Health Economics

Dr Pratap C Mohanty

Department of Humanities and Social Sciences,

Indian Institute of Technology Roorkee

Week – 09

Lecture 43- Non-monetary Valuation of Health: QALY

Welcome friends once again to our NPTEL MOOC module on Health Economics. I hope you have been going through it correctly, especially in understanding health evaluation and its economic approaches. We are in our second lecture on understanding the economic evaluations of health and healthcare. In this lecture, we again emphasise monetary and non-monetary health measures. Especially, in the last lecture, we discussed about the type of monetary approaches. We discussed two methods that are broadly understood in the context of monetary approaches. One is called the revealed preference approach, and the second is the stated preference approach.

While the revealed preference is based on economic logic, whereas the stated preference has a psychological foundation. There are various valuation techniques under stated preference, of which the most used are contingent valuation methods, and the second one is called the discrete choice experiment.

So, in this lecture, we emphasise the non-monetary evaluation or the valuation method. Why are non-monetary valuations, in most cases, considered to be dominating against the monetary outcomes? There are many ways of understanding the non-monetary indicators. As health status measures outcome, we will also study its evolution from incommensurable to commensurable approaches. Then, we will also discuss HRQOL (i.e., health-related quality of life), where we will discuss some generic description measures to developed measures such as QALY (Quality-adjusted life years). In this context, we will also emphasise the definition of QALY and its methods with the latest examples, including the standard questionnaire, which is followed.

So, let us go into the details. In the last lecture, we came across monetary measures of health, i.e., RP (revealed performance) and SP (stated preference) methods. Although, (monetary measures) are widely adopted in different areas of economics. This is not the case in health economics, where measures of health improvements dominate instead of the number of healthcare services.

The arguable reasons for this rejection of monetary measures include i) revealed preference measures are rare in the healthcare market, and use of SP (i.e., stated preference) is attached to some distrust. ii) Healthcare is indeed a special good, where willingness-to-pay criteria are widely rejected as a measure for resource allocation. iii) It is believed that the

CBA is not a practical tool for analysis, but CEA (which we already discussed earlier) is, as health is a more appropriate measure than money.

So, let us explore the measures of health outcomes that dominate the standard monetary approaches, such as revealed preference and stated preference approaches. As I already mentioned, we will discuss its evolution as part of the learning goal. This is necessary before learning the tools of health outcome measures.

In the long history of medical research, a wide range of output measures have been used. Most of them are certainly useful and sensible for specific purposes. These health outcome measures are largely non-monetary in nature and can be divided into two types- Incommensurable and commensurable.

For many years, health was measured using extreme indicators such as mortality, life expectancy, etc. However, over time, it is understood and clearly recognised that these indicate only the presence of disease and impact on an aggregate level, as they are often correlated with a general level of disease. However, there was a need to understand one important point, i.e.,- are there any effects of interventions on individuals? This led to the development of intervention-based measures.

One of the most widely used outcome measures of health is survival rate, which we will explain right now. Here, an intervention is counted as successful if patients survive a few more years; otherwise, it is simply considered a failure.

However, this is termed a crude measure as it does not distinguish between time period, such as- if the intervention caused the patients to survive one week or five years, both are considered to be successful. But the duration and the capture of time is missing, where improvement is required. An alternative to this is 'life years count'. In this case, when life years count is presented, it contradicts the survival rate (which we have taken and where something is missing). I am just mentioning here that the numeraire in the life years count case is no longer person, but the emphasis on years. The objective indicator is life data, which explains why an increased lifetime is an attractive outcome measure.

Although, with little adjustment, the survival rate and life years count commensurate in terms of mortality and facilitate comparisons across program areas. However, they are incommensurate in terms of morbidity as they neglect it completely. The morbidity aspect is very important because most clinical trials are related to morbidity. We can just attach some examples to these morbidity cases, such as- in any chronic disease that follows a detrimental progression; one would like to identify a specific diagnostic event. Hence, morbidity figures are essential.

