Assessing sufficient capability

Mitchell, Paul Mark; Roberts, Tracy E.; Barton, Pelham M.; Coast, Joanna

DOI:
10.1016/j.socscimed.2015.06.037

Accepted Manuscript

Assessing sufficient capability: a new approach to economic evaluation

Paul Mark Mitchell, Tracy E. Roberts, Pelham M. Barton, Joanna Coast

PII: S0277-9536(15)30013-7

DOI: 10.1016/j.socscimed.2015.06.037

Reference: SSM 10145

To appear in: *Social Science & Medicine*

Received Date: 7 July 2014

Revised Date: 2 June 2015

Accepted Date: 26 June 2015

Please cite this article as: Mitchell, P.M., Roberts, T.E., Barton, P.M., Coast, J., Assessing sufficient capability: a new approach to economic evaluation, *Social Science & Medicine* (2015), doi: 10.1016/j.socscimed.2015.06.037.

Manuscript Number: SSM-D-14-01943

Manuscript Title: Assessing sufficient capability: a new approach to economic evaluation

Order of Authors:

Author 1 Paul Mark Mitchell[1]

Author 2 Tracy E. Roberts[1]

Author 3 Pelham M. Barton[1]

Author 4 Joanna Coast[1]

[1] Health Economics Unit, School of Health and Population Sciences, Public Health Building, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK.

Corresponding author: Paul Mark Mitchell, paul.mitchell@kcl.ac.uk, please note current affiliation of Paul Mark Mitchell: Department of Social Science, Health & Medicine, King's College London, WC2R 2LS, UK. Research conducted in Birmingham.

Acknowledgment/s: We would like to acknowledge the patients and researchers involved in the provision of data from the Tayside joint replacement cohort, in particular, Beth Pollard, Diane Dixon, Marie Johnston and Paul Dieppe. Thanks must also to: Hareth Al-Janabi, Alastair Fischer, Emma Frew, Philip Kinghorn and Sabina Sanghera for providing constructive suggestions on earlier versions of this research; Aki Tsuchiya for discussing an earlier version of this paper at the UK Health Economists' Study Group (HESG) in Exeter, January 2013; and attendees at the HESG in Exeter, the "New Directions in Welfare III - Economics for a Better World" conference at the OECD headquarters in Paris, July 2013, and the University of Birmingham Health Economics Unit (HEU) Journal Club, where versions of this work were presented. We would also like to thank three anonymous reviewers of this article.

ABSTRACT

Amartya Sen's capability approach has been discussed widely in the health economics discipline. Although measures have been developed to assess capability in economic evaluation, there has been much less attention paid to the decision rules that might be applied alongside. Here, new methods, drawing on the multidimensional poverty and health economics literatures, are developed for conducting economic evaluation within the capability approach and focusing on an objective of achieving "sufficient capability". This objective more closely reflects the concern with equity that pervades the capability approach and the method has the advantage of retaining the longitudinal aspect of estimating outcome that is associated with quality-adjusted life years (QALYs), whilst also drawing on notions of shortfall associated with assessments of poverty. Economic evaluation from this perspective is illustrated in an osteoarthritis patient group undergoing joint replacement, with capability wellbeing assessed using ICECAP-O. Recommendations for taking the sufficient capability approach forward are provided.

Keywords: capability approach, multidimensional poverty, decision rules, economic evaluation, sufficient capability

INTRODUCTION

Economic assessment is concerned with determining how best to use resources to achieve particular ends. Most economic evaluations in healthcare are currently conducted from a particular interpretation of the "extra-welfarist" normative stance in which the aim is to maximise health gains resulting from interventions (Brouwer et al., 2008; Coast et al., 2008d). The quality-adjusted life year (QALY), which combines health-related quality of life (HRQoL) and life years into a single outcome (Drummond et al., 2005), has become the standard outcome in health economic evaluations in the UK and beyond, although disability-adjusted life years (DALYs) are also used within this "extra-welfarist" approach (Culyer, 1989; Murray & Lopez, 1996). However, there is a narrative that suggests both QALYs (Drummond et al., 2009) and more recently DALYs (Nord, 2013) are narrowly defined in terms of health, rather than something "extra" to welfare assessment used in traditional cost-benefit analysis (Birch & Donaldson, 2003).

The application of Sen's capability approach (Sen, 2009) within health economics enables a broader measurement of wellbeing to be considered compared to existing approaches. Research in this area has focussed on developing questionnaires that capture capability for use in health interventions (Al-Janabi et al., 2012; Anand et al., 2009; Coast et al., 2008a; Kinghorn et al., 2015; Lorgelly et al., 2008; Netten et al., 2012; Simon et al., 2013). However, there is currently little guidance on how such measures should be used to aid healthcare resource allocation decisions within the capability approach. One suggestion is that capability measures be adjusted for time, thereby enabling assessment of gains in terms of 'years of full capability equivalence' (Flynn et al., 2015). This option has been used by some authors (Henderson et al., 2013; Makai et al., 2014).

In existing approaches to economic evaluation there is a focus on maximising outcomes (Birch & Donaldson, 2003; Coast, 2009), irrespective of the distribution of outcomes within society (Hurley, 1998). Although there have been some attempts to consider different aspects of equity within a health maximisation approach, these are secondary to maximisation and are selective in choice of equity considerations, focusing either on adjusting for non-health inequities in life (for example, income and ethnicity (Asaria et al., 2015)) or proportional shortfalls in health (van de Wetering et al., 2013); they have also seldom been used in practice. The capability approach has equity at its core, both theoretically and empirically (Anand & Dolan, 2005; Coast et al., 2008c; Sen, 2009; Simon et al., 2013; Venkatapuram, 2011) . The capability approach is concerned with people's ability to do valuable things in their life, rather than only focusing on things they actually do. The approach has been used to target those worst off in capability terms rather than seeking an optimal social welfare function of capabilities (Alkire et al., 2008; Clark & Qizilbash, 2008). Indeed, Ruger has theoretically explored the idea of focusing on shortfall equality in health capability as an appropriate decision-rule (Ruger, 2010).

The theoretical focus on those who are worst-off is also apparent in much of the empirical capability research. Research into human development and international poverty assessment has developed multidimensional poverty indices (MPIs) referred to as the Alkire-Foster (AF) measures (Alkire & Foster, 2011a). Alkire and Foster (2011a) do not focus on a single indicator of poverty such as income, arguing that such a focus can be misleading in describing the true levels of poverty within a given society. Instead, they present a method that allows for additional factors to be considered. Since 2010, multidimensional poverty has been compared across countries using the MPIs generated by the United Nations and are reported in their human development reports. The measure is based on three dimensions (health, education and living standards) across ten indicators of states of poverty (UNDP, 2010). This multidimensional approach is not unique to the poverty measurement, as the Organisation for Economic Co-operation and Development Better Life Initiative also draws on multiple indicators of wellbeing, including health, to develop a better life index as their alternative to relying solely on GDP for assessing a nation's progress (Stiglitz et al., 2009).

This general approach to focusing on those who are worst-off in society may provide an alternative decision rule for economic evaluation conducted from a capability perspective. This means interpreting equity objectives in terms of absolute (rather than relative) shortfalls in capability, in which this is understood as an absolute deprivation in what a person is able to do and be in life. It is therefore worth considering the AF method in more depth. The AF measures focus on two key issues of poverty measurement: first the "identification method", which considers how an individual is identified as being poor or not poor; and second the "aggregation method" which measures the deprivation for those identified as poor (Alkire & Foster, 2011a). Each of these is discussed now in detail.

Identification Method

Within a one-dimensional poverty measure, there is a relatively straightforward process involved in defining whether a person is in poverty – although that is not to understate the complexity of the actual decisions involved. It is a case of determining the threshold on that one dimension below which a person is considered to be in poverty (e.g. the World Bank uses less than $1.25 a day to define an individual as poor (Ravallion et al., 2009)).

