Health Economists' Study Group, Exeter.

January 2013

A qualitative assessment of the content and face validity of the EQ-5D-5L and ICECAP-A.

Thomas Keeley[a,b], Hareth Al-Janabi[b], Paula Lorgelly[c], Joanna Coast[a,b]

[a] MRC Midlands Hub for Trials Methodology Research, School of Health and Population Science, The University of Birmingham, United Kingdom

[b] Health Economics Unit, School of Health and Population Science, The University of Birmingham, United Kingdom

[c] Centre for Health Economics, Monash University, Australia

Introduction

Economic evaluation is an important part of healthcare assessments and randomised controlled trials. In health, welfarism, the theoretical position that the goodness of a situation can be judged solely based on utility information, has largely been replaced by extra-welfarism as the predominant form of economic evaluation. Extra-welfarism allows outcomes other than utility, such as equity weights, freedom of choice, health state and capabilities, to be admitted into an analysis (1-3). In a truncation of the potential scope of the extra-welfarist approach, health economists have, in the main, treated health as the central focus of economic evaluation (1).

Many health economic evaluations combine length and quality of life to form the quality-adjusted life year (QALY). The QALY, a common metric, allows comparison between healthcare evaluations in different clinical areas and with different patient populations. However, the use of health status measures to judge the quality adjustment arguably narrows the scope of the QALY and reduces its appropriateness for use in a number of research areas, such as mental health (4), public health (5) and end of life care (6). For example, the benefit of an end of life care pathway may have a positive effect on an individual's independence or their social interaction. These benefits may be missed when assessing quality of life using health status measures.

The capability approach is an alternative framework derived from Amartya Sen's work on functioning and capability (7-10). In health, the capability approach may offer a fuller conceptualisation of the extra-welfarist approach by broadening the evaluative space and focusing on what a person is able to do and who they are able to be, rather than what they actually do and who they become (3). The approach admits information both directly and indirectly related to an individual's well-being, through the recognition of "agency goals" (goals that the hold value to the individual). Through the consideration of what a person is able to do (capability) rather than what they actually do (functioning), the approach recognises the importance of freedom of choice.

The ICECAP-A capability measure and EQ-5D-5L health related quality of life questionnaire, are two measurement instruments developed for use in (different interpretations of) extra-welfarist health economic evaluation. The ICECAP-A capability measure is a sister measure of the ICECAP-O capability measure, designed for use in the whole adult population. The descriptive system, formed through in depth interviews with the general public, defines quality of life as consisting of stability, attachment, autonomy, achievement and enjoyment (11). The measure assesses capability by asking whether a person "can" or is "able" to achieve particular states. Assessing a person's capability in this way has been shown to be feasible, and participants are able to comprehend and answer the questionnaire "in line" with the theory of the approach (12)(Under review). The content validity of the measure has not been assessed.

The EQ-5D-5L is a generic instrument based on the descriptive system contained in the EQ-5D-3L which defines health as consisting of: mobility, self-care, usual activities, pain and discomfort and anxiety and depression (13). The number of response options (levels) have been increased from three to five in an effort increase the sensitivity-to-change of the measure and reduce ceiling effects (14-16). The content and face validity of the measure has been assessed in healthy and patient populations (13).

Assessment of the psychometric properties of a measure is an important step in understanding the potential of an instrument to provide accurate and consistent measurement of a construct. For a measurement instrument to be useful it has to be reliable, sensitive to meaningful change and valid. Validity is a question of whether an instrument measures what it purports to measure and if changes in the construct produce changes in the measure (17).

Validation efforts divide predominantly into three groups: criterion validation, the correlation of a measure with a gold standard measure; construct validation, the "extent to which the dimension scores of an instrument correlate with...indicators of the health concept or concept of interest" $(18)_{p.43}$ and content validation. Content validation focuses on "the extent to which a scale represents the most relevant and important aspects of a concept in the context of a given measurement application" $(19)_{p.743}$. Content validation seeks to answer two closely related questions. First, does the instrument assess all the important domains of a concept? Second, does the measure appear to be sensible, relevant and appropriate? This second point is often termed "face validity" but is equally suitably considered as a subsection of content validity, which it will be herein.

The assessment of content validity is an important methodological and practical consideration. Assessment of the content coverage of a measure allows understanding of the inferences that can be drawn from the results of a measure. If a measurement instrument does not fully assess a concept this needs to be understood if appropriate inferences are to be drawn (20). Furthermore, the perception of researchers as to whether the measure is relevant and appropriate for use will often determine whether the measure is used. For example, if a principal investigator views a measure as crude or unwieldy, then it is unlikely to be incorporated into a trial questionnaire pack. Content validation in no way offers a complete assessment of the validity of a measure, rather it contributes important information to the "validity portfolio" of an instrument.

