for such comparisons. In CUA, the benefits of these interventions are commonly expressed in Quality-Adjusted Life-Years (QALYs), a utility-based health measure comprising both length and health-related quality of life (HrQoL). To assess HrQoL improvements, typically patients' health states are measured (using standardized instruments) using health dimensions such as mobility, pain and anxiety. Subsequently, these health states are valued (on a scale from 0 – dead – to 1 – perfect health). Such outcome measures are appropriate for curative services, where the goal is to improve health. However, in other fields of healthcare, such as mental health, social care, public health, and care for older people, a focus on health dimensions of quality of life (QoL) may be less appropriate if health improvement is not the only or even the main goal of the services provided (Al-Janabi, Flynn, & Coast, 2012). A relevant question is how to broaden the scope of

* Corresponding author.

E-mail address: Peter.Makai@radboudumc.nl (P. Makai).

outcome measurement within a CUA to include QoL domains that are intentionally affected by interventions in other fields of healthcare, in particular care for older people.

Current QALY measures using a quality adjustment factor that is based on domains of HrQoL only, may not be appropriate to evaluate interventions for older people such as long-term care. This holds since the latter interventions may be aimed at improving non-health aspects of QoL, such as maintaining independence, dignity, comfort or social interaction. Evaluating such interventions using HrQoL-instruments would likely undervalue the benefits. One of the most important challenges for performing CUA in the context of interventions aimed at older people thus concerns the availability of outcome measures attuned to the goals of services consumed by older people (Coast, Flynn, et al., 2008). The aim of this paper was to review the literature to investigate the existence of such appropriate outcome measures, which would facilitate CUA in the context of health and social care for older people.

Older people consume a variety of health and social services. These may be curative services such as hospital care, as well as long-term care services provided by nursing homes, residential homes, and home care. Often, elderly consume a combination of such services within an illness episode. The benefits of such a varied list of services should be evaluated using outcome measures that adequately capture the value of all services provided (Coast, Flynn, et al., 2008). This may be particularly difficult in long-term care. To illustrate this, consider an intervention aimed at reducing the frequency of restraining older people in a nursing home setting to prevent them from falling (Huizing, Hamers, Gulpers, & Berger, 2009). While reducing the use of physical restraints may not directly improve a patient's health (Huizing et al., 2009), such an intervention aims to restore dignity, freedom of movement, and control, outcomes that transcend health. If such an intervention were to be evaluated in a CUA, it is pivotal that outcome measures allow for capturing benefits 'beyond health' in order to provide adequate information on the costs and benefits of the intervention. Below we discuss some of the desirable characteristics of such instruments.

A first desirable characteristic of instruments attuned for evaluation of care for older people, is that such instruments should capture QoL dimensions transcending health. HrQoL instruments commonly used in CUAs measure health as a multi-dimensional construct minimally measuring psychological, physical and social dimensions (WHO, 1948), while for economic evaluation of services aimed at older people, particularly in long-term care other dimensions may also be relevant, such as affection or control. Instruments covering such dimensions 'beyond health' can be labeled as wellbeing instruments. There are two main conceptualizations relevant for the scope of wellbeing instruments. The first one focuses on wellbeing as an inherently subjective concept and thus holds that wellbeing does not contain health dimensions (Morgan, Grootendorst, Lexchin, Cunningham, & Greyson, 2011). By distinguishing between functional HrQoL dimensions and subjective wellbeing dimensions, both HrQoL and wellbeing are components of the overarching concept of QoL. The second conceptualization treats wellbeing as representing individuals' welfare (Nussbaum, 1993), which is dependent on individuals' functioning, thus encompassing HrQoL dimensions (see Fig. 1). In this view, wellbeing can be seen as synonymous with overall QoL. In this paper, wellbeing will be referred to in the latter meaning.

This second conceptualization may offer the opportunity to jointly explore treatment effects on health with other impacts on wellbeing. By broadening the evaluative space of a CUA (Coast, Flynn, et al., 2008), wellbeing instrument are, in principle, better equipped than HrQoL measures to capture the full benefit of interventions aimed at older people, also when these aim at

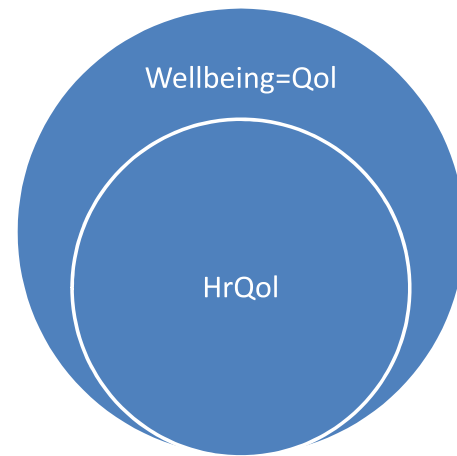


Fig. 1. Conceptualization of wellbeing and Quality of life.

outcomes beyond health. However, wellbeing instruments based on the subjective notion of wellbeing may not explicitly or completely capture health. This deserves attention, since the aim must be to adequately capture all relevant outcomes of interventions in order to come to a complete comparison of costs and benefits in an economic evaluation. While some wellbeing instruments may include health as an underlying concept (Hyde, Wiggins, Higgs, & Blane, 2003), it remains unclear whether existing outcome measure capture all wellbeing domains adequately and in such a way that allows inclusion in CUAs. To overcome this problem, it has been suggested that combinations of HrQoL and wellbeing instruments could be used in economic evaluations in older people (Davis, Liu-Ambrose, Richardson, & Bryan, 2013). Moreover, the lines between HrQoL and wellbeing measures may not always be easy to draw nor have been consistently drawn (when definitions of HrQoL or wellbeing differ between measures). Therefore, in reviewing measures that may be useful in economic evaluation of services aimed at older people, particularly in long-term care, we will include both measures labeled as HrQoL as well as measures of wellbeing. This allows an open and consistent categorization of instruments.

A second desirable characteristic of outcome measures for application in CUA in older people is that the classification system of health or well-being states is combined with a preference-based scoring system, as is the case for popular HrQoL instruments like the EQ-5D and SF-36. Preference-based instruments normally consist of (1) a descriptive system defined by the dimensions and answer categories of the instruments (states), and (2) a (pre-scored) weighting system reflecting the valuation of the states described with element (1) by a relevant population (e.g. general public) (Brazier, Ratcliffe, Salomon, & Tsuchiya, 2007). The weighting system thus allows particular states as described with the descriptive system to be transformed into a 'utility score', commonly reflecting the average strength of preference for the various states described. In case of HrQoL, these scores are typically anchored to a standardized scale, with 1 representing the utility of the best imaginable health state, and 0 representing the value for the state 'dead'. Negative values relate to health states valued as 'worse than dead'. For wellbeing instruments, anchoring on a 0–1 scale is also possible, 1 representing the best imaginable wellbeing instead of best imaginable health, while 0 can represent 'dead' or, more logically perhaps, the value for the worst level of all included domains in the descriptive system. Additionally, negative values for wellbeing instruments may also be allowed depending on the theory behind the instrument. Here, we will not limit our search to

preference based instruments of wellbeing, as it is imaginable that utilities are attached in a later stage to promising measures of wellbeing that are currently not preference based (similar to development of SF-6D from SF-36 (Brazier, Roberts, & Deverill, 2002)).

