Developing a health, social care and carer-related quality of life instrument to measure QALYs

Authors: Tessa Peasgood[1], Jill Carlton[1], Janice Connell[1], Nancy Devlin[3], Monica Hernandez[1], Karen Jones[2], Clara Mukuria[1], Bhash Naidoo[4], Stacey Rand[2], Juan Carlos Rejon[4], Donna Rowen[1], Aki Tsuychiya[1], Barend van Hout[1], John Brazier[1]

[1]School of Health and Related Research, University of Sheffield
[2]Personal Social Services Research Unit, University of Kent
[3]Office of Health Economics, London
[4]National Institute for Health and Care Excellence

ABSTRACT

Common QALY instruments (EQ-5D, SF-6D etc.) measure and value changes in health or *health-related* quality of life. As such, they have limited ability to capture the wider impact on carers' quality of life, or outcomes in social care. Sector specific QALY instruments (such as ASCOT for social care, and CarerQol for carers) can be used within a sector to measure quality of life. However, the use of sector-specific measures is problematic when the impact of an intervention includes sector specific and health outcomes. This also reduces comparability of evaluations undertaken across different sectors using different measures.

Here we present an alternative approach: developing a generic QALY instrument that is sufficiently broad to capture important outcomes from health and social care interventions. We present a case for deriving the domains for a new instrument from the voice of patients, carers and social care users. Qualitative research that has explored the impact of health conditions, disability, social care needs and caring on people's lives offered a rich source of data, which was sampled to gain an understanding of frequently occurring and important themes.

The data extraction and the initial selection of domains from a review of the qualitative literature was aided by a flexible theoretical framework which supports clarity on 1) the breadth of domain coverage and 2) whether a domain can reasonably be considered to be of value for its own sake, rather than as instrumental to other valued, and adequately captured outcomes. Psychometric analysis, of both existing and bespoke data, along with extensive consultation with stakeholders, experts and the public was used to refine the domain structure and identify high performing items to measure each selected domain.

This paper (1) sets out the case for this approach and (2) describes how the approach has been operationalised through the Extending the QALY project - a collaboration between the Universities of Sheffield and Kent, the Office of Health Economics, and NICE.

1. Background

Economic evaluation combines information on costs and benefits to inform priority setting and the allocation of scarce health care resources. In the context of health care, one of the most commonly used methods is to estimate the incremental cost per Quality Adjusted Life Year (QALY) of new health technologies. Life years gained are assigned a value on a scale anchored to one at full health and zero as equivalent to being dead which reflects the preferences associated with different levels of health-related quality of life (HRQoL) or health states. This value represents the 'Q' part of the scale and measures such as EQ-5D and the Short Form-Six Dimensions (SF-6D) can be used to measure and generate this value. HRQoL can cover aspects related to symptoms such as pain or anxiety and functioning such as hearing or mobility but can also include the impact on activities such as self-care. These HRQoL measures may have limited ability to capture the impacts of health care interventions on others who are indirect beneficiaries such as informal carers. These measures are also limited to health and would not capture important outcomes in sectors such as social care e.g. independence, confidence, safety or indeed overall wellbeing/quality of life (Forder et al, 2007).

Other sector specific QALY instruments (such ASCOT for social care, and CarerQol-7D for carers) offer one solution to this limitation. Within sector QALYs can support within sector resource allocation (although even here sector specific outcomes may be too narrow). For between sector resource allocation, such as interventions that impact upon patients and their informal carers, an assumption would need to be made that a carer-QALY is of the same social value as a patient-QALY or a social care-related QALY. However, where outcomes of an intervention (such as a befriending club for the elderly) include sector specific (e.g. increased independence) and health outcomes (e.g. reduced depression) it is not possible to combine sector specific instruments (such as ASCOT and EQ-5D). Measuring both would lead to double counting yet measuring just one would exclude potentially important benefits.

One approach to addressing this difficulty is to develop a new quality of life preference-based measure that covers all domains that are relevant and important across health and social care service users and for carers, anchored to a QALY scale. This would be a generic instrument, relevant to health and social care, that could then be used to evaluate publicly funded interventions, enabling a single measure to be used across different sectors thus improving comparability of evaluations undertaken across sectors and the consistency of resource allocation decisions.

Development of a new measure requires identification of the relevant domains/dimensions. There are choices to be made regarding how this should be done. The aim of this paper is to consider the options available for identifying domains within the context of an ongoing research project entitled 'Extending the QALY' which is a collaboration between the Universities of Sheffield and Kent, the Office of Health Economics, and NICE.

This project is, in part, a response to growing call for a preference based measure that captures broader wellbeing rather than just HRQoL (Brazier and Tsuchiya, 2015; Johnson et al, 2016). However, we adopt the term 'quality of life' in favour of 'wellbeing' due to the reduced potential for confusion over the term 'wellbeing'. For some, 'wellbeing' aligns closely with emotional health or positive feelings (either directly experienced or in relation to an evaluation of one's life) yet for others 'wellbeing' captures all that matters in making a judgement about how well a life is going. The

term 'quality of life' is used as a placeholder for the specific theoretical judgement around what makes for a good life, and the relevant information for the purpose in which the term is being used. Here, we are interested in 'how *good* an individual's life is' or the 'overall judgement of how well an individual's life is going' without any restriction to either subjective or objective criteria (Andrews and Withey, 1976). Neither the content of quality of life, nor the justification of why or how something improves quality of life, is included within this definition.

At the outset, we are restricting the new measure to 'health, social-care and carer-related quality of life' rather than the whole of quality of life. This covers all aspects of life that could be affected by health conditions (physical and mental), health care treatments, self-management, disability, social care need or use, and the experience of caring.

2.   Options for identifying domains

There are a number of different approaches that could be taken to establish the constituents or domains of interest of any new quality of life instrument. This includes drawing from existing measures, drawing directly from theories or asking relevant stakeholders to identify domains.

2.1 Draw from existing HRQoL, Carer-related QoL and Social care-related QoL measures

A reasonable starting point for the development of a new instrument is an evaluation of what is currently available and used - and their strengths and weaknesses. This ensures that previous work is capitalized upon, including empirical work around the validity and sensitivity of current instruments and how to improve their measurement properties (e.g. Longworth et al, 2014). Looking for overlap between existing measures was the approach adopted by the CarerQol-7D, which drew upon commonalities between nine key carer burden instruments (Brouwer et al, 2006). Whilst this may be expedient it faces a number of problems.