An emphasis is placed on qualitative methods as a way of assessing content validity due to the strengths of this form of research (21). This is counter to the mainly numerical evaluations that currently dominate validity assessment (22;23). Content validation can be completed with the public, patients and experts in relevant fields acting as informants. Patients have firsthand, personal experience of the concept being discussed and how it might be affected by different situations. They are in a position to provide an insider's

– emic – perspective: for instance, a cancer patient will have in depth knowledge of how *their* quality of life was reduced by *their* disease. Experts, such as doctors, researchers or clinical experts can provide an outsider – etic- perspective. They have the advantage of observing a number of individuals in different situations and how the construct of interest manifests itself in different individuals (19): an oncologist may have seen patients who had their quality of life reduced due to unmanaged pain, inability to care for themselves or periods of depression.

A longstanding challenge of content assessment, especially when measuring concepts such as quality of life, where a tangible concept is not being measured, is the question of what content is relevant and should be measured (24;25). This is particularly important in quality of life, where a universally accepted definition does not exist. This is a notable methodological challenge for researchers in this area; in an analysis which seeks to assess whether a measure samples the important and relevant dimensions of a construct, a lack of clarity as to what the important and relevant dimensions are, is a challenge.

This paper reports a qualitative assessment of the content validity of the EQ-5D-5L and ICECAP-A measures using professionals involved in health research. To our knowledge this is the first content validity analysis of either measure using research professionals. Results pertaining to the informants' conceptualisation of quality of life are presented, followed by informants' overall and item-by-item opinions of the measures under consideration. Data presented refers to the content coverage, relevance and suitability of the measure. Discussion focuses on these points.

Methods

Methods reported below were approved by the University of Birmingham Science, Technology, Engineering and Mathematics Ethical Review Committee on 8th December, 2011, application number 9ERN_11-0575.

Informant selection

Using theoretical sampling informants were purposively selected based upon the characteristics of professional role. Informants were selected from four professional groups: 1) clinical and public health trial experts (referred to as "trialists"); 2) medical doctors involved in research; 3) researchers with regular patient or participant contact (referred to as "frontline researchers"); 4) health economists working within a trial setting. Informants both with and without experience of using the ICECAP measures were selected and where they had experience an effort was made to select those with a priori positive and negative opinion of the measures. The use of theoretical sampling allowed the initial iterations of analysis to inform the later selection of informants (26).

Practicalities of recruitment

Potential informants were identified through use of online trial centre and university staff lists, in conjunction with author knowledge of potential informant roles and experience. An invitation email stating the aim of the research, the potential burden upon the informant and the ethical approval gained for the study was sent to potential informants. Only three potential informants approached were not interviewed. For two of these individuals a convenient time could not be found, while one individual declined. Snowball recruitment, whereby previously interviewed informants were asked to give recommendations of other potential informants, was used to recruit three frontline researchers. Recruitment was stopped when data saturation was identified.

The interview process

Prior to the interview informants were sent an information sheet in a timely manner to allow them to consider the content before choosing to sign the consent form at the start of the interview. Interviews were recorded using a digital voice recorder.

The interviews were semi-structured and designed to allow breadth and depth of discussion. A topic guide was used as an *aide memoire* to encourage a consistent structure to the interviews and adequate coverage of key topics. Content mapping and content mining questions, which facilitate the direction of in depth interviews (27), were used. Content mapping questions were used to open up the conversation and encourage narrative on a subject. Content mining questions were used to encourage informants to elaborate on key areas.

Interviews were broadly partitioned into two parts. The first part of the interview assessed the informants understanding of quality of life as a concept, their views on existing quality of life measures used in their research and the importance of quality of life measurement to the work they do. During the second part of the interview, informants were presented with a copy of the ICECAP-A and the EQ-5D-5L measures, one after the other. The order in which the measures were shown to the informants was random except when it was beneficial to the flow of the interview to consider one first. For example: if towards the end of the first part of the interview the informant was discussing his concerns with the sensitivity of the EQ-5D-3L, the EQ-5D-5L would be given to the informant first. In 9 of the interviews the ICECAP-A was considered first.

Informants were encouraged to discuss the measures at length. A focus was maintained by the facilitator on the content coverage of the measures and the appropriateness, suitability and usefulness of the measures in the research the informant was involved in. The informant was also encouraged to consider each of the items of the measure individually and comment on their completeness and relevance.

Analysis

Interviews were transcribed verbatim. Interview transcripts were coded using the Atlas-Ti computer-assisted qualitative data analysis software. A hierarchical coding structure was formed during the analysis of the first batch of interviews and refined throughout the subsequent batches of analysis. An iterative, constant comparative, thematic analysis of the transcripts was completed. In total 4 batches of analysis were completed. The constant comparative, thematic analysis allowed descriptive and explanatory accounts to be formed. Comparisons were drawn between informants and emerging themes were identified. Emerging themes were initially identified by the first author and then discussed with other named others, before being assessed in greater detail both through refinement of the interview topic guide and greater focus during analysis.

The two-part partitioning of the interview (described above) provided an opportunity for the development of an innovative two stage analysis of content validity, termed the *direct comparative approach*. This approach attempts to address the methodological challenge of content validation identified above; that of a lack of an agreed upon definition.

