Patient-Reported Outcomes

Principles of Measurement and Applicability in Economic Evaluation

Tamás Ágh PhD - András Inotai PhD

ELTE
Faculty of Social Sciences

Patient-Reported Outcomes

Principles of Measurement and Applicability in Economic Evaluation

Tamás Ágh PhD – András Inotai PhD

Authors of the book: Tamás Ágh PhD, András Inotai PhD

Reviewed by Gábor Kovács MD

PATIENT-REPORTED OUTCOMES: PRINCIPLES OF MEASUREMENT AND APPLICABILITY IN ECONOMIC EVALUATION

Tamás Ágh PhD – András Inotai PhD

Contents

Preface

Measuring health outcomes of a clinical intervention received by the patient have been gaining increasing interest in the last few decades, since societies aim to invest their limited healthcare budget into good value-for-money treatments. Although the traditional diagnostic approach relies on measuring physical, physiological or biochemical data of the patient, such as blood cholesterol level or blood pressure; it is not sufficient to collect data about only these parameters. Subjective data - i.e. information from the patient's perception of his own point of view - can only be obtained directly from the patient. In the modern medical practice both objective clinical data and subjective information on a patient's health condition are important to enhance disease management and improve clinical outcomes. The perspective of the patient and the physician maybe different as well as complementary, therefore both are necessary for a complete picture.

A patient reported outcome (PRO) is any report of the status of a patient's health condition that comes directly from the patient, without any interpretation of the patient's response by a clinician or anyone else. The field of PROs, which is one of the fastest-growing areas in health outcomes research, can be described as an interdisciplinary science that involves psychometrics, biostatistics, and health-economics as well.

This textbook is a part of the blended learning material prepared for the Patient Reported Outcomes module at Eötvös Loránd University (ELTE), MSc in Health Policy, Finance and Analysis. This textbook has been designed to support the improvement of student knowledge through self-study. In addition to introducing the core concept and measurement methods of PROs, including quality of life, adherence and patient satisfaction, this blended learning material also walks students through the health economic and health policy relevance of measuring PROs. Key messages are summarized at the end of each topic. To fully benefit from the textbook, students are encouraged to complete the self-check questions and read the compulsory and suggested readings for each section.

The authors are grateful to Gábor Kovács who reviewed and commented on the textbook at various stages and improved it, to Kirk Lederhaas for help with proof-reading, and to Zsuzsanna Kun who coordinated the production of the final textbook. The authors are also thankful to Prof. Zoltán Vokó for the opportunity to prepare the blended learning materials for the Patient Reported Outcomes module at ELTE, MSc in Health Policy, Finance and Analysis.

<div align="right">Tamás Ágh PhD and András Inotai PhD</div>

1. Module overview and introduction to PROs

Paradigm shift in medicine - background

There has been a paradigm shift in healthcare recently. According to Socrates, 'We should set the highest value not on living but on living well'. This means that besides increasing life longevity (quantity of life), improving quality of life is also becoming more and more important. This is even highlighted by the positive definition of health, proposed by the World Health Organization (WHO) in 1948: 'Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity' [1]. Why is the patient's perception of their own health state becoming more and more relevant in healthcare decision-making?

- First, the disease structure is changing. The acute, high mortality diseases are associated with a smaller incidence than a century ago, but parallel with the increase in life expectancy, the prevalence of chronic conditions with disabilities has grown. One may say, that patients, who died from the plague, diphtheria, or other devastating diseases a century ago, now live long enough to develop prostate cancer, rheumatic conditions, osteoporosis and many other diseases. Many of these conditions can now be treated, but also have a great impact on a patient's quality of life [2, 3]. In these conditions even a small deterioration in the patient's subjective perception of his/her own health status may be of great health loss, since these conditions affect life-length [4]. According to Lincoln: 'Live a good life. In the end it is not the years in a life, but the life in the years.' Valuing health states based on patient reported outcomes (PROs) is relatively new in the field of medicine. Psychometrics is a discipline in psychology that assists the measurement of PROs.
- The second key factor is a paradigm shift in medicine. The knowledge a patient has of his/her own health state is improving. There are many sources for the patient to access for easy-to-understand information, from the Internet or patient organizations to printed lay media. Patients are familiar with new treatment options already available in high income countries. Therefore they are acting as informed customers, not as patients, and consequently their role in their own medical treatment is increasing. Whereas the general practitioner's (GP) workload is very high and the time they have for their patients is restrictive and they only have limited time to keep up-to-date with recent scientific findings and best practices. It is challenging for the GP to encounter the patient who acts as an informed customer - since GPs are not trained to handle such situations. As a consequence, the patient-physician relationship and interaction are becoming more balanced, contrary to the previous paternalistic approach.

This can also be recognized in the improved terms and definitions of patient's adherence to the drug therapy.

- The third key factor is the different perception and perspective of the patient and physician on the patient's health state [5]. The physician generally relies on his/her own observation, which is based on physiological parameters, such as blood pressure, serum cholesterol level. On the other hand, patients are interested in their own perceptions, symptoms of the condition which influence their quality of life, adverse events of the medication, which in the end may affect their satisfaction with the treatment and their adherence to the medication. There are different methods for patient observation, from objective biologic parameters to clinician reported outcomes (ClinRO), caregiver reported and PRO (Figure 1.1). Certainly, the latter are associated with the highest subjectivity. These are not just clinical outcome assessment (COA) approaches from different perspectives, but also complementary and thus all of them are necessary to have a complete picture of the patient's health state.

Clinical Outcomes Assessment (COAs)