Hence, some subsequent development and the need to assess both mortality and morbidity led to the widely accepted and more general concept in this case, 'health-related quality of life' (HRQOL). So, let us explore what HRQOL is and how the authors defined it. In definition, the HRQOL is defined as "the value assigned to the duration of life as modified by the impairments, functional status, perceptions and social opportunities that are influenced by

disease, injury, treatment or policy", mentioned in the writing of Patrick and Erickson (1993).

The measurement here, in this case, addresses health-related aspects. No specific health variable is given since the phrase is written as health-related. In other words, since the related word is attached, different countries use different descriptive systems, which are nothing but standard questionnaires that define the quality of health using various attributes and their respective levels. So, one point to be clarified is that these HRQOL models use different descriptive systems, and they follow their own standard questionnaires or attributes.

Different descriptive systems in healthcare specifically define health differently. This is because the healthcare cases are different, and their purpose is differently defined. So, the questions that are attached or the descriptive system that is attached in the healthcare again varies from country to country.

There are several thousands of such currently existing descriptive instruments. A searchable database for these instruments and associated literature is also available at the Big Data Institute of Oxford. You can just follow the link (there are some changes, so please try to find out).

However, the descriptive system can be classified into two categories. One is called 'condition specific', and another is called 'generic'. So far as condition-specific descriptive systems are concerned, these descriptive systems are the questionnaires designed to measure health within a particular condition or disease group, such as questionnaires like AQLQ (Asthma Quality of Life Questionnaire) by Juniper et al. (1993), another one is Montgomery-Asberg Depression Rating Scale (MADRS), that is specifically on depression rating scale and usually has a 10-dimensions or attributes and each described in 7 levels. We will discuss everything in a specific context while referring to QALY and DALY. Another index and specific method is the Oswestry Disability Index (ODI), which is specifically for back pain patients. Again, it has 10 dimensions, and there are 6 levels for each. These are widely used descriptive system questionnaires for their respective diseases. As I told you, different scales or indexes are applied for different morbidity indicators.

Although these are useful, condition-specific descriptive measures are inapplicable to aid resource allocation, especially in healthcare. For resource allocation, outcome measures should be able to: i) compare the differences in improved mortality, and ii) compare different types of improvement in morbidity, then iii) this should account for and compare the improvements in both, i.e., mortality and morbidity, and iv) valuation of improvement should be based on the preference of affected people. These are all insights that were a little difficult in the previous case. I think we will be using those aspects in the later models.

To meet all these four requirements, one needs descriptive systems that are more generic in terms of measurement (i.e., Generic descriptive systems). These descriptive systems are largely, the questionnaires designed to measure and compare health status across a range of

different dimensions belonging to different diseases (as I already mentioned). Here, we will give you the real examples of how authors applied and consider these questionnaires with their respective dimensions.

Here, we have mentioned as most famous descriptive systems-

Some of the most famous descriptive systems are listed below

Descriptive system	Country of origin	Dimensions	Levels	Health states (L^D)
EQ-5D (formerly EuroQol)	UK	Mobility, self-care, usual activities, pain/discomfort, anxiety/depression	3	243 e.g.: $3^5 = 243$
Quality of well-being (QWB)	US	Mobility, physical activity, social functioning 27 symptoms/problems	3 2	1170
SF-6D (derived from SF-36)	UK (US)	Physical functioning, role limitations, social functioning, pain, mental health, vitality	4–6	18,000
Health utilities index: (HUI-III)	Canada	Vision, hearing, speech, ambulation, dexterity, emotion, cognition, pain	5–6	972,000
AQoL (Assessment of quality of life; or Australian QoL)	Australia	Illness, independent living, social relationships, physical senses, psychological well-being; each consists of three sub-dimensions	4	16.8 m
15D	Finland	Mobility, vision, hearing, breathing, sleeping, eating, speech, elimination, usual activities, mental function, discomfort/symptoms, depression, distress, vitality, sexual activity	5	30,518 m

Worldwide used-
EQ-5D, HUI-III, SF-6D
Locally used-
15D (Finland)

Source:- Olsen, J. A. (2017)

We will also mention their health states and how many possibilities are possible. Here, we have mentioned EQ-5D (formerly known as EuroQol), whose country of origin is the UK. It has five dimensions: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. Each has three levels. So, this is what we referred to. We said mobility, for example, has three levels: first, second, and third.