The first stage of the analysis, the *comparative* stage, used data from the initial part of the interview, where the informant defined their understanding of quality of life as a concept. This provides a reference point for the analysis in this informant; they have described what they define quality of life to be and what influences it. This description of quality of life is *compared* by the researcher with the content of the EQ-5D-5L and ICECAP. For example: if an informant held mobility, absence of pain and ability to go about one's life as important influences on quality of life, then it can be concluded that the informant held at least three of the items in the EQ-5D to be relevant. Whereas, if an informant identified sensory function and the ability to communicate as an important influence on quality of life, then a logical conclusion would be that the EQ-5D did not offer content coverage of this. This part has the advantage of the informant not being influenced by seeing the measure in the interviews.

The second stage, the *direct* stage, uses data from the second part of the interview where the informant has the measures in front of them. The informant discussed the relevance of the items and the overall coverage of the measure. In this part of the interview the informant was asked to think back to what they defined quality of life to be in the initial part of the interview, and assess whether they felt the measure covers *their* conceptualisation of quality of life. Using the reference point established earlier in the interview is important. It address the methodological issue of there being no widely accepted definition of quality of life, by allowing understanding of the informants conceptualisation, and allows a better understanding of the informants' opinions on the coverage of the measure.

Verbatim quotes from informants have been selected to be illustrative of how informants' accounts were linked to emerging themes. Ellipses (…) were used to denote missing speech; 'umm', 'err' and repeats of

words, which do not add meaning, were removed without the use of ellipsis. Square brackets are used to clarify informants' meaning.

Results

Interviews

17 informants were purposively sampled to achieve a diversity of professional roles and clinical areas. Interviews, lasting between 45 and 90 minutes, were completed between February and September 2012 at the informant's place of work. Data saturation was identified by interview 14. 3 additional interviews were conducted to check saturation and ensure adequate numbers were sampled from each professional role. Table.1 shows the characteristics of interview informants.

	Number interviewed (n=17)
Sex	
Male	7
Female	10
Location	
Australia	8
UK	9
QoL measure experience	
EQ-5D	15
ICECAP	5
Professional role	
Frontline researcher	4
Trialist	7
Health economist	3
Research doctor	3

Table.1) Informant characteristics

The findings of the research are reported below in two sections: 1) a description of how informants conceptualised quality of life in the first part of the interview and 2) general reactions to and an item-by-item assessment of the content of the EQ-5D-5L and ICECAP-A measures.

Informant conceptualisation of quality of life

Physical health was recognised as being an important determinant of an individual's quality of life. There was a high level of agreement amongst informants that poor physical health or physical disability reduced quality of life.

> *I still see health as important in that [quality of life]. I think when someone has got ill health…it is quite a big determinant.* [Frontline researcher, Australia]

> *…you could have someone with a bilateral amputation below the knee that's in a wheel chair and their quality of life is zero.* [Frontline researcher, UK]

Pain and side-effects of treatment were identified as aspects of physical health which were considered particularly important in determining quality of life. The language and tone used by informants indicated that pain was considered a particularly pervasive influence on quality of life.

> *Nothing worse for quality of life in many ways than chronic discomfort and pain.* [Research doctor, Australia]

Side-effects were a particular concern for informants who worked in the area or cancer research, using treatments which have severe effects upon the health of an individual.

> *They could be perfectly well...and then us give them lovely cytotoxic chemotherapy and a bit of radioactive isotopic one off injection which reduces their quality of life completely.* [Frontline researcher, UK]

Although considered a major determinant, physical health was not viewed as either the only determinant of quality of life or as holding an overriding influence. Informants recognised that people can have an adequate or even high quality of life, despite being in poor physical health states.

> *All you see is ill health and states that you don't want to get into, but there are people that get into those states and have a fantastic time.* [Trialist, UK]

Informants' viewing physical health as an incomplete determinant was part of a strong theme of quality of life being a concept which reached beyond physical health. Informants regularly used the words "general" and "broad" to define the term.

> *I would see it as being a pretty broad, much, you know beyond health.* [Health economist, UK]

> *...things that can impact on a person's [quality of] life other than being in your trial. Life is going on, having children, they are getting married, they are dying, they are living, having operations, they're going broke, they are making a fortune, and whatever else is happening.* [Research doctor, Australia]

Determinants other than physical health which influenced quality of life were discussed at length. These determinants divided into three areas: psychological health, friends and family and the ability to lead a normal, valued life. Psychological health was one of the first determinants discussed, by the majority of informants. Many of the informants discussed how psychological problems such as depression, can stem from a physical condition. Informants understood psychological health as having a notable impact on quality of life:

...it is so pervasive into everyday life, probably almost more than anything else. [Health economist, UK]

Relationships with friends and family were viewed as important. This appeared to be motivated by considerations of enjoyment, through spending time with loved ones, and the support, which loved ones can provide through regular contact.