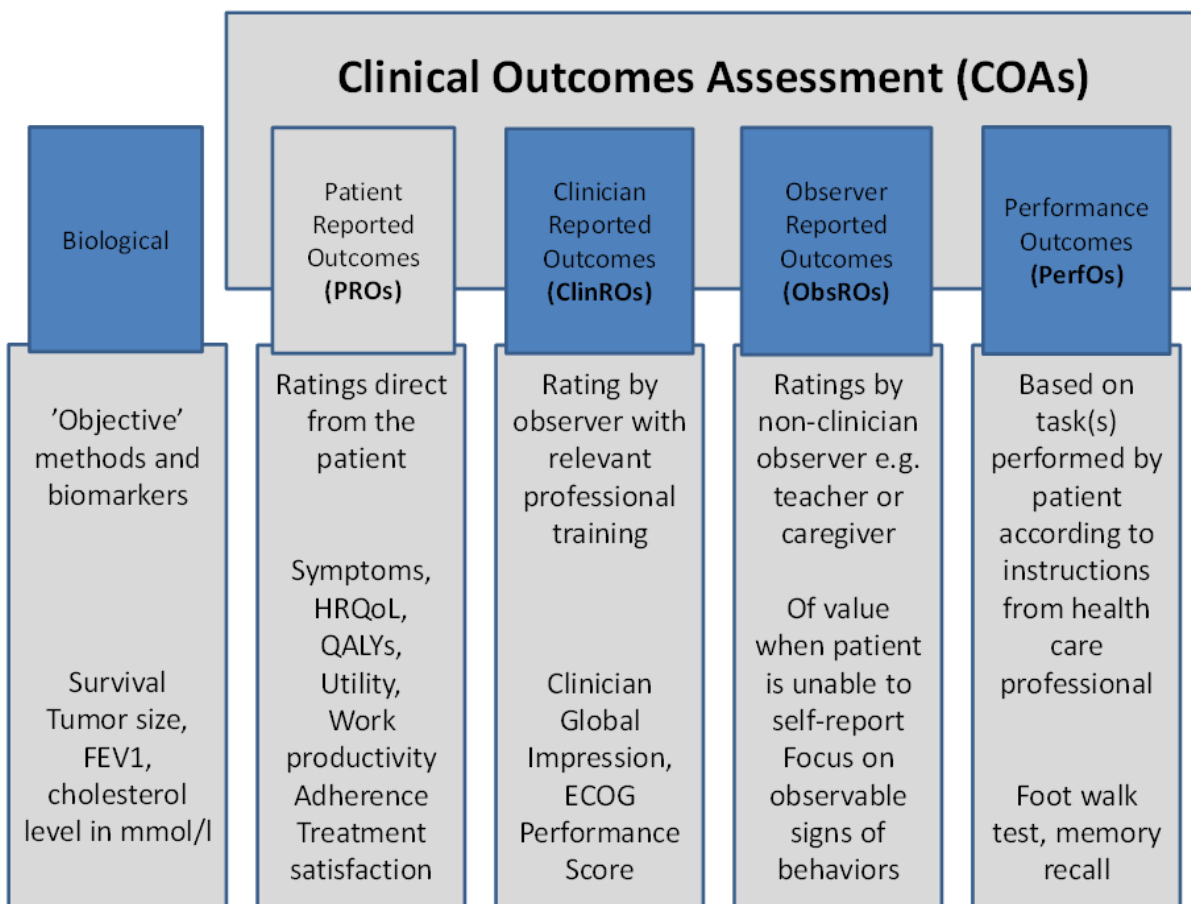| Biological	Patient Reported Outcomes (PROs)	Clinician Reported Outcomes (ClinROs)	Observer Reported Outcomes (ObsROs)	Performance Outcomes (PerfOs)
'Objective' methods and biomarkers	Ratings direct from the patient	Rating by observer with relevant professional training	Ratings by non-clinician observer e.g. teacher or caregiver	Based on task(s) performed by patient according to instructions from health care professional
Survival Tumor size, FEV1, cholesterol level in mmol/l	Symptoms, HRQoL, QALYs, Utility, Work productivity Adherence Treatment satisfaction	Clinician Global Impression, ECOG Performance Score	Of value when patient is unable to self-report Focus on observable signs of behaviors	Foot walk test, memory recall

Figure 1.1. Methods for patient observation – sources of health outcomes [6]

According to the guidance of the United States (US) Food and Drug Administration (FDA) a 'PRO is any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else. The outcome can be measured in absolute terms (e.g., severity of a symptom, sign, or state of a disease) or as a change from a previous measure. In clinical trials, a PRO instrument can be used to measure the effect of a medical intervention on one or more concepts (i.e., the thing being measured, such as a symptom or group of symptoms, effects on a particular function or group of functions, or a group of symptoms or functions shown to measure the severity of a health condition)' [7].

Module objectives

This module is going to introduce the term of PROs. As an umbrella term, PROs contain many more terms, including quality of life (QoL), adherence, and patient satisfaction [8], which will also be discussed during the module. Students will be introduced to how we measure different PROs. Advantages and disadvantages of using interviews and surveys are going to be discussed. Also, different classification systems of PRO/QoL measures will be introduced, including generic and specific instruments, index and profile type measures [9]. Students will be shown how PRO/QOL measures are developed. Since these instruments need to measure exactly what they were designed for, they have to fulfill strict psychometric criteria, and they need to be validated [10]. Students will be introduced to key evaluation criteria of PRO/QoL measures [10, 11]. Since development and validation of a new instrument is time and resource consuming, for lower income countries with limited resources, the cultural adaptation and translation of existing measures could be the feasible approach. Locally adapted measures, however, also need to be validated. Measuring PROs in special patient groups, including children will also be discussed in the module [12]. Further PRO terms, such as definition and measurement of medication adherence and patient satisfaction will also be introduced [13].

In the second part of the module, the health economic and health policy relevance of measuring PROs will be discussed. A lecture will introduce the heterogeneity of health outcomes related to a wide spectrum of health technologies to be measured and compared. To address these challenges and to meet economic requirements of health outcome measures, the Utility construct has been developed [14]. These are index type scores with reference points within the instrument to death (0.0) and perfect health (1.0) and may be used to combine changes in quality and quantity of life due to the possibility of linking, comparing and trading off these different aspects. Students will be shown how to measure utility with direct and indirect measures [15]. Students will also be

shown how utility is applied in the quality adjusted life year (QALY) concept [16]. QALY is the universal health outcome of cost-utility analysis, the most widely applied type of economic analysis used to inform healthcare decision-makers [17]. Of course, QALY is far from being unproblematic, considering that in the critique literature there are a wide spectrum of critique references on ethical, methodological and conceptual grounds [18, 19]. An important objective of the module is to help students acquire a balanced view on the advantages and disadvantages of the use of QALY in healthcare decision-making from different perspectives.

Cost utility analyses are not capable of prioritizing among different disease areas. Burden of disease (BoD) studies are suitable for identifying unmet needs and vulnerable patient groups in healthcare [20]. Students will be introduced to the two core components of BoD studies, the health loss and economic burden (cost of illness). QALY is not an appropriate construct to measure health loss in burden of disease studies; therefore the WHO applies disability adjusted life years (DALY) in the global burden of disease studies [20]. Students will be introduced to the term DALY, and its policy relevance. The use of DALY has been criticized even in BoD studies [21]. The QALY and DALY construct will be compared in the module; similarities and differences will be discussed. Different utility measures may result in different utilities – even when applied to the same patients with the same health state [22]. These differences may influence health policy decision-making. Therefore students will be introduced to the implications of transferability of health outcomes among different countries. Specific measures are very often not suitable for measuring utility, consequently their capacity to be used in healthcare decision-making is limited. Mapping enables specific measures to estimate utility. Therefore the concept of mapping will also be introduced. The core education material of the module is going to be supplemented and substantiated with small group exercises.

References

1.	Constitution of the World Health Organization. In: World Health Organization: Basic documents. 45th ed. Geneva: World Health Organization; 2005.

2.	Inotai A, et al. (2012) Health related quality of life and utility in patients receiving biological and non-biological treatments in rheumatoid arthritis. Rheum Int, 32: 963-969.

3.	Ágh T, et al. (2011) Factors associated with medication adherence in patients with chronic obstructive pulmonary disease. Respiration 82: 328-334.

4.	Inotai A, et al. (2012) Quality of life, utility and health burden in asthma, chronic obstructive pulmonary disease and rheumatoid arthritis. International Journal of Person Centered Medicine, 2: 505-510.

