BRIEFING PAPER

DEPARTMENT OF HEALTH PROPOSALS FOR INCLUDING WIDER SOCIETAL BENEFITS INTO VALUE BASED PRICING: A DESCRIPTION AND CRITIQUE

NICE DECISION SUPPORT UNIT

August 2013

Alec Miners,[1] John Cairns[1] and Allan Wailoo[2]

[1] London School of Hygiene and Tropical Medicine (LSHTM)
[2] School for Health and Related Research (ScHARR), University of Sheffield

Decision Support Unit, ScHARR, University of Sheffield, Regent Court, 30 Regent Street
Sheffield, S1 4DA

Tel (+44) (0)114 222 0734
E-mail dsuadmin@sheffield.ac.uk
www.nicedsu.org.uk

ABOUT THE DECISION SUPPORT UNIT

The Decision Support Unit (DSU) is a collaboration between the Universities of Sheffield, York and Leicester. We also have members at the University of Bristol, London School of Hygiene and Tropical Medicine and Brunel University.

The DSU is commissioned by The National Institute for Health and Care Excellence (NICE) to provide a research and training resource to support the Institute's Technology Appraisal Programme. Please see our website for further information www.nicedsu.org.uk

The production of this document was funded by NICE. The views, and any errors or omissions, expressed in this document are the authors only. NICE may take into account part or all of this document if it considers it appropriate, but it is not bound to do so.

EXECUTIVE SUMMARY

The aim of this document is to provide a critique of the Department of Health's (DoH) proposal for including Wider Social Benefits (WSBs) into economic evaluations submitted to NICE's Technology Appraisals Programme - the concept that treatments not only impact an individual's health, but also on wider considerations such as their ability to return to work / contribute to society. More specifically, and subject to the Terms of Reference (ToR, see Section 4) from Ministers, the Decision Support Unit (DSU) has been requested to:

- Review the proposal and the WSBs dossier produced by the DoH
- Summarise key aspects of the research / evidence submitted (methods and findings) and the proposed approach for incorporating WSBs
- Evaluate the proposed approach / methods
- Identify any issues associated with the proposed approach, and present alternatives where appropriate
- Identify any gaps in the evidence
- Consider the impact on relevant concepts included in the Guide to Methods of Technology Appraisal 2013 (including 'equalities') [1]
- Consider opportunities for use of evidence currently submitted to the Technology Appraisal Programme, or otherwise the consequences for the submission of evidence

The DoH has defined WSBs as the difference between the amount of resources a patient contributes to society (production) and the amount they utilise (consumption). That is, the net economic contribution of an individual (to society) – the term 'net' emphasising not only the change in WSBs attributable to the treatment at hand, but also the value of notionally displaced WSBs as a results of a positive guidance. If a treatment generates more WSBs than it offsets (i.e. it produces a net gain in WSBs), then all else being equal its cost-effectiveness increases. The suggestion from the DoH is that this could then be reflected in deliberations and calculations by increasing the 'standard' threshold cost-effectiveness level. On the other hand, treatments that produce fewer WSBs than they offset (a net loss in WSBs) should be compared against a lower than standard cost-effectiveness threshold.

The DoH defines production as the addition of paid labour (employment) and unpaid labour (e.g. childcare). Consumption is defined as the combination of informal care, formal care, personal paid consumption (e.g. food), personal unpaid consumption (e.g. cleaning), informal child care and government services unrelated to health. Each of these categories of production and consumption are referred to as an 'element', many of which are adjusted for age, gender, International Classification of Diseases (ICD) code and quality of life (Qol). For example, the calculations suggest that men aged 50 years are associated with higher levels of paid production compared with elderly females. Each

one of these elements is based on an often complex series of analyses; this document contains a review of the construction of each element although they have not been reviewed in absolute detail due to their number.

The suggestion from the DoH is that adjustments for WSBs should be undertaken after the production of the usual cost-effectiveness estimates. There are a number of reasons for advocating this approach but a major reason is that WSBs are often non-linear with respect to patient characteristics such as age. This means that including WSBs in routine economic evaluations becomes a non-trivial task; the Terms of Reference require the approach to be relatively straight forward. WSBs are incorporated in routine cost per additional QALY calculations using an 'exchange rate'. For example, a net increase in WSBs of £17,000 using an exchange rate of £60,000 per additional QALY, is equivalent to (£17,000/£60,000) = 0.28 'adjusted QALYs', which can be added to the overall number of QALYs gained from the standard economic evaluation. The size of this exchange rate is likely to be crucial in terms of including WSBs in routine economic assessments but there is not an obvious value to choose.

Because of the difficulties of directly including WSBs calculations in the basic programming of routine economic evaluations, the DoH provides a 'Template' spreadsheet which can be used to incorporate WSBs after production of the standard cost-effectiveness estimate. The value of the notional displaced QALY is also calculated in this Template using the Reference Dataset, which is described in the previous burden of illness working document [2].

The DSU accepts that the broad approach taken by the DoH to include WSBs is appropriate i.e. the difference between the average production and average consumption of a particular patient group. However, we highlight a number of important issues that we believe warrant particular discussion, including the following:

- Each part of the WSB carries an equal weight but it may not be that changes in the different components are valued equally by the public.
- The introduction of WSB is broadening the perspective of the assessment of cost-effectiveness beyond the health of the patient and the costs to the NHS so as to include consequences for the rest of society. However, that this is not the same as taking a societal perspective.
- There are a number of important issues about the relationship between the implementation of WSB and the current, and ongoing, standard economic evaluation methodology.
- Many of the statistical analyses and assumptions that have gone into the construction of each element have not been reviewed or been subject to external consultation / debate.

- Given the number of parameters used to calculate the WSBs, the level of uncertainty around the estimate could be significant. However, the degree of uncertainty around the parameters is not presented in the Template.
- The non-linearities in the WSBs sometimes mean that explainable, but seemingly illogical, results appear. For example, increases in quality of life sometimes lead to reductions in WSBs when all else is held equal.
- There is concern that the ICD categories might not be sufficiently refined to capture important patient characteristics defined in NICE's scope documents (eg. stage of disease or the number of previously failed treatments). Consequently, the calculated WSBs might misrepresent the 'true' WSBs associated with a specific patient population.
- All of the elements are adjusted for at least one patient-level characteristic such as gender. The implicit outcome of this approach is that these characteristics define 'value' all else held equal. While this might be justifiable in pure economic terms, it is unclear whether this is acceptable from a public decision-making perspective.
- There does not appear to be an obvious method of defining the exchange rate to transform WSBs into QALY equivalent values.

A number of important areas for further research are highlighted at the end of the document

CONTENTS

TABLES & FIGURES

1. BACKGROUND

All health care interventions referred to NICE's Technology Appraisal Programme are subject to an economic evaluation, in order to help the Appraisal Committees determine whether they are cost-effective and should be funded by the NHS. There are a number of methodological considerations that must be considered within any economic evaluation. For example, the most appropriate source of estimating the relative treatment effect and the time horizon over which a technology should be evaluated.

One of the difficult tasks that confronts NICE is to be consistent across technology appraisals in terms of decision-making. One of the approaches it has developed to help with this is a 'reference case', in which the Institute specifies a set of methods that it believes is the most appropriate given its objectives.

Two of the defining features of NICE's current reference case is that outcomes should reflect *health* gains, whether they are for patients or carers, and costs should only be included if they relate to the National Health Service (NHS) or Personal Social Services (PSS) [1]. However, when there is a clear rationale to do so, non-reference case analyses are permitted, for example, where there are clear costs and benefits to other (non-health) government departments. The current technology appraisals methods guide [1] states that these issues should be identified during the scoping stage of an appraisal and that they should be presented separately from the reference-case analysis. Thus, Appraisal Committees can currently take wider issues into account when it is considered appropriate to do so but only in specific, and typically uncommon, circumstances.