Similarly, when we say self-care, we will find first, second and third level in the questionnaires. Maybe, the first and second are the extreme choices: not happy, happy or maybe neutral. We will give you the exact example here, we are carrying with the questionnaire. I will just show you (the questionnaire of EQ-5D from PPT). This is how the levels and the specific questions are defined. I will come to it. I am just presenting how one can think of different levels by different dimensions, and that is why I am just mentioning that sometimes you can go by different scales.

Another type of descriptive system is called 'quality of well-being' (QWB) by the US. Lets understand health states, especially using the EQ-5D case. So, in EQ-5D there are five dimensions and three levels. So, each dimension (case) comes three times, meaning each indicator has three choices. So, the possibilities of health states in a questionnaire worksheet would be 243. This is because it is 3 to the power 5 ($3^5 = 3 \times 3 \times 3 \times 3 \times 3 = 243$), i.e., 5 times you multiply, you will find all possibilities of 243). Similarly, we can compute for other cases, especially in Finland, where health states are considered bigger in terms of number (30,518 million).

When I say why these are all possibilities possible for health states, I will just show it to you with an example (using Euro-Qol). When I say- 'mobility', there are three indicators. When I say- 'self-care', there are three indicators. Likewise, each one has three indicators. For one single person, one might declare that I have no problems walking about. And, in another case, one might pick up this (indicating in ppt). In another case, the same person might pick up this, same person might pick up this, same person might pick up this. So, in that case, this is just one combination.

Similarly, out of all five, you will find a person is carrying so many combinations. So, each combination is presented, like every time it will be (since five types are there) 3 to the power 5. And in total, it accounts for a total of 243 health situations or states. Similarly, in the 15-dimension case of the Finland, it will be 5 to the power 15 (5^{15}). So, 5 to the power 15 is expected to be of this number (30,518 million). Other famous descriptive systems are mentioned as well; one is from Canada, and the other is from Australia. We have explained all these things here.

Now the question comes, is just having a generic descriptive system as a health state major enough for resource allocation or is there more to it? Certainly, there is more to it. Two problems with generic descriptive systems are- i) the generic system provides only states (i.e., quality of life) and ignores quantity (i.e., the duration). However, any meaningful health metric would have to include quality and quantity. ii) If the health state major simply consists of a set of dimensions that have different levels, it is simply termed health profiling or health profile.

Health profiles are better than condition-specific systems, as they allow inter-level comparisons (within a dimension). However, they do not provide comparisons between the dimensions. An example would be the analysis of the 'sickness impact profile', which you can refer from the paper of Bergner et al. (1981). Therefore, health profiles provide strong descriptive values, but the problem is profiles may not identify whether one health state is better than another.

We are citing the example here. If a researcher uses a generic descriptive questionnaire, say- EQ-5D (which I have already cited), that contains five dimensions starting from mobility, self-care, usual activity, pain or discomfort, and, anxiety or depression. Each contains three levels. These levels might be no problem, some problem, or extreme problem.

Here, if someone's health profile states- (21111), where, each number represents each level of an attribute. So, if the first aspect is related to mobility, that is picked up as 2, so the person has 'some problem' with mobility. As all other aspects or attributes are 1, that means the individual has no problem with other health attributes except mobility. So, (21111) refers to the person having (mobility) problems per his health profile.

If we have another case with another number, maybe- 22111, that means this individual is having some problem in his mobility and self-care, however he is good in other health

attributes. Hence, (22111) indicates that this (2nd person) has some self-care and mobility problems. Therefore, we can say that this person (2nd person) is worse off than the earlier individual with only mobility issues (i.e., 21111).

However, we cannot say that the given profile (21111) here, is better than another person who also has problem in one attribute but different attribute than the first individual. For example, suppose the first individual has an EQ-5D health profile (21111), while another person has a health profile (12111), Here. In that case, first individual has problem of mobility while another individual has some problem in self-care. So, these two cases cannot be compared by any magnitude. We cannot say which one is better because they have some qualitative reflection in different variables.