5.   Jachuck SJ, et al. (1982) The effects of hypotensive drugs on the quality of life. J Roy Coll Gen Pract, 32: 103–105.

6.   FDA. Clinical Outcome Assessment (COA): Glossary of Terms. http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualification Program/ucm370262.htm [Accessed online: 22th October, 2015]

7.   FDA. (2009) Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims http://www.fda.gov/downloads/Drugs/.../Guidances/UCM193282.pdf [Accessed online: 22th October, 2015]

8.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Patient Reported Outcomes (PRO). ISPOR, USA, 2003: 175-176.

9.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Health Related Quality of Life (HRQoL). ISPOR, USA, 2003: 129-131.

10.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Validity. ISPOR, USA, 2003: 246-248.

11.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Reliability. ISPOR, USA, 2003: 210-212.

12.   Matza LS, Swensen AR. (2004) Assessment of Health-Related Quality of Life in Children: A Review of Conceptual, Methodological, and Regulatory Issues. Value in Health, 7(1): 79-82.

13.   Cramer JA, et al. (2008) Medication compliance and persistence: terminology and definitions. Value in Health, 11(1): 44-47.

14.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Utility. ISPOR, USA, 2003: 241-242.

15.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Utility measurement. ISPOR, USA, 2003: 242-244.

16.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Quality Adjusted Life Year (QALY). ISPOR, USA, 2003: 195-197.

17.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Cost-Utility Analysis. ISPOR, USA, 2003: 45-48.

18.   Wiliams, A. (1996) QALYS and ethics: a health economist's perspective. Soc sci med, 43(12): 1795-1804.

19.   Harris J. (1987) QALYfying the value of life. J Med Ethics, 13: 117-123

20.   Murray CLJ. (1994) Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bulletin of the World Health Organisation, 72 (3): 429-445.

21.   Lyttkens CH. (2003) Time to disable DALYs? On the use of disability-adjusted life-years in health policy. European Journal of Health Economics, 4(3): 195-202.

22.   Nord E, et al. (2009) QALYs: Some Challenges. Value health, 12(Suppl1): S10-5.

2. Concept of PRO measurement

Data collection methods

PRO data collection can be done either using interviews or questionnaires. Both methods have their own positive and negative aspects.

Interviews involve direct verbal questioning of participants where private aspects of behavior can be gathered. Structured interviews allow for replication of the interview with others. A structured interview has a rigorous structure whereby all interviewers ask the same questions from all participants. Semi-structured interviews are conducted with a fairly open framework where not all questions are phrased beforehand; many questions are created during the interview which allows for more flexible communication. Interview questions can be either questions with multiple-choice answers or open questions which have no given answers. The important thing to remember when choosing this method is that interviewers may misinterpret the information received from the participant, which can lead to bias.

Written questionnaires can be given to a large number of people simultaneously, thus it is likely to be less costly compared to interviews, particularly in terms of the time spent collecting the data. Another advantage of this method is that the format of the questionnaire is standard for all participants and is not dependent on the mood of the interviewer; therefore, questionnaires can collect more reliable information for PRO research.

It is important to highlight the difference between a PRO questionnaire and a PRO instrument. A PRO questionnaire refers to a specific questionnaire used to collect PRO data. However, a PRO instrument has a wider meaning. The US FDA defines a PRO instrument as follows [1]: 'A means to capture data (i.e., a questionnaire) plus all the information and documentation that supports its use. Generally, it includes clearly defined methods and instructions for administration or responding, a standard format for data collection, and well documented methods for scoring, analysis, and interpretation of results in the target patient population.'

Types of PRO instruments

PRO instruments can be classified as per the measurement of the interested concept (horizontal-vertical coverage) as generic or specific PRO instruments.

Generic PRO instruments are designed to measure very broad aspects of health. The major advantage of generic PRO instruments is that these are suitable for a wide range

of patient groups and thus allow for a comparison across conditions as well. Moreover, generic instruments can be used to generate normative data for the general population. The major limitation of these instruments is that they may not be sensitive enough to detect the condition-specific, small changes of the given PRO.

The other type of PRO instruments is the specific PRO instruments. These instruments can be specific to a disease (e.g., diabetes, asthma), population (e.g., children, elderly) or function, symptom or problem (e.g., disability in daily living, pain). Disease specific PRO instruments are developed to measure the patient's perceptions of a specific disease or health problem (e.g., St George's Respiratory Questionnaire in COPD). Population specific PRO instruments are designed to be appropriate for particular demographic groups (e.g., Child Health and Illness Profile-Child Edition). Problem specific PRO instruments are developed to assess one particular aspect of health status (e.g., Beck Depression Inventory).

Disease specific PRO instruments address particular diseases; therefore they are sensitive to clinically important changes in health in their target disorders. The major disadvantage of disease specific PRO instruments is that they cannot be used in samples which do not have the relevant health problem. Therefore, disease specific PRO scores cannot be compared with those for the general population or patients with other disorders.

The pros and cons of generic and disease specific PRO instruments are summarized in Table 2.1.

	Pros	Cons
Generic instrument	Can be used with healthy populations to generate normative data Comparison across conditions possible	May not focus adequately on area of interest May not be responsive
Disease specific instrument	Clinically sensitive May be responsive	Does not allow for cross-condition comparison

Table 2.1. The pros and cons of generic and disease specific PRO instruments

A PRO instrument can be categorized according to the type of the result presentation as well. Index type PRO instruments (e.g., EuroQoL 5 dimensions (EQ-5D)) express health states as a single index score. In contrast, profile type PRO instruments (e.g., Nottingham Health Profile) provide a profile score for each dimension of health state, thus they can be used to detect differential effects on different aspects of health status.

Based on the complexity of the PRO instruments we can differentiate single-, or multi-item and single-, or multi-domain instruments. An item can be defined as "an individual question, statement, or task (and its standardized response options) that is evaluated by the patient to address a particular concept [1]." Domain is 'a sub-concept represented by a score of an instrument that measures a larger concept comprised of multiple domains; domains can be subdivided into items [1].' The relationship between the questionnaire or items in a PRO instrument and the concepts measured can be shown in a conceptual framework diagram (Figure 2.1).
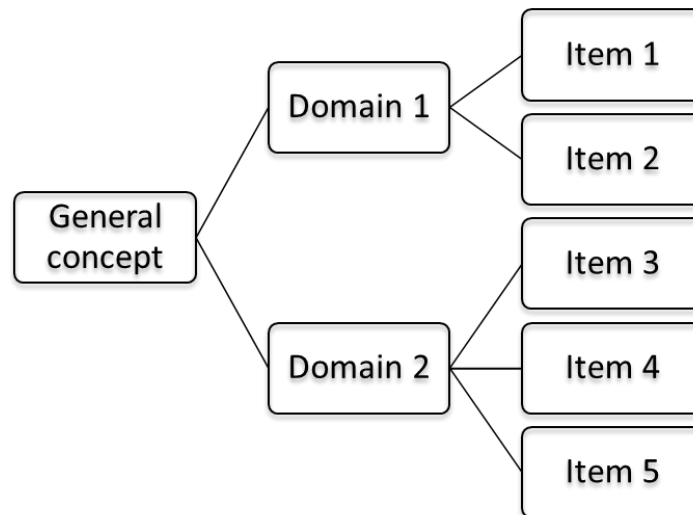
Figure 2.1. Conceptual framework

A simple concept can be assessed with a single-item scale on a particular concept of interest (the thing that is measured by the PRO), whose score is estimated by a single response to a single question (e.g. pain). Multi-item scales are scales formed by more than one item and are used chiefly for concepts considered complex to measure (e.g., physical function). Where multiple items are needed and they are referring to different aspects of the same concept or attribute, they form multiple-domain instrument (e.g., SF-36) [2].