A third desirable characteristic of instruments to be used in evaluations in older people is that their feasibility of use and psychometric properties are well-established (Brazier et al., 2007). Instruments which measure what they intend to measure and those which do so with a smaller error seem to be more preferable to instruments lacking such properties. This aspect will also be considered in the review.

The number of instruments developed specifically to address and evaluate outcomes of healthcare services targeted at older people is growing. However, guidance is lacking on which instruments can or should be used for CUA of (long-term care) services aimed at older people. Such guidance depends on knowledge regarding the existing instruments, their ability to capture relevant outcomes and their feasibility and validity. The results of this review allow us to formulate (preliminary) advice on the choice of outcome measure for conducting economic evaluation of services aimed at older people.

Hence, we set out to perform a systematic literature search to identify generic outcome measures used in older people, which are applicable to all people irrespective of the type and nature of diseases they have, thus, in principle, facilitating comparisons between people, treatments and services. Thus, we excluded disease specific instruments, which specifically aim to measure HrQoL in well-defined populations. Instruments were included irrespective of whether they were labeled as HrQoL or wellbeing instruments, and whether they were preference-based or not. When possible, the usefulness of these instruments for CUA was assessed. For the review we used a structured, three-step approach. First, we extracted the relevant generic HrQoL and wellbeing instruments used in the studies. Second, we assessed their current and potential degree of suitability for economic evaluations in older people, particularly in long-term care. Finally, we examined the most promising instruments in more detail, with specific attention paid to their psychometric properties.

Method

Search methods

Database sources

We searched Pubmed, EMBase and CINALH for the English-language literature using the same keywords. In addition to standard medical and healthcare databases, we searched major psychological, sociological, and economic databases, namely Psycinfo, and Econlit, and Social Science Citation Abstracts to account for the multidisciplinary nature of QoL research (Davis et al., 2013).

Search terms

We used the following search terms, their synonyms and their combinations to find validated, generic, preference-based HrQoL and wellbeing measures: “elderly”, “older”, “geriatric”, “quality of life”, “HrQoL”, “wellbeing”, “validated questionnaire”, “validated measure”, “utilities”, and “preference-based valuation”. For a complete description of the search terms see Appendix 1. The search strategy was customized for all databases.

Selection criteria

Selected articles met the following inclusion criteria. First, HrQoL or wellbeing was an explicit outcome measure of an empirical article or validation study and its target population was

characterized by ‘older persons’ or ‘elderly’, above 65 years. Second, the articles were written in English, measured HrQoL or wellbeing through a questionnaire, and were published after 2000. Third, we excluded studies (1) using an instrument measuring only symptoms or instruments measuring only one dimension, (2) reported decision-analytical modeling, or (3) used disease-specific measures. Finally, we investigated the selected articles’ reference lists to identify the original development articles of the instrument used in the identified article.

Evaluation of the instruments

The following aspects of the instruments were evaluated in June 2012 for all identified instruments: scope of the instrument (HrQoL or wellbeing), dimensions measured, availability of utility tariffs, and frequency of use. PM evaluated the titles and abstracts of all the studies, assessed articles for inclusion and exclusion; AN has independently checked the accuracy of this assessment. PM and AN then classified instruments independently. Differences in opinion were resolved by consensus.

We examined all selected instruments to identify those potentially useful for economic evaluations in older people, particularly in long-term care. Since there is no consensus definition of wellbeing or HrQoL across disciplines (health science, psychology, sociology, economics) in order to classify instruments we examined their operationalizations. For our purpose, we classified instruments as HrQoL or wellbeing according to the following definitions: HrQoL instruments measure health as minimally measuring psychological, physical and social dimensions (WHO, 1948), while wellbeing instruments measure broader QoL domains (as well). We also classified the instruments by the availability of utility tariffs and frequency of use.

Finally, we investigated the most promising instruments (preference-based or commonly used wellbeing instruments) for economic evaluation in more detail, looking at feasibility of use in older people, paying attention to age-related cognitive decline and psychometrics in order to determine which instrument had the most potential for actual usage in economic evaluations in older people. Psychometrics were evaluated according to the criteria outlined for the critical appraisal of psychometric properties (McDowell & Newell, 2006). McDowell distinguishes between the thoroughness and the results of instrument validity and reliability. Validity can be defined as the extent to which a test measures what it

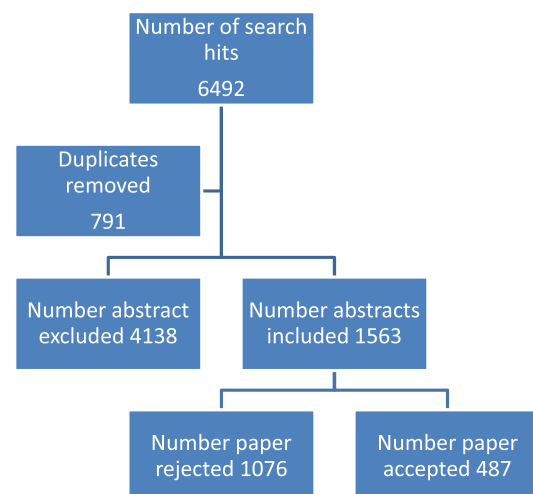


Fig. 2. Search strategy.

Table 1
HrQoL and wellbeing instruments with their classifications.