In such comparisons, an index is required to overcome this problem. The index will provide a singular value instead of a series of indicative information, as mentioned in different dimensions. Such an index will not only solve the two problems found in a generic descriptive system, but it will also fulfil four requirements needed for resource allocation. A sample of that EQ-5D is presented below. You can also go through the website, which we have suggested for reference. Otherwise, you can just search Google and you will find these details.

RATE OWN HEALTH: EURO-QOL

---

*Instructions: Please check the answer that best describes **your** current health today. (Please note that we are asking about your health and not the health of any of the above hypothetical cases given before.*

---

1. Mobility

- ○ I have no problems walking about
- ○ I have some problems walking about
- ○ I am confined to bed

2. Self-Care

- ○ I have no problems with self-care
- ○ I have some problems washing and dressing myself
- ○ I am unable to wash or dress myself

3. Usual Activities

- ○ I have no problems with performing my usual activities (e.g., work, study, housework, and family or leisure activities)
- ○ I have some problems with performing my usual activities
- ○ I am unable to perform my usual activities

4. Pain/Discomfort

- ○ I have no pain or discomfort
- ○ I have some pain or discomfort
- ○ I have extreme pain or discomfort

5. Anxiety/Depression

- ○ I am not anxious or depressed
- ○ I am moderately anxious or depressed
- ○ I am extremely anxious or depressed

One such index is 'quality-adjusted life years' (QALY). It will give you to have a singular commensurable value. Quality adjusted life years (QALY) is initially referred from the work of Late Alan Williams, who was a leading proponent of QALY and presented this sophisticated health measure. The initial developers (of QALY) were Fanshel and Bush (1970), and also Torrance (1970) (both developed it during same year). For any individual, there are three possible health gains. Either we say- improved HRQOL, or longer lifetime, or increased probability of survival. The rationale for this QALY, is that it account for all these three possibilities related to health gains. QALY will also account for these three possibilities in one currency or unit, making it commensurable.

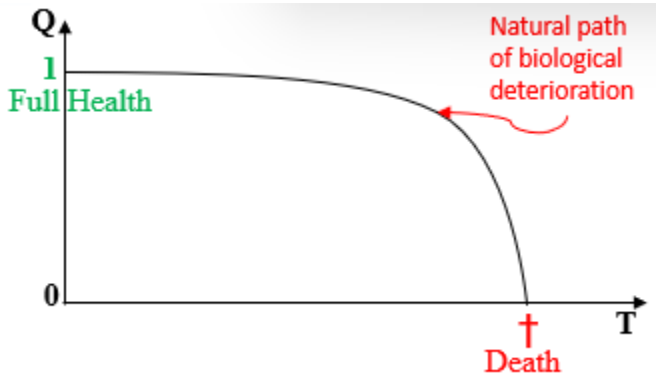
So, what are the steps for QALY? QALY, basically, counts for the product of $Q \times T$. Where Q stands for the 'quality of life' and T stands for the 'time in that health'. The steps we are just mentioning in part 1, part 2. We first calculate the 'length of time' (T), and then we calculate the 'quality of life' (i.e., Q). So, from the different descriptive systems, we will get the indicator and its profile. Then, we will select the appropriate non-monetary valuation method. Then, we will compute Q, and finally, we will use the obtained values of T and Q in the given formula. So, we will use it for the QALY index. While calculating T, depends on available data, Q calculation requires using different valuation approaches (those will be discussed later).

QALY is measured as a product of 'average health-related quality of life' (Q) and 'lifetime in that health' (T). Hence,

$$QALY = Q \times T$$

where,
QALY= Quality adjusted life years
Q= average health related quality of life ($\overline{HRQOL}$)→ **measured on a scale of 1 to 0**, &
T = Lifetime in that health → **counted in years**

Here, one year in full health refers to one QALY. Further, using both Q and T, QALY can be represented- using the expected life span.