Firstly, past limitations may be repeated, which includes missing out important domains or including those that are of minimal importance to service users. The research methods adopted by past instrument developers may not meet current standards since the methods of development of measures has developed significantly over the last 30 years with an increased focus on bottom-up rather than top-down approaches. Secondly, social circumstances (such as social attitudes or the physical environment) may have changed over time resulting in the impacts of certain conditions altering with time. Thirdly, the approach lacks theoretical rigour. Previous instruments may have different aims and objectives, and they may not have been designed to evaluate publicly funded interventions.

That said, understanding the domain content of existing measures is important as a content validity check. If the domains of the new instrument do not overlap with older instruments we would hope that there would be a reasonable explanation as to why this might be.

Table A.1 in the Appendix shows the domains covered within six commonly used HRQoL instruments, two social care-related or elderly quality of life instruments and three carer quality of life instruments. The breadth of coverage and the divergence between instruments, even within a sector, is substantial. The WHOQoL measure (WHOQoL Group, 1994) and the AQoL-8D (Richardson et al., 2013) are particularly different to the other HRQoL measures despite similar aims for use in

the evaluation of health interventions. This disagreement in the descriptive systems of HRQoL measures creates considerable difficulty in understanding the scope of the concept of HRQoL, and suggests a need for clarity on why each attribute should be considered a constituent of HRQoL and why it is sufficiently important to be included within the measure.

2.2 Draw upon a theory of quality of life

The second approach is to base instrument development around a theory of quality of life. For example, CASP-19 derived its content from adopting a theoretical approach to measuring quality of life in early old age in which quality of life is "assessed as the degree to which human needs are satisfied" (Hyde et al, 2003: 187), where those needs (control, autonomy, self-realization and pleasure) were grounded in the theoretical work of Maslow (1968), Giddens (1990) and Doyal and Gough (1991).

There are a number of reasons why having a clear theory of quality of life that drives the development of the measure is useful: (1) any normative judgements or values implicit within the theoretical framework can be made transparent and subject to scrutiny; (2) the scope of what is to be measured is known in advance. Whether something is included as part of quality of life is a question of whether it is compatible with the chosen theory. It is therefore possible to be explicit and transparent about the scope of measurement; (3) theories may be amenable to empirical testing e.g. testing the dimensionality of any model through psychometric analysis and (4) future users of the instrument are able to see its theoretical basis and judge whether this is in line with their own needs.

A clear theoretical basis is therefore useful both during development and for future use of an instrument. However, there are many competing theories about what makes for a good life – both in terms of *what* makes a life go better and *why* it does so. Most prominent contenders have also faced prominent criticism.

Proponents of a desire fulfilment account, in which life is improved if and only if individuals have more of the things they desire in their life, struggle to adequately deal with ill-informed, meaningless or unworthy preferences, nor measure preferences beyond the proxy of goods and services, nor address the role of expectations in forming desires (Sen, 1985).

Proponents of hedonism, in which the good life is one that has the most pleasure and the least pain (Crisp, 2006), struggle to address the critique that happiness is arguably not *'the only thing that we have reason to value, nor the only metric for measuring other things that we value'* (Sen, 2008: p26).

Similarly, proponents of subjective wellbeing (SWB), which is typically measured through life satisfaction, satisfaction with different domains of life, happiness and other positive and negative affects, struggle to address concerns that an individuals' subjective reports of their life relate strongly to their personal expectations and frames of reference. As noted by Felce and Perry, these frames of reference are

> *"…shaped by experience. One, cannot assume that a person's frame of reference will embrace all possibilities; it is affected by the judgment of what is possible and typical for a person in that situation."* (Felce and Perry, 1996: p65)

This paper does not attempt to give an overview of the defences against these criticisms but to simply note that they are unresolved and, as such, these approaches would be difficult to adopt as a theoretical basis to support public resource allocation.

If neither mental state accounts (hedonism, or SWB) nor preference satisfaction accounts are suitable, the remaining contender would be some form of objective list account, "according to which all instances of a plurality of basic objective goods directly benefit people" (Rice, 2013: p197). An objective list account of what makes a life go well does not provide an explanatory reason as to why something makes a life better. There is no single and common explanatory reason, such as being desired, or bringing happiness or satisfaction, that justifies any attributes presence on the list. Hence each attribute on the list requires additional justification (Fletcher, 2013).

Lists are derived from many different perspectives. Some are based on goods that are deduced from fundamental and moral reasons for actions (Murphy, 2010; Finnis, 1980), others based theories of justice and entitlement (Nussbaum, 2001), others from theories of the actualization of human potential (Maslow, 1987), others on perfectionism or the development and exercise of essential human capacities (Hurka, 1993), others on drawing together commonality across theories of psychological functioning (Ryff, 1989), others through bringing together approaches on quality of life from different disciplines - physiology, philosophy, economics (Kenny and Kenny, 2006).

Whilst there is much overlap, there is also disagreement. For example, in the inclusion of religion and practical reasonableness (Finnis, 1980), virtue (Fletcher, 2013), or material welfare (Kenny and Kenny, 2006). These disagreements arise from a number of sources including: a focus on the good life or a life that is good *for* the individual, different opinions on human flourishing, the role of human rights within a concept of a good life, and different reasoning around whether a good can be taken as having primary or instrumental value. Finnis, below, notes commitment to removing instrumental goods from his list:

> *"Now besides life, knowledge, play, aesthetic experience, friendship, practical reasonableness, and religion, there are countless objectives and forms of good. But I suggest that these other objectives and forms of good will be found, on analysis, to be ways or combinations of ways of pursuing (not always sensibly) and realizing (not always successfully) one of the seven basic forms of good, or some combination of them."* (Finnis, 1980: 90)

Whilst it might be possible to seek some overlap between different objective list accounts, theoretical agreement is highly unlikely. Consequently, any *purely* theoretically driven approach may not have sufficient legitimacy to support public resource allocation; it will struggle to meaningfully answer a challenge as to why one list is privileged over another.

2.3 Draw upon direct qualitative interviews with relevant future subjects

The third approach to selecting the content of the new measure is to draw directly from the voice and views of stakeholders of the new instrument.

Whom do you ask when designing a generic measure?