This "identification method" becomes more complex in multidimensional poverty measurement. Here, the classification of an individual as poor requires a decision about the number of dimensions in which a person has to fall below the threshold. Atkinson (2003) outlines two common identification approaches in poverty assessment. First, the "union approach", whereby a person is classified as poor if they fall below the threshold on ***any*** single dimension. Second, the "intersection approach", whereby a person is poor only if they fall below the threshold in ***all*** included dimensions (Atkinson, 2003).

Alkire and Foster (2011a) found flaws with both identification methods and developed an alternative "in between" method for AF measures, referred to as the "dual cutoff" method. The dual cutoff method operates by first identifying a cutoff for each dimension below which a person is classed as deprived in that dimension, and second determining the number of dimensions in which an individual must be deprived to be classified as poor. These cutoffs can vary with context, enabling flexibility for specific purposes, whether it be a cross-national comparison of multidimensional poverty or a more specific policy question (Alkire & Foster, 2011b).

Aggregation Method

The AF measures provide a number of different aggregation methods, depending on the complexity of the poverty measurement required. Common to all is a "censoring" step, whereby those who do not meet the criteria for poverty (i.e. individuals not deprived in the required number of dimensions) are censored from the remainder of the poverty measurement exercise. Four methods of aggregation capture four different AF measures of multidimensional poverty. These consider (i) whether a person is poor or not poor (Headcount ratio (H)), (ii) for those identified as poor, how many dimensions they are poor, accounting for the *breadth* of poverty over dimensions considered (Adjusted headcount ratio (M_0)), (iii) how far away an individual is from the threshold on each dimension in which they are deprived, accounting for poverty *depth* within a dimension (Adjusted poverty gap (M_1)), and (iv) whether different weights across dimensions are attached to the same levels below the thresholds on dimensions, accounting for *severity* of poverty across dimensions (Adjusted Foster-Greer-Thorbecke (FGT) measure (M_2) (Foster et al., 1984)). The formulae for the four AF measures are presented in *Appendix A* [INSERT LINK TO ONLINE FILES].

Alkire and Foster (2011a) illustrated their AF measures for the United States, by measuring multidimensional poverty between three ethnic groups (African-American, Hispanic and White) across four dimensions (income, health status, health insurance, education level), with a cutoff of deprivation in two dimensions for a person to be considered poor. Using the Headcount ratio (H), Alkire and Foster (2011a) found the African-American population in their sample were the most impoverished when analysing income only. However, when using their dual cutoff approach for the Adjusted headcount ratio (M_0), the Hispanic population group were the most deprived overall, because health insurance and education attainment indicators were considerably worse for Hispanics. An example of the AF measure calculations is presented in *Appendix B* [INSERT LINK TO ONLINE FILES]. Although these methods are proving popular in the human development literature (Alkire & Santos, 2013), to utilise this approach within economic evaluations requires modification to account for the longitudinal effectiveness of health interventions.

The aim of this paper is to develop and illustrate a new approach for using capability instruments to inform health and social care decision-making, drawing from the literature on health economic evaluation and the Alkire-Foster measures for assessing poverty. We refer to this method as the “sufficient capability approach”. By adapting and further developing the Alkire-Foster (2011a) multidimensional poverty methods, we demonstrate how capability deprivation can be measured in a health setting (although the method is not limited to this setting) and across time.

First, the paper focuses on the development of these novel methods for conducting economic evaluation from a capability perspective by generating the sufficient capability approach. Three steps are considered: defining a threshold of sufficient capability; developing a sufficient capability score; and generating methods for assessing sufficient capability over time. Second, the application of the approach is then illustrated using a joint replacement dataset for osteoarthritis patients. A discussion on the sufficient capability approach, in light of previous attempts to use the capability approach concludes this paper.

DEVELOPMENT OF THE SUFFICIENT CAPABILITY APPROACH

If we are interested in focusing resources on bringing individuals up to a minimum threshold of capability, then different methods of analysis are needed to those used in economic evaluation focused on maximisation. As the measures developed by Alkire and Foster focus on broad indicators, produced at a national level and used in a cross-sectional manner, it is not immediately clear how they can be directly applied in the health and social care decision-making context. Here, methods are developed that can be applied in such a context, with a focus on those people who do not reach what is considered a minimum acceptable level of capability – a level of capability that is here deemed "sufficient". This sufficient capability approach involves three steps. The first step is to define a threshold for sufficient capability. The second and third steps adapt and further develop the AF method for use in economic evaluation; in the second step, methods are developed for generating a 'sufficient capability score' (SCS) and in the third step, SCS is combined with length of time to produce a capability outcome over time.

The approach is developed for use with an index of capability. The illustration uses the ICECAP-O due to data availability, but other capability instruments could potentially be used (Coast et al., 2015; Lorgelly, 2015). The ICECAP-O is a simple, self-reported questionnaire and is designed to measure capability wellbeing for older people (Coast et al., 2008a). It was developed primarily for use in resource allocation decisions across both health and social care for people aged 65 years and older (Grewal et al., 2006). The instrument consists of five attributes of capability wellbeing, each measured across four levels (Table 1). The five conceptual attributes are attachment, security, role, enjoyment and control, and within the instrument these are, respectively, expressed as the ability to: have love and friendship; think about the future without concern; do things that make you feel valued; have enjoyment and pleasure; and be independent (Grewal et al., 2006).

The values within ICECAP-O were developed using best worst scaling (Flynn et al., 2007) and are on an interval scale anchored at 1 (equivalent to full capability on all attributes (44444)) and 0 (equivalent to having no capability on each of the five capability dimensions (11111)) (Coast et al., 2008a). Construct validity for the measure has been established for general populations in the UK (Coast et al., 2008b; Flynn et al., 2011) and Australia (Couzner et al., 2013b). In a number of patient populations, the validity (Horwood et al., 2014; Makai et al., 2012; Makai et al., 2013), responsiveness (Comans et al., 2013; Couzner et al., 2012; van Leeuwen et al., 2015) and relationship with health status (Couzner et al., 2013a; Davis et al., 2012; Davis et al., 2013; Mitchell et al., 2013) have been tested.

Defining the Threshold of Sufficient Capability (TSC)

Using the AF methodology we set a minimum threshold level of capability that a person must achieve to be considered to have a "sufficient" level of capability wellbeing. We refer to this minimum level as the "threshold of sufficient capability (TSC)" and it is defined as the level of capability at or above which a person's level of capability wellbeing is no longer a concern for policy. TSC is defined as the level of capability deemed sufficient for each dimension under consideration. TSC also accounts for the number of dimensions (cutoff number, k) which a person needs to fall below to be classed as not having sufficient capability.

A person who has reached a level of 'sufficient' capability across all attributes will have reached an adequate level of wellbeing, indicating that further allocation of resources to this individual is no longer a priority. A person who does not reach the sufficient level on k attributes will fall below TSC. The focus of evaluation becomes the movement of people closer to TSC (as efficiently as possible).

To apply this approach in practice, the first step in defining the TSC using the AF measures is the "identification method" (Alkire & Foster, 2011a). This requires determining the threshold level on each dimension (dTSC) below which there is considered to be a shortfall in sufficient capability. The capability measure employed here, the ICECAP-O, has four levels of capability for each of its five attributes, conceptually ranging through full capability (level 4) , a lot of capability (level 3), a little capability (level 2) and no capability (level 1) (Coast et al., 2008a). In theory, a large number of different thresholds could be implemented as there is no need for the level in which a person is considered to be in capability poverty to be consistent across attributes. Here, we propose two possibilities for ease of interpretation:

- Option 1: assuming that if a person has at least 'a lot' of capability (i.e. level 3) on each attribute they have sufficient capability ***("33333")***.
- Option 2: assuming that if a person has at least 'a little' capability (i.e. level 2) on each attribute they have sufficient capability ***("22222")***.