...even though he has poor health...he gets huge enjoyment from the family. [Frontline researcher, UK

...their daughter is not helping them, or sort of something like that probably has an influence. [Trialist, Australia]

Informants often described the importance of friends and family from the perspective of losing loved ones, or life events which make contact with friends and family harder:

...there is often a lot of umm unsettling things that are happening in their life, like they have had to move into assisted care or their spouse has died. [Frontline researcher, Australia]

The ability of an individual to lead their normal life was discussed by a sizable minority of informants. Informants attached importance to individuals being able to fulfil social roles which they had previously and the ability to achieve things they value.

...that they are able to perform social roles that they would normally perform I guess [is important]. [Trialist, Australia]

I suppose I would think that someone is doing what they want to be able to do. [Frontline researcher, Australia]

Social roles were normally defined as an activity which consisted of interaction with other people, such as looking after the Grandchildren or going to work. Achievements were normally focused on day-to-day things that the person may want to do, such as reading, looking after animals, or other activities that provided a sense of satisfaction, rather than big one-off achievements such as running a marathon.

In summary, informants viewed health as a significant, but not overriding influence on quality of life. A broad conceptualisation was described with psychological and social factors being considered important influences.

ICECAP-A measure

Informants viewed the ICECAP-A measure as a broad assessment of quality of life, appropriate for use in the research fields in which they worked. The measure was viewed as a short, uncomplicated measure, suitable for use with participants in a busy research environment:

> *It is a lovely length...because they haven't, you don't have the time to spend with a long questionnaire.* [Frontline researcher, UK]

The ICECAP-A was viewed as a broader, more general measure. Informants noted that it was different to existing health related quality of life measures; focusing on the emotional determinants of quality of life.

> *Yeh I guess this one is more general...and focuses mostly on the emotional.* [Trialist, Australia]

When asked to think back to how they defined quality of life in the first part of the interview and comment on whether the ICECAP measure captured that definition, the majority of informants felt that it did. While a number of informants noted that it did not directly assess the health of an individual and some felt this was an important dimension that the measure did not assess.

> *It is just how far away from health it gets I suppose...I think it is just the distance from health [which is a concern].* [Health economist, UK]

A tension was present in the data between informants who viewed the measure as assessing dimensions that matter to patients and a small minority who felt the subject matter were too sensitive and not appropriate to ask patients. The perception of inappropriateness was motivated by a concern over the questions being upsetting for people who had low levels of the dimensions being assessed, rather than being inappropriate per se. Informants who felt the measure was patient-focused were more likely to be frontline researchers or research doctors, while those who held concerns over the measure were more likely to be trialists.

> *I think they would be [happy to answer], because it sums up the kind of conversations you have with patients and I think they would be quite comforted with it.* [Frontline researcher, UK]

> *I find them [the questions] less appropriate...* [Trialist, UK]

The majority of informants felt that the measure was appropriate for use in the research area in which they worked, with only a small number providing a contrasting opinion. There was a consensus that the measure would be favoured in addition to, rather than as a replacement for existing health related quality of life measures. Informants felt that something which provided more information about the source of the problems and maintained a focus on health related quality of life was also required. This was motivated by a perception that the ICECAP-A measure was not measuring health related quality of life.

I think it would definitely [be of use]...I probably wouldn't replace the EQ-5D but I would think quite seriously about using it in addition. [Health economist, UK]

I like it. I love it. That is my initial feeling. I love it. [Research doctor, Australia]

So it would compliment, I wouldn't see it as a replacement for an EQ-5D, but it would certainly complement an EQ-5D type instrument. [Health economist, UK]

A number of informants discussed the capability wording of the ICECAP-A measure. Informants showed a level of cognitive struggle in understanding the focus of the question and a level of concern existed about whether the wording would be understandable for participants in the studies. However, it is observed that in the majority of the discussions about the capability wording informants reached an understanding that would be broadly in line with the capability theory (i.e. they understood the question correctly).

I don't like the "I am able" or "I can", you know I don't know, it feels as if in some way you are the person with the control , so I CAN have a lot if I want to I can have a lot of love and friendship. [Trialist, UK]

Stability

The stability item was designed to assess an individual's ability to live a life of continuity, without feeling concern or uncertainty (11). Prior to viewing the measure, informants identified stability as important dimension of quality of life. Living with fear and uncertainty due to a physical condition or illness and the concern that unemployment due to illness can cause, was identified by a number of informants.

You get frightened of taking your medicine. You get frightened of going to sleep, in case you don't wake up. [Research doctor, Australia]

...getting back to work. The worries of losing a job, not getting back, not being able to get jobs back. [Health economist, Australia]

Upon considering the measure, there was a broad acceptance that the stability item was relevant to the assessment of quality of life. Informants indicated that the item would resonate with participants of different ages, and would be influenced by both health and non-health factors. A number of informants recognised that the item was assessing a construct that they had identified as important in the earlier part of the interview.