Types of data and measurement scales

There are four different types of measurement scales: (1) nominal, (2) ordinal, (3) interval, and (4) ratio scales [2]. The four scale types are ordered in that all later scales have all the properties of earlier scales plus additional properties.

Nominal scales satisfy only the identity property of measurement, are used for labeling variables without any quantitative value [2]. Nominal scales are qualitative scales. For

example gender can be measured with a nominal scale. Individuals may be classified as male or female, but neither value represents more or less than the other. A sub-type of nominal scale with only two categories is called "dichotomous".

Ordinal scales have the property of both identity and direction [2]. With ordinal scales, it is the order of the values that is important, but the differences between the values are uninterpretable. An example for an ordinal scale is how the subject feels today (unhappy, OK, happy). It is important to note that the median and the mode are the only measures that can be used to assess the central tendency on a set of ordinal data.

The third category of measurement scales is the interval scales. Interval scales have the properties of identity, direction, and equal intervals [2]. Interval scales are numeric scales (quantitative scale) in which we know not only the order, but also the exact differences between the values. Interval scales do not have a "true zero" point. Many PRO scales are assumed to be interval in practice [2]. Celsius temperature scale is such an example. This scale is made up of equal temperature units, the difference between 10°C and 20°C is equal to the difference between 20°C and 30°C. Moreover, the Celsius scale does not have an absolute zero point. The zero on the Celsius scale was arbitrarily selected as the temperature point, representing the temperature when water is freezing.

Ratio scales satisfy all four of the properties of measurement: identity, direction, equal intervals, and an absolute zero [2]. Ratio scales tell us about the order, they tell us the exact value between units, and they also have an absolute zero point. Absolute zero point can be defined as a point where none of the qualities being measured exist. An example for the ratio scale is how many hours a day the subject spends on a computer.

Likert scales and visual analogue scales (VASs) are special scales commonly used in PRO questionnaires.

A Likert scale is generally a five (or seven) point scale which uses fixed choice response formats and are designed to measure attitudes. Well-designed Likert scales exhibit "symmetry" and "balance". "Symmetry" means that the scale contains equal numbers of positive and negative positions whose respective distances apart are bilaterally symmetric about the "neutral" value. "Balance" means that the distances between each candidate value are the same, which allow quantitative comparisons across items containing more than two candidate values. Examples for Likert scales items are:

- Agreement: strongly agree / agree / undecided / disagree / strongly disagree;
- Frequency: very frequently / frequently / occasionally / rarely / never;
- Likelihood: almost always true / usually true / occasionally true / usually not true / almost never true.

VAS is also a psychometric response scale. Respondents specify their level of agreement to a statement by indicating a position along a continuous line between two end-points. The continuous aspect of the VAS differentiates it from discrete scales such as the Likert scale. The simplest VAS is a straight horizontal line of fixed length, usually 100 mm. The ends are defined as the extreme limits of the parameter to be measured (e.g., symptom, pain) orientated from the left (worst) to the right (best).

References

1. Food and Drug Administration. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009

2. Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

3. Development of PRO instruments

Background: PRO instrument vs. PRO concept

A PRO instrument is a means to capture concepts related to the health experiences of individuals - how patients feel or function in relationship to their disease, condition, or treatment- and is designed to collect data about a PRO concept. A PRO instrument is not only a questionnaire but also entails all the information and documentation that support its use. It includes clearly defined methods and instructions for administration or responding, a standard format of collection, and well-documented methods for analysis, and the interpretation of results [1, 2].

In contrast, a PRO concept is what a PRO instrument measures. The PRO concept represents the aspects of how patients function or feel related to a health condition or its treatment [1, 2].

Developing a new PRO instrument

In the US FDA Guidance on PRO measurement, five iterative steps of PRO development are distinguished: (1) hypothesize conceptual framework, (2) adjust conceptual framework and draft instrument, (3) confirm conceptual framework and assess other measurement properties, (4) collect, analyze and interpret data, and (5) modify instrument (Figure 3.1).
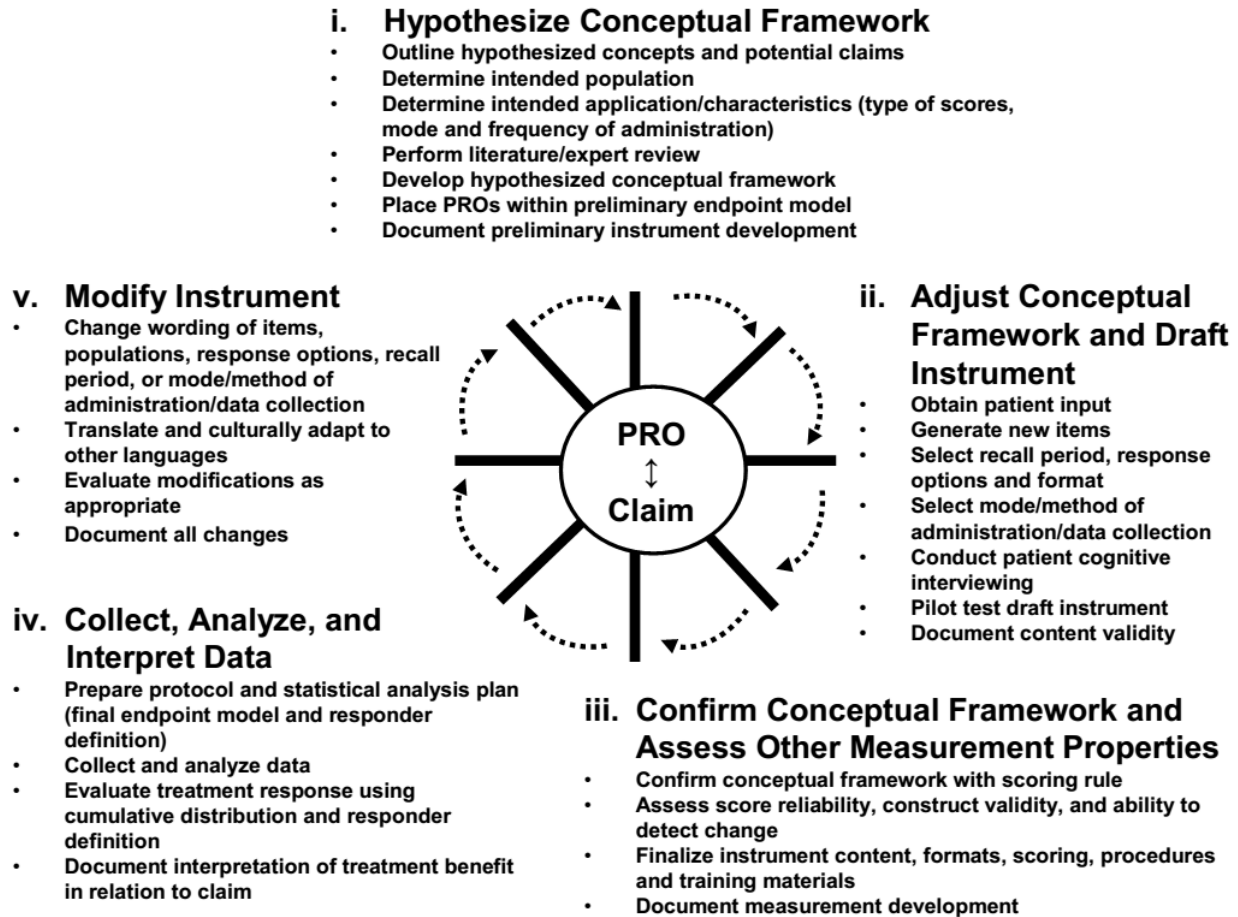
i. Hypothesize Conceptual Framework