To use the threshold, the original values of ICECAP-O are re-scaled so that 1 is equal to the TSC. This new objective only gives priority to those below sufficient capability. Our formula for calculating values for each threshold dimension is presented below:

$$V_{tx} = \frac{v_{tx}}{(v_{t1}+\ldots+v_{tn})} \quad (1)$$

Here V_{tx} = new value on threshold dimension level t in dimension x, v_{tx} = original value of the threshold level t in dimension x, $v_{t1}+\ldots+v_{tn}$= sum of threshold level t values across all dimensions before transformation (i.e. original ICECAP-O values). For example, in Table 1, level 3 on the attachment attribute for threshold Option 1 "33333", $v_{tx} \approx 0.2325$; $(v_{t1}+\ldots v_{tn})$ when TSC "33333" ≈ 0.868; thus $V_{tx} \approx 0.2679$.

Any scores higher than the threshold of sufficient capability for each option will be given the value of the sufficient capability threshold (i.e. 1). Any shortfalls in capability below this threshold are then allocated a shortfall value according to both:

- ***the extent of that shortfall*** (whether at the level of ‘no capability’ or ‘a little capability’ for option 1; not applicable for option 2 as only one level ‘no capability’ below threshold);

and

- ***the rescaled ICECAP-O population values***. The ICECAP-O general population based value set is additive and on a linear scale, such that the numerical value is meaningful and the values across all attributes can be summed to give an overall index between 0 - representing no capability – and 1 - representing full capability. For options 1 and 2, the index score of 1 will now represent TSC in each scenario depending on each threshold level respectively, i.e. a value of 1 represents sufficient capability. Rescaled values for both threshold options can be seen in Table 1.

INSERT TABLE 1

Determining the Sufficient Capability Score (SCS)

Once the threshold has been determined, values for levels below the threshold need to be calculated to reflect societal values for these states of capability wellbeing. For values below the threshold for a dimension (dTSC), the following method is used to calculate the capability value compared with the threshold level:

$$V_x = \frac{v_x}{v_{t1}+v_{t2}+...v_{tn}} \text{ ; unless } v_x > v_{tx,} \text{ then } V_x \equiv V_{tx} \text{ for TSC } = 1 \quad (2)$$

This calculation generates a new 0-1 value scale for ICECAP-O, with 0 still reflecting the "no capability" response levels on all ICECAP-O attributes. However, 1 no longer reflects "full capability" across all attributes, instead this reflects the threshold of sufficient capability (TSC). To calculate an overall reflection of sufficient capability, individual responses are summed across this new scale to calculate an individual's Sufficient Capability Score (SCS). To calculate SCS for an individual, the values attached for each threshold option proposed in the previous section are presented in Table 1. For example, using threshold option 1 "33333" as the level where sufficient capability is reached, an ICECAP-O profile of "43233" has the same SCS score as an ICECAP-O profile of "44244" (an example of calculating sufficient capability is given in *Appendix C* [INSERT LINK TO ONLINE FILES]). This aggregation approach is comparable to the adjusted FGT (M_2) AF measures, albeit with higher scores reflecting improvement in capability.

Estimating Sufficient Capability over time: Years of Sufficient Capability (YSC)

The third step involves a key calculation in health evaluations, which is to consider both wellbeing (however defined) and changes in wellbeing over time, as interventions are aimed at morbidity and mortality reductions. This is something which has not been tackled within the capability literature and has been identified as an issue for practical evaluations (Alkire et al., 2008). This aspect of the sufficient capability approach aligns itself with current methods applied to generate health economics outcomes that combine quality with length of time in the state, like QALYs.

At this stage it is important to note the different anchors on HRQoL measures such as the EQ-5D (Brooks, 1996) and the ICECAP-O. Extra-welfarist HRQoL measures are anchored on a 0-1, dead to full health scale, where it is possible to have states worse than dead depending on the valuation method used (Brazier et al., 2007). The ICECAP-O is anchored on a 0-1 scale, where the anchors represent no capability to full capability (Coast et al., 2008a). The ICECAP-O is anchored differently to HRQoL measures used to produce QALYs. In interpreting the zero value on the ICECAP-O index it is important to note that:

> *"A number of states may produce such a zero value: assessment of capabilities as being non-existent in relation to all attributes; unconsciousness; and death"* (Coast et al., 2008a, p. 878)

SCS is a flexible measure which can be applied to maximise capability levels to the TSC, or inversely minimise shortfalls from the TSC. Since the focus in this paper is to compare the sufficient capability approach with current practice within the UK and other developed nations where the QALY is the primary economic outcome, only maximising SCS over time is explored here. To achieve this, Years of Sufficient Capability (YSC) are generated to give a longitudinal measure representing gains in sufficient capability over time. In cases where SCS remains constant over a period of time, the equation for YSC is:

$$YSC = SCS \times T \quad (3)$$

where SCS = Sufficient Capability Score and T = time. A year of life in sufficient capability has the value one; a year with no capability (whatever the reason for this) has a value of zero. If SCS varies over time, YSC is calculated using the area under the curve approach, as used in QALY calculations (Drummond et al., 2005). For illustrative purposes, a comparison is made with YSC and an outcome using the original ICECAP-O valuation dataset, which has been referred to as Years of Full Capability (YFC) equivalence (Flynn et al., 2015).

ILLUSTRATION OF THE SUFFICIENT CAPABILITY APPROACH

A dataset from the clinical orthopaedic area of joint replacement is used to illustrate the potential of the sufficient capability approach within a clinical context. The dataset is a subset of the Tayside Joint Replacement cohort (Pollard et al., 2009). The data applied here were collected as part of the UK Medical Research Council "MOBILE" Health Services Research Collaboration and received ethics approval from the Tayside Committee on Medical Research Ethics.

Between September 2006 and June 2007, 107 patients about to undergo primary joint replacement surgery at Ninewells Hospital, Dundee, UK, were asked to complete the ICECAP-O at baseline, one year and three years. Although a relatively small dataset for health economic analysis, it is the first context in which pre-intervention and post-intervention data are available for the ICECAP-O. It provides adequate information for illustration purposes of the sufficient capability approach.

The demographics of this sample are summarised in Table 2. At baseline, the average age of this population was 69.72 years. The ICECAP-O scores at baseline for the 106 patients who completed the ICECAP-O was 0.773, less than the ICECAP-O average from an over 65 UK population sample of 0.832 (Flynn et al., 2011).

INSERT TABLE 2

Defining the Threshold of Sufficient Capability

In the first analysis, the results for patients who completed the ICECAP-O at all three time-points (n=42) are presented for both threshold options at “a lot of capability 33333” and “a little capability 22222”. The Alkire-Foster measures for poverty, in our case capability poverty, are calculated at all possible cutoffs (k) for the three time periods for both TSC options. This information is then used to justify the choice of cutoff for calculating SCS in the succeeding analysis. Table 3 presents the AF measures for the “33333” TSC and Table 4 presents the AF measures for the “22222” TSC.

From both Tables 3 and 4, it is clear that the choice of TSC and cutoff is crucial in measuring the level of poverty in terms of sufficient capability for a given population. Using the simplest AF measures, the Headcount Ratio (H), and the cutoff (k) = 1, for shortfalls in sufficient capability to occur, comparing Table 3 and 4 shows that whilst almost three quarters (74%) of the population have shortfalls in sufficient capability at the TSC of “33333”, only one in six (17%) of the population have shortfalls at the lower threshold of “22222”. Also, it can be seen in Tables 3 and 4 that the cutoff (k) can play an important role in identifying those who are below sufficient capability, as k=4 leads to everyone having sufficient capability at threshold “22222” and 1 out of 6 below sufficient capability at “33333”. Given the small numbers who report shortfalls in sufficient capability for the “22222” threshold prior to treatment, changes in levels of sufficient capability are unlikely to be captured by the intervention using this threshold. Although this should not be used as a basis for defining a poverty level, for the purposes of illustrating this method, this means that the “33333” threshold is more helpful. Therefore, only the “33333” is employed for the calculation of the SCS in the remainder of this paper, where k = 1 is employed for ease of interpretation.