...it makes sense because...the phase one patients I see are very palliative and they don't have a lot of time. But you can still be settled and secure with months to live. [Frontline researcher, UK]

Attachment

The attachment item is designed to assess the importance of support, social contact and relationships (11). Prior to considering the measure, informants identified the ability to function in a social context as an important consideration. Relationships were identified as important both for the enjoyment and support they provide. A loved one dying was often given as an example of the importance that relationships can have on quality of life. The significance to people suffering from illnesses to achieve social contact and the limiting effect that illnesses can have upon one's ability to achieve social contact was discussed at length.

> *And in the last year of his life, he died by the cancer, he said...this has been the best year of my life, because until this moment I never realised how loved I've been.* [Trialist, UK]

> *...people who are incontinent tend to become completely reclusive, they don't want to mix, they are terrified. They don't want to go out there and pee in front of everyone.* [Research doctor, Australia]

Upon considering the measure informants recognised attachment as a relevant item to ask. It was also noted that it is an area that is not often assessed.

> *Well things like love, friendship and support. It is all that thing around social connectiveness and support and intimacy. We as a research group are very interested in that in people with HIV* [Research doctor, Australia]

A level of concern was expressed by a number of informants, over two particular points. Firstly, there was a concern that the item was assessing a number of different concepts (love, friendship and support) within one question. The possibility that a person may have one, but not all of the concepts being assessed was a concern.

> *...I could just imagine, particularly with love friendship and support, I could imagine someone saying they have got one but not the other things. Might be difficult.* [Trialist, UK]

The second point of concern raised was about asking a question on what they perceived to be a sensitive subject. A lot of this concern was focused on the prospect of asking the question to an individual who had recently lost a loved one or in the process of a relationship breakup. Some informants who had identified attachment as important before viewing the measure held this concern.

> *...i don't think number two is very appropriate. There is a person here, they may, they haven't got any love friendship. That is something that is completely outside their control.* [Trialist, UK]

Autonomy

The autonomy item is designed to assess the ability to be independent, both in the practical sense of being able to look after oneself and being able to make decisions (11). A small number of informants discussed

independence as a dimension of quality of life prior to seeing the measure. The term "independence" was not used, rather discussion by these informants focused on the ability of an individual to do day-to-day things, such as shopping, which was closely linked with a person's mobility.

> *...they can't get down to the shops to do their shopping. Or, they want to do something and they can't do it. It is hard, they have got to think about is it feasible to do something that they want to, based on how mobile they are.* [Trialist, Australia]

In comparison to the limited discussion prior to viewing the measure, the majority of informants identified the autonomy item as of central importance to the assessment of quality of life. There was a consistent opinion that the item was more important to elderly people.

> *...especially with older people that independence is hugely important to them, and that's one of the depressing things for them when they lose that independence I think.* [Frontline researcher, UK]

Achievement

The achievement item is a measurement of the degree to which an individual can attain their goals and move forward in life (11). The influence upon quality of life of being able to achieve and attain personal goals was not discussed by many informants prior to viewing the measure. Gaining a sense of achievement through work was discussed briefly by a small number of informants. The importance of being able to look back with a sense of achievement was noted as important.

> *...i think he [young cancer sufferer] has kind of condensed it all to "Yeah, I am 25 and I have achieved everything I want"...and he is perfectly sane in what he is saying, be he has just reflected back and gone, "Yeah, I have achieved".* [Frontline researcher, UK]

Informants provided greater discussion after seeing the measure. The item was considered to be relevant, although disagreement existed as to whether it was relevant for older people. The breadth of the item was considered by a number of informants as being too broad and for some this raised the question about whether the top item was really achievable.

> *I mean I don't think that I can achieve and progress in all aspects of my life, I would love to be able to. BUT.* [Trialist, UK]

The use of the word "progress" was questioned. For some it focused on the area of paid employment while, those who worked in cancer noted that patients could misunderstand the question as assessing their illness.

I think some of them [patients] might think progress as in is the treatment working. Or maybe that is my cancer background. Is there any progress, has the cancer progressed? [Frontline researcher, UK]

Enjoyment

The enjoyment item assesses the enjoyment gained from fun and exciting things, as well as the simple pleasures in life (11). Enjoyment was discussed a few informants from the perspective of people with illnesses or disabilities enjoying life in spite of their condition. It was normally identified through providing examples, rather than stating explicitly that enjoyment was a construct of quality of life.

You have people that have an enormously great quality of life who can't walk anywhere...because they have this great social structure and play cards all day. [Trialist, Australia]

On considering the item, informants were split between those that felt the attribute was important and relevant, and those that did not. For those that felt the item was not relevant, a motivating factor appeared to be that the measure was too broad to be relevant. Field notes taken by the interviewer noted a level of surprise by some informants upon seeing the enjoyment and pleasure item.

I think it is very important to have certain feelings...like enjoyment and pleasure. [Frontline researcher, UK]

What do you mean by enjoyment and pleasure?...I suppose not vague, but possibly ambiguous. [Health economist, UK]

EQ-5D-5L

Informants viewed the EQ-5D-5L measure as simple and straight forward measure of health. The length and simplicity of the measure was viewed positively and a number of informants noted that the language used was appropriate.