Instrument	Scope	Dimensions	Utilities	Number of studies validated in elderly
Health utility index 2 (Torrance et al., 1996) ^a	HrQoL	Sensation, mobility, emotion, cognitive, self-care, pain, fertility	Yes	Many 32
Health Utility Index 3 (Feeny et al., 2002)	HrQoL	Vision, hearing, speech, ambulation, dexterity, emotion, cognition, pain	Yes	Many 11
Quality of well-being (Kaplan et al., 1978)	HrQoL	Mobility, physical activity, Social activity, symptoms	Yes	Several 5
Herdecke Quality of life questionnaire (Ostermann et al., 2005)	HrQoL	Initiative power and interest Social interaction Mental balance Mobility Physical complaints Digestive well-being	No	Single 1
Duke health profile (Novella et al., 2001)	HrQoL	physical, mental, social, general, perceived health, self-esteem, anxiety, depression, pain, disability	No	Single 1
Nottingham health profile (Baro et al., 2006)	HrQoL	Energy level, pain, Emotional reaction, sleep, social isolation, physical abilities	No	Several 5
Sickness Impact Profile (Rothman et al., 1989)	HrQoL	sleep and rest, emotional behavior, body care and movement, home management, mobility, social interaction, ambulation, alertness behavior, communication, work, recreation and pastimes, and eating	No	Few 3
Assessment of Quality of life (Osborne et al., 2003)	HrQoL	illness, psychological wellbeing, physical senses, social relationship, independent living	Yes	Several 5
SF-6D (Brazier et al., 2002)	HrQoL	physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, mental health	Yes	Many 189
(SF-36, Sf-12 SF-8)				
EQ5D (The EuroQoL, 1990)	HrQoL	anxiety/depression, mobility, self-care, usual activities, pain, discomfort	Yes	Many 117
15D (Sintonen, 2001)	HrQoL	Mobility, vision, hearing, breathing, sleeping, eating, speech, elimination, usual activities, mental function, discomfort and symptoms, depression, distress, vitality, sexual activity health and functioning, social and economic, psychological/spiritual, family	Yes	Many 10
Ferrans and Powers QLI (Ferrans & Powers, 1992)	Wellbeing	attachment, security, role, enjoyment, control	No (but weighted)	Several 6
ICECAP-O (Coast et al., 2008)	Wellbeing	food and nutrition, personal care, safety, social participation and involvement, control over daily living	Yes	Few 4
OPUS (Ryan et al., 2006)	Wellbeing	Control over daily life, Personal cleanliness and comfort, Food and drink, Accommodation cleanliness and comfort, Safety, social participation, Occupation, Dignity, Living situation	Yes	Single 1
ASCOT (Malley et al., 2012)	Wellbeing	Control, autonomy, self-realization and pleasure	No	Single 1
CASP-19 (Hyde et al., 2003)	Wellbeing	Life overall Health and functioning Social relationships Leisure and social activities Independence, control over life, freedom, Home and neighborhood Psychological and emotional well-being Financial circumstances Religion and culture	No	Single 1
Older people quality of life profile (Bowling, 2009)	Wellbeing		No	Single 1

Table 1 (continued)

Instrument	Scope	Dimensions	Utilities	Number of studies validated in elderly
WHO-Qol Old (Peel et al., 2007)	Wellbeing	sensory abilities, autonomy, past, present, future activities, death and dying	No	Several 7
WHO-Qol Bref (Harper, 1998)	Wellbeing	Overall Qol, general health, physical, psychological, social relationships, Environment	No	Many 13
WHO-Qol 100 (Power, 1998)	Wellbeing	physical health, psychological health, social relationships, independence, environment, spirituality	No	Many 5
Comprehensive quality of life Scale (Gullone, 1999)	Wellbeing	material well-being health, productivity, intimacy, safety, place in community, emotional well-being	No	Single 1
Personal Well-being index (Cummins et al., 2003)	Wellbeing	satisfaction with health, personal relationships, community, overall satisfaction, standard of living, achievement, safety, future security, spirituality	No	Single 1
Interactive Computerized Quality of life Scale (ICQOL)(SF) (Jamison et al., 2007)	Wellbeing	overall life satisfaction, day-to day functioning, self-esteem, health status, energy level, Work, Home Life, Social Life, Interacting with others, Body Image, Mood, Effect of Stress/Fear, Sense of Achievement, Life Expectancy, Aches/Pains, Sleep/Rest Comfort, Illnesses, Activity Level, Sex Life, Stamina, Pleasures	No	Few 2
MANSA (Manchester Short Assessment of Quality of Life) (Bjorkman & Svensson, 2005)	Wellbeing	satisfaction with life as a whole, job, financial situation, friendships, leisure activities, accommodation, personal safety, people that the person lives with, family, health, mental health	No	Single 1
SPF-IL (Nieboer et al., 2005)	Wellbeing	Behavioral confirmation, affection, status, comfort, stimulation	No	Single 1
McGill quality of life scale (Cohen et al., 1995)	Wellbeing	Physical well-being Physical symptoms Psychological Existential, Support	No	Few 2
Quality of Life Questionnaire (Ruiz & Baca, 1993)	Wellbeing	Social support, General satisfaction, Physical well-being, free time	No	Single 1
Quality of life inventory (Frisch et al., 1992)	Wellbeing	self-esteem, goals/values, health, learning, work, creativity, play, helping, friends, neighborhood, community, home, children, love, money, relatives	No	Few 2
Social Wellbeing of Nursing home residents-scale (Gerritsen et al., 2010)	Wellbeing	affection, behavioral confirmation, status social wellbeing	No	Single 1
Quality of Life Scale (Flanagan, 1982)	Wellbeing	Material and physical wellbeing, relationships with other people, social, community and civic activities, personal development and fulfillment, recreation	No	Single 1
Quality of life in elders with multiple morbidities (Holzhausen et al., 2010)	Wellbeing	Family Own health Friendship Cognitive abilities Mobility and physical functioning Hobbies Social contacts Cultural and aesthetic matters Developing new abilities	No	Single 1

(continued on next page)

Table 1 (continued)

Instrument	Scope	Dimensions	Utilities	Number of studies validated in elderly
		Own abode (Social) participation Finances Spouse (Social) commitment Well-being and sensual experience Autonomy and self-determination Weltanschauung/philosophy Travel Incontinence/continence		
Life in General scale (Campbell et al., 1976)	Wellbeing	General satisfaction, Fatigue, fear, anxiety, unhappy, depressed, shaking/trambling loneliness, friends, social life, world too complicated	No	Single 1
Personal wellbeing index ((IWbG), 2006)	Wellbeing	standard of living, health, achievements in life, relationships, safety, community connectedness, and future security, spirituality/religion	No	Few 2
National wellbeing index (Cummins et al., 2003)	Wellbeing	country's economic situation, state of the environment, social conditions, government, business, and national security	No	Single 1
Satisfaction with Life scale (Diener et al., 1985)	Wellbeing	General satisfaction, life conditions, life close to ideal, goal achievement, lack of regret	No	Single 1

Few: 1–4 studies; Several: 5–8 studies; Many: more than 9 studies.

^a Full references of the HrQoL and wellbeing instruments development papers can be found in [Appendix 2](#).

intends to measure. The major forms of validity are **content validity**, which is assessed by the dimensions present; **construct validity**, which is tested in terms of **convergent** and **discriminant validity**; and **sensitivity to change**. **Reliability is the capacity of a measure to provide consistent and stable estimates.** Reliability has two major forms relevant here: responsiveness over time (**test retest reliability**), and **internal consistency** in case of multiple factor variables.

Results

Search outcome

Fig. 2 shows the search strategy. From the initial 6492 hits, we removed 791 duplicates. From the remaining articles, 1563 met the inclusion criteria. Of these, 1076 were subsequently assessed as meeting the exclusion criteria. The systematic search therefore resulted in 487 included articles. We identified 34 generic (non-disease specific) HrQoL and wellbeing instruments in these articles, which were subsequently evaluated.