INSERT TABLE 3 AND 4

Determining the Sufficient Capability Score (SCS)

Using “33333” as the TSC, the average levels of SCS at baseline, 1 year and 3 years post-intervention are presented in Tables 5, 6 and 7 respectively. Overall, the SCS for the baseline population (n=106) is 0.857, which is lower than the average of the sample from the ICECAP-O valuation dataset (Coast et al., 2008a), whose average SCS score is 0.894. Table 5 shows the attributes in which shortfalls in sufficient capability occurred pre-intervention, with the “enjoyment” attribute reporting the highest number of shortfall responses below the “33333” threshold (46%). In Table 6, SCS is calculated for patients who completed the ICECAP-O at one year post-operation (n=58). This resulted in an increased SCS score of 0.05 (0.88→0.93) from baseline. Patients who completed ICECAP-O at 3 years post-intervention (n=55) also reported an improved SCS from baseline by 0.031 (0.881→0.912), which can be seen in Table 7. In Tables 6 and 7, the improvements in the lower levels of capability come predominantly from the “role” and “enjoyment” ICECAP-O attributes.

INSERT TABLES 5, 6 and 7

Estimating Years of Sufficient Capability (YSC)

The final analysis involves calculating the change in SCS over time to generate Years of Sufficient Capability (YSC). Only the 42 individuals who completed the ICECAP-O at all three time-points are included in this illustrative calculation. SCS for these patients at baseline is 0.871, which is assumed to stay constant if the intervention is not provided to the patients (although in practice the condition would be expected to worsen). SCS at one year (0.923) and SCS at three year post-intervention (0.902) are used to calculate the intervention group, with each SCS assumed to be connected linearly. Figure 1 shows this calculation graphically, with the darker area displaying the YSC gain from treatment, calculated as follows:

$$[0.5\times (0.871+0.923)+(0.923+0.902)]-[3\times 0.871] = 0.11.$$

INSERT FIGURE 1

Comparing with Years of Full Capability equivalence

As a comparator for YSC, Years of Full Capability (YFC) equivalence are reported, where the original ICECAP-O values are employed. YFC are calculated using a similar under the curve approach as for YSC. At baseline, ICECAP-O original values for the patients are equal to 0.789. ICECAP-O values at year one (0.851) and at year three (0.824) result in YFC gained of 0.13 when compared with the ICECAP-O baseline score (0.789) over a three year period:

$$[0.5\times (0.789+0.851)+(0.851+0.824)]-[3\times 0.789] = 0.13.$$

DISCUSSION

In this study, a novel method for applying the capability approach to inform healthcare decision-making has been developed. This involved extending the Alkire-Foster methods in the context of a measure with index values and accounting for duration. We have shown how to; define a threshold for sufficient capability, generate a sufficient capability score and use these scores to produce a capability outcome over time. Drawing from multidimensional poverty methodology (Alkire & Foster, 2011a), there are two options for measuring what constitutes sufficient capability: cross-sectional and longitudinal. From our illustrative example of the application, the cross-sectional multidimensional poverty assessment showed that large differences occur depending on which threshold of sufficient capability is employed (Tables 3 and 4). Years of sufficient capability and years of full capability equivalence were calculated as outcome measures, combining capability wellbeing with duration. The osteoarthritis example employed here showed little difference in moving from full capability (YFC gained=0.13) to sufficient capability (YSC gained=0.11). In other interventions that have a greater impact on those below the sufficient capability threshold, there may be bigger differences in the resource allocation priorities if this approach was adopted.

This study uses an illustrative example to show how the capability approach can be used to assess capability deprivation applying and modifying methods from the multidimensional poverty literature. We also demonstrate how sufficient capability can be calculated and generated within economic evaluations.

Whilst it has been stated elsewhere that years of full capability equivalence can be calculated by using capability questionnaires (Flynn et al., 2015), it is argued here that years of sufficient capability, with its greater focus on those in capability poverty, is more closely aligned to the theory underpinning the capability approach. Our approach represents a method of tackling equity by shifting the evaluation space from health to capability and moving from maximisation to sufficiency. Decision-makers may be interested in using this approach alongside QALYs or in isolation.

The osteoarthritis example here was used only for illustrative purposes to demonstrate how to apply the sufficient capability approach. The limitations of our study include the small sample size and the illustrative control group as a comparator. However, when applying this method in a comparable way to how standard economic evaluations are used in practice, it would be expected that such data limitations would not be an issue. In terms of strengths, for policymakers, this is the first study to show how to conduct a capability-based analysis within a health economics setting using capability questionnaires as the basis for measuring change in individual wellbeing. Whilst the theoretical arguments of the capability approach are compelling, the practicality of applying capability analysis has been lacking (Sugden, 1993). This study attempts to help bridge this gap between theory and practice. Unlike previous studies that have argued for the adoption of QALYs as a best estimate of capabilities (Bleichrodt & Quiggin, 2013; Cookson, 2005), we have developed the sufficient capability approach for assessing capability directly.

In this paper we have developed a method that enables researchers to apply the capability approach in practice. However, this study focuses only on the benefit side of economic evaluation. Further consideration of the costing perspective when adopting a capability approach needs to be addressed. Furthermore, to enable the use of capability measures on a wider scale it is important for the threshold for a sufficient capability level for assessing capability deprivation to be agreed and this may vary between jurisdictions. The English health and social care guidance body, the National Institute for Health and Care Excellence (NICE), has recently recommended the use of capability measures in economic evaluations assessing social care (NICE, 2014). More research is required as to how a national guidance body, such as NICE, who may adopt the sufficient capability approach, would assign a threshold level for their population of interest. Most thresholds and weighting systems in multidimensional wellbeing indices are judged arbitrarily by the researcher in question

(Decancq & Lugo, 2013), this *ad hoc* approach is unsuitable for practical decision making. There is some evidence of quantitatively setting core poverty thresholds that could potentially be adopted (Clark & Qizilbash, 2008). Alternatively, it might be more appropriate to pursue qualitative participatory methods for setting such a threshold for a given society (Coast, 1999).

Once a threshold is agreed, future research could focus on using the sufficient capability approach in studies with larger datasets and also to compare results across patient groups/interventions using this type of analysis. Previous research has shown, for example, where willingness-to-pay could lead to different priorities than using the cost-per-QALY approach (Olsen & Donaldson, 1998) and a similar comparative analysis could be conducted for the sufficient capability approach.

Table 1 The ICECAP-O Questionnaire (Coast et al. 2008a) and response values

ICECAP-O dimension	ICECAP-O Values	TSC "33333"	TSC "22222"
Attachment (Love and Friendship)			
(4) I can have all of the love and friendship that I want	0.2535	0.2678	0.2412
(3) I can have a lot of the love and friendship that I want	0.2325	0.2678	0.2412
(2) I can have a little of the love and friendship that I want	0.1340	0.1545	0.2412
(1) I cannot have any of the love and friendship that I want	-0.0128	-0.0147	-0.0230
Security (Thinking about the future)			
(4) I can think about the future without any concern	0.1788	0.1234	0.1189
(3) I can think about the future with only a little concern	0.1071	0.1234	0.1189
(2) I can only think about the future with some concern	0.0661	0.0761	0.1189
(1) I can only think about the future with a lot of concern	0.0321	0.0370	0.0578
Role (Doing things that make me feel valued)			
(4) I am able to do all of the things that make me feel valued	0.1923	0.2066	0.2332
(3) I am able to do many of the things that make me feel valued	0.1793	0.2066	0.2332
(2) I am able to do a few of the things that make me feel valued	0.1296	0.1494	0.2332
(1) I am unable to do any of the things that make me feel valued	0.0151	0.0174	0.0272
Enjoyment (Enjoyment and Pleasure)			
(4) I can have all of the enjoyment and pleasure that I want	0.1660	0.1893	0.2132
(3) I can have a lot of the enjoyment and pleasure that I want	0.1643	0.1893	0.2132
(2) I can have a little of the enjoyment and pleasure that I want	0.1185	0.1365	0.2132
(1)I cannot have any of the enjoyment and pleasure that I want	0.0168	0.0193	0.0302
Control (Independence)			
(4) I am able to be completely independent	0.2094	0.2129	0.1936
(3) I am able to be independent in many things	0.1848	0.2129	0.1936
(2) I am able to be independent in a few things	0.1076	0.1240	0.1936
(1) I am unable to be at all independent	-0.0512	-0.0590	-0.0922

