A multi-criteria analysis approach to commissioning: practical dilemmas in priority setting.

Paper for presentation at HESG, Glasgow, July 2004

Authors & Affiliation:

Ed Wilson[1], Richard Fordham[1]

1. Health Economics Support Programme, Health Economics Group, University of East Anglia

Corresponding author & contact details

Ed Wilson
Health Economics Support Programme
Health Economics Group
School of Medicine, Health Policy & Practice
University of East Anglia
Norwich
NR4 7TJ

ed.wilson@uea.ac.uk

01603 591444

Abstract

Issue being addressed: The use of cost-per-QALY league tables is of limited use in prioritising health care expenditure as QALYs do not capture all the relevant goals of the NHS. An alternative means of measuring the benefits is a multi-criteria weighted benefit score. Dividing the net cost of a programme by the weighted benefit score generates a cost-value ratio, by which programmes can be ranked. This ranking is, however, biased due to the benefit score being bounded and costs being unbounded and leads to cost (not cost-effectiveness) dominating the rankings. In this paper, we present possible pragmatic solutions when applying the technique to prioritise expenditure in a PCT in the East of England.

Method: The overall benefit of each proposed health care programme is scored using a multi-criteria weighted benefits scoring system. These can be combined with programme costs in a number of ways: a simple cost-value ratio (dividing net cost by benefit score), using cost indices in place of net cost to calculate the ratio, using cost per patient affected, or ranking by cost and by benefit separately. Each method has its own strengths and weaknesses which will be discussed in turn.

Results: The multi-attribute utility scoring system captures all important dimensions of benefit with a weighted benefit score between 0-10. When combined with cost, proposals are ranked in order of cost-value ratio.

Implications/ Discussion: Because benefits are bounded by the nature of the scoring system and costs are open-ended, a crude cost-value ratio ranking is potentially dominated by cost rather than cost effectiveness and fails to take into account the number of patients affected and/or scale of each programme. The alternative methods described may mitigate this, but all have drawbacks.

Key issues for Discussion by HESG audience: Identification of the best practical method of combination of costs and benefits.

Status of work: Likely to be substantive changes

Introduction

As resources are limited, the demand for health care will always exceed the capacity of the available resources. Therefore decisions have to be made over which treatments and services to prioritise over others to meet the aims of the PCT most closely. These decisions are currently undertaken in a PCT in the East of England by a 'cluster group' consisting of chief executives, finance directors and clinical directors of both the PCT and the local major acute trust. Both the PCT and the acute trust arrive at the meeting with a list of desired purchases (e.g. new staff and new equipment from the acute trust, and a mixture of health care packages from the PCT). The meeting then agrees to trade off against these lists until all available funds have been allocated.

The advantages of this system are that it is relatively quick and cheap to operate, and it is clear exactly where the decisions are made. Criticisms of the process are that it is supplier-led with the acute trust holding much of the power, almost entirely input-focussed with no consideration of outcomes, there is no evidence base used to support any of the bids, and the decision making process is opaque making it difficult to justify why a particular decision was made.

The authors were requested by the PCT to devise a tool that would enable the fair and equitable commissioning of new services. The tool would require bids for funds to be output / outcome based (thus making it possible to value the relative effectiveness of different programmes of care), making use of the literature and other evidence in support, and rank bids according to best value for money.

A technical approach to prioritisation is the use of a cost per QALY league table (Maynard, 1991). A low cost per QALY represents better value for money. Thus by funding in increasing order of cost per QALY, the biggest QALY gain for the population can be obtained for the given budget.

There were several experiments globally in the 1990s to implement such systems and similar prioritisation methods, most notably in Oregon (Coast 1996). However, the

deficiencies of the purely technical solution became apparent: data were often simply not available or became out of date before they could be incorporated into the system (Holm 1998). In some cases, cost-effectiveness ranking did not produce socially acceptable priorities and other considerations had to be incorporated ex post.

The problem may lie in the use of QALYs themselves. QALYs are sensitive to the tool used to value them, and may not capture all dimensions of quality of life relevant to patients. Furthermore, ranking by cost per QALY assumes that the sole objective of the health care system is to maximise QALYs gained (Brinsmead & Hill, 2003). Wider societal objectives such as access to services, equity and exogenously determined national priorities are not necessarily captured by the QALY. Additionally, PCTs are obliged to directly involve the public in "the planning…, development and consideration of proposals for changes in the way their services are provided" (HMSO 2001). It is highly questionable whether prioritisation solely on the basis of a cost per QALY sufficiently fulfils this requirement.