Dimensions measured

The systematic search uncovered **34 generic** (i.e. non-disease-specific) instruments of which **23 were classified as wellbeing** instruments and **11 as HrQoL instruments**. The HrQoL instruments operationalized health with an average of eight dimensions, which pertained to only physical, psychological, and/or social functioning. Therefore, the HrQoL instruments in **Table 1** did not meet our definition of wellbeing.

The wellbeing instruments usually measured some health dimensions (**WHO, 1948**) also included in HRQoL instruments, next to broader domains of QoL. The exception was CASP, an instrument exclusively measuring non-health dimensions of QoL (**Hyde et al., 2003**). These additional, non-health dimensions included in the identified wellbeing instruments could be classified into four main concepts: (1) **purpose in life and achievement** (wishes, goals, values, spirituality, self-realization, activity

level, achievements, work); (2) **worries about security and safety** (present and future); (3) **financial well-being** (money, financial situation, standard of living); and (4) **personal freedom** (control, autonomy, independence). Less frequently mentioned dimensions were related to **pleasure; creativity and play**, or **related to the environment such** as physical environment, community, and neighborhood.

Preference-based instruments

We found five HrQoL measures (EQ5D, AQoL, Quality of Well Being, SF-36 and HUI) and two wellbeing measures (ASCOT, formerly the OPUS, and ICECAP-O) for which utility scores existed (**Coast, Flynn, et al., 2008; Drummond et al., 2005; Forder & Caiels, 2011; Malley et al., 2012; Osborne, Hawthorne, Lew, & Gray, 2003; Potoglou et al., 2011; Ryan, Netten, Skatun, & Smith, 2006**). The ASCOT was developed based on earlier experience with the OPUS instrument. The development of the latter instrument was not without problems (**Ryan et al., 2006**), e.g. in a first valuation study no differences in utility scores for the different levels of OPUS's safety dimension were detected. For the ASCOT, however, the development, validation and valuation of the instrument seemed successful with the following eight dimensions included: control over daily life, personal cleanliness and comfort, food and drink, accommodation cleanliness and comfort, safety, social participation, occupation, dignity and an additional question on living situation. The ICECAP-O was developed in several steps, jointly leading to the currently available instrument (**Coast, Flynn, et al., 2008**) with five dimensions: attachment, security, role, enjoyment, control. The ASCOT (with the exception of the dichotomous 'living at home' dimension) and the ICECAP-O both use four answering levels per included domain (**Coast, Flynn, et al., 2008; Forder & Caiels, 2011**). As for the preference elicitation techniques and anchoring used in the two wellbeing instruments, both used discrete choice experiments (DCE), with the ICECAP-O using Best-Worse Scaling (BWS), while the ASCOT used both BWS as well as the more traditional DCE to elicit preferences for health and wellbeing states (**Netten et al., 2012**). The ICECAP-O

Table 2
Potentially most relevant instruments for CUA in elderly care.

	ICECAP-O	ASCOT	Ferrans and Powers QLI	WHO-Qol old + WHO-Qol Bref
&Dimensions				
Physical	–	+	+	+
Psychological	+	–	+	+
Social	+	+	+	+
Purpose in life and achievement	+	+	+	+
Financial	–	–	+	–
Security	+	+	–	–
Personal Freedom	+	+	–	+
Psychometrics and other criteria				
Validation	+	–	+	+
*Reliability – Thoroughness	0	0	+++	++
**Reliability – Results	0	0	+++	+++
*Validity – Thoroughness	++	+	+++	++
**Validity – Results	++	++	+++	+++
Item number	5	9	64	24 + 25
Useful in cognitively declined populations	+	+	–	–
Utilities	+	+	–	–

&*+ = included in the instruments, – = not included in the instrument.

*In case of thoroughness, 4 categories are distinguished: 0 = no reported evidence of reliability or validity, + = very basic information only, ++ = several types of tests, or several studies have reported reliability or validity +++ = all major forms of reliability/validity tested.

**In case of results of the validation, the categories are: 0 = no numerical results reported, + = weak reliability/validity, ++ = adequate reliability/validity, +++ = excellent reliability/validity.

utility scores were normalized with 0 indicating no capabilities, while 1 denotes full capabilities (Coast, Flynn, et al., 2008). Dead and states worse than dead are not defined on the scale. The ASCOT is similarly anchored at 1 and 0, although here 0 is anchored to 'dead' and negative values (states worse than dead) are possible. A detailed discussion on preference elicitation approaches for HrQol instruments can be found elsewhere (Drummond et al., 2005).

Validation

Based on the studies included in this review, the most widely used HrQol instruments were the EQ5D and the SF-36, which have been extensively validated across a wide range of conditions and countries. The most widely used wellbeing instruments were Ferrans and Powers Qol index (QLI) and the WHO-Qol instruments. To date, the ICECAP-O preference-based instrument has been more widely validated than the ASCOT. A more thorough comparison of the two preference-based and the two most frequently used questionnaires is given in Table 2.

Comparison of the most promising instruments

The most extensively used instruments, Ferrans and Powers QLI and the WHO-Qol OLD, both have as important disadvantages for economic evaluations that preference based weights for outcomes are lacking. These preferences could, however, be obtained in future studies, which would improve their usefulness in economic evaluations. Both instruments have been tested for reliability and validity with good or excellent results. In older people, Ferrans and Powers QLI had excellent reliability, with an internal consistency of 0.86–0.96 (Ferrans & Powers, 2013). Test–retest reliability was tested in the general population with a test–retest correlation of 0.87 using a two-week interval (Ferrans & Powers, 2013). The instrument showed good validity; it was moderately to highly and

positively correlated with life satisfaction and general health perception. Moreover, it was moderately negatively correlated with disease burden, and showed sensitivity to change in 27 intervention studies (Ferrans & Powers, 2013). Ferrans and Powers QLI, however, misses some of the frequently measured dimensions of wellbeing. A more comprehensive instrument is the WHO-Qol OLD, which has shown good reliability in older people (internal consistency of 0.88–0.89 and test–retest reliability of 0.91 after two weeks) (Peel, Bartlett, & Marshall, 2007). Good validity was suggested with medium to strong negative correlations with different depression measures and moderately positive correlations with general health perception. Additionally, there is increasing evidence for its sensitivity to change for a number of conditions (Figueira et al., 2012; McDowell & Newell, 2006). A major disadvantage of both instruments was that they are relatively long. The WHO-Qol OLD, in fact, is an extension of the WHO-Qol BREF, having no less than 24 additional questions (Harper, 1998). Furthermore, proxy versions were not available for either instrument.