I think the great beauty of this is that you can do this in two minutes flat. [Trialist, UK]

...this is very black and white, I have no problems, I have problems. [Research doctor, Australia]

When asked to think about how the EQ-5D-5L assessed the conceptualisation of quality of life which they had previously described, a conflict of opinion existed. A number of informants viewed it as offering sufficient coverage of quality of life:

...for assessing a general quality of life it is actually quite good. [Trialist, Australia]

A roughly equal number of informants viewed it as a measure of health or functioning, but not the broad conceptualisation of quality of life described above.

> *...that one is more broadly health.* [Research doctor, Australia]

> *...it is not capturing how they FEEL about their LIFE. It's, they are not saying "I have a good life" or not...This one is what you can do and what problems do you have.* [Trialist, UK]

Despite this disagreement around the measures coverage of quality of life, there was a broad recognition that the measure was of use in the research areas that informants worked in. This appeared to be motivated partly by awareness that there was a strong precedent of use of the euro-qol measure and recognition by funding and rationing bodies.

> *It is hard to beat the EuroQol in terms of sort of NICE guidance and everything that is out there already.* [Health economist, Australia]

Informants who had previously used EQ-5D-3L, noted that the EQ-5D-5L was an improvement. In discussion before the measures were considered the EQ-5D-3L was described as "crude", not being sensitive enough and being hampered by a "huge" ceiling effect. It was thought that the increase in levels would improve the ability of the measure to record change in health state. While the descriptive system of the measure remaining the same was criticised, the changes in the EQ-5D-5L measure made it more attractive to informants interviewed.

> *I think I would prefer it to the euro-qol that we are using now.* [Frontline researcher, Australia]

Mobility

Mobility was identified as an important influence on quality of life. In the prior discussion informants identified the ability to walk, as well as the ability to move upper and lower limbs as impacting on quality of life. Mobility was not valued for itself (i.e. the enjoyment of being able to walk); rather it was valued because it allowed individuals the independence to access their normal everyday life.

> *...it is independence, it is to do with mobility, it's the getting to the shops, being able to do what you want to do, when you can do it...* [Trialist, Australia]

Disagreement existed, on viewing the measure, as to whether the item fully assessed mobility Informants noted that the item assessed a persons' ability to walk, not their ability to be independently mobile. This was considered to limit the scope of the item.

> *...that's just about walking, whereas [if] people can be independently mobile in a wheel chair, they can actually have quite a high quality of life.* [Trialist, Australia]

Self-care

Prior to viewing the measure, very few informants identified self-care as an important determinant of quality of life. Those informants who discussed self-care, did so through use of examples of people they know or patients they have worked with. No informant directly raised it as a consideration of quality of life.

Upon seeing the measure informants felt the item was narrow, and arbitrary. Many felt it should be considered as part of usual activities, rather than as a dimension in and of itself.

> *...as a sort of category of assessment...it is only one action, like making a cup of tea. It is a bit arbitrary really.* [Trialist, UK]

Usual activities

A number of themes relating to usual activities were discussed. Informants discussed the importance of people being able to complete normal (the word "normal" was used as opposed to "usual") activities, with going to work and having social contact regarded as being important. In the broad conceptualisation of quality of life offered by informants in their prior discussion, a large number of the non-health dimensions could be considered to relate usual activities.

> *...in terms of their participation, that they are not able to do things that they normally do...whether it is looking after the grandchildren or cooking meals or something like that.* [Trialist, Australia]

On considering the usual activities item, informants found it to be very broad and noted the need for the clarifying statement that comes with the item. Only a few informants directly stated that this breadth was a problem, but there was a notable hesitance in the language used referring to the breadth of the item.

> *Whereas usual activity, work, home work, leisure is massively broad.* [Trialist, UK]

> *I think it is good that they explain what usual activities is though.* [Health economist, Australia]

Pain

Pain was identified as a particularly important, even pervasive health-related influence on quality of life. Informants with clinical training indicated that pain could be managed to reduce its influence on quality of life.

> *...if we get on top of the pain with the treatment, obviously their quality of life is going to be different.* [Frontline researcher, UK]

The pain item was noted by informants as being an important construct to measure in the questionnaire. However, concern existed about the phrasing of the question, with a small number of informants noting that the item was assessing two distinct dimensions: pain or discomfort.

You got to wonder why you'd bother asking pain or discomfort. Wouldn't you ask one, because they are so different... [Trialist, Australia]

Anxiety and depression

The affect of the psychological state of an individual upon their quality of life was identified by a large number of informants. Many identified the psychological state as being influenced heavily by the physical state. Worry, concern, fear and anxiety were identified as psychological dimensions, along with depression.