Thus a more pragmatic approach to achieving allocative efficiency and valuing outcomes is sought. A desirable technique might be characterised by "value systems, transparency and accountability" (Holm 1998). An example of this is a multi-criteria analysis (which forms the basis of a standard option appraisal), and has been widely employed in both health and non-health sectors (Chappel et al 2001, Scott & Lees 2001, Dodgson et al. undated).

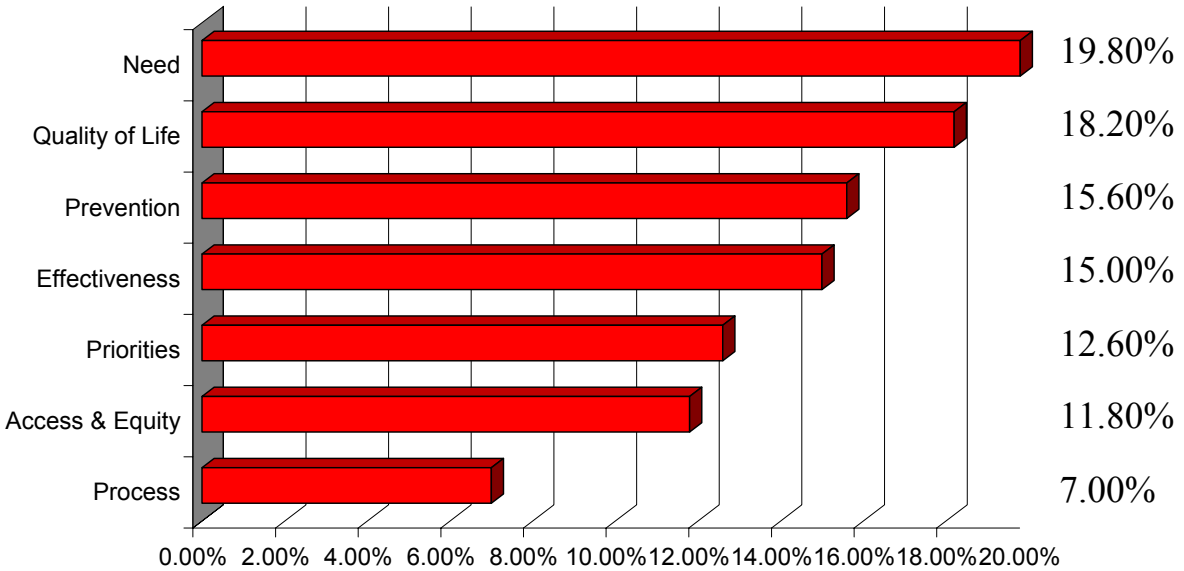
The remainder of this paper consists of a description of the method adopted in an East of England PCT. During development of the method, an interesting theoretical issue arose (although this problem has not been realised in practice as evaluation of the method is still underway). The description of this, and a number of potential solutions form the main focus of this paper, and follow on from the description of the method.

Description of Method

Full details of the method are published elsewhere (Wilson & Fordham 2004). The method adopted in the study PCT is based on a multi-criteria analysis approach, similar to that used routinely in option appraisal processes. This generates a single weighted benefit score between 0 and 10 from a set of explicit pre-determined criteria. The net cost impact of the bid is divided by this score to generate a 'cost-value'[1] ratio. Ranking all bids from lowest to highest cost-value ratio generates a prioritised list with those at the top being most efficient at achieving a point of benefit compared with those lower down.

The framework for calculating the benefit score of bids was devised, tested and analysed at a series of workshops held between October 2003 and February 2004. Around 20 representatives from across the local health economy, including NHS clinicians, PCT and acute trust managers, social services and the voluntary sector attended the workshops. The outcome was a list of seven criteria, each of which was weighted relative to one another. Participants refined these criteria during and after the workshops, developing precise definitions and testing the framework with a number of mock bids. Figure 1 below shows the criteria and weights (see Appendix I for full definitions of criteria).

Figure 1 – Criteria and weights



[1] The term 'cost-value' ratio is used throughout to refer to the ratio of cost to weighted benefit score, so as not to confuse with conventional definitions of cost-effectiveness (i.e. cost per QALY or other unit of effectiveness).