This is what is presented in this diagram (you can just see). The Q ranges from 0 to 1. The T is in the horizontal axis, Q is in the vertical axis, and this is where we are mentioning the expected life span, till this ends with the time of death. The natural path of biological deterioration occurs after certain period of time. Although, every time, we will have some deterioration, but the major fall occurs at the later age. Hence, there has been a sharp fall. The shape of life span for an average person varies across countries, and the gains accordingly are measured through QALY.

Let us consider an intervention that improves the health state by ΔQ. Is it possible to be improved? Yes, of course, it is possible if you have some intervention either through 'quality of life' (i.e., Q) or through longevity (i.e., T).

So, if we are changing the ΔQ, it will be precise will be equal to $QALY_{G\text{-}Q}$, presented as change in or improvement in the QALY, where G refers to gain due to policy change. Let us move further. This is how it improves when we have given the T (T is not changing when we are emphasising or intervening through the change in Q):

$$\boldsymbol{QALY_{G-Q} = \Delta Q \times T}$$

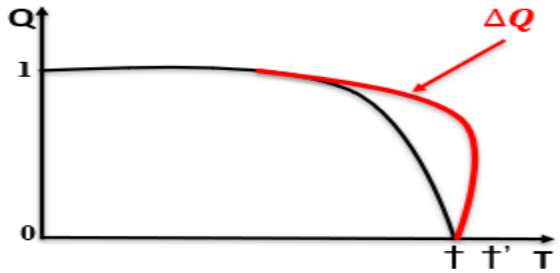


Fig.-Expected Life span

So, our life span might be shifted and improved (that is what ΔQ implies). Alternatively, we can improve longevity through change in time (in terms of lifetime). So, in that case, it will be ΔT. Hence, there will be a paradigm shift in terms of the lifespan, and we might find better possibilities, that is corroborated by the change in T.

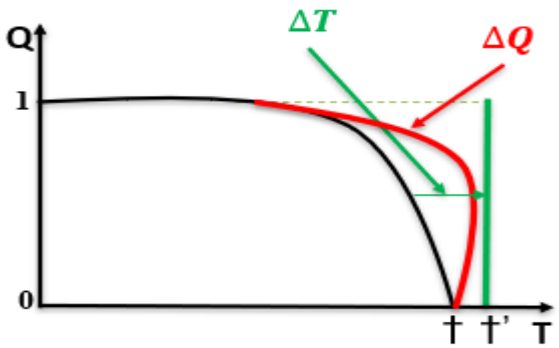
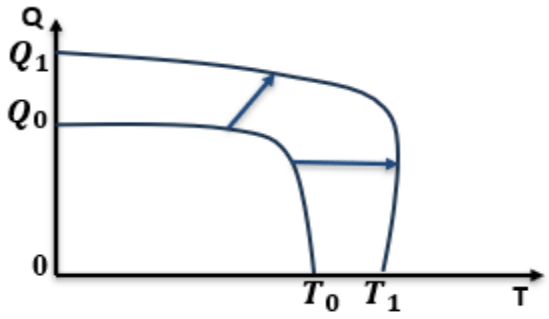


Fig.-Expected Life span

$$\boldsymbol{QALY_{G-T} = Q \times \Delta T}$$

Hence, what we are concluding from this example is that- a program that improves health quality (i.e., ΔQ) is in terms of morbidity (health quality in terms of morbidity), becomes commensurable in terms of mortality (i.e., ΔT). So, ΔQ becomes commensurable in terms of health outcomes with a program that improves longevity (i.e., ΔT).

There are also interventions that improve both ΔT and ΔQ as explained in the diagram (both quality and length of life).