TSC "33333" & "22222" - Threshold levels of sufficient capability

Table 2 Baseline Descriptive Statistics for Tayside replacement dataset

	N	n	Mean	SD	Minimum	Maximum
Sample size	107					
Missing data	106	1				
Males	106	55				
Employed	106	13				
Living Alone	106	28				
Age (mean)	106	106	69.720	8.854	48.000	89.000
ICECAP-O (base)	106	106	0.773	0.167	0.159	1.000
ICECAP-O (1 year)	106	58	0.862	0.132	0.516	1.000
ICECAP-O (3 year)	106	55	0.832	0.138	0.481	1.000
ICECAP-O complete (base)	42	42	0.789	0.132	0.368	0.998
ICECAP-O complete (1 year)	42	42	0.851	0.134	0.516	1.000
ICECAP-O complete (3 year)	42	42	0.824	0.146	0.481	1.000

N,total population; n,sub population from N; SD, standard deviation;

Table 3 AF poverty measures applied to ICECAP-O levels using "33333" as the capability poverty threshold (n=42)

33333	BASELINE						1-YEAR POST-OP						3-YEAR POST-OP				
CUTOFF(k)	k=1	k=2	k=3	k=4	k=5		k=1	k=2	k=3	k=4	k=5		k=1	k=2	k=3	k=4	k=5
AF						**AF**						**AF**					
H	0.738	0.500	0.333	0.167	0	**H**	0.357	0.310	0.191	0.119	0.024	**H**	0.476	0.333	0.262	0.191	0
M_0	0.348	0.300	0.233	0.133	0	**M_0**	0.200	0.191	0.143	0.100	0.024	**M_0**	0.252	0.224	0.195	0.152	0
M_1	0.093	0.103	0.095	0.065	0	**M_1**	0.063	0.066	0.061	0.048	0.014	**M_1**	0.078	0.088	0.087	0.069	0
M_2	0.061	0.070	0.066	0.043	0	**M_2**	0.043	0.045	0.042	0.034	0.010	**M_2**	0.052	0.060	0.060	0.045	0

AF, Alkire and Foster multidimensional poverty methods; k, cutoff in number of dimensions for individuals to be poor; H, headcount ratio; M_0, adjusted headcount ratio; M_1, adjusted poverty gap; M_2, adjusted-Foster-Greer-Thorbecke measure.

Table 4 AF poverty measures applied to ICECAP-O levels using "22222" as the capability poverty threshold (n=42)

22222	BASELINE						1-YEAR POST-OP						3-YEAR POST-OP				
CUTOFF(k)	k=1	k=2	k=3	k=4	k=5		k=1	k=2	k=3	k=4	k=5		k=1	k=2	k=3	k=4	k=5
AF						**AF**						**AF**					
H	0.167	0.048	0.024	0	0	**H**	0.119	0	0	0	0	**H**	0.167	0.024	0.024	0	0
M_0	0.048	0.024	0.014	0	0	**M_0**	0.024	0	0	0	0	**M_0**	0.043	0.014	0.014	0	0
M_1	0.048	0.024	0.014	0	0	**M_1**	0.024	0	0	0	0	**M_1**	0.043	0.014	0.014	0	0
M_2	0.010	0.010	0.008	0	0	**M_2**	0.003	0	0	0	0	**M_2**	0.007	0.006	0.006	0	0

AF, Alkire and Foster multidimensional poverty methods; k, cutoff in number of dimensions for individuals to be poor; H, headcount ratio; M_0, adjusted headcount ratio; M_1, adjusted poverty gap; M_2, adjusted Foster-Greer-Thorbecke measure.

Table 5 Responses below Threshold of Sufficient Capability "33333"(baseline)

ICECAP-O responses (baseline)

Below TSC	ATTACHMENT	SECURITY	ROLE	ENJOYMENT	CONTROL	TOTAL RESPONSES
level 2	9	29	38	38	27	27%
level 1	2	12	7	11	6	7%
Below TSC per attribute	10%	39%	42%	46%	31%	

n = 106; SCS, Sufficient Capability Score; TSC, Threshold of Sufficient Capability; SCS = 0.857; TSC "33333"

Table 6 Responses below Threshold of Sufficient Capability "33333"
(baseline and 1 year post operation; n=58)

ICECAP-O responses (baseline)						
Below TSC	ATTACHMENT	SECURITY	ROLE	ENJOYMENT	CONTROL	TOTAL RESPONSES
level 2	3	15	21	22	14	26%
level 1	1	6	2	5	1	5%
Responses below TSC per attribute	7%	36%	40%	47%	26%	
ICECAP-O responses (1 year)						
Below TSC	ATTACHMENT	SECURITY	ROLE	ENJOYMENT	CONTROL	TOTAL RESPONSES
level 2	10	8	11	15	4	17%
level 1	0	6	0	0	1	2%
Responses below TSC per attribute	17%	24%	19%	26%	9%	

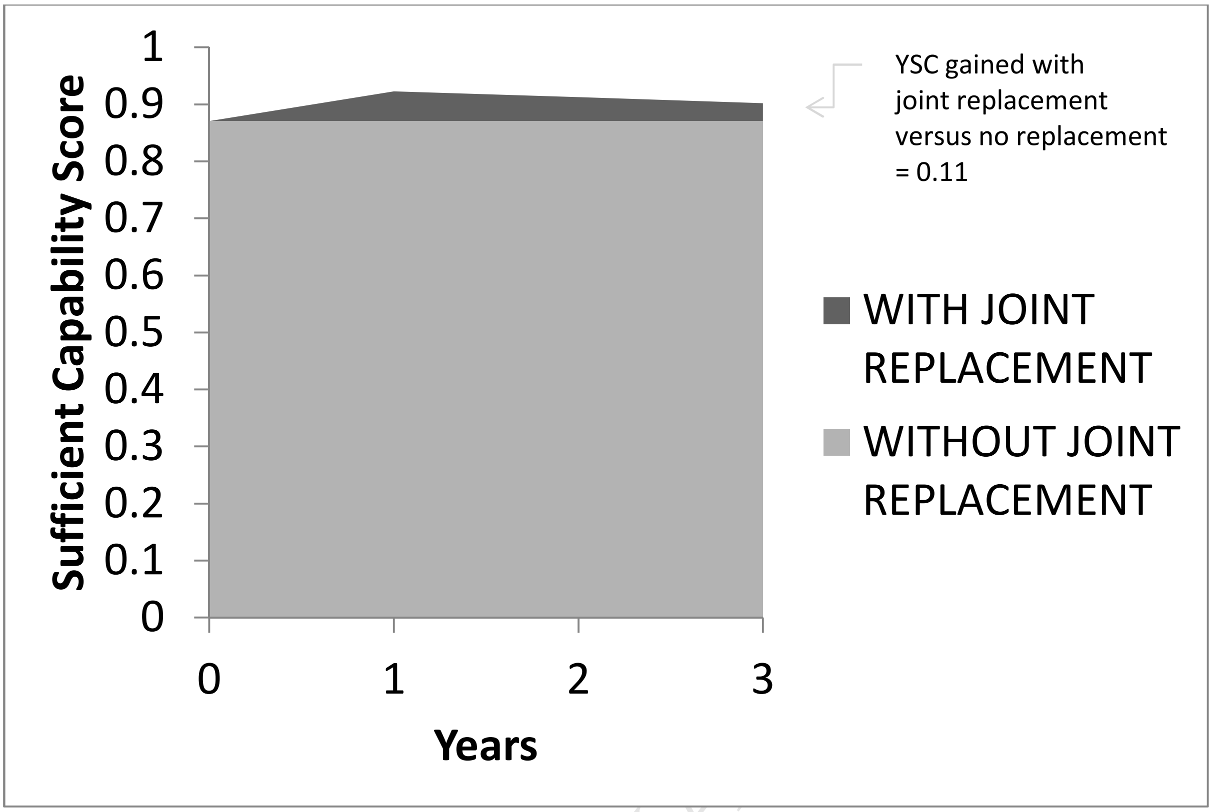
SCS, Sufficient Capability Score; TSC = Threshold of Sufficient Capability; SCS (baseline "33333") = 0.88; SCS (1 year "33333") = 0.93.