Bids for PCT funds are submitted on a standard form which relates directly to the seven criteria. The net cost impact for the bid is calculated over a three year time horizon. Costs included are direct NHS costs (split by capital and revenue), 'wider health economy costs' (e.g. costs borne by social services or the voluntary sector), and patient out of pocket costs. Savings (e.g. from discontinuation of current service provision, sale of equipment etc) relevant to NHS, wider health economy and patients are subtracted from this. Therefore the cost component attempts to encapsulate the incremental cost to society of the bid.

Each bid is scored against each criterion by a scoring panel which meets 4 times a year. The group arrives at a subjective score between 0 and 10 against each criterion, where 5 represents no change in current provision, greater than 5 is an improvement and below 5 is a deterioration. In this way, the score represents the incremental improvement (or deterioration) against each criterion compared with current service provision. For each bid there will therefore be seven scores out of 10 (one for each criterion). These raw scores are multiplied by the weights (as per Figure 1) and summed to generate a weighted benefit score for the bid.

The next stage is to divide the net cost impact of the bid by this score to give a cost-value ratio. The lower the cost-value ratio, the less the cost per 'point of benefit'. Therefore, bids are ranked from lowest to highest cost-value ratio. This forms the recommendation to the PCT board for which bids to fund first.

Once the Board receives the recommendation, the list is reviewed, and, if deemed appropriate, changes are made to the ranking. This is because, as with any prioritisation framework, there are imperfections in the mechanism: it is not possible to be certain that every conceivable dimension of benefit is included in the criteria. Furthermore, for practical purposes, the time horizon was limited to 3 years, thus undervaluing bids with benefits expected after this time (likewise over-valuing bids with substantial costs expected after this time). Therefore, there is a risk of biasing against longer term bids, and obtaining perverse rankings which do not make 'intuitive sense'. It is crucial, however, that changes to the recommended ranking are

not arbitrary and must be justified and documented, thus retaining the transparency of the system.

The Problem

During the development of the process, it was discovered that measuring the benefit score on a bounded scale (between 0 and 10) and measuring cost on an unbounded scale presented an interesting dilemma. Since cost can vary over much wider values than benefit, it is possible that when comparing cost-value ratios of two or more bids, cost could have a much greater influence than benefit score. The result is that the ranking may be dominated by cost rather than cost-value, and thus bids end up ranked in cost order only. This risk is illustrated by the results of a test scoring session on four mock bids (W, X, Y and Z) drawn up to test the process developed for the PCT (Table 1).

Table 1 Results of mock scoring exercise

Bid	Net Cost Impact	Score	Cost/Point
W	-£160,000	6.380	-£25,078
X	£21,600	7.543	£2,863
Y	£95,000	6.720	£14,137
Z	£168,000	5.012	£33,522

In this case, the ranking by cost-value is the same as the ranking by net cost. This may just be a coincidence, but a theoretical example suggests otherwise.

Suppose there were two bids for funds from the PCT, A and B. Assume they are absolutely identical in every way except that one is 10 times the scale of the other. It is therefore exactly 10 times as costly, and 10 times as beneficial (assuming constant returns to scale) on every possible dimension (10 times the QALY gain, 10 times the 'access' etc).

Suppose the benefit of bid A was 10 QALYs, and the cost £100. Project B would obtain 100 QALYs at a cost of £1,000. These have equal cost-utility ratios of £10 per QALY (Table 2).

Suppose using the weighted scoring system described above, project A was scored at 6. As project B has 10 times the benefits in every possible dimension, project B should be scored at 60, but the maximum possible score is 10. The cost-value ratio of A is now £100/6 = £17 per point. The cost-value of B is now £1,000/10 = £100 per point. It therefore looks like A is much better value for money, when in fact it is just cheaper. Unless we have an inherent preference for smaller projects, these two should be ranked equally.

Table 2 Illustration

Bid	QALYs	Cost	**£/QALY**	Points	Cost	**£/point**
A	10	£100	**£10**	6	£100	**£17**
B	100	£1,000	**£10**	10	£1,000	**£100**

So, assuming that this problem is a real problem (which the theoretical example above appears to demonstrate), what should be done?

Potential Solutions

There are five potential solutions. These are:

- Benefit re-scaling;
- Cost re-scaling;
- Unbounding benefits;
- Bounding costs;
- Taking average rankings

These are each considered in turn below, along with a demonstration of their impacts on the mock bids shown in Table 1, and a discussion of their relative strengths and weaknesses.