The current process of determining cost-effectiveness assesses whether the benefits to patients of a technology, measured in Quality-Adjusted Life Years (QALY) gains, are sufficient to offset the losses to patients elsewhere in the NHS when funds are re-allocated to the treatment under evaluation. If they are, then a treatment can be considered cost-effective (as it generates positive net health benefits). If not, then the introduction or continued use of an intervention is much less likely to be recommended for use on economic grounds (as the net health benefits are negative). In principle, the decision regarding cost-effectiveness can be made by comparing an estimated incremental cost-effectiveness ratio (ICER) for a given intervention, against a threshold value of a QALY, as the latter represents the value of the notional displaced QALY in the NHS, or opportunity cost of a decision and the amount of health that is offset. In other words, the cost-effectiveness threshold defines the tipping point at which positive / negative net health benefits are defined. In theory, this approach ensures that new treatments do not displace more health gain than they provide, and will lead to decisions which do not diminish the overall value of benefits gained from the NHS budget.

Broadly speaking cost-effectiveness is taken into account by Appraisal Committees by assessing the robustness of the submitted ICERs comparing them to the Institute's stated cost-effectiveness threshold range [1] and by taking into account other issues that are considered important but are not necessarily captured in the cost-effectiveness assessments. The latter can be thought of as more of a 'deliberative process' given remaining issues of importance.

2. THE DEPARTMENT OF HEALTH PROPOSALS AND BROADER FRAMEWORK

This Working Party has been assembled to debate the proposal from the Department of Health (DoH) that changes should be made to NICE's current technology appraisals methods. More specifically, the DoH has put forward the proposals that economic evaluations submitted to NICE should reflect:

- **'Burden of illness'** (BoI) – the belief that society places higher values on QALYs gained by individuals with relatively high 'burden of illness'.
- **'Wider societal benefits'** (WSB) – the concept that treatments not only impact on an individual's health, but also on wider considerations such as their ability to return to work / contribute to society.

3. DECISION SUPPORT UNIT (DSU) REMIT

BoI and the broader VBP framework were discussed in the previous briefing paper [2] and associated Working Party meeting. Therefore, the main purpose of this document is to outline and critique the proposal for WSBs. Note that we do not debate whether it is appropriate for WSBs to be included in the first instance. This is taken as set given the Terms of Reference (ToR) from Ministers. Rather the discussion and critique relate to the strengths and weaknesses of the proposed approach, and the subsequent 'weighting' that should be attached to each element given the various levels of uncertainty. More specifically, for WSBs the DSU has been tasked to:

- Review the proposal and the WSBs dossier produced by the DoH
- Summarise key aspects of the research / evidence submitted (methods and findings) and the proposed approach for incorporating WSBs
- Evaluate the proposed approach / methods
- Identify any issues associated with the proposed approach, and present alternatives where appropriate
- Identify any gaps in the evidence
- Consider the impact on relevant concepts included in the Guide to Methods of Technology Appraisal 2013 (including 'equalities')
- Consider opportunities for use of evidence currently submitted to the Technology Appraisal Programme, or otherwise the consequences for the submission of evidence

4. TERMS OF REFERENCE TO NICE FROM MINISTERS

The methods for value assessment of branded medicines under VBP should:

1. Be applied to medicines within the scope of the VBP system, and incorporated into the methods for other categories of guidance at NICE's discretion
2. Adopt the same benefit perspective for all technologies falling within the scope of VBP, and for displaced[1] treatments
3. Be as transparent and predictable as possible
4. Be informed by the best available evidence
5. Include a simple system of weighting for burden of illness that appropriately reflects the differential value of treatments for the most serious conditions
6. Encompass the differential valuation of 'End of Life' treatments in the current approach within the system of Burden of Illness weights
7. Include a proportionate system for taking account of Wider Societal Benefits
8. Not include a further weighting for Therapeutic Innovation and Improvement
9. Produce guidance for patients and the NHS which describes the clinical and cost effectiveness of the technology and its position in clinical practice

Also note the following text that accompanies the ToR:

'The perspective adopted for measuring WSBs should, in principle, be as set out in the HMT Green Book for Appraisal and Evaluation in Central Government - which specifies the cross-Government approach for evaluating costs and benefits of spending decisions. However in practice it will be important to reflect uncertainties in the evidence for the magnitude of WSBs, the novelty of the approach, and the degree of consensus among stakeholders. Options may in practice include constraining the weight given to different elements of WSBs in the valuation of treatments, or initially taking a selective approach to the types of benefit included in the assessment framework, in order to support incremental broadening of the value perspective. It will be important to ensure that the approach to incorporating WSB is applied systematically and consistently'

[1] That is, the value of a new treatment is considered net of the value of what is displaced

5. THE DEPARTMENT OF HEALTH'S PROPOSAL FOR WSBS

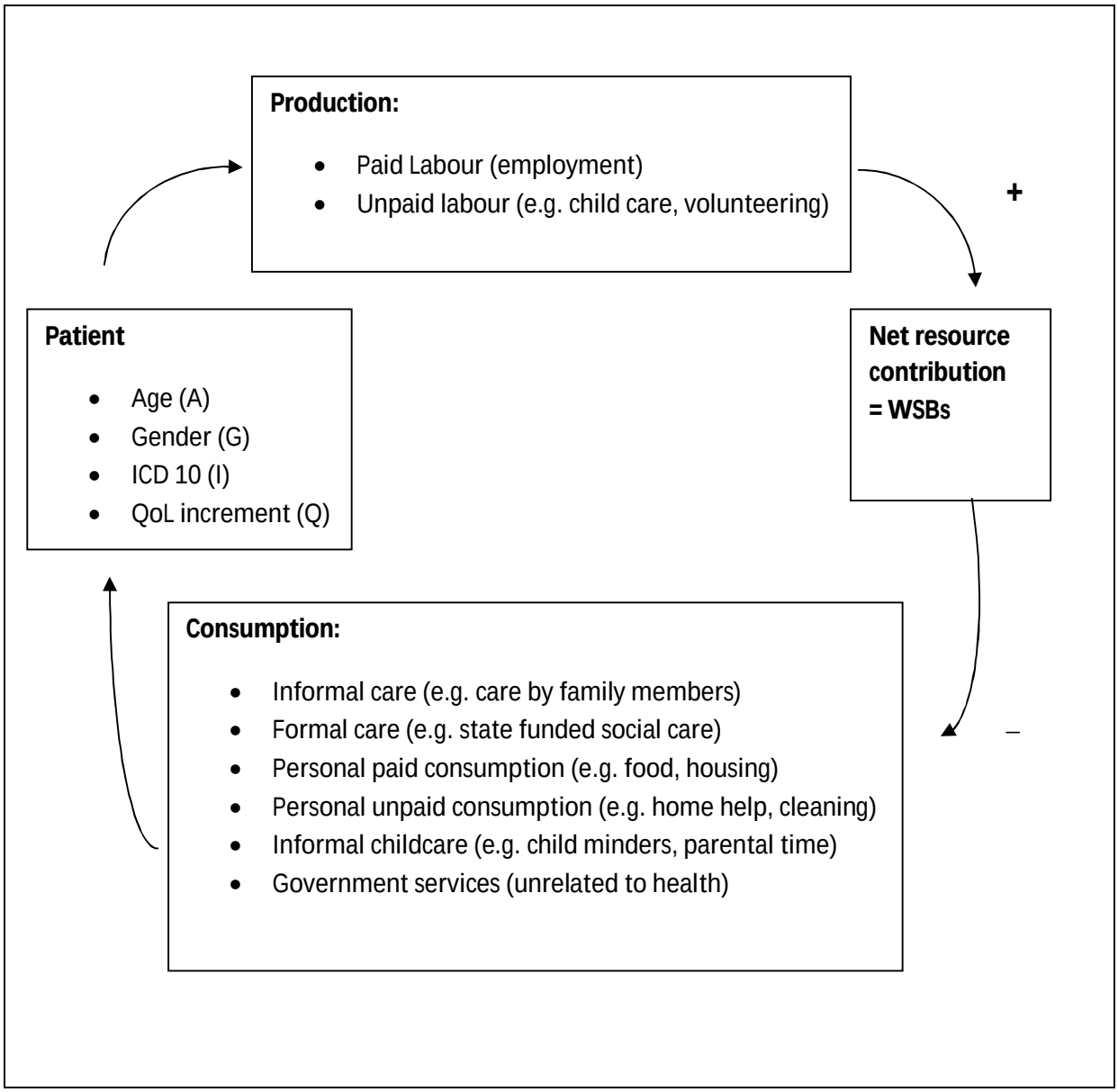
5.1. Defining WSBs

The DoH's proposal states that treatments not only impact on the immediate patient, but also on a patient's family and the rest of society [3]. While the definition / labelling of WSBs has been criticised,[2] [4] the DoH defines it as the difference between the amount of resources a patient contributes to society (production) and the amount they utilise (consumption). That is, the net economic contribution of an individual (to society) (see

[2] Some commentators have stated a preference for the term 'net resource impact' as it more accurately reflects the DoH's proposal. To some 'WSB' implies extended health benefits rather than inclusion of non-health related costs and benefits.