Table 7 Responses below Threshold of Sufficient Capability "33333"
(baseline and three year post-operation; n=55)

ICECAP-O responses (baseline)						
Below TSC	ATTACHMENT	SECURITY	ROLE	ENJOYMENT	CONTROL	Total Responses
level 2	3	17	22	22	13	28%
level 1	0	5	1	4	2	4%
Responses below TSC per attribute	5%	40%	42%	47%	27%	
ICECAP-O responses (3 year)						
Below TSC	ATTACHMENT	SECURITY	ROLE	ENJOYMENT	CONTROL	Total Responses
level 2	4	16	13	11	10	20%
level 1	0	5	3	1	1	4%
Responses below TSC per attribute	7%	38%	29%	22%	20%	

SCS, Sufficient Capability Score; TSC, Threshold of Sufficient Capability; SCS (baseline "33333") = 0.881; SCS (3 year "33333") = 0.912.

Figure 1 Example of Years of Sufficient Capability (YSC) for Osteoarthritis Patients (n=42)



References

Al-Janabi, H., Flynn, T., & Coast, J. (2012). Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. *Quality of Life Research,* 21, 167-176.

Alkire, S., & Foster, J. (2011a). Counting and multidimensional poverty measurement. *Journal of Public Economics,* 95, 476-487.

Alkire, S., & Foster, J. (2011b). Understandings and misunderstandings of multidimensional poverty measurement. *The Journal of Economic Inequality,* 9, 289-314.

Alkire, S., Qizilbash, M., & Comim, F. (2008). Introduction. In F. Comim, M. Qizilbash, & S. Alkire (Eds.), *The Capability Approach: Concepts, Measures and Applications* pp. 1-25). Cambridge, UK: Cambridge University Press.

Alkire, S., & Santos, M. (2013). A Multidimensional Approach: Poverty Measurement & Beyond. *Social Indicators Research,* 112, 239-257.

Anand, P., & Dolan, P. (2005). Equity, capabilities and health. *Social Science & Medicine,* 60, 219-222.

Anand, P., Hunter, G., Carter, I., Dowding, K., Guala, F., & Van Hees, M. (2009). The Development of Capability Indicators. *Journal of Human Development and Capabilities,* 10, 127.

Asaria, M., Griffin, S., Cookson, R., Whyte, S., & Tappenden, P. (2015). Distributional Cost-Effectiveness Analysis of Health Care Programmes – A Methodological Case Study of the UK Bowel Cancer Screening Programme. *Health Economics,* 24, 742-754.

Atkinson, A.B. (2003). Multidimensional Deprivation: Contrasting Social Welfare and Counting Approaches. *The Journal of Economic Inequality,* 1, 51-65.

Birch, S., & Donaldson, C. (2003). Valuing the benefits and costs of health care programmes: where's the 'extra' in extra-welfarism? *Social Science & Medicine,* 56, 1121-1133.

Bleichrodt, H., & Quiggin, J. (2013). Capabilities as menus: A non-welfarist basis for QALY evaluation. *Journal of Health Economics,* 32, 128-137.

Brazier, J., Ratcliffe, J., Salomon, J.A., & Tsuchiya, A. (2007). *Measuring and Valuing Health Benefits for Economic Evaluation*. Oxford, UK: Oxford University Press.

Brooks, R. (1996). EuroQol: the current state of play. *Health Policy,* 37, 53-72.

Brouwer, W.B.F., Culyer, A.J., van Exel, N.J.A., & Rutten, F.F.H. (2008). Welfarism vs. extra-welfarism. *Journal of Health Economics,* 27, 325-338.

Clark, D.A., & Qizilbash, M. (2008). Core Poverty, Vagueness and Adaptation: A New Methodology and Some Results for South Africa. *The Journal of Development Studies,* 44, 519-544.

Coast, J. (1999). The appropriate uses of qualitative methods in health economics. *Health Economics,* 8, 345-353.

Coast, J. (2009). Maximisation in extra-welfarism: A critique of the current position in health economics. *Social Science & Medicine,* 69, 786-792.

Coast, J., Flynn, T.N., Natarajan, L., Sproston, K., Lewis, J., Louviere, J.J., et al. (2008a). Valuing the ICECAP capability index for older people. *Social Science & Medicine,* 67, 874-882.

Coast, J., Kinghorn, P., & Mitchell, P. (2015). The Development of Capability Measures in Health Economics: Opportunities, Challenges and Progress. *The Patient - Patient-Centered Outcomes Research,* 8, 119-126.

Coast, J., Peters, T.J., Natarajan, L., Sproston, K., & Flynn, T.N. (2008b). An assessment of the construct validity of the descriptive system for the ICECAP capability measure for older people. *Quality of Life Research,* 17, 967-976.

Coast, J., Smith, R., & Lorgelly, P. (2008c). Should the capability approach be applied in Health Economics? *Health Economics,* 17, 667-670.

Coast, J., Smith, R.D., & Lorgelly, P. (2008d). Welfarism, extra-welfarism and capability: The spread of ideas in health economics. *Social Science & Medicine,* 67, 1190-1198.

Comans, T., Peel, N., Gray, L., & Scuffham, P. (2013). Quality of life of older frail persons receiving a post-discharge program. *Health and Quality of Life Outcomes,* 11, 1-7.

Cookson, R. (2005). QALYs and the capability approach. *Health Economics,* 14, 817-829.

Couzner, L., Crotty, M., Norman, R., & Ratcliffe, J. (2013a). A Comparison of the EQ-5D-3L and ICECAP-O in an Older Post-Acute Patient Population Relative to the General Population. *Applied Health Economics and Health Policy,* 11, 415-425.

Couzner, L., Ratcliffe, J., & Crotty, M. (2012). The relationship between quality of life, health and care transition: an empirical comparison in an older post-acute population. *Health and Quality of Life Outcomes,* 10, 1-9.

Couzner, L., Ratcliffe, J., Lester, L., Flynn, T., & Crotty, M. (2013b). Measuring and valuing quality of life for public health research: application of the ICECAP-O capability index in the Australian general population. *International Journal of Public Health,* 58, 367-376.

Culyer, A.J. (1989). The normative economics of health care finance and provision. *Oxford Review of Economic Policy,* 5, 34-58.

Davis, J., Bryan, S., McLeod, R., Rogers, J., Khan, K., & Liu-Ambrose, T. (2012). Exploration of the association between quality of life, assessed by the EQ-5D and ICECAP-O, and falls risk, cognitive function and daily function, in older adults with mobility impairments. *BMC Geriatrics,* 12, 65.

Davis, J., Liu-Ambrose, T., Richardson, C., & Bryan, S. (2013). A comparison of the ICECAP-O with EQ-5D in a falls prevention clinical setting: are they complements or substitutes? *Quality of Life Research,* 22, 969-977.