- Outline hypothesized concepts and potential claims
- Determine intended population
- Determine intended application/characteristics (type of scores, mode and frequency of administration)
- Perform literature/expert review
- Develop hypothesized conceptual framework
- Place PROs within preliminary endpoint model
- Document preliminary instrument development

v. Modify Instrument

- Change wording of items, populations, response options, recall period, or mode/method of administration/data collection
- Translate and culturally adapt to other languages
- Evaluate modifications as appropriate
- Document all changes

PRO ↕ Claim

ii. Adjust Conceptual Framework and Draft Instrument

- Obtain patient input
- Generate new items
- Select recall period, response options and format
- Select mode/method of administration/data collection
- Conduct patient cognitive interviewing
- Pilot test draft instrument
- Document content validity

iv. Collect, Analyze, and Interpret Data

- Prepare protocol and statistical analysis plan (final endpoint model and responder definition)
- Collect and analyze data
- Evaluate treatment response using cumulative distribution and responder definition
- Document interpretation of treatment benefit in relation to claim

iii. Confirm Conceptual Framework and Assess Other Measurement Properties

- Confirm conceptual framework with scoring rule
- Assess score reliability, construct validity, and ability to detect change
- Finalize instrument content, formats, scoring, procedures and training materials
- Document measurement development

Figure 3.1. Development of a PRO instrument [2]

Instead of presenting all steps in the development of a PRO instrument, only its key steps will be discussed. For a more detailed description of the development please see the FDA guidance [2] and the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Good Practices [1, 3].

The key steps in the development of a PRO instrument are:

- To determine goals,
- To determine the context of use,
- To generate items,
- To reduce and word items,
- To conduct cognitive interviews for evaluation of content validity.

The development of a PRO instrument starts with the identification of the key goals. These targets should be considered throughout the instrument development process.

The next step is to determine the context of use. This includes an understanding of the disease or condition in the target population and considerations related to specific aspects of the target population. It is important to understand the pathophysiology of the disease, the potential risk factors, the process of the diagnosis, the signs and symptoms, and the possible effects and side-effects of the treatment. Literature as well as clinical experts can help to understand the disease. A hypothesized disease model can help us organize and visualize the key features of a disease including PRO and non-PRO outcomes (Figure 3.2).
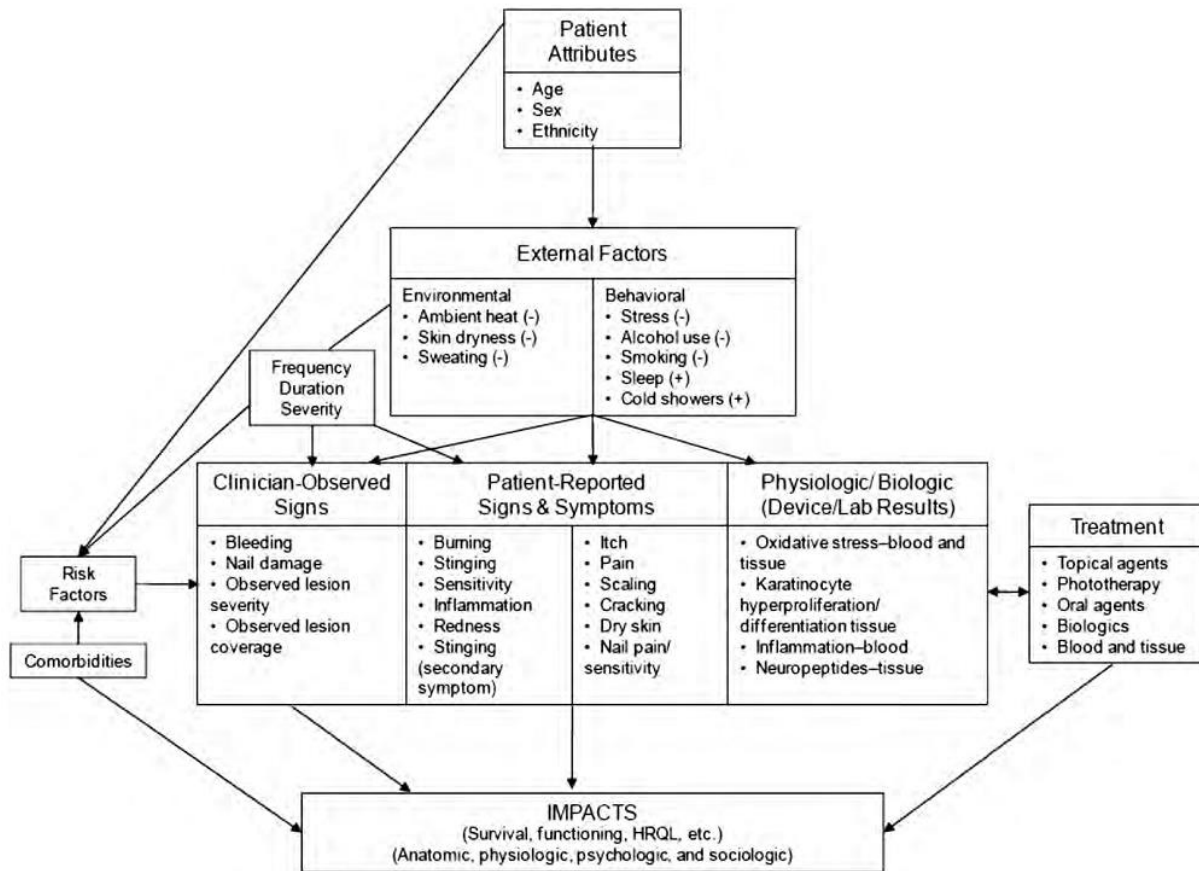
Figure 3.2. Hypothesized disease model for psoriasis for use in discussing context of measurement for the disease, target population, and treatment [1]

During the development of an instrument thought should be given to the characteristics of the target population, including language, culture, age and other characteristics of subjects, such as cognitive functions.