Figure 1). More importantly, the emphasis is on how this balance of consumption and production changes as a consequence of treatment – i.e. the net change in WSBs. The principle behind this definition is that any resource generated by a patient that is not consumed (positive WSBs) is available for someone else to use and benefit from – even if that individual is not actually 'known'. On the other hand, resources that a patient uses cannot be available for use by another member of society and this therefore represents a 'cost'. To illustrate this point, the DoH gives the example of a treatment that enables a patient to return to work and reduces their need for social care. The resources they contribute through their labour in this instance increases while their use of social care resources decreases. All else being equal, this would result in a net increase in WSBs following treatment meaning there would be more resources available for use by someone else. The net impact of this change would be to increase the treatment's cost-effectiveness.

Figure 1: Schematic of the DoH's proposal for including WSBs



The WSB framework proposed by the DoH estimates 'net' WSBs as a function of age, sex, ICD code and quality of life (QoL). The term 'net' refers to the difference in WSBs between the treatment and comparator *and* the potentially displaced treatments in the NHS if guidance is positive. For example a hypothetical treatment might generate an additional £1,000 of WSBs compared with a comparator treatment, but if £2,000 of WSBs are simultaneously displaced, the overall net change in WSBs is -£1,000.

The DoH's approach to including WSBs is (necessarily) detailed. Therefore, to 'simplify' the process they have put forward a 'Template spreadsheet' that can be used ex ante for each appraisal given information on ICD10 code, the QoL increment as a result of treatment (derived from a standard

economic evaluation) and the overall number of QALYs gained (also derived from the standard economic evaluation). The DoH suggests that the Template could provide 'Base Case' estimates of WSBs that can be made more appraisal-specific if required. In theory, the existence of this Template means that manufacturers are not required to collect any new data, although it might be desirable for them to do so.

There are two other important reasons for the existence of the Template. First, as with BoI, it is important that the number of potential WSBs displaced in the NHS as a results of positive guidance is estimated. This can only be achieved on a practical basis using the 'Reference Dataset' (which is in the Template). It is referred to in more detail in the BoI briefing paper but it is based on the UK element of the WHO Global Burden of Disease [5] project. Second, it is the DSU's understanding that a previous version of the WSB 'Base Case' spreadsheet allowed stakeholders to directly input information regarding starting ICD (chapter) code, gender, sex and average age in order to calculate overall WSBs. However, a significant problem with this suggestion was identified. More specifically, WSBs are *non-linear* with respect to many patient-level characteristics and particularly with age. This requires that a mean WSB is calculated over the distribution of a patient population with respect to demographics such as age, not an estimate of WSBs for the average patient.

To illustrate the importance of this non-linearity issue, consider the following simplified example. A population, all with full health, where 50% of individuals are 20 years of age and the remaining 50% are 60 years of age. Thus, the populations mean age is 40 years, and assume at this level WSBs are £1,700 per month. However, also assume that the average WSBs for a 20 and 60 year old are -£1,200 and £800 per month respectively. The more technically accurate mean WSB for this population is therefore -£200 ([-£1,200+£800]/2) not the aforementioned £1,700.

The existence of these non-linearity issues is a *significant* barrier to calculating WSBs in 'routine' cost-effectiveness models submitted to NICE. An approach is needed that reflects both the non-linear relationship between age and WSBs and the fact that age and QoL will change over the time where QALYs accrue. The DSU has previously explored the extent to which results vary according to different simplifications and approximations that could be made either using the summary information that results from a cost-effectiveness model (i.e. mean costs and QALYs for each treatment and comparator), or by incorporating WSB estimates within the cohort models that are typically submitted [6]. Whilst this work was by no means exhaustive, our tentative conclusions were that there would be a substantial risk of bias from any of the approaches we tested. Our belief is that the only feasible alternative to using the Template approach, given the existence of these non-linearities is to use patient-level (microsimulation) models: anything less than this could lead to a large degree of 'inaccuracy' in the projected WSBs. However, models based on these approaches are relatively

uncommon in submissions to NICE's Technology Appraisal Programme and are often associated with substantial disadvantages. For example, they can be complex in terms of the level of programming required, and it is not always feasible to include a probabilistic sensitivity analysis using NICE recommended software such as Microsoft Excel, which is NICE's preferred approach for assessing parameter uncertainty.

5.2. The Template

The Template requires the 'user' to enter the following for *each appraisal:*

1. ICD10 code
2. Aggregate change in QoL (over the relevant time horizon) as a result of treatment [3]
3. The mean number of QALYs gained also as a result of treatment [4]

Other *non-appraisal* specific information is also required:

4. An 'exchange rate' between WSBs and QALYs (currently set at £60,000 per QALY by the DoH)
5. A value for the 'standard' NICE cost-effectiveness threshold (currently set at £25,000 per QALY for convenience as it is the midpoint in NICEs 'usual' £20,000 to £30,000 per additional QALY threshold range
6. Information regarding the notional number of displaced QALYs in the NHS
7. BoI weights (discussed at the previous Working Party meeting and not discussed further in this document)

The exchange rate is a mechanism for converting WSBs into a 'QALY-worth equivalent' so the two benefits of treatments, WSBs and QALYs – the latter based on standard approaches to valuing health gains - can be added together. For example, imagine a treatment is associated with a net increase in WSBs of £17,000. Using an exchange rate of £60,000 per additional QALY, this is equivalent to (£17,000/£60,000) = 0.28 'adjusted QALYs', which can be added to the overall number of QALYs gained from the standard economic evaluation. So if the latter showed an increase in QALYs of 1.0 due to improvements in health, then the overall increase in QALYs would be 1.28 (1.0+0.28). Note that the DoH provides a justification for using a £60,000 per QALY exchange rate, and states that it is a value used across all government departments. More specifically, the value has been derived using

[3] Aggregate QoL is defined as the sum of the aggregate increases in QoL as a result of treatment over the relevant time horizon. For example, in the absence of treatment an individual might experience a QoL of 0.8 over 2 years. If, as a result of treatment, QoL increased to 0.9 and 0.95 in years 1 and 2 respectively, then the aggregate QoL would be 0.25 ([0.9-0.8] + [0.95-0.8]).

[4] Since the total WSBs are a function of QALY gains, and therefore proportional to them, information regarding the QALY gained per patient is not actually required to calculate the WSBs. It is included in the Template as a method of calculating the potential value based price of a drug.

the ‘value of a prevented fatality’ employed by the Department of Transport and other government departments. The value is approximately £1.6 million per prevented fatality, divided by a discounted number of potential QALY gains from a prevented accident, which equates to approximately £60,000 per additional QALY.

Although this QALY value is twice as large as NICE’s usual £20,000 to £30,000 per additional QALY range, it is not immediately obvious that the two should be the same. The NICE threshold can be interpreted as reflecting the marginal value of a QALY given the NHS budget, whereas the WSB is an estimate of the net contribution to the economy which is or can be devoted to a very wide range of activities. One interpretation might be that since the government does not seek to increase the NHS budget they believe that the marginal value of these resources is higher elsewhere. Another might be that the DoH believes the marginal value of a QALY (taking into account BoI and WSB) is closer to £60,000.

In any case, one simple approach to ‘down-weighting’ the contribution of WSBs in terms of their economic importance, if it is seen as appropriate to do so, is to have a high exchange rate. Conversely a lower exchange rate would imply that WSBs are valued more highly. For example, an exchange rate of £30,000 per additional QALY would increase the previous estimate of 0.28 to 0.57 (£17,000/£30,000). Also note that WSBs can be negative meaning that the overall affect of including WSBs can be to lower the total number of adjusted QALYs.