The ICECAP-O instrument has only five items, while the ASCOT has nine items, making their use in older people quite feasible. The ICECAP-O measures five of seven most frequently identified dimensions of wellbeing, but it has the (potential) disadvantage of not directly measuring a physical health dimension. It is possible that physical health is captured indirectly by the other dimensions, which is suggested by several empirical findings (Coast, Peters, Natarajan, Sproston, & Flynn, 2008; Flynn, Chan, Coast, & Peters, 2011; Makai, Brouwer, Koopmanschap, & Nieboer, 2012; Makai, Koopmanschap, Brouwer, & Nieboer, 2013). The ASCOT also measures five of the seven most frequently identified dimensions of wellbeing. The ASCOT has as a potential disadvantage that it does not explicitly measure a psychological dimension.

Another advantage of the ICECAP-O is its more widespread validation as compared to the ASCOT. It has been applied in different settings and cultures such as the UK, Australia and the Netherlands and Canada (Coast, Peters, et al., 2008; Comans, Peel, Gray, & Scuffham, 2013; Couzner, Ratcliffe, & Crotty, 2012; Couzner, Ratcliffe, Lester, Flynn, & Crotty, 2012; Davis et al., 2012, 2013; Flynn et al., 2011; Henderson et al., 2013; Makai et al., 2012, 2013; Mitchell, Roberts, Barton, Pollard, & Coast, 2013). Although clearly related to its early stage of development and use, this implies that the validity of the ASCOT is more uncertain, especially in different settings and cultures than used in so far (Malley et al., 2012). Its psychometric properties also require further testing. Disadvantages of both instruments are that they lack explicit assessments of their reliability and sensitivity to change.

Discussion

Key findings

This study reviewed the literature to search for outcome measures which can be used in economic evaluations of interventions in older people, particularly in long-term care. To avoid leaving out potentially useful and relevant outcome measures, we included instruments labeled as HrQol instruments in the first stage of the review. This was deemed important since the classification of instruments as HRQoL or wellbeing need not have been done consistently. We retrieved 34 generic instruments, of which 23 were classified as wellbeing instruments and 11 as HrQol instruments. Additional dimensions of wellbeing instruments that emerged from the review included purpose in life and achievement, security, financial well-being and personal

freedom. Of the wellbeing instruments, two had utility scores available, allowing use in economic evaluations: the ICECAP-O and the ASCOT. The two most widely validated wellbeing instruments, the WHO-QoL OLD and Ferrans and Powers QLI, do not have utility scores.

Wellbeing instruments enable researchers to evaluate a wider range of benefits of services for older people, thus more closely conforming to the goals of some interventions especially in long-term care (e.g. less restraints or a better living environment in nursing homes). They typically go beyond measuring HrQoL alone and can measure the benefits of interventions that aim to produce value beyond HrQoL domains.

Methodological issues

There are some limitations of this review worth considering. Any classification of instruments based on the dimensions measured is inherently subjective. To overcome this problem as much as possible, the instruments were classified by two reviewers in a structured manner using standardized criteria, while the additional authors were consulted for additional expert guidance. Nonetheless, other categorizations than the one presented here are possible, especially in light of the inclusive nature of this review, which allowed an exceptionally broad range of instruments to be included. Another limitation is that the review excluded grey literature. Therefore, instruments still in development may have been missed.

Choice of instruments

A first noteworthy point is that even though wellbeing or QoL is a difficult theoretical concept (Morgan et al., 2011; Wilson & Cleary, 1995), its actual measurement converges to a limited number of dimensions. Such a convergence could form the basis of an operational definition of wellbeing, although there is no consensus at this point.

Most wellbeing instruments measure a combination of health and non-health consequences, making them potentially suitable for evaluating interventions that result in a combination of health and non-health consequences. The exception is the CASP, which exclusively measures non-health outcomes (Hyde et al., 2003). Nonetheless, even though it seems to be rooted in a more subjective notion of wellbeing which is distinct from health, the CASP may still capture health consequences indirectly, if the measured domains are influenced by health status.

A thorough exploration of how the individual dimensions of HrQoL and wellbeing relate to each other is an important yet difficult conceptual and empirical puzzle beyond the scope of this paper. For such work, additional conceptual and integrative reviews based on qualitative studies may be necessary as well. A few important features of wellbeing instruments are nonetheless worth noting for potential users and developers.

First, different instruments measure the dimensions of wellbeing on different levels. We can attempt to classify the dimensions according to Wilson's taxonomy (Wilson & Cleary, 1995), where outcome measures are placed on a continuum from medical variables to overall QoL. Outcome measures have five levels: physiological, symptomatic, functional, perceptive, and overall QoL. For example, ICECAP-O and ASCOT both measure a dimension of control, but seem to do this on different levels. The ICECAP-O asks respondents if they are able to be independent, which can be viewed as measuring on the perceptive level. The ASCOT asks if they have control over daily lives, which can be viewed as measuring on the functional level. Such distinctions have an influence on how the measure aims to

capture benefits in a comprehensive manner. While wellbeing measures on the perception level may be more abstract attempting to capture benefits with broad dimensions, wellbeing measures on the functioning level often are more specific and may be aimed at explicitly capturing the dimensions relevant for (i.e. influenced by) health and social care. By measuring on the functional level, a wellbeing instrument may be more comparable to current HrQoL measures which also typically measure the dimensions on this level. At the same time, comprehensiveness may then require a large number of dimensions. In contrast, dimensions measured on the perception level may reach comprehensiveness through a smaller list of dimensions. The latter may improve the feasibility of use in elderly populations.

With respect to the relationship between health and non-health dimensions, certain HrQoL dimensions underlie wellbeing dimensions completely (as may be the case for the CASP) or partially, as may be the case with ICECAP-O, where the physical health dimension can be thought of as underlying, for instance, the dimensions of control and role. Even though validation work shows that the ICECAP-O reflects and captures all three health dimensions (Coast, Peters, et al., 2008; Flynn et al., 2011; Makai et al., 2012, 2013) there is also some indirect evidence that the ICECAP-O may not measure physical health as fully as a HrQoL measure like the EQ5D does, but more research on this issue remains necessary (Davis et al., 2013). Similarly, further research is especially encouraged on the ASCOT, to investigate whether the lacking psychological dimension is indirectly captured by some other dimension(s) (Forder & Caiels, 2011).

The majority of the reviewed wellbeing instruments did not appear to be directly useful for economic evaluations in older people consuming health and social care. Many are not preference-based, and would thus require a utility-elicitation procedure to be more readily useful to CUA, following the example of the SF-36. This might be relevant to Ferrans and Powers QLI and the WHO-QoL OLD, since both have been widely used and extensively validated. If utility scores would be derived for the states described by these instruments, the findings of previous studies using these instruments could also be revisited. As development of utility scores for existing lengthy questionnaires typically involves reducing the number of included items in the descriptive system (Brazier et al., 2002), the feasibility of including these instruments (in shortened form) in economic evaluation would improve as well. However, these shortened instruments would require additional validation in order to ascertain that they retain their psychometric properties.