Depression is a frequent co-morbidity of severe physical illness. [Researcher doctor, Australia]

...they worry about recurrence...anxiety is a major issue [Health economist, Australia]

The anxiety and depression item attracted the most discussion of any of the items on either measure. Informants felt it to be a highly relevant item, but noted a problem with the use of the word "depression". For some, the concern focused around the stigma that is attached to the word and how this might shape participants answers. For others, the use of the words "depression", and "anxiety", as a summary for psychological health, was lacking in scope; they felt that this medicalised or technical word was not a good summary measure.

...anxious and depressed...people don't like that word depression. You know, "don't tell me that" [I'm depressed]... [Frontline research, Australia]

Well you can just ask if someone is sad or not. It is kind of medicalising things isn't it... [Health economist, UK]

Discussion

This paper has assessed the content validity and relevance of two outcome measures suitable for use in health research and extra- welfarist economic evaluation. To our knowledge this is the first qualitative assessment of the content validity of the EQ-5D-5L or the ICECAP-A, using health research professionals. Informants were selected to ensure a representative spread of professional roles involved in health research.

The EQ-5D-5L and ICECAP-A measures were considered to be of use in the research areas in which these research professionals worked. Both questionnaires were viewed as short and simple to complete and this was considered to be an advantage when using them in a busy research environment. The increase in levels of the EQ-5D-5L was viewed as a positive development that made the measure more attractive. This is in line with the basis for development. The perceived "understandbility" of the capability wording used in the ICECAP-A measure was a concern for informants. This finding should be considered in light of research that found the general public were largely able to understand and answer the questions (12).

Health research professionals described quality of life as a broad concept. Central to this concept was physical health, but for those interviewed, quality of life was not solely determined by physical health. Psychological health and emotions, social relationships and the ability to lead a normal life were considered important influences. Therefore, based on this initial discussion a comprehensive assessment of quality of life would be broad and comprehensive and not limited to, or excluding completely, health indicators. This description of quality of life included a number of dimensions that are assessed in the EQ-5D-5L and the ICECAP-A measures. Pain, mobility, psychological health and normal activities were identified as important dimensions and these are assessed in the EQ-5D descriptive system. Relations with friends and family, concern and fear and independence were also identified as important considerations and these are assessed in the ICECAP-A descriptive system.

The item by item discussion showed that for both measures the majority of content was considered relevant. There was considerably less discussion by informants of the EQ-5D-5L items compared to the ICECAP-A items. This may have been due to a greater familiarity with and acceptance of the questions in the EQ-5D-5L, given that many informants had experience of using the three level version. Or, it may be that informants felt more "challenged" by the questions in the ICECAP-A. In the EQ-5D-5L the relevance of the content of the self-care item was questioned; while in the ICECAP-A measure the content of the enjoyment item, and to a lesser extent achievement, were questioned.

The wording, phrasing and scope of the items were assessed. Concern was raised over walking being used as a proxy for mobility, the use of the term "depression" and assessment of two separate concepts in the pain and discomfort question in the EQ-5D-5L. In addition to the concerns around the capability wording of the ICECAP-A measure, concern over the assessment of multiple concepts in the attachment item and the suitability of that item was notable. Concern also existed around the use of the term "progress" in the achievement item, although this did appear to be isolated to cancer research professionals.

Overall, informants perceived of the EQ-5D-5L as being a health focused measurement, while the ICECAP-A measure was viewed as broader more general assessment of quality of life. There was an recognition that the EQ-5D-5L did not fully assess quality of life as informants defined and that the ICECAP-A measure did not assess health, which informants noted as an important construct of quality of life, in detail. This could be considered as broadly in line with the aims of each measure: the EQ-5D-5L is designed to be a health related quality of life measure, while the ICECAP-A capability measure's theoretical grounding is that focuses on wellbeing more broadly defined. Use of the ICECAP-A in addition to the EQ-5D-5L, or another health related quality of life measure, was viewed as a positive step in the assessment of quality of life.

Several themes within the work appeared to be related to the professional role that informants held. Those who worked in cancer research were more likely to identify side-effects as a notable influence on

quality of life and show concern about the use of the word “progress” in the achievement item of the ICECAP-A. In contrast, those who were frontline researchers or researcher doctors (i.e. those who have or have had contact with research participants) were likely to hold a positive view of the content and relevance of the ICECAP-A measure.

This work has its limitations. Researchers working in oncology have been (slightly) oversampled, with five of the informants working in this field. Linked to this limitation is the observation that the use of snowball sampling resulted in the recruitment of informants who had professional relationships with previously interviewed informants. This may have resulted in less breadth and variation in the sample. A final limitation was that three of the interviews were conducted under time constraints, which was partly a result of interviews being conducted at the informant’s place of work. When looking for the informant to fully describe their conceptualisation of quality of life, this time pressure presented a methodological problem.

In conclusion, this work has shown the EQ-5D-5L and ICECAP-A measures to hold a degree of content validity. Both measures assess content which health research professionals define as relevant to the assessment of quality of life in their fields. The assessment of the content validity of quality of life measures hold significant challenges, which this work has sought to address.