5.3. The Reference Dataset

The Template is programmed so that when an ICD 10 code is inputted, an array of information based on the Reference Dataset is called. The data is divided into 1,281 disease areas (ICD10 codes) and further subdivided into broad age (eight categories) and sex categories meaning that a total of 20,496 cells or ‘bins’ are defined. For each of these bins, the Dataset includes information on:

1. The average QoL
2. The overall burden of illness measured in QALYs as an aggregate figure and broken down into i) loss due to lower QoL, ii) loss due to length of life (LoL) and iii) the average LoL gain from treatments in this disease category.

The information in these bins is combined in the spreadsheet with a substantial number of other parameters, assumptions and formulae / equations to calculate a ‘rate’ per unit of time for each WSB element (see

Figure 1). The total amount of WSBs in monetary terms is a function of this rate and the length of time the rate is applicable for, summed across all elements. For example, consider only the element 'Paid Production', and an individual who *gains* 0.2 QALYs in terms of length of life (LoL) and that the QoL during this time was 0.5. This implies that the individual must have *gained* 0.4 years in LoL (0.2 / 0.5).[5] If the rate of paid production, given age, sex and the QoL increment, is say £20,000 per annum, this would imply total *gains* in WSBs due to paid production of £8,000. The exchange rate is then applied to turn the value into a QALY equivalent (£8,000/£60,000) to produce a *gain* of 0.13 QALYs. Thus, the Template can be seen as a calculator used at some point between (or as part of the) routine cost-effectiveness submissions put before (or as part of) the initial Appraisal Committee meeting.

5.4. Valuing displaced treatments

The proposed overall method of incorporating WSBs (and BoI) into a cost-effectiveness analysis is to adjust the cost-effectiveness threshold. As previously explained in the Briefing Paper for BoI, this approach is logical since it reflects the value of a notional displaced QALY in the NHS. For example, assuming NICEs current cost-effectiveness threshold is £20,000 per additional QALY, and a treatment produces more WSBs than it offsets, then all else being equal the 'hurdle' of £20,000 per QALY will be raised – making it more likely a treatment would be acceptable on economic grounds. Conversely, a treatment that offsets fewer WSBs than it produces will ultimately be compared against a cost-effectiveness threshold lower than this value, making a positive recommendation less likely.

$$Adjusted\ ICER\ threshold = Current\ threshold\ ICER \times \left(\frac{1 + BoI\ weight + WSBs\ per\ QALY}{1 + dBoI\ weight + dWSBs\ per\ QALY}\right)$$

Where BoI weight (expressed as a %) is the additional premium given to a QALY as the result of valuing 'severity' and WSBs per QALY are the total number of QALY equivalent WSBs for a given treatment divided by the total number of QALYs gained. The '*d*' represents the equivalents for displaced treatments. Both terms for the displaced treatments are calculated as a weighted average across all ICD 10 codes in the Reference Dataset and are constant across appraisals since the average of what is potentially displaced is considered to be unaffected by the specific appraisal topic. An example of the adjustment to the ICER threshold is given in

[5] While information for LoL, expressed in terms of the % QALYs attributable to LoL, is taken from the Reference Dataset for each ICD 10, age and gender. In theory this is something that could be made appraisal specific but the practicalities of doing so are unclear.

Table 1 below.

Table 1: Example of ICER threshold adjustment given all necessary information

		Treatment	Displaced Treatment
Standard approach	Incremental cost	50,000	50,000
	Incremental QALYs	2	2
	ICER	£25,000	£25,000*
With VBP adjustments	BoI weight	+30%	+20%
	Weighted QALYs	2.6	2.4
	WSBs	£12,000	£30,000
	QALY worth**	0.2***	0.5
	Total gain / loss in QALYs	2.8	2.9
Adjusted ICER threshold	**£24,138^**		

*centre of range given no adjustments; **using an exchange rate of £60,000 per additional QALY. *** 0.1 as this is per QALY (0.2/2); ^the adjusted threshold ICER is calculated as £25,000 x [(1+0.3+0.1) / (1+0.2+0.25)] using the formula on the previous page

In this example, in the absence of any VBP considerations, the treatment is precisely on the threshold ICER if a value of £25,000 per additional QALY is used, and it can be considered cost-effective. However, this changes when both VBP elements are applied. Note that the aggregate number of QALYs, or more precisely, QALY equivalents, becomes higher for the displaced treatments (2.9 QALYs) than the treatment under evaluation (2.8 QALYs). For this reason, the adjusted ICER threshold has reduced from £25,000, as more QALYs have been offset than gained. It follows that the treatment is by definition no longer cost-effective. This is simultaneously illustrated by the fact that the ICER (£25,000) is higher than the adjusted threshold ICER value (£24,138). However, *if* the total gain in QALYs from treatment had been higher than 2.9, say 3.0, then the adjusted ICER would have increased as more QALY equivalents were being offset than gained.

6. CALCULATING THE MONETARY VALUE OF WSBS

The preceding text outlined the broad framework for taking into account WSBs, including how the monetary valuation is 'adjusted' into a QALY equivalent through an exchange rate. This section of the report provides more detail as to how the monetary values of the various elements (see Figure 1) are estimated in the first instance. It is extremely important to note that the DoH has undertaken extensive work and undertaken numerous calculations to arrive at these 'rates' for each element. The following text, therefore, represents a brief précis of the entire approach, highlighting a number of the key issues and uncertainties.

To set the scene, Table 2 summarises which patient characteristics are taken into account when calculating the individual components of production and consumption. Where A refers to eight age groups (0-4, 5-14, 15-29, 30-44, 45-59, 60-69, 70-79, 80+), G to gender, I to ICD10 classification and Q to QoL.

Table 2: Summary of modelling of WSB

Component of production or consumption	Patient characteristic	Note
Paid labour	AGQ	AG determines hourly rate, AQ determines proportion working
Unpaid labour	AGIQ	AG determines hours & hourly rate, AGIQ determines proportion working
Informal care	AGIQ	Q determines hours per day, AGIQ determines days of informal care
Formal care	AQ	AQ determines probability of using residential care & probability of using non-residential care
Personal paid consumption	AQ	Tends to decrease with age & increase with Q
Personal unpaid consumption	A	Only dependent on age
Informal childcare	AG	Hours of paid & unpaid childcare & hourly cost of paid childcare depend on age, hourly cost of unpaid childcare depends on AG
Government services	A	Consumption of government services determined solely by age

6.1. Production

Production is divided into two elements, paid and unpaid labour

6.1.1. Paid production

Paid production is defined as the value of labour provided by the patient. The rationale for including this is that if a treatment increases an individual's ability to work, society values this 'benefit' and it should be reflected in an evaluation. More specifically, the DoH calculates it as the product of an age-quality of life determined probability of being in paid employment and average gross wage (which varies by age and gender):

$$Paid\ Labour_{AGQ} = Productivity_{AQ} \times Wage\_pcm_{AG} \times oncosts$$

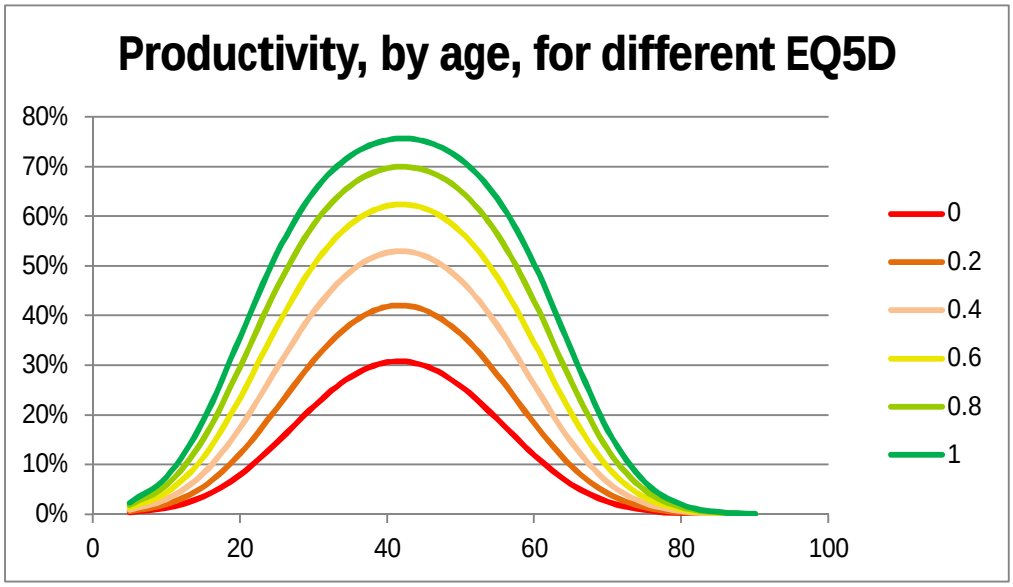
On-costs (employer overheads) are presumably included in order to more closely reflect the value of the production of the worker. Age features on all the relevant terms (defined as eight age bandings), but ICD code is not included and QoL only impacts on the rate of productivity. Gross wages, by age and gender, were taken from the Annual Survey of Hours and Earnings (ASHE, 2011) [7] whereas on-costs are provided by Eurostat 2012 data at 16.4% [8]. Productivity (or rather 'level of

productivity') is a measure of the proportion of working time spent actually working and is designed to encompass all possible reasons for not working, including unemployment, retirement, not being of working age, as well as ill-health. Figure 2 shows how the level of productivity varies by age and QoL. In essence it shows that productivity 'peaks' around the age of 50 and is unsurprisingly higher for people who are relatively well in terms of QoL.