The preference-based instruments ICECAP-O and ASCOT have been developed more recently, and thus have not been extensively validated. A major drawback of the current preference-based measures is that they do not integrally measure health and non-health consequences to the extent that for instance the Ferrans and Powers QLI and the WHO-QoL OLD do. This can be a problem for interventions that, while not exclusively aimed at non-health dimensions, have an effect on health as well.

Currently, the ICECAP-O and the ASCOT may be useful in the context of economic evaluations. While both are promising, the validation of the ASCOT at this point lags behind that of the ICECAP-O. If further studies provide support for their sensitivity to change, and clarify the relationship between the ICECAP-O, ASCOT and various health dimensions both instruments could be a suitable wellbeing instrument for economic evaluations in older people, particularly in long-term care. In fact, currently the Social Care Institute for Excellence (Francis, 2009) and the NICE (NICE, 2013) guidelines recommend using these two instruments for measuring and valuing effects in the United Kingdom. Sensitivity

to change of wellbeing instruments seems particularly relevant, as interventions are one factor amongst many which influence wellbeing. Therefore, further research on sensitivity to change of wellbeing instruments to a number of interventions is particularly encouraged. In the context of sensitivity to change, the general design of instruments may also matter. The ASCOT for example measures the effect of particular services more specifically than the ICECAP-O does. Hence, one might expect the ASCOT to potentially be more sensitive to changes in the provision of these services. On the other hand, the ICECAP-O may be more sensitive to changes in the provision of other social care services or the general care context, which are not included in the ASCOT dimensions. Such hypotheses have to be rigorously tested, using pre-specified hypotheses (Mokkink et al., 2010; Terwee et al., 2007). This may also shed more light on the question when to use which instrument.

Additionally, given the fact that, at present, the ability of the ICECAP-O and ASCOT to also (completely or adequately) capture all relevant health dimensions remains unclear, it seems advisable to use a health measure such as the EQ-5D or SF-6D along with the ICECAP-O or ASCOT instrument in economic evaluations of interventions aimed at older people in order to explicitly capture health benefits alongside broader benefits. We note that health-related and wellbeing-based utilities should not be condensed into a single utility index. First of all because they relate to two different scales and concepts that cannot simply be added. Secondly, because it is currently unclear which dimensions would be double-counted by the different instruments. More research is required to investigate the potential degree of double-counting when using these measures simultaneously and the degree of missed health effects when using only wellbeing measures. If and when it becomes clear which if any health dimensions may be missed by the wellbeing measures, then a common valuation of different measures may be attempted (Alava, Brazier, Rowen, & Tsuchiya, 2013), resulting in an instrument capturing all relevant dimensions. However, this requires much research, potentially involving the development of methods for combining capabilities and functionings. So far, CUAs using such broad measures assessing the full benefits of elderly care have yet to be published. We should emphasize that such studies would be using a far broader concept of utility than one that is solely health-related, better suiting the aims of many interventions aimed at older people, particularly in long-term care.

Choice of maximand

When attempting to support optimal allocations of scarce resources for older people, it is pivotal to include all costs and all benefits of interventions, and this can be achieved with wellbeing instruments. One theoretically well-developed approach towards quantifying wellbeing in that context stems from the capability approach. The capability approach is claimed to underlie the ICECAP-O as well as the ASCOT (Forder & Caiels, 2011). While most health measures seek to measure functionings, the capability approach focuses on capabilities, which are two different concepts (Sen, 1982). According to capability theory, functionings can be defined as beings or doings of the individual, while capabilities are potential beings and doing, or potential functionings. For example, a classic distinction is made between a person who is starving or fasting. While both are equal in terms of functionings, the former lacks capabilities while the latter simply chooses not to engage in a functioning (Sen, 1982). It is important to recognize such differences in choosing outcome measures, as they (implicitly) define the maximand of interventions.

In the context of receiving health and social services, services can be seen to expand peoples' capabilities through either directly allowing people to function (for example washing them) or indirectly through mitigating an impairment (Forder & Caiels, 2011). Following Forder's reasoning, outcome measures in older people should be able to measure improvement in wellbeing even if personal functioning is not improved, as long-term care services allow individuals to achieve outcomes that they would not be able to achieve themselves.

Although in this review we have limited ourselves to a review of wellbeing instruments in older people, particularly in long-term care, the problem of the evaluative space goes beyond services for older people, and is applicable to the whole healthcare sector. In order to maintain the possibility to evaluate interventions across the whole healthcare sector, ideally comparable wellbeing measures (or better still, one overall wellbeing measure) should be available for the entire healthcare sector. It appears that there is a great need for appropriate wellbeing instruments, since different fields of healthcare, such as mental health, social care, public health may not be directly (solely or mainly) aimed at improving health. In fact, there are preference-based wellbeing instruments being developed for the general adult population, for example the ICECAP-A (Al-Janabi et al., 2012). It is an interesting area of future research to investigate whether such measures adequately capture all dimensions relevant for older people as well or whether specific measures for them remain necessary. Such issues should be considered and explored further in development and validation of preference-based wellbeing instruments.

Conclusion

The development and use of wellbeing instruments for CUA in older people aiming to capture the benefits of both health and social care is a new, developing and important area of research. In the short-run, two preference-based instruments may be useful in the context of economic evaluations: the ICECAP-O and the ASCOT. The validation of the ASCOT at this point lags behind that of the ICECAP-O, although both require substantial validation. During this validation work, attention should be paid to the exact relationships between the ICECAP-O, ASCOT and various health dimensions contained in widely validated preference-based HrQoL measures, such as the SF-6D or the EQ-5D. An alternative direction forward would be to develop scoring algorithms for extensively validated non-preference-based measures that encompass more dimensions, following the example of the SF-36 (Brazier et al., 2002). Thus, utility weights could be attached to the results of earlier studies with these instruments. Irrespective of future direction, the conceptual puzzle of which dimensions need to be covered by wellbeing instruments for CUA remains unsolved. This also holds for how the dimensions should be measured and at what level. While further instrument validation and development remain crucial to capture the benefits of all services aimed at older people within CUA, with the availability of preference-based wellbeing instruments, reaching such a goal has become more feasible. This, in turn, has the potential of allowing a more optimal and fair allocation of services aimed at older people.

Acknowledgments

This study was funded by the Netherlands Organization for Health Research and Development (ZOnMw) grant numbers 53200005, 60-60900-95-002.

Appendix 1

Search terms for all databases.