There are four potential stakeholder groups: members of the public; policy makers; experts and practitioners; and future respondents (i.e. health and social care users and carers).

All members of society are potentially future users of health and social care, current and future payers of health and social care services, and beneficiaries of public health initiatives. However, the average member of the public has less direct experience of health conditions, disability, social care needs, and caring than current health and social care users and carers.

The views of policy makers, or those who may wish to use a future instrument to aid their decision making, are important for instrument acceptability. Policy makers are (hopefully) capable of taking a slightly detached and broader view than immediate service users who are arguably too implicated to be involved in public resource allocation. A role with a responsibility for advising upon or making resource allocation decisions is likely to have encouraged considerable thought as to the relevant factors influencing decisions. Similarly, experts and practitioners in the relevant sector are likely to have given thought to their aims and what makes for a successful intervention. The social care-related quality of life measure, ASCOT, for example, was based initially on the views of policy makers and social care providers (extracted through interviews) and confirmed with service users (Netten et al., 2012). Whilst this is likely to generate outcomes that have practical value to policy makers, there is a concern that practitioners and policy makers may have opinions and life experiences that are not representative of the service user population and they may make different judgements about what is important to the service users' quality of life.

Focusing only on current users of health and social care and current carers has the advantage that it based on actual rather than hypothetical or indirect (i.e. experienced through others) or past experience. Indeed, best practice for the development of Patient Reported Outcome Measures (PROMs) is to begin with the perspective and voice of the patients (Rothrock et al, 2011). The perspective of current service users and carers will also ensure that domains will be appropriate for the time (and potentially place) in which the new measure is to be used.

How do you frame the questions/discussion about quality of life?
There are four different approaches to framing questions to ask respondents about quality of life in order to identify the themes or domains of importance:

a. Ask – is X important to your life?

One option is to identify candidate domains based on theory, existing literature, or views of experts, and then try and establish content validity through qualitative work to gather views of respondents towards these domains. Respondents may endorse or otherwise domains, but the activity will generate focusing effects. Disentangling whether an endorsement is genuine or a reluctance to challenge what is presented or the perceived wisdom of the researchers will be difficult.

b. Ask – what matters to your life?

Another option is to ask general questions about what is important to an individual's life or quality of life, or what matters to them. In 2010/11 the ONS ran a consultation around the UK which asked 'what things in life matter to you?', similarly within the development of the WHOQoL respondents were asked 'what matters to your life?' (WHOQoL, 1994).

This approach has three potential problems. Firstly, respondents may be unduly influenced by current issues, whether that be frustration with Council services (Evans, 2011) or distrust of politicians (Canavarro et al., 2009). There is a danger that the question is interpreted as 'what is bothering you in your life now'. Respondents may not raise important domains that are not at the forefront of their mind; they may not even be aware of the value of a particular good (such as mobility) until they face a limitation in that domain. Secondly, respondents may report what matters on the surface but not the underlying, more universal, concepts unless they are probed to do so.

Lastly, asking broader questions about what makes for a good life may include domains that are not likely to be sensitive to change following public policy intervention, and particularly those in health and social care. There is a potential empirical solution to this – start with a broad concept of quality of life and narrow the focus to those areas that are most sensitive to change following interventions based on empirical work.

c. Ask – 'what matters to your life' indirectly with additional content mining

Another option is to foster a discussion with respondents (e.g. patients or social care users) in which what matters to an individual's quality of life arises indirectly and the respondent is probed to consider the underlying values behind their views. A good example of this type of additional content mining around this question comes from the ICECAP instrument development. The interview process for the ICECAP-O is explained as:

> "*In-depth interviews were informant-led, opening with broad questioning about what was important to the older people, what they enjoyed, got pleasure from, or valued in their lives. …. As interviews progressed, the researchers used responsive questioning to probe underlying attributes of quality of life. So, for example, if an informant said that they valued their faith, the researcher would ask them first of all for further (factual) details, covering issues such as the faith followed, what this involves and so on. The researcher would then explore with the informant what it is about faith that brings quality to their lives, asking questions such as: What is it about your faith that is important to you? How does your faith make a positive contribution to your life? What is important to you about attending worship?*" (Grewal, et al, 2006).

This exploratory content mining approach (Al Janabi et al, 2012) begins with questions that are easier to answer e.g. *How do you spend your time at moment?*, and builds with the respondent a sense of what matters to them in their life through probing questions such as: "*What it is about these factors that is important?*". This is cognitively less demanding for the participants, and allows the interviewer to uncover underlying important aspects of what the respondent thinks makes their life go well. However, in this search for primary goods there may be a risk that a good, which has a

causal relationship to another good, is treated as purely instrumental – even if it may additionally hold non-instrument value to the individual.

      d.   Ask - how does your health condition, disability, health care, self-management, caring role, use or need for social care impact upon your life?

The last approach restricts the discussion to health, social care and carer-related experience. There is a clear distinction between asking broadly about an individual's quality of life versus asking about how a particular circumstance impacts on their quality of life. The interview may also use probing and less direct discussion about a respondent's experience of their circumstances or condition to draw out underlying themes.

Framing around the impact of health conditions, treatments, caring or social care use focuses the discussion onto domains that are likely to be sensitive to health and social care interventions. However, it also limits discussion to those facets in life that the respondent can cognitively attribute to a particular circumstance or condition. It may be difficult for a respondent to know whether their inability to concentrate or feel joy, for example, is a consequence of their health condition or caring responsibility.

3.   The approach adopted by the Extending the QALY project

The next section will detail the approach to instrument development adopted by the Extending the QALY project. The aim of the project is to develop a new instrument to measure quality of life in order to 'extend' the QALY beyond HRQoL to also capture the benefits of interventions on carers and those of relevance in social care. The new instrument needs to be fit for the purpose of supporting resource allocation whilst accurately reflecting the voice of the service user and carer. Figure 1 reflects these requirements, which are discussed in detail below.