The third key step in the development of an instrument is the item generation. In order to generate the content of a PRO instrument two qualitative data collection approaches

can be used: individual interviews and/or group discussions with patients (typically called focus group). Generally, a mixture of these two methods is beneficial. The major advantages and disadvantages of individual interviews and focus groups are summarized in Table 3.1.

	Focus groups	Interviews
Pros	Rich source of data Allows the use of ideas from others as clues to express their own views Participants can compare their experiences with others Able to reach many participants at once	Get more in-depth and detailed information about an individual's experience Can be useful for sensitive topics Data can be easier to analyze Scheduling can be easier
Cons	Data can be tough to analyze because the discussion will also contain reactions to the comments of other group members Moderators need to be highly trained and able to lead the group One strong group member can sway the tone of the entire group	It may take longer to collect data Limited to one participant's view at a time Interviewers need to be trained with excellent communication skills May be more costly

Table 3.1. Pros and cons of individual interviews and focus groups [1]

Individual interviews are ideal for concepts that are sensitive and where subjects are unlikely to share information in a group setting. Focus groups, on the other hand, can stimulate discussion of topics and comparison of experiences across participants that cannot be captured in individual interviews. However, focus groups also have disadvantages. For example, one strong group member can sway the tone of the entire group. When such a participant dominates, the collected data do not necessarily represent the group as a whole. After analyzing the qualitative data captured with interviews and/or focus groups, the items and the conceptual framework of the instrument can be developed. The conceptual framework is an explicit description or diagram of the relationships between the questions or items in a PRO instrument and the concepts measured [2].

The fourth key step is the item reduction. Items are generated from the language used by subjects from the individual interviews or focus groups. Numerous items are

formulated per concept/domain, often with significant overlap in wording, and it is not always clear which terminology is the most appropriate. Further interviews can help developers choose between the different options [4]. At all times, items should be worded carefully and clearly. Double negation must be avoided. The character size is also important since small characters should be avoided and using effects for highlighting words (e.g., bold, italic, underline) can improve the readability of the items.

The final key step of the development of a PRO instrument is to conduct cognitive interviews for the evaluation of content validity (please find below more information on content validity) [3]. During a cognitive interview patients are asked to complete the newly developed draft questionnaire, and while doing so, they are instructed to share what they are thinking and to explain how they are interpreting the content of the measure. Individual interviews are recommended rather than focus groups. If there are issues with any part of the PRO, the interviewer can ask how they would reword a question to make it clearer. Once the content of the questionnaire has been confirmed via the cognitive interviews, the draft is ready for psychometric testing.

Validity and reliability

Validity assesses the extent to which an instrument measures what it is meant to measure. Validity has three main types: (1) content, (2) construct, and (3) criterion validity [4]. Content validity demonstrates that the instrument measures the concept of interest including evidence that the items and domains of an instrument are appropriate and comprehensive relative to its intended measurement concept, population, and use [2]. Construct validity provides evidence that relationships among items, domains, and concepts conform to a priori hypotheses concerning logical relationships that should exist with other measures or characteristics of patients and patient groups [2]. If there is a mismatch between the targeted PRO scale and its intended construct, then the problem could be that the scale is good but the theory is wrong, the theory is good but the scale is not, or both the theory and the scale are useless or misplaced. Criterion validity shows the extent to which the scores of a PRO instrument are related to a known gold standard measure of the same concept.

Reliability assesses how precise or stable the instrument measures what it measures and is typically discussed in terms of reproducibility. It has two main types: (1) internal, and (2) repeatability reliability. Internal reliability is based on item-to-item correlations and the numbers of items in the questionnaire. Repeatability reliability is based on the analysis of variances between repeated measurements on the same subjects.

References

1.   Patrick DL, et al. (2011) Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1--eliciting concepts for a new PRO instrument. Value Health, 14(8): 967-977.

2.   Food and Drug Administration. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009

3.   Patrick DL, et al. (2011) Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 2--assessing respondent understanding. Value Health, 14(8): 978-988.

4.   Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

4. Patient adherence

Terminology

There are a number of terms used to describe the extent to which a patient undertakes the recommendations of healthcare providers. The most commonly used terms are compliance, adherence and persistence. There is a lack of uniformity in standards of definitions used to describe the concepts of compliance, adherence and persistence, which causes many difficulties when comparing or combining results of different studies. The definitions of the WHO [1] and the ISPOR [2] are the most widely accepted in the literature.

Although most studies have focused on medication adherence, adherence encompasses numerous other health-related behaviors as well. The WHO definition also reflects this concept. According to the WHO definition adherence refers to 'the extent to which a person's behavior - taking medication, following a diet, and/or executing lifestyle changes - corresponds with agreed recommendations from a healthcare provider' [1]. Medication adherence, as defined by the ISPOR, 'refers to the act of conforming to the recommendations made by the provider with respect to timing, dosage and frequency of medication intake' [2]. According to the definition of ISPOR, medication persistence can be defined as 'the duration of time from initiation to discontinuation of therapy' [2] (Figure 4.1).

Recently, medication adherence has become the preferred term instead of medication compliance. The primary difference between compliance and adherence is that compliance reflects the patient as a passive recipient of the recommendations of healthcare providers. The definition of adherence accurately highlights the importance of the patient's active role in their own healthcare, which emphasizes that the relationship between the patient and the healthcare provider should be based on a partnership, instead of a one-sided paternal relationship.
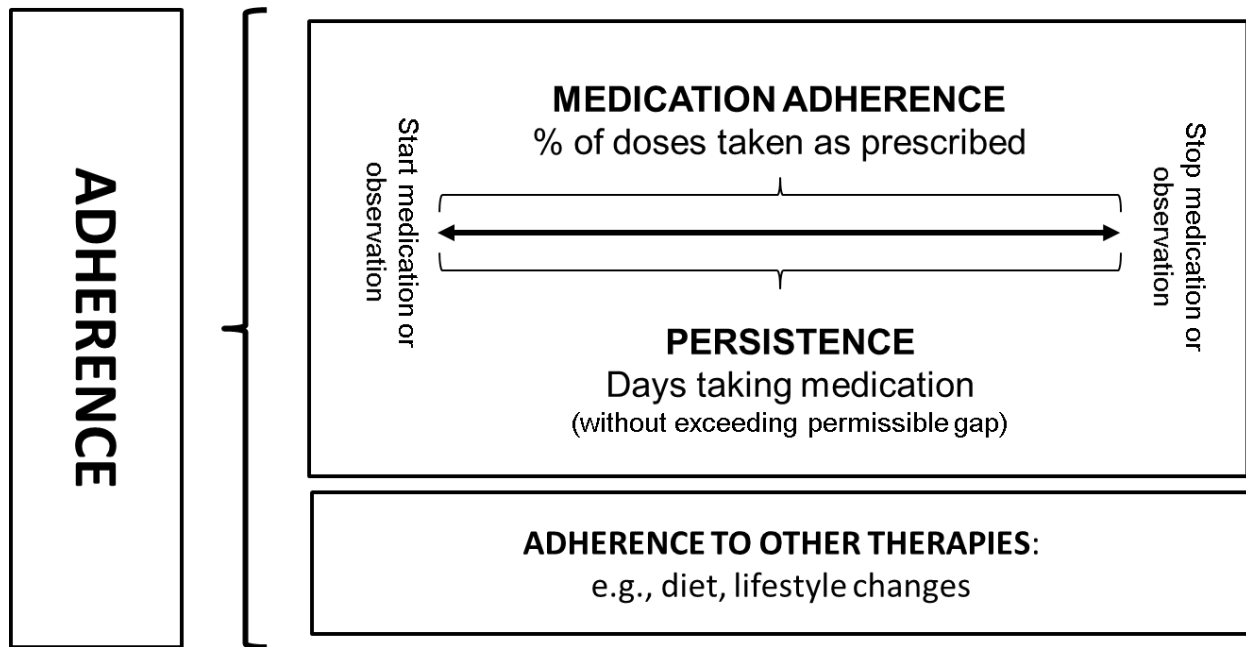
Figure 4.1. Definitions of adherence, medication adherence and persistence [2]