The applied hourly wage-rate for men is higher than women, and assumed to plateau around the age of 60 years for both sexes; note it is not assumed to change with ill-health. Applying a differential rate for men and women implies that women are valued less than men holding everything else constant. While this might be justifiable in terms of reflecting current labour market statistics, there is a question as to whether using differential values is merely perpetuating (unacceptable) wage inequalities. The same holds true elsewhere in the current WSB proposal wherever a gender-adjusted wage-rate is used.

The sum of these assumptions is that currently people who are female, relatively ill, young, or old, are valued less in terms of paid production compared with relatively well middle aged men.

Figure 2: The level of productivity given age and QoL (productivity as a percentage is on the y-axis and age on the x-axis) [3]



6.1.2. Unpaid production

It is important to reflect unpaid labour in the calculations, since this helps to ensure more of an individual's contribution to society is reflected in the WSB estimates. The estimate of the value of unpaid labour is broken down into three components – 1) the production of childcare, 2) informal care

(unpaid care for others who are unwell) and 3) all other unpaid production. For each of these three elements the following is estimated:

$$Hours\ at\ full\ health_{AG} \times Net\ hourly\ wage \times (1 - Sick\ rate_{AQ})$$

Net hourly wage is used rather than gross hourly wage as it better reflects the opportunity cost of time to the individual. The estimate of unpaid childcare uses the Time Use Survey (2000) [9] to identify which patients are likely to have children and provide care, given their age and gender. Both 'active' (actual care) and 'passive' child and informal care ('just being there') are included, and it is assumed that the provider of childcare spends the full potential working day providing care (net of any formal care received by the child, either at school or in other paid childcare).

For unpaid care (for the sick or elderly), the average hours of informal care provided by patients at full health is estimated, given their age and gender, using the Time Use Survey. For carers who live with the person they provide care for, it is assumed that additional time is spent providing 'passive' care, with estimates informed by the Survey of Carers in Households (2009) [10]. For all other unpaid production (e.g. domestic work), the Time Use Survey is also used. The value (wage-rate equivalent) of unpaid labour was estimated using the average net wage of people in work. It is unclear why this calculation (unlike many others) does not adjust for age and gender (AG) groups. The sick rate is then applied.

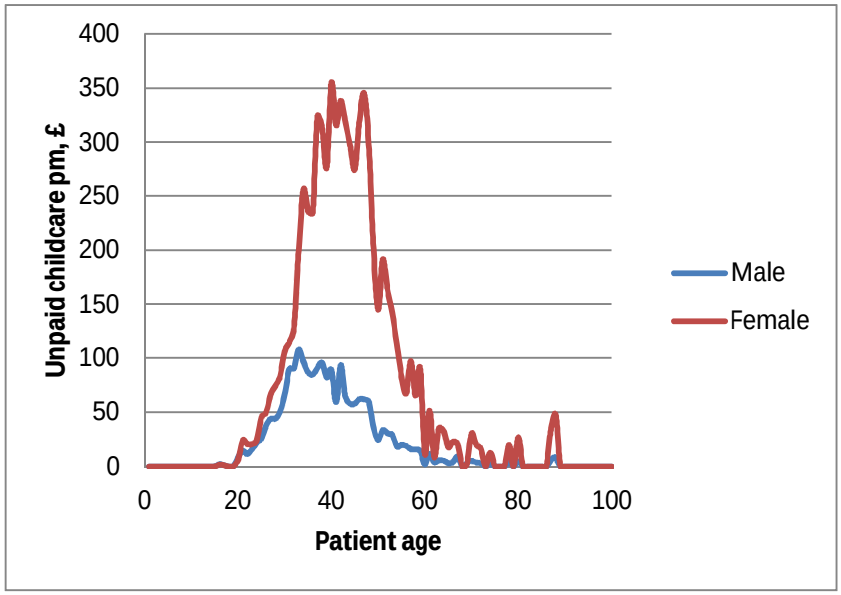
There are many assumptions, calculations and transformations that go into determining these values. Here we focus on one aspect, the production of childcare, the results of which per calendar month can be seen in Figure 2 (although note that values are also ICD chapter and QoL adjusted via the sick rate). In terms of observations, the values are clearly non-linear, and 'jump around' noticeably, potentially highlighting deficiencies with the data. There may be an argument for 'smoothing' the relationship prior to use. For example, the value drops from the age of 40 years for women, before rising, then again falling, without a clear logic. While in the Template age is categorised into eight bands, meaning that these problems might be negated, the data if taken at face value, shows steep variation in the values within each band – for example between 30 and 45 years. This could imply that the bands are too wide to accurately reflect the observed values. Another potentially important aspect of the calculation is that value is only attributed to the youngest child, the marginal value associated with older / other children is said to be zero. Moreover, the value of childcare production is only valued inside of potential working time, as it is stated that childcare outside of this time does not represent an opportunity cost to patients. The justification for this is that it offers the same value to the patient as the second best alternative leisure activity; some parents may disagree

with this assumption. As a final comment, also noted by the DoH, the Time Use Survey used to estimate a number of parameters, is at least 13 years old and there is some concern as to whether the values taken from it are contemporaneous.

One of the parameters that has proved the most difficult to estimate, is the sick rate adjusted for a variety of patient characteristics. In a previous approach it was estimated for age, gender and ICD code using Health Outcomes Data Repository (HODaR) data. However, it was heavily criticised at a previous stakeholder meeting. There are a number of reasons why but they include the facts that the dataset consists of individuals who had recently been discharged from hospital, the QoL instrument (the EQ-5D) asks individuals about their health 'today', when the study recall period was six weeks and the ICD coding was at chapter level – which only consists of 19 categories. The use of QoL, measured as utilities using the EQ-5D, to be indicative of productivity is questionable. For example, previous commentators have noted that people with acute and chronic diseases may have the same EQ-5D loss but different effects on paid labour. The suggested solution was to use the domains on the EQ-5D, rather than the composite utility score to assess how productivity varies with quality of life.

This HODaR based analysis has since been replaced using evidence from the Understanding Society study [11]. More specifically, the sick rate is calculated from the previously described productivity analysis (Productivity_{AQ}), and is defined as the ratio of Productivity_{AG}, to productivity at full health, given a person's age. Note that at the time of writing, the DSU has not had the opportunity to examine this analysis in any further detail. Moreover, it is our understanding that further attempts to refine the 'sick rate' analysis are currently being investigated.

Figure 3: The value of the production of child care by sex [3]



6.2. Consumption

6.2.1. Formal and informal care

Consumption of formal care in the proposal includes the consumption of residential and non-residential care whether it is paid for by the individual or the state and informal care is defined as 'support' provided by family and friends. The latter is estimated separately for children and adults.