1. Covers aspects that have been identified by service users and their carers as important to their quality of life

2. Meets predefined criteria based on being fit for purpose

E-QALY

3. Ideally, also in line with what policy makers in health and social care think is important to their decision making
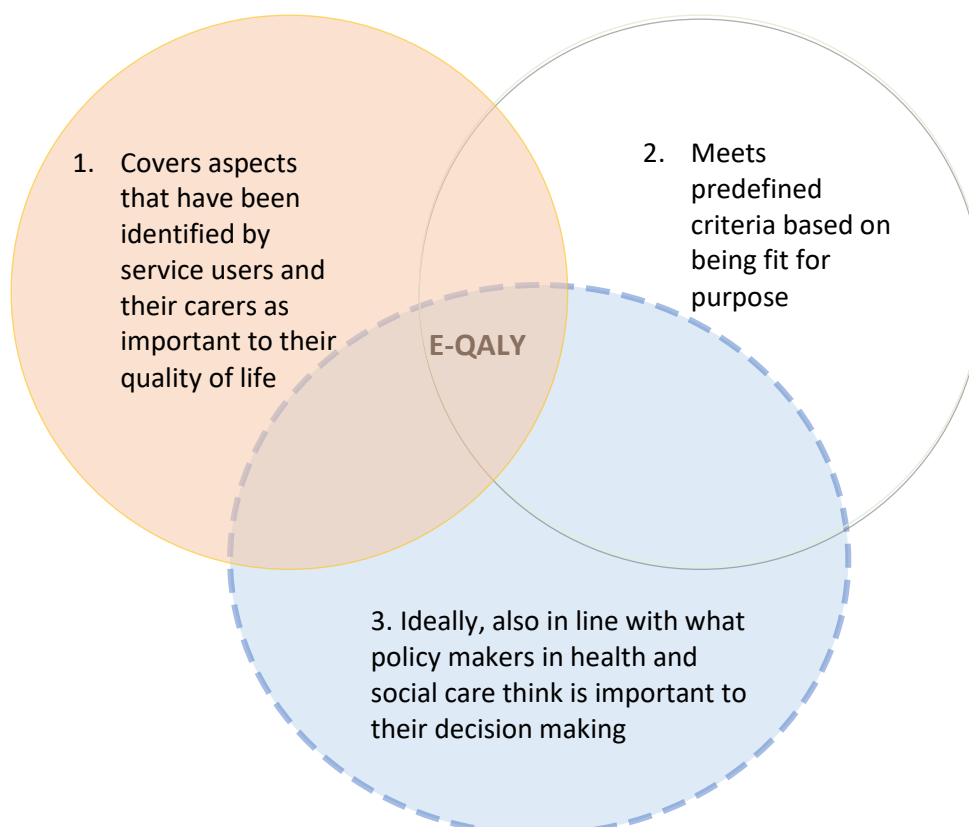
Figure 1: Requirements of the new instrument

Requirement 1: Covers aspects that have been identified by service users and their carers as important to their QoL

To be useful for resource allocation any new instrument must have broad public support and support within key subject groups (e.g. those who will be impacted upon by the type of decisions for which the instrument will be used, such as social care users, patients, carers). To facilitate this support the instrument should visibly be based on the views of service users and also align with common sense ideas of quality of life and individual wellbeing.

Requirement 2: Meets predefined criteria based on being fit for purpose

The instrument must be fit for the purpose of supporting resource allocation decisions in health care and social care. This intended use of the instrument imposes a number of constraints, including the need to conform to good measurement properties (see Fitzpatrick et al., 1998; Reeve et al., 2013). More specifically, the instrument should meet the following criteria:

1. hold legitimacy with the public, service users, carers, service providers, decision makers for use in supporting public resource allocation
2. provide a measure of quality of life as a single figure within a cardinal scale anchored on 0 (equivalent to dead) and 1 (equivalent to full quality of life) which reflects a social judgement on the value of that state
3. be able to be treated as having interpersonal and inter-temporal comparability with a reasonable degree of confidence
4. be inclusive and able to represent quality of life for all potential future respondents (i.e. different ages (above 18 years), different genders, those in or out of work, with or without children or close family, those close to death), or where this is not possible be transparent about its limitations to inclusivity
5. be amenable to routine use and inclusion in research studies and clinical trials, hence respondents should be able to complete questions tapping into each domain with minimal burden (in terms of time, and emotional and cognitive effort).
6. have good measurement properties including: high response rate; content validity (the process for development should be transparent and there should be broad agreement across stakeholders that meaningful outcomes are captured); face validity (there should be evidence of shared interpretation of items); be precise and able to distinguish between meaningful differences in quality of life across the full range of the scale; be responsive and sensitive to meaningful change; be reliable with high test-re-test reliability
7. avoid double counting where possible to avoid both instrumental and higher-level domains, and be transparent about limitations in its ability to do so
8. be translatable into a variety of languages and ideally be relevant across cultures.

Requirement 3: In line with what policy makers in health and social care think is important to their decision making

Figure 1 shows the third criteria, broad support across policy makers and practitioners to be an ideal rather than an essential requirement. This is based on a premise that whilst alignment between the views of policy makers and service users and their carers is ideal, the basis for the instrument content is primarily the voice of the service user and their carers. Should an attribute consistently be raised as important from the perspective of the service user (carer/patient/social care user) it would be included for consideration even if it was not also a concern of service providers/policy makers.

In relation to domain selection, the E-QALY project adopted a multi-pronged approach, drawing upon the most appropriate components of the approaches discussed above (2.1, 2.2 and 2.3). We aimed to learn from and apply the experience of past instrument development across health, social care, carers, disability and quality of life. We also set out a clear theoretical framework for the positioning of the instrument. Although the theoretical approach adopted does not dictate the content of the domains, it does commit to them being drawn from the future stakeholders. In addition, as discussed below, we also developed a 'working' conceptual model to help frame more detailed discussions around the domains and how they relate to each other and how individual components relate to the bigger picture.