1. Benefit re-scaling

The weighted benefit scores are currently on a scale of 0 to 10. A transformation of scores to, for example, between 0 and 1000 would make the actual numbers on the cost per point scale smaller, therefore there would be a psychological perception that there was not so much difference between the cost per point values (Table 3).

Table 3: Benefits Rescaling

Bid	Net Cost Impact	Score	Cost/Point
W	-£160,000	6,380	-£25.08
X	£21,600	7,543	£ 2.86
Y	£95,000	6,720	£14.14
Z	£168,000	5,012	£33.52

This might have some relevance in a post hoc review of rankings through influencing perceptions on the relative difference between cost-value rankings (the difference between bids X and Y on this scale is only £11.28 per point versus £11,274 in Table 1), but this is purely a psychological 'trick', is a very simplistic procedure, and completely fails to tackle the problem. The ranking is still in cost order.

2. Cost re-scaling

An alternative to rescaling the benefit scores is to rescale the costs. Two possible techniques for doing this are explored below. Firstly, the impact of taking the log of cost is examined. Secondly, conversion of cost to a cost per patient affected is explored.

2.1 Log of cost

Taking the log of cost would have the effect of bringing the tails of the cost distribution closer together: high cost proposals would be 'brought in', showing a lower log(cost)-value ratio. This may mean higher cost proposals stand more chance of appearing 'cost-effective'. The theoretical justification for this is unclear, and it would not alter the fundamental problem of bounded benefits versus unbounded costs. This also creates problems when handling negative costs (where the log is undefined).

2.2 Cost per patient

The method as presented takes the net financial impact of a bid over three years as the numerator in the ratio. An alternative to this is to use the net cost per patient affected (simply dividing the net financial impact by the number of patients affected / treated / targeted as appropriate). This method has been used before (Scott & Lees 2001), and Table 4 shows the impact of doing this to the running example.

Table 4: Cost Rescaling

Bid	Patients affected	Net Cost Impact	Net cost per patient	Score	Cost/Point
X	800	£21,600	£27	7.543	£3.58
W	5,000	-£160,000	-£32	6.380	£5.02
Z	150	£168,000	£1,120	5.012	£223.46
Y	30	£95,000	£3,167	6.720	£471.28

Dividing by the number of patients factors out the scale effect of bids. This gives a fairer comparison of the options. In this case, the result is to switch bids W & X and Y & Z.

The major drawback of this method is that, to be strictly valid, as the costs relate to the individual patient, the benefits too should relate only to the individual, and not to wider societal issues such as access to services, equity and progress towards targets. Yet these factors cannot be excluded as they are relevant to the PCT in its decision making.

3. Unbounding benefits

This is an alternative approach where costs are left unbounded, but the benefits valuation method is changed such that the benefit scores become unbounded. This method is described in a case study in the Department for Transport, Local Government and the Regions Multi-Criteria Analysis Manual (Dodgson et al. undated). Briefly, instead of subjectively valuing each bid against each criterion on a scale of 0 to 10, the first bid is selected as the reference bid. This is scored 10 against the first criterion. The second bid is compared with the first against this criterion. If it is 'twice as good' as the reference bid, it scores 20. If it is 'half as good', it scores 5 etc. This is repeated for all bids against all criteria individually.

This method neatly solves the bounded benefits problem, and makes the calculation of the cost-value ratio valid, but generates problems of its own in that scores between criteria are not readily comparable, thus the weighting of the criteria relative to one another becomes meaningless. The only solution is to leave out the weighting, and calculate total scores rather than weighted averages (which then implies equal weighting across all criteria!).

4. Bounded costs

This solution is the inverse of that described above, leaving benefits bounded and adding constraints to costs.

There are a number of ways to do this. The first involves the creation of a cost index (of which there are two possible permutations). The alternatives involve expressing the cost of a bid either in terms of the percentage of available funds or derivation of an index based on the average cost of a funded bid from the previous year.

4.1 Cost Indices

This consists of converting the cost of all bids into a cost index where the cheapest bid would be scored 0 and the most expensive, 1000. The costs of all the other bids are converted to the index according to their relative cost.