Methods for measuring and calculating medication adherence

There is no gold standard of measuring medication adherence. Several methods have been proposed in the literature, but each method has its strengths and limitations [3]. The easiest way to assess medication adherence within clinical settings is to collect information from the patients themselves through questionnaires or patient diaries. However, it should be noted that self-reporting methods may overestimate adherence. Another commonly used method is the retrospective analysis of pharmacy dispensing data records. This approach is rapid and inexpensive; nevertheless, pharmacy dispensing records can be used to evaluate adherence to prescription refill, but not the medication intake directly. Thus, it may be an inaccurate means to assess medication adherence. Pill count and canister weighing are widely used methods of adherence assessment in clinical trials. Pill counts are limited to oral medications, but canister weighing can be used to monitor inhaled or liquid drugs as well. These approaches assess only the quantity of the medication removed from the canister, but provide no evidence on medication intake. The cap of the medication canisters can be equipped with a microchip that stores data about each opening of the canister. Therefore electronic monitors may provide an accurate measure of dosing history but also cannot confirm medication intake. The major disadvantage of this method is its price; it is relatively costly. Medication adherence can also be measured with direct assessments, such as direct observation of the medication intake, or evaluation of blood levels, urinary excretion of the drug, its metabolite or drug-marker. However, these methods

are unpleasant for the patient. Interestingly, therapeutic drug monitoring may overestimate the actual adherence rate because patients tend to comply shortly before the drug test but not during the whole observation period. Furthermore, drug tests are insensitive to inhaled medications. Another direct method is to equip pills with microchips that can report back exactly when, what kind of and how much medicine the patient has taken. The main goal of this new method is to help patients adhere to their medication regimen and to get accurate data on medication intake in clinical trials.

One of the most commonly used approaches for calculating adherence to mono-pharmacotherapy from pharmacy dispensing records is medication possession ratio (MPR) [2]. In the model of MPR, the total day's supply dispensed over the observation period is divided by the number of days of the observation period. Results of MPR can be expressed as a continuous or dichotomous variable. If MPR is given as a continuous variable, it can range from 0 (no medication dispensed) to 1 (maximal adherence). It is important to highlight that medication oversupply (MPR>1) can be considered as medication non-adherence as well. The cut-off point used to categorize patients into adherent and non-adherent groups should be determined in a disease- and therapy-specific way. However, the cut-off point is generally set at 80%, independent of whether this adherence rate is adequate for disease control or not. To be able to calculate persistence, a permissible gap period must also be defined, which specifies the maximum allowable time period between refills without discontinuation of the therapy [2]. Although most studies use a 60-day permissible gap, a permissible gap should be determined in a disease- and therapy-specific way. The length of a persistence period can be counted in days. But, persistence can also be given as the percentage of the number of persistent patients at the end of a predefined time period.

Prevalence of non-adherence

Non-adherence to medication is common and poses a significant barrier to optimal disease management. According to the WHO, adherence to long-term therapies averages only 50% [2]. Adherence and persistence rates differ between chronic disorders. In hypertension, dyslipidemia and diabetes adherence is around 67-76% and persistence is 63% [3], and in asthma and in COPD adherence ranges between 20-60% and persistence 7-16% [4]. Non-adherence reduces the clinical benefit of the therapy and accounts for many of the observed differences between the efficacy reported in randomized controlled trials (RCTs) and the effectiveness of the drug treatment achieved in real-world conditions. In RCTs, the stringent follow-up protocol limits the occurrence of medication non-adherence; therefore, non-adherence rates derived from RCTs do not reflect an objective picture. For example the average rates of adherence to

COPD medication in RCTs have been estimated to be 70–90%; however, in real-world conditions these rates are only in the range of 20–60% [4, 5].

Consequences of non-adherence

Non-adherence may have clinical and cost consequences and may affect health-related quality of life (HRQoL) as well. Increased adherence results in better health outcomes. Persistent anti-hypertensive therapy is associated with a 40% increased chance of blood pressure attainment [6], and non-adherence to statin therapy in cardiovascular patients increase the incidence of myocardial infarction by two times [7].

The effect of non-adherence on medical costs work in two ways: (1) it has an immediate and direct impact on drug costs and (2) it has a less immediate and indirect impact on health service utilization / healthcare costs. In general, non-adherence is likely to reduce drug costs, but increase subsequent overall health service utilization / healthcare costs. But its effect on drug and healthcare costs is highly dependent on the condition, the therapy and the time frame of the analysis. The impact of medication adherence on drug costs is highly affected by the extent of non-adherence, but does not always results in decreased drug costs. For example this may occur in cases when non-adherent patients overuse their medications or dispense prescriptions but do not use them and stockpile medications. The impact of non-adherence on overall healthcare utilization is determined primarily by the clinical effectiveness of the medicine. The impact of medication non-adherence on resource use is large, where health service use is highly associated with the extent of the management of the condition and the medication has a key role in the management of the condition.

Association between medication adherence and HRQoL is dual: not only adherence can affect HRQOL, but HRQoL may also impact medication adherence [5]. The effect of adherence on HRQoL may be a consequence of the effectiveness of therapy and the negative effects that it can generate (i.e., side-effects, daily life limitation of therapy, social stigma). Dynamics between adherence and HRQoL may differ over time and the negative effects of medication non-adherence may become more and more dominant in the long-term. A patient's decision to adhere or not and to what extent is a personal trade-off between the benefits and the negative effects of the therapy on HRQoL at any given time [5]. Psychiatric comorbidities (e.g., depression) may influence the relationship between medication adherence and HRQoL.