3.1 The theoretical approach adopted

The theoretical underpinning for the instrument is extra-welfarist both in the commitment to a multi-dimensional measure of benefit (drawing upon an objective list account of quality of life) and the role of social preferences in judging the value attributed to different states. This is in line with Culyer (1990) who has argued that when evaluating health care in addition to utility the 'characteristics of people', should be taken into consideration. This incorporates the

> *"…characteristics of individuals (like whether they are happy, out of pain, free to choose, physically mobile, honest). Extra-welfarism thus transcends traditional welfare: it does not exclude individual welfares from the judgements about the social state, but it does supplement them with other aspects of individuals' (Culyer, 1991, p.67)*

The theoretical approach adopted also draws on the capabilities framework (Sen, 1985, 1993, 2008) in which the evaluative space is *"people's functioning's (their beings and doings) and capabilities (their real or effective opportunities to achieve those functionings)"* (Robeyns, 2005: p192). Sen recommends that the exact choice of functionings and capabilities be determined by deliberative process with stakeholders:

> "*when the capability approach is used for policy work, it is the people who will be affected by the policies who should decide on what will count as valuable capabilities in this policy question*" (Robeyns, 2005: 196)

However, unless the concept of capability or the value of having the opportunity for a functioning is raised as important by carers, health and social care service users themselves and can be measured

in such a way that the instrument remains fit for purpose, the focus here is on *actual* functioning rather than capability. Fleurbaey (2006) has argued that including actual functioning in addition to capabilities is still compatible with the capabilities approach, hence the difference in approach, particularly at the measurement level, may not be as large as it at first seems.

Whilst the lack of commitment to incorporating capability or freedoms, in addition to functioning distinguishes our approach from a capabilities approach, it is still worth acknowledging that this work draws on, and benefits from, the academic lineage of the capabilities framework.

3.2 Domain selection approach

The approach taken to select attributes of importance was to base these directly on the views of service users and carers identified through a large-scale qualitative literature review. Additional qualitative work was considered unnecessary based on breadth of existing high quality qualitative research (spanning a range of disciplines) that focuses upon health, social care and carer-related quality of life.

Drawing on existing literature enables coverage of many different health conditions, different types of social care users and carers across different age ranges, different ethnicities, and geographic locations (the review included literature from Europe, USA, Canada, Australia and New Zealand). It also enables the material to be drawn from many different qualitative interviewers and researchers, who may adopt different interview styles and questions, and approach their data with different assumptions and perspectives.

This breadth of coverage would simply not be feasible with primary qualitative work. Indeed, the relevant high quality published material is so extensive that for many areas of interest it was possible to rely upon existing qualitative reviews. The full details of the literature review will be reported elsewhere.

Data extraction and synthesis was conducted by four reviewers and applied the 'framework' method. This is a structured approach to organizing and analysing data that begins with a conceptual framework that is modified and expanded upon as new themes emerge from the data extraction (Ritchie and Spencer 1994). This required an *a priori* conceptual model to support data extraction. Setting out a provisional conceptual model, which was subject to on-going revisions, not only supported the data extraction during the review, but also created a framework for understanding and communicating the relationships between domains/sub-domains.

The structure of the initial conceptual model developed for the framework analysis was based on the widely known Wilson and Cleary's (1995) model of health-related quality of life[1]. Their model links biological and physiological variables to symptom status and then to functional health, then

---

[1] Alternative models were considered: including the WHO framework for health and disability, Felce and Perry 1995, Ferrans et al, 2005 revision to Wilson and Cleary (1995), Maslow's hierarchy of needs (1987), Ryan and Deci's Self-Determination Theory (2000), Cummins (2005). The use of the Wilson and Clearly (2005) framework was chosen because it provided a simple structure for direct and indirect impacts of circumstances and physiological variables and a useful way of thinking about the connection to environmental and individual mediators, without over-imposing model complexity that would be incompatible with our main approach.

health perceptions and finally overall quality of life. These pathways are placed in the context of personal and environmental factors, which act as mediators.

Given the aim of moving beyond HRQoL, we amended the Wilson and Cleary model to include aspects beyond biological functioning such as being a carer or being in receipt of health or social care (Figure 2). These circumstances may impact upon physical or mental symptoms or impairments, in addition to daily circumstances (such as time spent caring or undertaking treatment). These symptoms and circumstances were seen in turn as impacting on functioning and activity, social connectedness, feelings and physical sensations, and identity.

Figure 2. Initial conceptual model
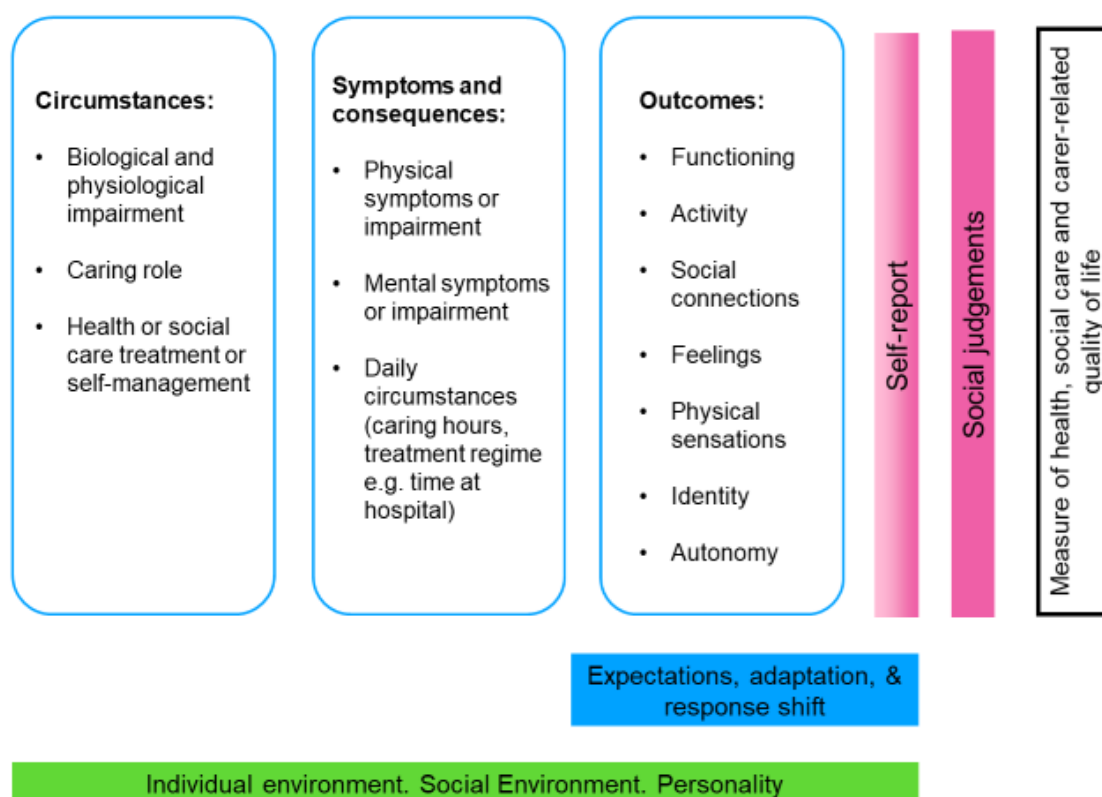
In line with Wilson and Cleary (1995) we placed the model within the context of a social and individual environmental and personality factors as mediators. Given that the instrument will be self-report we also flag the potential for expectations, adaptation and response shift to influence both the evaluation the individual makes about their functionings and other outcomes, and their self-reports of that assessment. Our model does not incorporate a stage of 'health perception' that is included in the Wilson and Cleary model since the value of the state will be based on social judgement rather than individual perception. A social judgement will be made about how these components combine to make up the overall quality of an individual's life.