Population	Outcome	Questionnaire	Exclusion
Elderly	Quality of life	Validation	Gene
Older	Overall Quality of Life	Validation questionnaire	Genetic
Older people	Health-related Quality of Life	Validated measure	Clinical utility
Geriatric	Health Related Quality of Life	Validated measures	Utilization
Gerontology	Generic Quality of Life	Validated questions	Risk factor
Nursing	Health Perceptions	Validity and reliability	Risk factors
Old age	Functional Status	Reliability and validity	Semi-structured interview
Old-age specific	Functional State	Utilities	
Old age specific	Environment	Utility	
Older adults	Health Status	Util	
Older people	Well-being	Preference based valuation	
Frail elders	Welfare	Valuation	
Frail elderly	Wellbeing	Preference-based health measure	
Older patient	Quality weight	Capabilities	
	States	Choice experiment	
	State	Choice experiments	
	Health outcome measurement	Discrete choice experiments	
	GEN-QOL	Preferences	
	GEN-QOLQ	Economic evaluation	
	OQOL	Health outcome measure	

Appendix 2

Full references of HrQoL and wellbeing development papers in Table 1

- (IWbG), I. W. G. (2006). Personal wellbeing index – adult, manual 4th edition. Manual. Melbourne: Australian Centre on Quality of Life, Deakin University.
- Baro, E., Ferrer, M., Vazquez, O., Miralles, R., Pont, A., Esperanza, A., et al. (2006). Using the Nottingham Health Profile (NHP) among older adult inpatients with varying cognitive function. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 15, 575–585.
- Bjorkman, T., & Svensson, B. (2005). Quality of life in people with severe mental illness. Reliability and validity of the Manchester Short Assessment of Quality of Life (MANSA). *Nordic Journal of Psychiatry*, 59, 302–306.
- Bowling, A. (2009). The Psychometric properties of the older people's quality of life questionnaire, compared with the CASP-19 and the WHOQOL-OLD. *Current Gerontology and Geriatrics Research*, 2009, 298950.
- Brazier, J., Roberts, J., & Deverill, M. (2002). The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics*, 21, 271–292.
- Campbell, A., Converse, P. E., & Rodgers, W. L. (1976). *The quality of American life: Perceptions, evaluations, and satisfactions*. New York: Russell Sage Foundation.
- 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 & Medicine* (1982), 67, 874–882.
- Cohen, S. R., Mount, B. M., Strobel, M. G., & Bui, F. (1995). The McGill Quality of Life Questionnaire: a measure of quality of life appropriate for people with advanced disease. A preliminary

- study of validity and acceptability. *Palliative Medicine*, 9, 207–219.
- Cummins, R. A., Eckersley, R., Pallant, J., van Vught, J., & Misajon, R. (2003). Developing a national index of subjective wellbeing: the Australian unity wellbeing index. *Social Indicators Research*, 64, 159–190.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49, 71–75.
- Feeny, D., Furlong, W., Torrance, G. W., Goldsmith, C. H., Zhu, Z., DePauw, S., et al. (2002). Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Medical Care*, 40, 113–128.
- Ferrans, C. E., & Powers, M. J. (1992). Psychometric assessment of the quality of life index. *Research in Nursing & Health*, 15, 29–38.
- Flanagan, J. C. (1982). Measurement of quality of life: current state of the art. *Archives of Physical Medicine and Rehabilitation*, 63, 56.
- Frisch, M. B., Cornell, J., Villanueva, M., & Retzlaff, P. J. (1992). Clinical validation of the quality of life inventory: a measure of life satisfaction for use in treatment planning and outcome assessment. *Psychological Assessment*, 4, 92–101.
- Gerritsen, D. L., Steverink, N., Frijters, D. H. M., Ooms, M. E., & Ribbe, M. W. (2010). Social well-being and its measurement in the nursing home, the SWON-scale. *Journal of Clinical Nursing*, 19, 1243–1251.
- Gullone, E. C. R. A. (1999). The Comprehensive Quality of Life Scale: a psychometric evaluation with an adolescent sample. *Behaviour Change*, 16, 127–139.
- Harper, A. P. N. (1998). Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine*, 28, 551–558.
- Holzhausen, M., Kuhlmeier, A., & Martus, P. (2010). Individualized measurement of quality of life in older adults: development and pilot testing of a new tool. *European Journal of Ageing*, 7, 201–211.
- Hyde, M., Wiggins, R. D., Higgs, P., & Blane, D. B. (2003). A measure of quality of life in early old age: the theory, development and properties of a needs satisfaction model (CASP-19). *Ageing & Mental Health*, 7, 186–194.
- Jamison, R. N., Fanciullo, G. J., McHugo, G. J., & Baird, J. C. (2007). Validation of the short-form interactive computerized quality of life scale (ICQOL-SF). *Pain Medicine (Malden, Mass.)*, 8, 243–250.
- Kaplan, R. M., Bush, J., & Berry, C. C. (1978). The reliability, stability and generalizability of a health status index. In *Proceedings of the American Statistical Association, Social Statistics Section*.
- Malley, J. N., Towers, A. M., Netten, A. P., Brazier, J. E., Forder, J. E., & Flynn, T. (2012). An assessment of the construct validity of the ASCOT measure of social care-related quality of life with older people. *Health and Quality of Life Outcomes*, 10.
- Nieboer, A., Lindenberg, S., Boomsma, A., & van Bruggen, A. C. (2005). Dimensions of well-being and their measurement: the Spf-II scale. *Social Indicators Research*, 73, 313–353.
- Novella, J. L., Ankri, J., Morrone, I., Guillemin, F., Jolly, D., Jochum, C., et al. (2001). Evaluation of the quality of life in dementia with a generic quality of life questionnaire: the Duke Health Profile. *Dementia and Geriatric Cognitive Disorders*, 12, 158–166.
- Osborne, R. H., Hawthorne, G., Lew, E. A., & Gray, L. C. (2003). Quality of life assessment in the community-dwelling elderly: validation of the Assessment of Quality of Life (AQoL) Instrument and comparison with the SF-36. *Journal of Clinical Epidemiology*, 56, 138–147.
- Ostermann, T., Bussing, A., Beer, A. M., & Matthiessen, P. F. (2005). The Herdecke Questionnaire on Quality of Life (HLQ): validation of factorial structure and development of a short form

within a naturopathy treated in-patient collective. *Health and Quality of Life Outcomes*, 3.

Peel, N. M., Bartlett, H. P., & Marshall, A. L. (2007). Measuring quality of life in older people: reliability and validity of WHOQOL-OLD. *Australasian Journal on Ageing*, 26, 162–167.

Power, N. K. W. T. W. G. (1998). The World Health Organization quality of life assessment (WHOQOL): development and general psychometric properties. *Social Science & Medicine*, 46, 1569–1585.

Rothman, M. L., Hedrick, S., & Inui, T. (1989). The Sickness Impact Profile as a measure of the health status of noncognitively impaired nursing home residents. *Medical Care*, 27, S157–167.

Ruiz, M. A., & Baca, E. (1993). Design and validation of the “Quality of Life Questionnaire” (“Cuestionario de Calidad de Vida”, CCV): a generic health-related perceived quality of life instrument. *European Journal of Psychological Assessment*, 9, 19–32.

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.

Sintonen, H. (2001). Comparing properties of the 15D and the EQ-5D in measuring health-related quality of life. *Archives of Hellenic Medicine*, 18, 156–160.