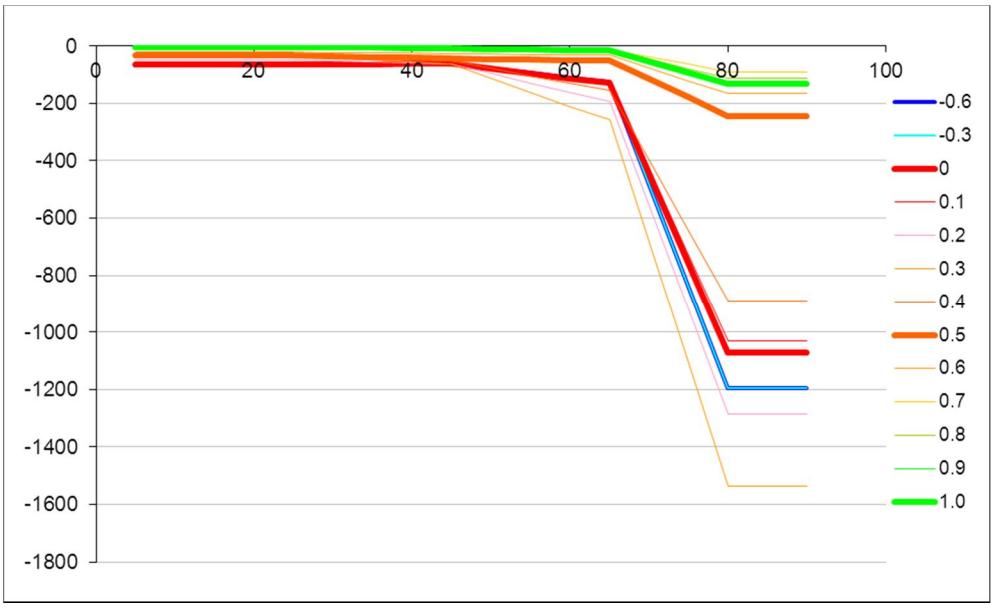
$$(Prob\ of\ using\ residential\ care_{AQ} \times Cost\ of\ residential\ care)$$
$$+ (Prob\ of\ using\ non-residential\ care_{AQ} \times Cost\ of\ non-residential\ care)$$
$$+ (Days\ of\ informal\ care_{AGIQ} \times Hours\ of\ informal\ care\ per\ day_{Q} \times Net\ wage)$$

Once again, there are numerous parameters and assumptions that underlie the above formula. One of the required estimates is the distribution of individuals in residential / non-residential care by QoL. This was assessed using EQ-5D data from the Adult Social Care Survey [12] (it only contained the EQ-5D questions relating to pain and anxiety /depression) and other data from it which were said to have been linguistically mapped to the remaining three EQ-5D domains. The uncertainties inherent in this process are unclear. Adjustments were also said to have been made to the annual cost of treating dementia and stroke patients, as a commissioned literature review from the Personal Social Sciences Research Unit (PSSRU) [13] identified these two conditions as they were found to be associated with systematically higher levels of care use. However, neither the details of this review process are provided nor the potential problems with it. For example, it could be that some relatively costly

conditions were not identified simply because research into these areas has been conducted less frequently. 'Multipliers' are applied to increase the costs of formal care for both conditions; 8.41 and 5.88 for dementia and stroke care respectively. While the steps to calculate these multipliers are described in some detail, their level of uncertainty and its importance are both unclear. For example, average annual costs for both patient groups were calculated based on average QoL and age distributions, but it is unclear if non-linearities in these distributions were taken into account. If they have not then the average estimate is likely to be biased, although the size and importance of this bias is unclear.

Figure 4 shows (what is believed to be) the average monthly formal consumption costs by age and QoL. Note that individuals with a QoL of 0.7 have the lowest costs, those with the highest mean cost have a QoL of 0.3 or 0.6 – the graph is unclear as to which but the point is it is not individuals who are in the lowest QoL grouping. The next most costly QoL strata are 0.2 and states lower than 0 respectively. Thus, the rank ordering of the QoL strata in terms of cost shows no logical pattern even if one allows for some degree of non-linearity. Following this logic, in some circumstances increases in QoL will lead to increases in the threshold ICER since it will lead to higher formal care costs. The DoH possibly could explain such an observation, however the implications of this 'problem' could be substantial in terms of sometimes generating seemingly inconsistent and large changes in the threshold ICER since formal care costs are often a large component of total WSBs. Although we have said less with regards to informal care, we note similar, if not more exaggerated, problems (see Figure 5). The adult proportion of the data (>18 years of age) is taken from a prospective survey of inpatients, thus they are unlikely to be representative of the general population.

Figure 4: Estimates of the average monthly level of formal care consumption (£ on the y-axis) by age (x-axis) and QoL



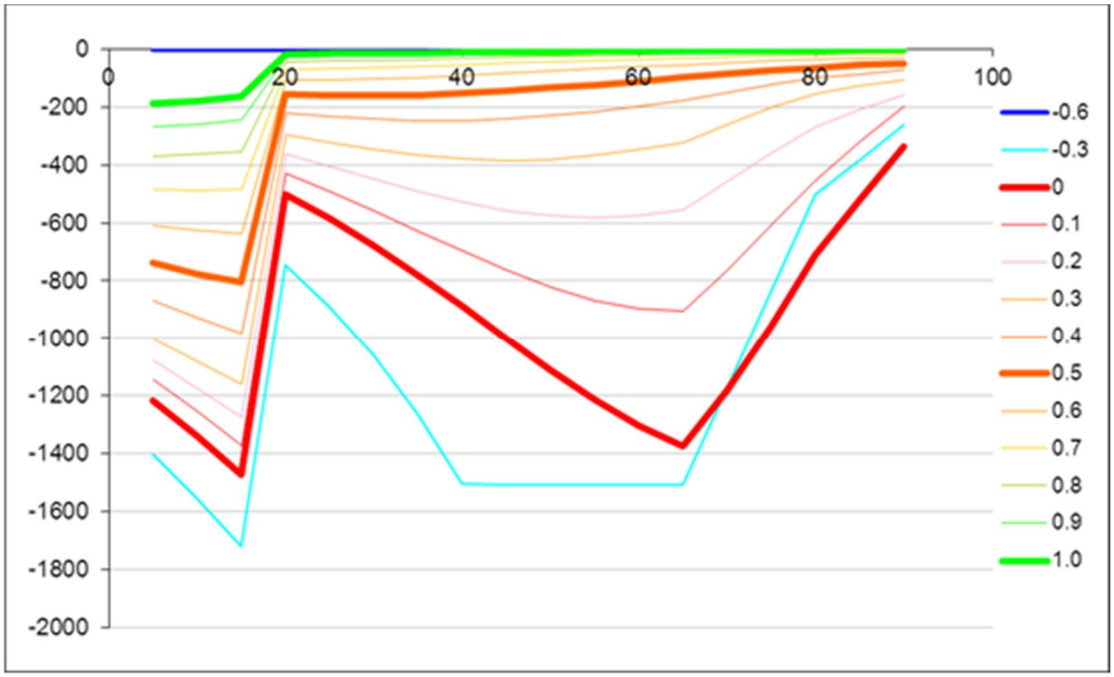
6.2.2. Consumption - Personal unpaid

Personal unpaid consumption (e.g. domestic work - but not related to health care) is estimated solely with respect to the patient's age. The average value of unpaid consumption per month is estimated at £1,210 per month, based on data from the Time Use Survey and assumptions regarding gross weekly wages to value time.

6.2.3. Consumption - Personal paid

The level of personal paid consumption (e.g. purchasing of food, housing and leisure activities), is dependent on QoL and age. Note that if QoL goes up, so does the level of personal paid consumption which therefore has the overall effect of reducing WSBs everything else held constant. The relationship between QoL and level of consumption was assumed to be linear, hence the seeming inconsistencies compared with other categories of consumption are not seen (Figure 6). However, no evidence to support this assumption is presented and the study used to estimate how private consumption varies with patient age was based on an analysis of Swedish data (a 2002 Phd thesis by Ekman Mattias is referenced, but no further details are provided).