This can be formulated as:

$$\boldsymbol{QALY} = (\boldsymbol{Q_1} \times \boldsymbol{T_1}) - (\boldsymbol{Q_0} \times \boldsymbol{T_0})$$

where,

$\boldsymbol{Q_0}$= $\overline{HRQOL}$ without treatment

$\boldsymbol{Q_1}$= $\overline{HRQOL}$ with treatment

$\boldsymbol{T_0}$= Lifetime without treatment

$\boldsymbol{T_1}$ = Lifetime with treatment

We are giving an example here for your better clarity. Suppose an individual (aged 50) whose life expectancy is 65 and has a health issue that keeps his health state below full health (let say- 0.7) for 5 years. After 5 years, his health state deteriorates further to a level of 30 percent (or 0.3) for the next 3 years before succumbing to death. So, then, let us calculate QALY. So, in that case, simply $QALY = Q \times T$ (Q times T) and we have already mentioned that it is in the fraction 0 to 1 (Q varies from 0 to 1) and T in the absolute years of lifespan. So, if we compute the initial situation, we get-

$QALY = Q\ x\ T => 0.7\ x\ 5 = 3.5\ QALYs$

Here, we see that due to a health state of 0.7, the next 5 five years equals only 3.5 QALY (quality-adjusted life years). In other words, 5 years with health issues, equals 3.5 quality adjusted life years.

Then, after 5 years his health deteriorates to 0.3 (i.e., Q after 5 years is 0.3), so on computation, we get-

$QALY\ =\ Q\ x\ T\ =\ 0.3\ x\ 3\ =\ 0.9\ QALYs$

So, for the next 3 years with 0.3 Q equates to 0.9 QALYs. So, for this entire 5 +3 years (that is 8 years), adding these two is 3.5+ 0.9= 4.4 QALYs. Hence, it will be of 4.4 years. This is one clarification we have given; you can calculate and know how authors calculate QALY. I think it will be useful for you.

I am just giving another example for you to calculate. We are also specifying the answers for you. That is, suppose after the first 5 years of health issue (from the same example), the individual finally decides to consult a doctor. Now, we are taking an intervention, and this intervention has resulted in an increase in his health status to 0.9. So, instead of 0.3, now 0.9 for the remaining time or his expected life. Calculate the new QALY and the total QALY. The answer would be actually 12.5 for total QALY(you can just see). Compute the new QALY on your own, using the same formula.

We are now discussing the theory to practice. It all seems fine in theory, but how do we obtain reliable figures to put into this formula in practice? To answer this, we need to understand the steps in calculating QALY. As stated in part 1, we calculate T (i.e., length of life), where data is normally obtained from mortality tables or the survival rates. While operational mortality rates might give the probability of success outcomes, the measurement of T enables comparison on an interval scale (e.g., 10 = 2 x 5).

Similarly, we need Q that comes in part 2, to be measured on with interval properties. This requires using the descriptive systems. And, for the choice of system, economists usually go with the one recommended by their national evaluation guidelines. For example, researchers in UK prefer EQ-5D as the descriptive system. It is the most important descriptive system per the National Institute for Health and Care Excellence (NICE) guidelines. So, you can also prefer EQ-5D, which has 5 dimensions and 3 levels each.

Once the generic descriptive system is decided, we will receive the health state indicators or health profile. Then, the next step is to determine how do we measure and value various combinations of health (such as the non-monetary values) with the selected generic health systems. There are several alternative techniques available to do so. Moreover, like the descriptive systems and monetary valuation techniques (like revealed preference and stator preference), these (the non-monetary health valuation techniques) are also questionnaire-based methods. Here, using a particular technique, respondents are asked to place a value on one or more health states.

As these techniques are interesting and useful, hence we will be touching upon this in our next lecture. So that is all for the lecture. Let me summarise it (in short). In this lecture, we discussed why willingness to pay using revealed preference (RP) and stated preference (SP) were rejected as a health measure. You can also prepare for your assignment questions. We discussed the non-monetary valuation of health. From health status as a measure, we can look at their history, development and their types. We discussed incommensurable and commensurable outcomes, health descriptive systems, their calculations, limitations and even their profile measurement, which gives estimation and limitations etc. We also touched upon the discussion of QALY, their objective, methods and examples.

In the following lecture, we will discuss - health state valuation techniques for computation of QALY and the alternatives to QALY such as DALY and HLYE etc. will be emphasised. So, these are the readings. So, I suggest, please follow as there will be a number of questions derived from these units. With this, I must thank you for attending the lecture. Thank you.