Whilst there are some dominant causal pathways that run left to right (such as from a biological variable, through to symptom, through to ability to achieve a particular area of functioning), there

are also many complex bidirectional relationships and interconnections (social relationships for example may influence pain or mental health symptoms). Consequently, we have not included arrows in the conceptual model.

This conceptual model was initially populated with concepts that we expected to find in the review based on initial knowledge of the literature. Figure 3 shows the more detailed working conceptual model that was used for the framework analysis. As expected, some modifications were made during extraction to better suit the data. For example, the separation of mental health symptoms from feelings and emotions was not consistent with the data, and autonomy/control, which was initially within 'self-identity' was found to be conceptually quite distinct. The working conceptual model was repopulated with content drawn from the review.

Figure 3: Working conceptual model for data extraction
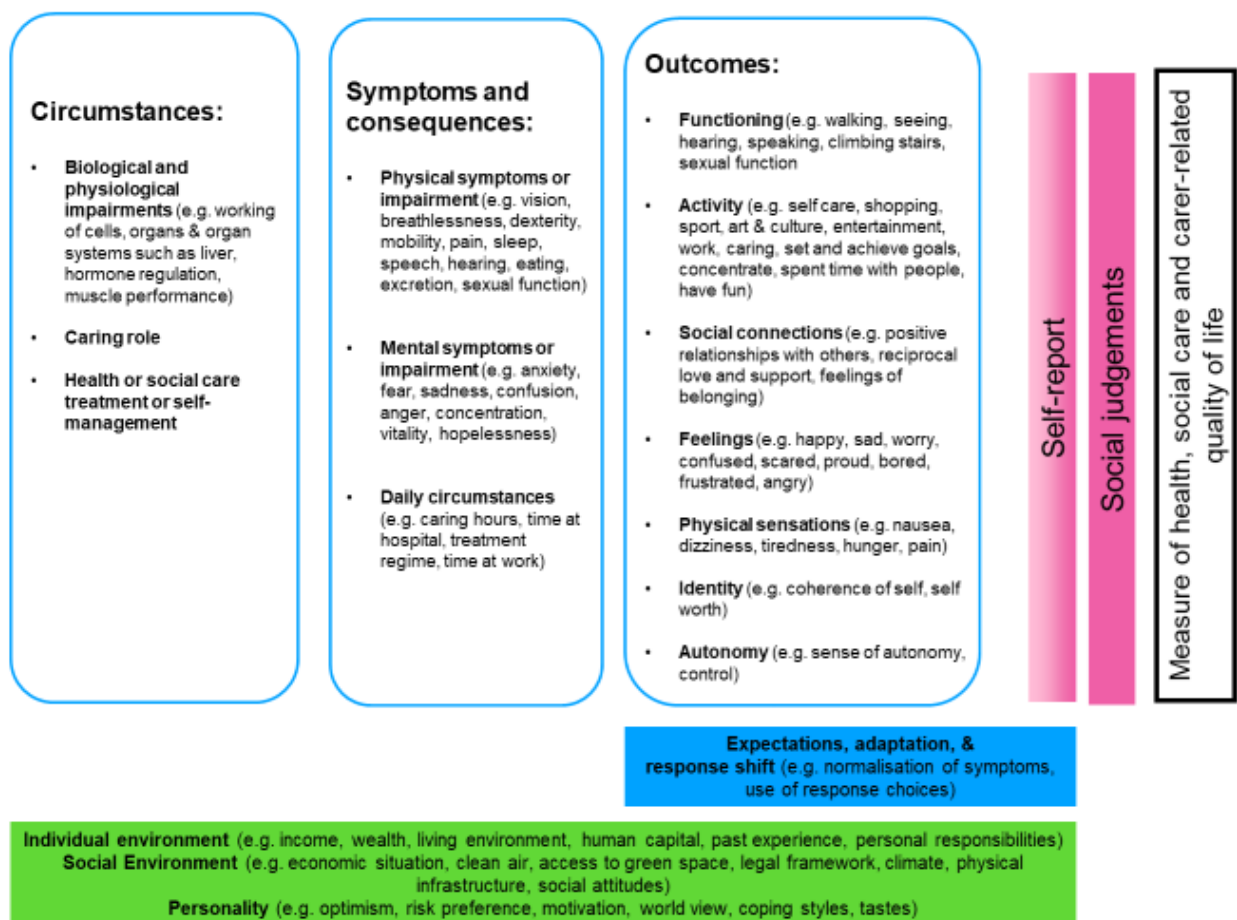
The synthesised data was used to identify the themes/domains and sub-themes/domains that are relevant for quality of life in the context of health, social care and caring. This resulted in a large number of themes and sub-themes each of which was required to meet certain pre-agreed suitability criteria before they could be recommended for consideration for inclusion with the measure.

There were four criteria applied. Firstly, there had to be reasonable consensus that this was an important aspect of quality of life for most people. In this case, the domain/sub-domain had to be raised as important for the quality of life of *most* patients, social care users and carers for the conditions or situations that were assessed in the review. The theoretical approach adopted provides no justification for consideration of inclusion beyond consensus or regular occurrence within the qualitative literature. The lack of an explanatory theory as to why an attribute makes life go better means there is no separate recourse to deal with conflicting evidence. Where an attribute is raised as a key component of quality of life by very few people, or is only relevant for particular groups, pragmatically, it does not fit well within a generic measure. For example, themes around work, sexual function and fertility were important to some people in some conditions. Whilst they would give rise to problems of relevance if kept within a standard generic quality of life measure, their exclusion is still problematic and recorded as a flag for future testing.