Decancq, K., & Lugo, M.A. (2013). Weights in Multidimensional Indices of Wellbeing: An Overview. *Econometric Reviews,* 32, 7-34.

Drummond, M., Brixner, D., Gold, M., Kind, P., McGuire, A., Nord, E., et al. (2009). Toward a Consensus on the QALY. *Value in Health,* 12, S31-S35.

Drummond, M., Sculpher, M.J., Torrance, G.W., O'Brien, B.J., & Stoddart, G.L. (2005). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford, UK: Oxford University Press.

Flynn, T.N., Chan, P., Coast, J., & Peters, T. (2011). Assessing quality of life among British older people using the ICEPOP CAPability (ICECAP-O) measure. *Applied Health Economics and Health Policy,* 9, 317-329.

Flynn, T.N., Huynh, E., Peters, T.J., Al-Janabi, H., Clemens, S., Moody, A., et al. (2015). Scoring the ICECAP-A Capability Instrument. Estimation of a UK General Population Tariff. *Health Economics,* 24, 258-269.

Flynn, T.N., Louviere, J.J., Peters, T.J., & Coast, J. (2007). Best–worst scaling: What it can do for health care research and how to do it. *Journal of Health Economics,* 26, 171-189.

Foster, J., Greer, J., & Thorbecke, E. (1984). A Class of Decomposable Poverty Measures. *Econometrica,* 52, 761-766.

Grewal, I., Lewis, J., Flynn, T., Brown, J., Bond, J., & Coast, J. (2006). Developing attributes for a generic quality of life measure for older people: Preferences or capabilities? *Social Science & Medicine,* 62, 1891-1901.

Henderson, C., Knapp, M., Fernández, J.-L., Hirani, S.P., Cartwright, M., Rixon, L., et al. (2013). Cost effectiveness of telehealth for patients with long term conditions (Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial. *BMJ,* 346.

Horwood, J., Sutton, E., & Coast, J. (2014). Evaluating the Face Validity of the ICECAP-O Capabilities Measure: A “Think Aloud” Study with Hip and Knee Arthroplasty Patients. *Applied Research in Quality of Life,* 9, 667-682.

Hurley, J. (1998). Welfarism, Extra-Welfarism and Evaluative Economic Analysis. In M.L. Barer, T.E. Getzen, & G.L. Stoddart (Eds.), *Health, health care and health economics: perspectives on distribution* pp. 373-395). Chicester, UK: John Wiley & Sons Ltd.

Kinghorn, P., Robinson, A., & Smith, R. (2015). Developing a Capability-Based Questionnaire for Assessing Well-Being in Patients with Chronic Pain. *Social Indicators Research,* 120, 897-916.

Lorgelly, P. (2015). Choice of Outcome Measure in an Economic Evaluation: A Potential Role for the Capability Approach. *PharmacoEconomics*.

Lorgelly, P., Lorimer, K., Fenwick, E.A., & Briggs, A.H. (2008). The Capability Approach: Developing an Instrument for Evaluating Public Health Interventions: Final Report. Glasgow, UK: Glasgow Centre for Population Health.

Makai, P., Brouwer, W.F., Koopmanschap, M., & Nieboer, A. (2012). Capabilities and quality of life in Dutch psycho-geriatric nursing homes: an exploratory study using a proxy version of the ICECAP-O. *Quality of Life Research,* 21, 801-812.

Makai, P., Koopmanschap, M., Brouwer, W.B., & Nieboer, A.A. (2013). A validation of the ICECAP-O in a population of post-hospitalized older people in the Netherlands. *Health and Quality of Life Outcomes,* 11, 57.

Makai, P., Looman, W., Adang, E., Melis, R., Stolk, E., & Fabbricotti, I. (2014). Cost-effectiveness of integrated care in frail elderly using the ICECAP-O and EQ-5D: does choice of instrument matter? *The European Journal of Health Economics*, 1-14.

Mitchell, P.M., Roberts, T.E., Barton, P.M., Pollard, B.S., & Coast, J. (2013). Predicting the ICECAP-O Capability Index from the WOMAC Osteoarthritis Index: Is Mapping onto Capability from Condition-Specific Health Status Questionnaires Feasible? *Medical Decision Making,* 33, 547-557.

Murray, C.J.L., & Lopez, A. (1996). *The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020*. Cambridge: Harvard University Press.

Netten, A., Burge, P., Malley, J., Potoglou, D., Towers, A.-M., Brazier, J., et al. (2012). Outcomes of social care for adults: developing a preference-weighted measure. *Health Technology Assessment,* 16, 1-166.

NICE. (2014). Developing NICE guidelines - the manual. London: National Institute for Health and Care Excellence (NICE).

Nord, E. (2013). Disability weights in the Global Burden of Disease 2010: Unclear meaning and overstatement of international agreement. *Health Policy,* 111, 99-104.

Olsen, J.A., & Donaldson, C. (1998). Helicopters, hearts and hips: Using willingness to pay to set priorities for public sector health care programmes. *Social Science & Medicine,* 46, 1-12.

Pollard, B., Dixon, D., Dieppe, P., & Johnston, M. (2009). Measuring the ICF components of impairment, activity limitation and participation restriction: an item analysis using classical test theory and item response theory. *Health and Quality of Life Outcomes,* 7, 1-20.

Ravallion, M., Chen, S., & Sangraula, P. (2009). Dollar a Day Revisited. *The World Bank Economic Review,* 23, 163-184.

Ruger, J.P. (2010). *Health And Social Justice*. Oxford, UK: Oxford University Press.

Sen, A. (2009). *The Idea of Justice*. London: Allen Lane.

Simon, J., Anand, P., Gray, A., Rugkåsa, J., Yeeles, K., & Burns, T. (2013). Operationalising the capability approach for outcome measurement in mental health research. *Social Science & Medicine,* 98, 187-196.

Stiglitz, J.E., Sen, A., & Fitoussi, J.P. (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress. Paris: Commission of the Government of France.

Sugden, R. (1993). Welfare, Resources, and Capabilities: A Review of Inequality Reexamined by Amartya Sen. *Journal of Economic Literature,* 31, 1947-1962.

UNDP. (2010). Human Development Report. New York: United Nations Development Programme.

van de Wetering, E.J., Stolk, E.A., van Exel, N.J.A., & Brouwer, W.B.F. (2013). Balancing equity and efficiency in the Dutch basic benefits package using the principle of proportional shortfall. *The European Journal of Health Economics,* 14, 107-115.

van Leeuwen, K.M., Bosmans, J.E., Jansen, A.P.D., Hoogendijk, E.O., van Tulder, M.W., van der Horst, H.E., et al. (2015). Comparing Measurement Properties of the EQ-5D-3L, ICECAP-O, and ASCOT in Frail Older Adults. *Value in Health,* 18, 35-43.

Venkatapuram, S. (2011). *Health Justice*. Cambridge, UK: Polity Press.

Highlights

- Aligning health economics and multidimensional poverty methods to measure wellbeing
- Sufficient capability offered as objective in line with capability theory
- Years of Sufficient Capability developed to assess capability wellbeing over time
- Illustration of the Years of Sufficient Capability method for decision making

APPENDIX A. FORMULAE FOR ALKIRE-FOSTER MEASURES

Box 1.A The four Alkire Foster measures[1]

1. Headcount ratio (H): p/P
   (where p = population who are poor; P = total population under consideration)
   H represents the proportion of people who are poor (p) in a population (P)

2. Adjusted headcount ratio (M_0) = H×(d_p/D)
   (where d_p = dimensions that individual is poor, D = total number of dimensions possible for poor population (p))
   M_0 shows on how many dimensions a person is classified as poor out of total poverty they could suffer across dimensions.

3. Adjusted poverty gap (M_1) = M_0×(l_p/L)
   (where l_p = levels below threshold where poor population (p) responded, L = total levels below threshold across dimensions for poor population)
   M_1 takes into consideration the breadth of deprivation amongst the poor (i.e. M_0), as well as the depth of poverty suffered (in terms of distance individual is from the poverty threshold).