Suggested points of discussion:

- Which sort of journal is likely to be most interested in this paper/work?
- Are there particular parts of the work that should be emphasised more, or less?
- We make no comparison between the EQ-5D-5L and ICECAP-A. Would this be informative?
- Is adequate information given about the informant selection, recruitment and achieved sample?
- Is the method of analysis and presentation of the results of this analysis clear and accessible to the reader?

Reference List

(1) Brouwer WB, Culyer AJ, van Exel NJ, Rutten FF. Welfarism vs. extra-welfarism. J Health Econ 2008 Mar;27(2):325-38.
(2) Culyer AJ. The Normative Economics of Health Finance and Provision. Oxford review of Economic Policy 1989;5(1):34-56.
(3) Coast J, Smith RD, Lorgelly P. Welfarism, extra-welfarism and capability: the spread of ideas in health economics. Soc Sci Med 2008 Oct;67(7):1190-8.
(4) Chisholm D, Healey A, Knapp M. QALYs and mental health care. Social Psychiatry and Psychiatric Epidemiology 1997 Feb;32(2):68-75.
(5) Chalkidou K, Culyer A, Naidoo B, Littlejohns P. Cost-effective public health guidance: Asking questions from the decision-maker's viewpoint. Health Economics 2008 Mar;17(3):441-8.
(6) Coast J, Flynn T, Sutton E, Al-Janabi H, Vosper J, Lavender S, et al. Investigating Choice Experiments for Preferences of Older People (ICEPOP): evaluative spaces in health economics. Journal of Health Services Research & Policy 2008 Oct;13:31-7.

(7) Sen A. Equality of what? In: Sen A, editor. Choice, Welfare and Measurement.Oxford: Blackwell; 1980.
(8) Sen A. Inequality Reexamined. Oxford: Oxford University Press; 1992.
(9) Sen A. Capability and Well-Being. In: Nussbaum M, Sen A, editors. The Quality of Life.Oxford: Oxford University Press; 1993. p. 30-53.
(10) Sen A. Health: perception versus observation. BMJ 2002 Apr 13;324(7342):860-1.
(11) Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. Quality of Life Research 2012 Feb;21(1):167-76.
(12) Al-Janabi H, Keeley T, Mitchell P, Coast J. Completion of self-report health and capability measures: a think aloud study. Soc Sci Med (under review) 2012.
(13) Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of Life Research 2011 Dec;20(10):1727-36.
(14) Janssen MF, Birnie E, Haagsma JA, Bonsel GJ. Comparing the standard EQ-5D three-level system with a five-level version. Value in Health 2008 Mar;11(2):275-84.
(15) Janssen MF, Birnie E, Bonsel GJ. Quantification of the level descriptors for the standard EQ-5D three-level system and a five-level version according to two methods. Quality of Life Research 2008 Apr;17(3):463-73.
(16) Pickard AS, De Leon MC, Kohlmann T, Cella D, Rosenbloom S. Psychometric comparison of the standard EQ-5D to a 5 level version in cancer patients. Medical Care 2007 Mar;45(3):259-63.
(17) Borsboom D, Mellenbergh GJ, van Heerden J. The concept of validity. Psychological Review 2004 Oct;111(4):1061-71.
(18) Brazier J, Deverill M. A checklist for judging preference-based measures of health related quality of life: Learning from psychometrics. Health Economics 1999 Feb;8(1):41-51.
(19) Magasi S, Ryan G, Revicki D, Lenderking W, Hays RD, Brod M, et al. Content validity of patient-reported outcome measures: perspectives from a PROMIS meeting. Quality of Life Research 2012 Jun;21(5):739-46.
(20) Streiner DL, Norman GR. Health Measurement Scales: a practical guide to their development and use. 4 ed. Oxford: Oxford University Press; 2008.
(21) Brod M, Tesler LE, Christensen TL. Qualitative research and content validity: developing best practices based on science and experience. Quality of Life Research 2009 Nov;18(9):1263-78.
(22) Sireci SG. The construct of content validity. Social Indicators Research 1998 Nov;45(1-3):83-117.
(23) Lasch KE, Marquis P, Vigneux M, Abetz L, Arnould B, Bayliss M, et al. PRO development: rigorous qualitative research as the crucial foundation. Quality of Life Research 2010 Oct;19(8):1087-96.
(24) Cronbach LJ, Meehl PE. Construct Validity in Psychological Tests. Psychological Bulletin 1955;52(4):281-302.
(25) Carmines EG, Zeller RA. Reliability and Validity Assessment. 1 ed. London: Sage Publications; 1979.
(26) Ritchie J, Lewis J, Elam G. Designing ans Selecting Samples. In: Ritchie J, Lewis J, editors. Qualitative Research Practice.London: SAGE Publications; 2009. p. 77-108.
(27) Legard R, Keegan J, Ward K. In-depth Interviews. In: Ritchie J, Lewis J, editors. Qualitative Research Practice.London: SAGE Publishing; 2009. p. 138-69.