Figure 5: The average monthly level of informal care consumption (£ on the y-axis) by age (x-axis) and QoL [3]



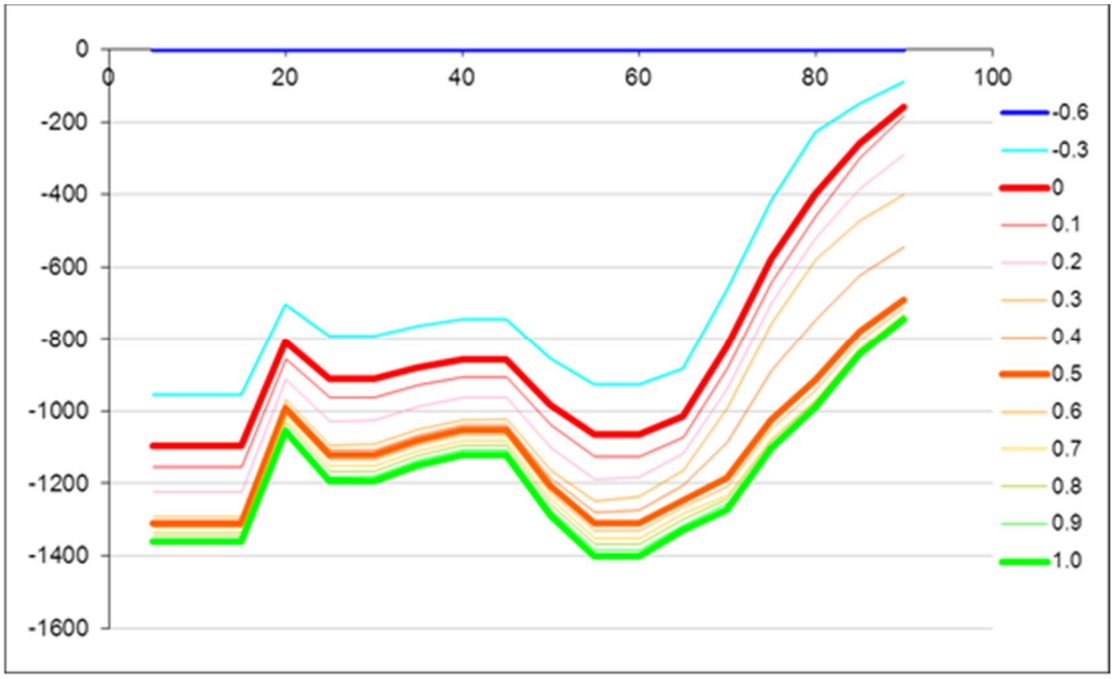
6.2.4. Consumption - Informal childcare

Informal childcare consumption is defined as the consumption of all paid childcare (e.g. from childminders) and unpaid childcare (e.g. from parents). Note that care received from children at

school is excluded since its cost is captured in the 'consumption of government services' element of net production.

The DSU notes that final levels of childcare consumption are presented by QoL and appear to be programmed in the Template, however no details of how this is done are contained in the DoH's dossier. It also appears to be the case that children at very severe levels of health (equivalent to a QoL of -0.6) do not incur any childcare consumption costs as they are assumed to be in hospital / formal care – however the choice of this value appears to be arbitrary.

Figure 6: The average monthly value of personal paid consumption (£ on the y-axis) by age (x-axis) and QoL [3]



6.2.5. Consumption - Government services

There are many assumptions and sources of evidence that have been used to estimate the consumption of government services; note that costs related to the specific treatment / disease are excluded since it is assumed they will be included in standard submissions, but all other health care costs are included.[6]

[6] We note that inclusion of so-called 'future unrelated health care costs' in economic evaluations was rejected by NICE in its 2008 Technology Appraisal methods guide update.

The DoH has defined this category as consumption of government services that can be related to a given person, given the patient's age.

$$(Gov't\ expenditure - Gov't\ expenditure\ on\ transfers\ payments - Gov't\ expenditure\ on\ public\ goods)_A$$

They include costs such as those relating to education, other health care, law courts and prisons. However, they exclude costs that are not specific to an individual – so called 'public goods' whose level of provision is unaffected by the level of consumption. These include categories of cost such as defence, street lighting and research / development. Government transfer payments (e.g. social security payments) are also excluded since they are said to be already taken into account under private paid consumption.

The total government expenditure per patient comes from the Public Expenditure Statistical Analysis (PESA, 2012) [14]. From this, the costs relating to transfer payments and public goods are subtracted. Spending has been assumed to differ by age for health and education. However, the DSU is unclear how this has been undertaken. The DoH's dossier firstly refers to a 2002 Swedish study but later refers to other sources of evidence such as population estimates from the UKs Office of National Statistics. For all other remaining government services, consumption is assumed to be constant across age groups (as combined they represent only a small proportion of overall consumption on government services). Note that the DSU is unclear whether the DoH proposal is to determine the level of government consumption as being dependent on QoL as well as age. This is not explicitly stated in the dossier or a methodology provided, yet it appears to be programmed in the Template.

7. ISSUES WITH THE PROPOSED APPROACH

Accepting the aim of incorporating WSB into the appraisal process, the first issue to address is whether the definition of WSB proposed by the DoH (the difference between the average production and average consumption of a particular patient group) is the most appropriate means to reflect WSB. Each part of the WSB carries an equal weight and in terms of the underlying logic of estimating net resource impact this must clearly be the case. However, it may not be that changes in the different components are valued equally by the public, for example, an increase in paid production (of £500 per month) may not be regarded as having the same value as reductions in care delivered by family members (of £500 per month).

The introduction of WSB is broadening the perspective of the assessment of cost-effectiveness beyond the health of the patient and the costs to the NHS so as to include consequences for the rest of

society. Note, however, that this is not the same as taking a societal perspective. The patient is part of society but we are only counting changes in their health. Suppose as a consequence of a treatment being adopted paid production was increased and these patients then increased their personal paid consumption by the same amount. There would be no change in their net resource contribution as a result of these two changes. If instead these patients do not increase their personal paid consumption then there will be a beneficial impact on WSB. Are the patients better off as a result of their increased consumption? The traditional view of economics would be that they have gained from their increased consumption. Thus increased personal consumption may be of value to patients but it is not valued by this appraisal process.

There are a number of issues about the relationship between the implementation of WSB and the current, and ongoing, standard economic evaluation methodology. While there are issues regarding sequencing and the relative importance to be given to these different elements of the appraisal there is also a striking difference in the level and depth of the underlying analysis. NICE's standard economic evaluation methodology has evolved considerably over time, and routinely involves a very detailed, careful analysis of the implications for health outcomes and resource use for narrowly defined groups of patients. The resulting guidance has been issued for well-defined groups in particular circumstances, for example:

Tocilizumab in combination with methotrexate is recommended as an option for the treatment of rheumatoid arthritis in adults if:

- *the disease has responded inadequately to disease-modifying anti-rheumatic drugs (DMARDs)* ***and*** *it is used as described for tumour necrosis factor (TNF) inhibitor treatments in Adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis (NICE technology appraisal guidance 130), specifically the recommendations on disease activity and choice of treatment* ***or***
- *the disease has responded inadequately to DMARDs and a TNF inhibitor and the person cannot receive rituximab because of a contraindication to rituximab, or because rituximab is withdrawn because of an adverse event,* ***and*** *tocilizumab is used as described for TNF inhibitor treatments in Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a TNF inhibitor (NICE technology appraisal guidance 195), specifically the recommendations on disease activity* ***or***
- *the disease has responded inadequately to one or more TNF inhibitor treatments and to rituximab*
- ***and*** *the manufacturer provides tocilizumab with the discount agreed as part of the patient access scheme. (para.1.1 TA 247)*

Looking back at Table 2 and comparing the level of aggregation and the opportunities to reflect the WSB associated with the treatment of a well-defined patient group in particular circumstances. Much of the estimation is at the level of AG groups. The main opportunity for the WSB associated with the

specific patient group is with respect to Q. But here it is likely that there is a mismatch in terms of health state values used. It appears that an ICD average will be used for the comparator and some average improvement (presumably from the standard economic evaluation). This average improvement in health state is not routinely produced but should be identifiable given access to the economic model (but not from the QALY gain and life year gain since these will be in present value terms). But is the improvement produced from the economic evaluation applied to the ICD average meaningful or a necessary and acceptable approximation? In the case of tocilizumab, 346,000 people are estimated to have rheumatoid arthritis, ten per cent of whom are eligible for second line treatment with a biological, and in turn about half of these would be eligible for treatment with tocilizumab.

Whether the apparent difference in robustness between the standard economic evaluation and the WSB calculation is an important issue may depend on their relative influence over the subsequent decision-making. It is also possible that the asymmetry influences how the standard economic evaluation is perceived.