4. Adjusted Foster- Greer-Thorbecke(FGT) (Foster et al., 1984) measure[2] (M_2) = M_0×(v_p/V)
   (where v_p = values attached to levels below threshold where poor population (p) responded, V = total value between lowest levels on all dimensions and threshold for p)
   M_2 allows greater weight to be attached to more severe response levels on dimensions. It enables the measurement of the frequency (H), breadth (M_0), and severity of deprivation suffered by the poor.

[1] This is a shortened explanation of Alkire-Foster measures. For more detail on measures, see Alkire & Foster (2011a).

[2] Please note, the M_2 applied in this study is a modification to the original Alkire-Foster measure. In the original, weight is put on the lower category score through mathematical weights (Alkire & Foster 2011a). In this study, we use ICECAP-O population weights (Coast et al., 2008a).

APPENDIX B. EXAMPLE OF CALCULATING ALKIRE-FOSTER MEASURES

BOX 1.B: EXAMPLE OF ALKIRE-FOSTER MEASURES OF MULTIDIMENSIONAL POVERTY

Take 3 individuals (X,Y & Z) assessed across 4 dimensions (D1,D2,D3,D4) which indicate poverty. All 4 dimensions are categorical with 5 (1-5, worst-best) responses possible for each dimension.

	Dimension 1	Dimension 2	Dimension 3	Dimension 4
Individual X	2	3	4	5
Individual Y	3	3	3	3
Individual Z	5	1	4	2

Let **poverty line** across 4 dimensions fall at **response level 3**.
Let **cutoff** for an individual to be classified as poor fall below poverty line on **any one dimension**.
Let **values** for levels below poverty line for **level 1 = 1; level 2 = 0.3.**

(1) **Headcount Ratio (H) = p/P**
p = individual X (poor on dimension 1) & individual Z (poor on dimension 2 and 4) = 2
P = total sample size = 3
H = 2/3 = 0.667

(2) **Adjusted Headcount Ratio (M_0) = H×(d_p/D)**
Individual X d_p = 1 out of 4 dimensions
Individual Z d_p = 2 out of 4 dimensions
d_p = 3 out of 8 dimensions
M_0 = 0.667×(3/8) = 0.250

(3) **Adjusted Poverty Gap (M_1) = M_0×(l_p/L)**
Individual X l_p = 1 out of 8 levels below poverty line
Individual Z l_p = 3 out of 8 levels below poverty line
l_p = 4 out of 16 levels below poverty line
M_1 = 0.250×(4/16) = 0.063

(4) **Adjusted FGT (M_2) = M_0×v_p/V**
Individual X v_p = 0.3 out of 4 for lowest value attached across all dimensions
Individual Z v_p = 1.3 (1+0.3) out of 4 for lowest value attached across all dimensions
pv = 1.6 out of 8 the values attached to lowest dimensions
M_2 = 0.250×(1.6/8) = 0.050

APPENDIX C. EXAMPLE OF CALCULATING SUFFICIENT CAPABILITY

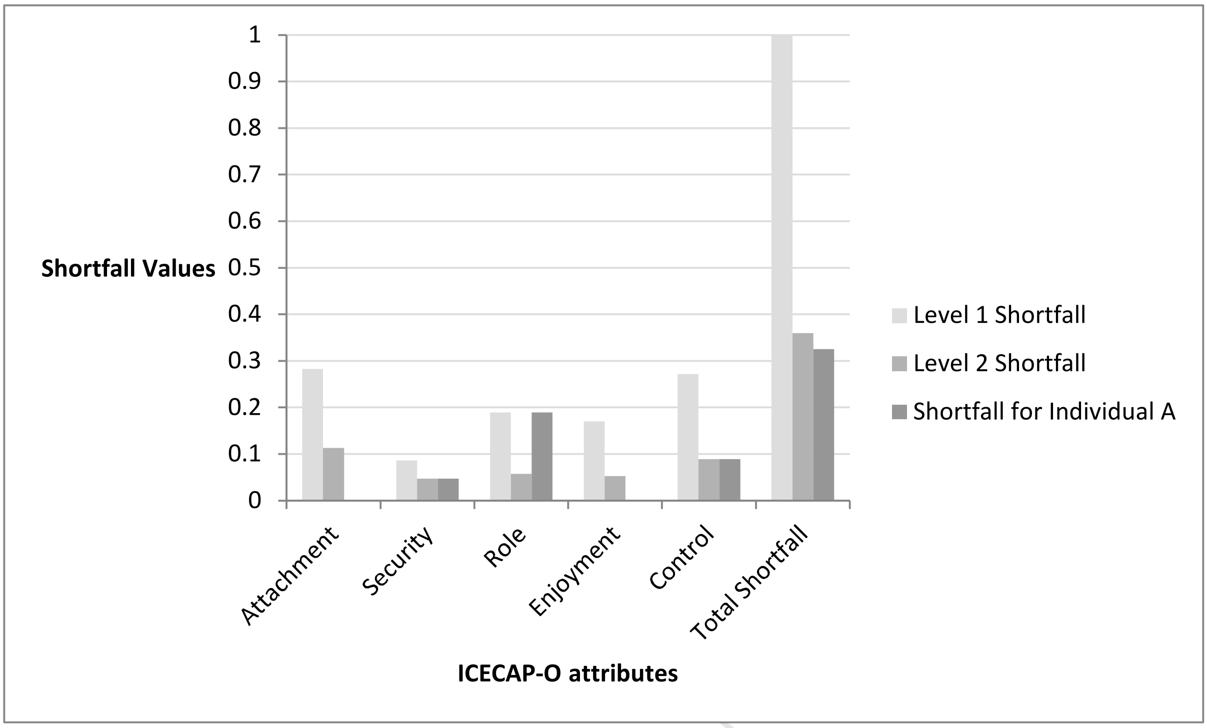
In Table 1.C and Figure 1.C, an individual example of how the sufficient capability score (SCS) is calculated for a given threshold (Option 1 – "33333") for an individual.

Table 1.C Shortfall in Sufficient Capability on ICECAP-O for individual A (threshold Option 1 "33333")

1. Love and Friendship			Tick one box only in each section
I can have all the love and friendship that I want	✓	4	
I can have a lot of the love and friendship that I want		3	
I can have a little of the love and friendship that I want		2	
I cannot have any of the love and friendship that I want		1	
2. Thinking about the future			
I can think about the future without any concern		4	
I can think about the future with only a little concern		3	
I can only think about the future with some concern	✓	2	
I can only think about the future with a lot of concern		1	
3. Doing things that make you feel valued			
I am able to do all of the things that make me feel valued		4	
I am able to do many of the things that make me feel valued		3	
I am able to do a few of the things that make me feel valued		2	
I am unable to do any of the things that make me feel valued	✓	1	
4. Enjoyment and pleasure			
I can have all of the enjoyment and pleasure that I want		4	
I can have a lot of the enjoyment and pleasure that I want	✓	3	
I can have a little of the enjoyment and pleasure that I want		2	
I cannot have any of the enjoyment and pleasure that I want		1	
5. Independence			
I am able to be completely independent		4	
I am able to be independent in many things		3	
I am able to be independent in a few things	✓	2	
I am unable to be at all independent		1	

Individual A ICECAP-O profile (42132); Highlighted Light Grey = sufficient capability for given attribute and given threshold. Highlighted Dark Grey = shortfall in sufficient capability for given attribute for given threshold.

Figure 1.C Shortfalls in Sufficient Capability for Threshold Option 1 ("33333")



Sufficient Capability Score (SCS) = Sufficient Capability (1) – Total Shortfall (0-1). Shortfall for individual A (42132) for threshold option 1 ("33333") = 0.325; SCS for individual A = 0.675. Capability instrument, ICECAP-O .