Secondly, again pragmatically, as a self-report instrument the measure can only contain domains/sub domains that are acceptable to respondents. Some respondents are likely to be unwilling to disclose information about sexual intimacy or suicidal ideation, for example. Domains that could be considered as judgemental, which may lead to social desirability issues, were also seen as impractical. For example, the literature review identified 'being a burden to others' as an important aspect of quality of life for some, however, this was not taken forward as it was considered too problematic in terms of acceptability.

Thirdly, domains/sub-domains were not taken forward where they were considered to be mostly instrumental, as long as their consequences could be captured elsewhere. The final list of domains should cover primary goods, or the things people care about for their own sake, otherwise some attributes will be double counted and the length of the instrument will be unmanageable. Instrumental/non-instrumental is not a straightforward categorization. The conceptual model helps in the consideration of what drives final outputs and where double counting may arise. However, many domains/sub-domains raised in the review can be thought of as both instrumental to something else of value (for example, pain has an impact on activity limitation) *and* of intrinsic value (it is just not nice to experience being in pain). Drawing the line between inclusion and exclusion based on instrumental value has a degree of arbitrariness. Dexterity was dropped on the basis that the majority of the impact of loss of dexterity would be picked up in the activities that the individual was able to do and their feelings of frustration and confidence. Knowledge and information were dropped on the basis that this would be picked up in feelings, activities, sense of control and coping. In both examples, a case could be made that they have intrinsic value.

Fourthly, domains/sub-domains which were considered to be covering very similar to concepts, or where one concept could be seen as a component of another concept, these were merged. For example, a construct around 'ability to keep things normal' was merged with 'coping', the constructs 'disclosure' and 'the reaction of others' were merged with 'stigma'.

The process of dropping and merging domains was iterative with extensive discussion. It included reflections back to the literature and to the terminology used by service users, and broad consultation across project governance and other consultation groups. These included a Public Involvement group, members of NICE staff, members of NICE Citizen's Council, the EuroQol Descriptive Systems Working Group, a project Steering Group and large, on-line Advisory Group.

Final selection of domains will draw upon face validity interviews and psychometric analysis of items developed to identify each sub-domain. This will help contextualise any theoretical differences between domains; something may be theoretically but not empirically distinguishable. A survey across a large sample of groups that represents future respondents will enable exploration of potentially redundant domains.

4. Discussion

There are a number of key strengths of the approach adopted by the E-QALY project. Firstly, the project applies best practice in instrument development in relation to explicit theoretical underpinnings, incorporation of qualitative work, the reliance upon the voice of the service user and the use of psychometric analysis.

Secondly, the project aims to be as transparent as possible in relation to decisions made throughout; including within the literature review, the criteria for the instrument overall and for selecting appropriate domains, and decisions made in matching potential domains to the criteria.

Thirdly, the project has, and continues to, engage in extensive consultation across range of stakeholders.

The focus on health, social care and carer-related quality of life whilst of benefit in terms of future sensitivity to interventions also potentially limits the instruments relevance for other sectors such as housing, prisoner wellbeing, community interventions – all situations where health may be one of a number of important outcomes. The validity and sensitivity of the new instrument can be tested in different groups in the future.

This approach may be seen as adopting a deficit model, in which problems and difficulties arising from caring, health or disability status are the focus rather than the positive effects upon human capacity and flourishing. There are two responses to this. Firstly, the domains/sub-domains identified in the review do capture attributes beyond problems and difficulties with basic living – such as self-worth, dignity, autonomy, happiness, hope, enjoyment of life, connectedness, and meaningful activity. Secondly, given that the opportunity cost of interventions evaluated using the new instrument may include additional years of life, a top anchor of a 'good' life is arguably more appropriate than a 'perfect' life.

The next steps of the project include face validity interviews across six countries (England, USA, Australia, Germany, China and Argentina) and a large UK psychometric survey, both of which will be used to support domain and item selection. This will be followed by valuation studies to enable the instrument to be scored in line with social values.

References:

Alborz A. The Nature of Quality of Life: A Conceptual Model to Inform Assessment. Journal of Policy and Practice in Intellectual Disabilities. 2017 Mar 1;14(1):15-30.

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 1;21(1):167-76.

Andrews FM, Withey SB. Social indicators of well-being: The development and measurement of perceptual indicators. New York: Plenum. doi. 1976;10:978-1.

Brazier J, Tsuchiya A. Improving cross-sector comparisons: going beyond the health-related QALY. Applied health economics and health policy. 2015 Dec 1;13(6):557-65.

Brouwer WB, van Exel NJ, van Gorp B, Redekop WK. The CarerQol instrument: a new instrument to measure care-related quality of life of informal caregivers for use in economic evaluations. Qual Life Res. 2006 Aug;15(6):1005-21.

Canavarro MC, Serra AV, Simões MR, Rijo D, Pereira M, Gameiro S, Quartilho MJ, Quintais L, Carona C, Paredes T. Development and psychometric properties of the World Health Organization Quality of Life assessment instrument (WHOQOL-100) in Portugal. International journal of behavioral medicine. 2009 Jun 1;16(2):116-24.

Crisp R. Reasons and the Good. 2006. Oxford: Oxford University Press.

Culyer AJ. Commodities, characteristics of commodities, characteristics of people, utilities, and the quality of life. 9-27 in Baldwin S, Godfrey C and Propper C. eds., 2002. *Quality of life: Perspectives and policies*. Routledge.

Cummins RA. Moving from the quality of life concept to a theory. Journal of Intellectual disability research. 2005 Oct 1;49(10):699-706.

Evans J. Findings from the national well-being debate. London: Office for National Statistics. 2011 Jul.

Felce D and Perry J. Assessment of Quality of Life in Schalock RL (ed) Quality of Life. 1996. Volume 1 Conceptualisation and Measurement, Washington: American Association on Mental Retardation.

Ferrans CE, Zerwic JJ, Wilbur JE, Larson JL. Conceptual model of health-related quality of life. Journal of nursing scholarship. 2005 Dec 1;37(4):336-42.

Finnis J. Natural law and natural rights. Oxford University Press; 2011 Apr 7.

Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess 1998;2: 1e74.

Fletcher G. A fresh start for the objective-list theory of well-being. Utilitas. 2013 Jun;25(2):206-20.

Fleurbaey M. Capabilities, Functionings and Refined Functionings. Journal of Human Development, 2006. 7: 299–310.