An early attempt to develop methods of evaluating the benefits of government investments may have resonance despite more recent changes in diet:
"One of them likened the problem to appraising the quality of a horse-and-rabbit stew, the rabbit being cast as the consequences that can be measured and evaluated numerically, and the horse as the amalgam of external effects, social, emotional, and psychological impacts, and historical and aesthetic considerations that can be adjudged only roughly and subjectively. Since the horse was bound to dominate the flavour of the stew, meticulous evaluation of the rabbit would hardly seem worthwhile." [15]

In relation to the above quote, it is important to note that there are a large number of statistical analyses and assumptions that have gone into the construction of each element (such as the level of paid production) that we have not reviewed. For example, we have noted that the level of productivity in the labour market is an important component of calculating paid production. However we have said relatively little about how this parameter in itself has been constructed. The same is true for most of the individual components in each of the existing elements. This is partly because this process would take a considerably longer amount of time than the DSU has had available but also because we do not believe this level of critiquing is necessary for the particular Working Party meeting. Having said this, we acknowledge that this level of critiquing will be required at a later date if the proposal to implement WSBs in their current form is taken forward. To illustrate this point we return to the construction of the productivity rate of the element 'paid production'. The 'Understanding Society' 2009-2010 dataset [11] was used to calculate the rate, which includes respondents from over 50,000 UK residents. Importantly, individuals were assumed to be productive

only if they had been in paid employment in the last week, no allowance is made for paid employment prior to this time. Moreover, the relationship between EQ-5D levels and productivity was estimated by 'mapping' SF-12 (a different quality-of-life instrument that was contained in the survey) to the EQ-5D by first calculating the SF-12's physical and mental component scores. However, important questions need to be raised over a number of issues, such as:

- the choice of covariates in the models (for example there is some suggestion that the level of production varies by gender, but it is unclear why this covariate has been omitted from the final analyses yet it is included in other components of the element i.e. the wage-rate)
- whether the models systematically underestimate the level of productivity particularly for those in good health
- whether alternative model specifications would provide a better fit to the data
- the level of uncertainty that must exist in the current productivity estimates as a combined result of all these analyses / assumptions.

A major issue to consider is what weight to give WSBs relative to other components, such as the standard economic evaluation. There is at least one easy way in which operationally to select how much weight the WSB is given – through the selection of the exchange rate between the QALYs and the monetary estimate of WSB. Clearly there are others, such as capping the magnitude of WSB for any patient group that will be included in the calculation of the adjusted ICER threshold, although they are not without problems as there is no obvious rate at which a cap or tapering system should be applied.

As it is standard practice with NICE methods and procedures, the detail of the WSB calculations needs to be fully shared with stakeholders in order to receive constructive criticism from a wider circle. Without giving all stakeholders opportunities to comment and potentially influence the detail of implementation it may prove difficult to obtain enthusiastic participation in what is quite a dramatic departure from previous practice. The calculation of WSBs involves many steps and assumptions and making the full details of its calculation readily available would appear to be an important step in facilitating further refinement and enhancing transparency.

7.1. Consequences for the submission of evidence

There are important decisions to be made regarding use of the Reference Dataset. One option is that all appraisals must use the Reference Dataset. Another would be that manufacturers can present their own estimates of WSB in addition to those from the Reference Dataset. Another possibility is that manufacturers must use the Reference Dataset if they do not present their own estimates. There are clearly some advantages in permitting non-Reference Dataset submissions. It would be a means of

encouraging research activity to refine the relatively crude estimates presented in the Reference Dataset. It might facilitate better decision-making by permitting use of more valid estimates of WSB, and it might reduce instances where a particular patient group is disadvantaged by the level of aggregation used when deciding their access to different health technologies. The disadvantages include the increased costs of scrutinising manufacturers' submissions, and potentially a loss of transparency (unless there is very clear guidance on how calculations are to be made and presented).

There are also issues about the introduction of WSB (and BoI) with respect to the standard economic evaluation. This could follow the approach taken with respect to end of life (EoL) [1], where, for the Committee at least, there was a two stage process with the standard evaluation coming before any consideration of EoL which then enters at the stage of deciding an appropriate threshold to use. Alternatively, WSB and BoI might become an integral part of the economic evaluation. The DoH proposal would appear to be consistent with a two stage approach. Indeed, a two stage approach may be unavoidable as the Committee's view and that of the manufacturer as to the incremental QALYs frequently differ and thus calculation of the WSB could not be finalised until the Committee had agreed on the relevant parameters from the standard economic evaluation. This could create a challenge for a manufacturer wishing to submit a non-Reference Dataset estimate of the WSB.

7.2. Gaps in the evidence

While there has been investigation of how the valuation of health-gains by the public is influenced by the BoI or EoL, how the public would prefer WSB to influence decisions regarding which health technologies are to be routinely provided to different patient groups has received much less attention.

Although it would not be appropriate to assess the merits of these proposals on the basis of what effect they would be likely to have on decision-making it would be useful to have some evidence on the likely consequences of their being implemented. One means of doing this would be to look back at some recent decisions and re-analyse them using the proposed methods. Rather than simply being able to gauge the likely magnitude and direction of the impact on decisions, this would provide some insight into the additional time and effort required to complete appraisals, for example, in generating the QoL information required.

The WSB calculation depends on many different relationships between the different elements of production and consumption and patient characteristics. Further research on all of these relationships would be valuable, whether it provides a more robust basis for the estimation of WSB or reassurance regarding the validity of the currently postulated relationships.

A further gap in the evidence concerns the absence of discounting from the VBP procedures (other than from the standard economic evaluation), for example, to what extent does it introduce systematic biases, and could discounting be readily introduced?

The issues around the non-linear relationship between age and estimated WSBs require further investigation, in particular in the context of current practices applied in standard economic evaluations with respect to patient age.

As previously noted, we believe that the individual items used to construct each element require further specific critiquing and refinement if required.

The calculation of the WSB involves an enormously large number of statistical analyses, assumptions and parameters, and accordingly involves substantial uncertainty. At this stage it is difficult to gauge the magnitude of the additional uncertainty introduced. Thought needs to be given to how to adapt the decision-making process to this increased uncertainty.

8. REFERENCES

1. National Institute for Health and Care Excellence, *Guide to the methods of technology appraisal*, 2013.
2. Miners, A.H., Cairns, J. and Wailoo, A., *Department of Health proposals for including burden of illness into value based pricing: a description and critique*, 2013, NICE Decision Support Unit.
3. Department of Health, *Documentation for Wider Social Benefits v3.2*, 2013.
4. Department of Health. *Feedback from technical workshop on the Wider Societal Benefits component of Value Based Pricing*. 2012.
5. Murray, C.J.L., Ezzati, M., Flaxman, A.D., et al. *GBD 2010: design, definitions, and metrics.* The Lancet, 2012. 380(9859): p. 2063-2066.
6. Griffin, S., Palmer, S., Soares, M., Spackman, E., Tappenden, P., Wailoo, A. and Whyte, S., *Incorporating Wider Societal Benefits into estimates of cost per QALY: Case studies*, 2012, NICE Decision Support Unit.
7. Office for National Statistics. *Annual Survey of Hours and Earnings*. 2011; Available from: http://www.ons.gov.uk/ons/rel/ashe/annual-survey-of-hours-and-earnings/2011-revised-results--soc-2010-/index.html.
8. Eurostat. *Total wages and salaries (tps00113)*. 2010; Available from: : http://epp.eurostat.ec.europa.eu/portal/page/portal/labour_market/labour_costs/main_tables
9. Office for National Statistics. *Time Use Study*. 2000; Available from: http://discover.ukdataservice.ac.uk/catalogue/?sn=4504.
10. Health and Social Care Information Centre. *Survey of carers in households (2009)*. 2010; Available from: http://www.hscic.gov.uk/pubs/carersurvey0910.
11. Understanding Society. *The UK household longitudinal study (wave 1)*. 2010; Available from: https://www.understandingsociety.ac.uk/.
12. NHS Information Centre. *Adult social care survey*. 2012.
13. Watters, S., Fernandez, J.-L., Knapp, M. and Wittenberg, R., *Estimates of social care costs for older people suffering from dementia, stroke, and musculoskeletal conditions*, Unpublished, Personal Social Sciences Research Unit.
14. Treasury, H. *Public expenditure statisical analysis*. 2012; Available from: www.gov.uk/government/publications/public-expenditure-statistical-analyses-2012.
15. Dorfman, R., *Measuring Benefits of Government Investments*1963: The Brookings Institution.