This creates issues where there are bids with a negative net financial impact (i.e. where the bid leads to a net financial saving to the NHS). To continue the example, this occurs in bid W where there is a net impact of -£160,000. There are two permutations of the cost index to handle this. The first is to value the most cost saving at 0. This gives a total range of cost of £328,000 (-£160,000 to +£168,000). The cost index of bid X then becomes £21,600/£328,000*1000 = 65.85 (Table 5).

Table 5 Cost Index with lower bound

Programme	Net Cost Impact	Cost index	Score	Cost/Point
W	-£160,000	0	6.380	n/a
X	£21,600	65.85	7.543	8.73
Y	£95,000	289.63	6.720	43.10
Z	£168,000	1000	5.012	199.52

The second permutation is to value zero cost at 0, and the highest cost item at 1000.

This gives a range of cost of £168,000. Thus the cost index of bid X is £21,600/£168,000*1000= 128.57 (Table 6)

Table 6: Cost Index with no lower bound

Programme	Net Cost Impact	Cost index	Score	Cost/Point
DS	-£160,000	-952.38	6.380	-149.28
TP	£21,600	128.57	7.543	17.04
RRS	£95,000	565.48	6.720	84.15
MHSI	£168,000	1000	5.012	199.52

The problem with this method is that whilst there is an upper limit on the cost index (1000), there is no lower limit, thus a very cost-saving programme could be valued at less than -1000. Therefore the first permutation is preferred.

Both cost and benefits are now bounded, but this has not altered the prioritisation order of the four bids. Creating a cost index would not be expected to solve the problem as the scale of the costs is still not being related to the scale of the bid itself or to the benefits. Simply transforming costs does not resolve the issue.

4.2 Percentage of available funds & proportion of average cost of funded bid from previous year.

These methods are alternatives to a cost index, which would result in costs being bounded, but, as for the cost indices, this does not tackle the issue of failing to take into account the scale of the project on the cost side.

A particular danger of using the average cost of a funded bid from the previous year is the risk of a return to commissioning based on historical patterns, rather than evidence of what works.

5. Average rankings

The final method has been used recently by a Scottish Health Board (Scott & Lees 2001). Bids are ranked firstly by benefit score, and then again by cost. The average ranking generates a unified ranking index ('Prioritisation Scoring Index', PSI), taking into account both cost and benefit.

This method takes equal account of both cost and benefit, and will lead to high benefit / low cost bids being prioritised above low benefit / high cost ones. The drawback of this system is that perverse rankings may occur in the middle of the table because the relative size of costs and benefits are not taken into account.

To illustrate this, suppose four bids, C, D, E & F had costs and benefits on some measure (this could be QALYs or some weighted benefits score) as per columns 2 and 3 of Table 7 ('cost' and 'benefit'). The ranking in terms of cost per point of benefit (or cost per QALY etc) is C, D, E and then F. But, when ranking by PSI, bid C (highest benefit, least cost) does indeed come in at the top, and bid F (relatively low benefit, highest cost) comes in last, but in between, the PSI ranks E above D. As the cost per benefit point for E is higher than for D, there is a risk that some health gain will be foregone by choosing to invest in E in preference to D.

Table 7: Prioritisation Scoring Index

Bid	Cost	Benefit	£/benefit	£ rank	Benefit rank	Av rank	PSI
C	£15,000	50	£300	1	1	1	1
D	£20,000	35	£571	2	4	3	3
E	£30,000	48	£625	3	2	2.5	2
F	£40,000	40	£1000	4	3	4	4

Conclusion

As demands for health care will always outstrip the resources available, prioritisation decisions have to be made in order to maximise the objectives of the PCT subject to the funding available. An explicit prioritisation procedure making use of the evidence base with a clear set of goals and objectives is more open and defensible than implicit decision making techniques.

There are practical restrictions in implementing a 'theoretically pure' priority setting mechanism (for example data availability, uncertainty and existence of competing objectives). Consequently any pragmatic scheme will necessarily stray from a theoretic ideal. Furthermore, attempts to correct for subsequent inconsistencies can introduce flaws of their own.

We devised a scoring system to measure benefits based on a multi-criteria analysis approach. Using this to calculate cost-value ratios and then ranking produced a list of bids in a recommended priority order. However, this created potential problems because whilst the cost of a bid will to some extent reflect its scale, the score is

independent of the scale. Therefore there is a risk of cost dominating the ranking when comparing bids of differing scale.