The EuroQol, G. (1990). EuroQol – a new facility for the measurement of health-related quality of life. *Health policy (Amsterdam, Netherlands)*, 16, 199–208.

Torrance, G. W., Feeny, D. H., Furlong, W. J., Barr, R. D., Zhang, Y., & Wang, Q. (1996). Multiattribute utility function for a comprehensive health status classification system: Health Utilities Index Mark 2. *Medical Care*, 34, 702–722.

References

- Al-Janabi, H., Flynn, T., & Coast, J. (2012). Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. *Quality of Life Research*, 21, 167–176.
- Alava, M. H., Brazier, J., Rowen, D., & Tsuchiya, A. (2013). Common scale valuations across different preference-based measures estimation using rank data. *Medical Decision Making*, 33, 839–852.
- Brazier, J., Ratcliffe, J., Salomon, J. A., & Tsuchiya, A. (2007). *Measuring and valuing health benefits for economic evaluation*. Oxford: Oxford University Press.
- Brazier, J., Roberts, J., & Deverill, M. (2002). The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics*, 21, 271–292.
- 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 & Medicine* (1982), 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.
- Comans, T. A., Peel, N. M., Gray, L. C., & Scuffham, P. A. (2013). Quality of life of older frail persons receiving a post-discharge program. *Health and Quality of Life Outcomes*, 11, 58.
- Couzner, L., Ratcliffe, J., & Crotty, M. (2012). The relationship between quality of life, health and care transition: an empirical comparison in an older post-acute population. *Health and Quality of Life Outcomes*, 10, 69.
- Couzner, L., Ratcliffe, J., Lester, L., Flynn, T., & Crotty, M. (2012). Measuring and valuing quality of life for public health research: application of the ICECAP-O capability index in the Australian general population. *International Journal of Public Health*, 1–10.
- Davis, J. C., Bryan, S., McLeod, R., Rogers, J., Khan, K., & Liu-Ambrose, T. (2012). Exploration of the association between quality of life, assessed by the EQ-5D and ICECAP-O, and falls risk, cognitive function and daily function, in older adults with mobility impairments. *BMC Geriatrics*, 12, 65.
- Davis, J. C., Liu-Ambrose, T., Richardson, C. G., & Bryan, S. (2013). A comparison of the ICECAP-O with EQ-5D in a falls prevention clinical setting: are they complements or substitutes? *Quality of Life Research*, 1–9.
- Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press.
- Ferrans, C. E., & Powers, M. (2013). *Reliability and validity of the Ferrans and Powers Quality of Life Index*. <http://www.uic.edu/orgs/qli/reliability/reliabilitymain.htm> Accessed 01.06.13.
- Figueira, H. A., Figueira, A. A., Cader, S. A., Guimarães, A. C., De Oliveira, R. J., Figueira, J. A., et al. (2012). Effects of a physical activity governmental health programme on the quality of life of elderly people. *Scandinavian Journal of Public Health*, 40, 418–422.
- Flynn, T. N., Chan, P., Coast, J., & Peters, T. J. (2011). Assessing quality of life among British older people using the ICEPOP CAPability (ICECAP-O) measure. *Applied Health Economics and Health Policy*, 9, 317–329.
- Forder, J. E., & Caiels, J. (2011). Measuring the outcomes of long-term care. *Social Science & Medicine*, 73, 1766–1774.
- Harper, A. P. N. (1998). Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine*, 28, 551–558.
- Henderson, C., Knapp, M., Fernández, J.-L., Beecham, J., Hirani, S. P., Cartwright, M., et al. (2013). Cost effectiveness of telehealth for patients with long term conditions (Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial. *BMJ*, 346.
- 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.
- Hyde, M., Wiggins, R. D., Higgs, P., & Blane, D. B. (2003). A measure of quality of life in early old age: the theory, development and properties of a needs satisfaction model (CASP-19). *Ageing & Mental Health*, 7, 186–194.
- Makai, P., Brouwer, W., Koopmanschap, M., & Nieboer, A. (2012). Capabilities and quality of life in Dutch psycho-geriatric nursing homes: an exploratory study using a proxy version of the ICECAP-O. *Quality of Life Research*, 21, 801–812.
- Makai, P., Koopmanschap, M. A., Brouwer, W. B., & Nieboer, A. A. (2013). A validation of the ICECAP-O in a population of post-hospitalized older people in the Netherlands. *Health and Quality of Life Outcomes*, 11, 1–11.
- Malley, J. N., Towers, A. M., Netten, A. P., Brazier, J. E., Forder, J. E., & Flynn, T. (2012). An assessment of the construct validity of the ASCOT measure of social care-related quality of life with older people. *Health and Quality of Life Outcomes*, 10.
- McDowell, I., & Newell, C. (2006). *Measuring health: A guide to rating scales and questionnaires*. New York: Oxford University Press.
- Mitchell, P. M., Roberts, T. E., Barton, P. M., Pollard, B. S., & Coast, J. (2013). Predicting the ICECAP-O capability index from the WOMAC Osteoarthritis index is mapping onto capability from condition-specific health status questionnaires feasible? *Medical Decision Making*, 33, 547–557.
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., et al. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*, 19, 539–549.
- Morgan, S., Grootendorst, P., Lexchin, J., Cunningham, C., & Greyson, D. (2011). The cost of drug development: a systematic review. *Health Policy*, 100, 4–17.
- Netten, A., Burge, P., Malley, J., Potoglou, D., Towers, A.-M., Brazier, J., et al. (2012). Outcomes of social care for adults: developing a preference-weighted measure. *Health Technology Assessment*, 16, 1–166.
- Nussbaum, M. (1993). *The quality of life*. Oxford: Clarendon Press.
- Osborne, R. H., Hawthorne, G., Lew, E. A., & Gray, L. C. (2003). Quality of life assessment in the community-dwelling elderly: validation of the Assessment of Quality of Life (AQoL) Instrument and comparison with the SF-36. *Journal of Clinical Epidemiology*, 56, 138–147.
- Peel, N. M., Bartlett, H. P., & Marshall, A. L. (2007). Measuring quality of life in older people: reliability and validity of WHOQOL-OLD. *Australasian Journal on Ageing*, 26, 162–167.
- Potoglou, D., Burge, P., Flynn, T., Netten, A., Malley, J., Forder, J., et al. (2011). Best–worst scaling vs. discrete choice experiments: an empirical comparison using social care data. *Social Science & Medicine*, 72, 1717–1727.
- 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.
- Sen, A. (1982). *Choice, welfare and measurement*. Cambridge, MA: Harvard University Press.
- Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60, 34–42.
- WHO. (1948). *Official records of the World Health Organisation*.
- Wilson, I. B., & Cleary, P. D. (1995). Linking clinical variables with health-related quality of life: a conceptual model of patient outcomes. *JAMA: The Journal of the American Medical Association*, 273, 59–65.