Forder J, Netten A, Caiels J, Smith J, Malley J. Theoretical concepts in assessing the outcomes of social care. PSSRU, University of Kent, Canterbury. 2007.

Grewal I, Lewis J, Flynn T, Brown J, Bond J, Coast J. Developing attributes for a generic quality of life measure for older people: preferences or capabilities?. Social science & medicine. 2006 Apr 1;62(8):1891-901.

Hyde M, Wiggins RD, Higgs P, Blane DB. A measure of quality of life in early old age: the theory, development and properties of a needs satisfaction model (CASP-19). Aging & mental health. 2003 May 1;7(3):186-94.

Hurka T. Perfectionism. Oxford University Press on Demand; 1996.

Johnson R, Jenkinson D, Stinton C, et al. Where's WALY?: A proof of concept study of the "wellbeing adjusted life year" using secondary analysis of cross-sectional survey data. *Health and Quality of Life Outcomes*. 2016;14(1):126. doi:10.1186/s12955-016-0532-5.

Kenny A. and Kenny C. Life, Liberty and the pursuit of utility: happiness in philosophical and economic thought. Andrews UK Limited; 2006.

Maslow AH. Toward a psychology of being (second edition). 1968. Princeton, NJ: Van Nostrand.

Murphy MC. Natural law and practical rationality. Cambridge University Press; 2001 Jun 11.

Netten A, Burge P, Malley J, Potoglou D, Towers AM, Brazier J, Flynn T, Forder J. Outcomes of social care for adults: developing a preference-weighted measure. Health Technology Assessment. 2012;16(16):1-66.

Nussbaum MC. Women and human development: The capabilities approach. Cambridge University Press; 2001.

Reeve, B.B. et al. (2013), "ISOQOL Recommends Minimum Standards for Patient-reported Outcome Measures Used in Patient-centered Outcomes and Comparative Effectiveness Research", Quality of Life Research, Vol. 22, No. 8, pp. 1889–1905, http://dx.doi.org/10.1007/s11136-012- 0344-y.

Richardson J, Iezzi A, Khan MA, & Maxwell A. Validity and reliability of the Assessment of Quality of Life (AQoL-8D) multi attribute utility instrument. *The Patient: Patient-Centered Outcomes Research*, 2013. doi: 10.1007/s40271-013-0036-x

Rice CM. Defending the Objective List Theory of Well‑Being. Ratio. 2013 Jun 1;26(2):196-211.

Robeyns I., Selecting capabilities for quality of life measurement, Social Indicators Research (2005) 74: 191–215

Rothrock N, Kaiser K, Cella D. Developing a Valid Patient-Reported Outcome Measure. *Clinical pharmacology and therapeutics*. 2011;90(5):737-742. doi:10.1038/clpt.2011.195.

Ryff CD. Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*. 1989. 57: 1069–1081.

Sen AK. Well-Being, Agency and Freedom: The Dewey Lectures 1984. The Journal of Philosophy, 1985. 82(4):169–221.

Sen AK. Capability and Well-Being. In M. Nussbaum and A. Sen (eds.), Quality of Life (pp. 30–53). Oxford: Clarendon Press. 1993.

Sen AK. The Economics of Happiness and Capability." In L. Bruni, F. Comim, M. Pugno (eds.), Capability and Happiness. New York: Oxford University Press. 2008.

WHOQoL Group. Development of the WHOQOL: Rationale and current status. International Journal of Mental Health. 1994 Sep 1;23(3):24-56.

Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life: a conceptual model of patient outcomes. Jama. 1995 Jan 4;273(1):59-65.

Table A.1. Commonly used instruments to measure HRQoL, SCRQoL, CarerQoL, QALY and QoL for the elderly

	Health related quality of life						Social care-related quality of life	Carer-social-care related quality of life	Carer-related quality of life		
	EQ-5D	SF-6D	HUI mark 3	15D	AQoL-8D	WHOQoL	ICECAP-O	ASCOT	ASCOT-Carer	Carer experience scale	CarerQol-7D
Mobility / walking / physical functioning	√	√	√	√	√	√					
Physical health											√
Self-care / Activities daily living	√				√	√					
Eating				√							
Elimination				√							
Household tasks					√						
Dexterity			√								
Breathing				√							
Usual activities / role limitation	√	√		√							
Pain/discomfort	√	√	√	√	√	√					
Cognition / memory / concentration /thinking			√	√		√					
Sexual activity / intimacy				√	√	√					
Vision			√	√	√						
Hearing			√	√	√						
Communication / speech			√	√	√						
Sleep				√	√	√					
Energy / vitality		√		√	√	√					
Control over daily life (or over caring) / autonomy / independence					√		√	√	√	√	
Personal cleanliness and comfort								√			
Food and drink								√			
Personal safety / security / freedom						√	√	√	√		
Occupation / work capacity /						√		√	√	√	

meaningful activity more i.e. leisure, hobbies, voluntary work, studying											
Accommodation cleanliness and comfort								√			
Looking after yourself well (sleep/diet)									√		
Time and space to be yourself									√		
Social participation and involvement / social functioning		√				√	√	√			√
Social isolation					√						
Social exclusion					√						
Community role					√						
Intimate relationships (close relationships)					√		√				
Enjoy close relationships					√						
Family role					√						
Getting on with the care recipient										√	√
Feeling supported or having support (from family/friends and/or external organisations)						√		√		√	√
Depression or anxiety or mental health	√	√									√
Depression/sadness				√	√						
Anxious / stressed /worried				√	√						
Contentment					√						
Happiness / unhappy / sadness			√		√	√					
Pleasure					√		√				
Enthusiasm					√						
Coping					√						
Feeling a burden					√						
Worthlessness / self-esteem					√	√					
Confidence					√						
Self-harm					√						
Despair					√						
Anger					√						
Tranquillity / calm					√						
Fulfilment (inc. with carer role)/ achievement							√			√	√
Financial						√					√
Body image and appearance						√					
Dependence on medicinal substances and medical aids						√					
Opportunities for acquiring new information and skills						√					

Transport						√				
Physical environment (pollution/noise/traffic/climate)						√				20
Participation in and opportunities for recreation/leisure						√				
Home environment						√				
Health and social care: accessibility and quality						√				
Religion/Spirituality/Personal beliefs						√				