We examined a number of potential solutions to this, including rescaling the benefit or cost estimates, re-scoring the benefits such that the score is unbounded and can therefore reflect the scale of the bid, creating bounded cost indices, or taking average rankings of both cost and benefit separately. All these solutions have their limitations. Therefore, in the absence of a 'first best' pragmatic scheme, there is a need to identify the 'second best' from the alternatives presented.

We feel that the most pragmatic and most easily implemented solution is to relate costs to the scale of the bid by dividing the net financial impact by the number of patients affected, and using this as the numerator of the cost/value ratio.

References

Brinsmead R, Hill S. *Use of pharmacoeconomics in prescribing research. Part 4: is cost-utility analysis a useful tool?* Journal of Clinical Pharmacy and Therapeutics 2003;28:339-346

Chappel D, Bailey J, et al. *Implementation and evaluation of local-level priority setting for stroke*. Public Health 2001;115:21-29

Coast J. *The Oregon Plan: technical priority setting in the USA*. In Coast J, Donovan J, Frankel S (eds) *Priority Setting: The Health Care Debate*. Wiley, Chichester, UK. 1996

HMSO. *Health & Social Care Act, 2001*. Chapter 15. Act of Parliament. Web Document. 2001. http://www.legislation.hmso.gov.uk/acts/acts2001/20010015.htm. Accessed April 2004

Dodgson J, Spackman M, Pearman A, Phillips L. *DTLR multi-criteria analysis manual*. Web document. Undated. http://www.odpm.gov.uk/stellent/groups/odpm_about/documents/pdf/odpm_about_pdf_608524.pdf. Accessed February 2004

Holm S. *Goodbye to the simple solutions: the second phase of priority setting in health* care. BMJ 1998;317:1000-2

Maynard A. *Developing the Health Care* Market. The Economic Journal 1991;101:1277-1286

Scott S, Lees A. *Developing a prioritisation framework: experiences from a Scottish Health Authority*. Health Expectations 2001;4:10-17

Wilson E, Fordham R. *Commissioning Health Services in West Norfolk Primary Care Trust: A tool for evidence based and cost-effective* commissioning. Web document. 2004. http://www.med.uea.ac.uk/research/research_econ/hesp/Reports.htm#WestNorfolkPrioritisation. Accessed March 2004

Appendix I: Criteria Definitions

1. Access & equity
   - Does this proposal increase or improve access to services for the target population?
   - Does this proposal have any impact on access to services for other populations or other NHS agencies (positive or negative)?
   - Is this a locally based service?
   - Is this service available to all who need it?
   - Is this patient-centred healthcare? Do they get a say in the delivery of their care? Is there demonstrable 'patient & public involvement'?
   - Does the proposal enable treatment in an appropriate environment?
   - Does the proposal raise the profile of an important but currently low profile disease / condition?

2. Effectiveness
   - Is the proposal proven to work? (what evidence is there for it working?)
   - What is the quality / grade of the evidence? (e.g. well conducted randomised controlled trial versus expert opinion).
   - What is the balance of risk and benefit to the patient?
   - Will the proposal result in enough activity to maintain quality? (clinical governance issues)

3. Local & National Priorities

   - How far towards meeting an explicit national or local target does this proposal go (for example, NICE, NSFs, LDPs etc – Nb: see glossary)?

4. Need
   - What is the prevalence / incidence of the disease or condition this proposal is intended to treat?
   - What is the current mortality or morbidity associated with this disease/condition? (note this should take into account the impact of existing treatments)
   - Does this proposal meet an identified health need (either local or national)?
   - Does it meet public expectations / does it meet a local health want?

5. Prevention
   - Does the programme focus or put greater emphasis on prevention of ill health? (For example through health promotion, screening/ immunisation or reduction in future morbidity.)

6. Process
   - Is the proposal achievable within a realistic timescale?
   - Does the proposal involve multi-agency working / partnership working across different areas of the NHS (and wider bodies)?

- Is the proposal acceptable politically?

7. Quality of life
   - What impact does the intervention have on different domains of quality of life (e.g. disability reduction, increase in independence, pain reduction, whether it allows a patient to play active role in society, social relationships, etc)?
   - What is the potential QALY (Quality Adjusted Life Years) gain from the intervention?
   - Does the proposal decrease (future) care needs for the patient, carer or family?
   - What evidence is there for the patient experience / satisfaction?