Determinants of adherence

Adherence to medication is a multidimensional phenomenon. Adherence has a number of factors, including socio-economic, patient-related, condition-related, therapy-related, and healthcare team and system-related factors (Table 4.1) [1, 8].

Socio-economic factors	Family/social support (emotional, financial), social stigma, co-payment, income, employment status, etc.
Healthcare team and system-related factors	Barriers to healthcare, prescription by a specialist, healthcare provider-patient communication and relationship, etc.
Condition-related factors	Disease severity, perseverance of symptoms, psychiatric condition, clinical improvement, etc.
Therapy-related factors	Adverse effects, number of drugs/daily doses, duration of the treatment, etc.
Patient-related factors	Age, forgetfulness, marital status, education, etc.

Table 4.1. Factors associated with adherence [1, 8]

Evidence suggests that family and/or social support has a positive effect on adherence, and the lack of such support has a negative effect [8]. Social stigma of a disease or therapy may also be responsible for non-adherence [8]. For example, in patients with COPD the perceived social stigma associated with using inhalers in public might affect adherence [5]. Economic factors such as unemployment, low income, or high out-of-pocket drug cost may contribute to non-adherence as well [8].

Healthcare system factors have an important impact on adherence. Adherence requires a good relationship between healthcare providers and patients. Quality of communication is related to adherence. The type of caregiver also influences adherence. Medication adherence may increase if the prescribing physician is a specialist instead of a general practitioner. Furthermore, periodic visits, closer follow-up and hospitalization may also have an increasing effect on patient cooperation [8].

Adherence is also related to condition. Disease severity and presence of symptoms influence adherence. For example the asymptomatic nature of a disease has a negative effect on adherence. Evidence suggests that non-adherence is more common in patients with psychiatric conditions [8].

Patient-unfriendly therapy, e.g. the high number of drugs, frequent dosing, long treatment duration, is associated with non-adherence. Other factors, such as adverse

effects are also important. For example, patients with COPD often confuse the side-effects of inhaled corticosteroids with those of anabolic steroids, which may decrease their cooperation willingness [8].

Adherence is related to age; older patients seem to be more adherent. Older patient are more likely to adhere to therapy that requires adjustments in daily life. However, memory loss and cognitive impairment, which are both associated with aging, may adversely affect adherence. Some studies found that those married had better adherence than those who were single or divorced and that patients with higher levels of education had better adherence [8].

Methods for integrating adherence in pharmacoeconomic evaluations

To evaluate the impact of non-adherence/non-persistence on both health outcomes and costs requires the use of health-economics models. Decision-analytic models, Markov models or discrete event simulation models can be appropriate for this purpose. The choice of the model type is dependent on the condition being treated (e.g., acute vs. chronic), data availability (individual vs. aggregate data) and the type of adherence data (i.e., adherence data vs. persistence data) [9].

Decision-analytic models may be appropriate particularly for modeling adherence in acute diseases. In most conditions, such models can be developed even from published sources. Branches of the decision tree may be used to represent different levels of adherence (e.g., adherent/ non-adherent). However, when there are numerous health states, including the possibility of transitions from one health state to another and back again, the decision tree may become far too complex but Markov models can be appropriate to handle this problem efficiently. A sample hypothetical example for integrating persistence in health economic evaluations by use of a Markov model can be seen in Figure 4.2. This simple hypothetical model has three health states "progressive", "remissive" and "death". Progression from the remissive state to progressive state is dependent on an annual transition probability, Pt1. For those patients who discontinue treatment, Pt1 is assumed to increase. The probability of death from progressive and remissive health states is represented by transitional probabilities Pt2 and Pt3, respectively [9].
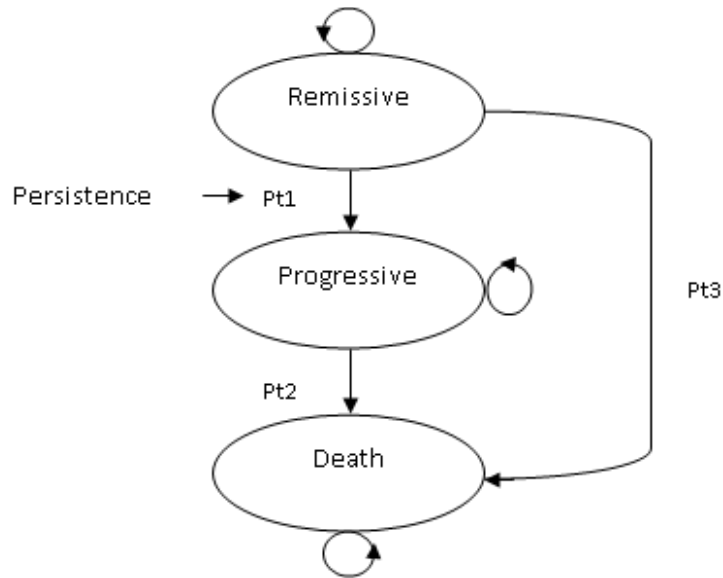
Figure 4.2. Schematic figure of a Markov model integrating persistence [9]

In a discrete event simulation (DES), the experience of individuals is modeled over time in terms of the events that occur and the consequences of those events. In a DES patients are specified as entities and treatment discontinuation as events. The pros of DES are that it facilitates interactions between adherence and time, as well as individual characteristics (e.g., adherence to drugs for asthma may be highly correlated with the severity of symptoms). However, developing a DES model requires more detailed, individual data than a typical Markov cohort model.

Adherence-enhancing interventions

Strategies for improving patient adherence have to be formulated based on factors related to adherence. Although non-adherence has often been perceived as a fault of patients, other stakeholders such as health policy, the pharma industry, the healthcare provider and the social environment may influence adherence as well. Medication non-adherence can be considered the effect of multiple determinants; consequently, multifaceted interventions may be the most effective. The main categories of adherence-enhancing interventions are interventions related to clinical innovation or patient education, patient reminders, cost-related approaches and other interventions (e.g., pharmacist programs) [10, 11] (Table 4.2). Successful adherence-enhancing programs include simplified treatment regimens, facilitation of the physician–patient relationship and patient education methods [10]. For the analysis of the cost–effectiveness of adherence-enhancing interventions, it is important to look at both costs

of the intervention and outcomes, not only in terms of adherence, but also in terms of the subjective value of the clinical outcome for the patient.

Clinical innovation	simplified regimen, long action medication, combination drug
Patient education	print materials, online communication, CD-ROMs
Patient reminders	tele-calling, e-mails, text messages, apps
Cost-related approaches	reducing co-payment, discounts, vouchers
Others	nurse education, pharmacist programs, patient organizations, self-monitoring

Table 4.2. Adherence-enhancing interventions [10, 11]

References

1.   World Health Organization. Adherence to long-term therapies: evidence for action. WHO, 2003. http://www.who.int/chp/knowledge/publications/adherence_ report/en/ [Accessed online: 22th October, 2015]

2.   Cramer JA, et al. (2008) Medication compliance and persistence: terminology and definitions. Value Health, 11(1): 44-47.

3.   Cramer JA, et al. (2008) The significance of compliance and persistence in the treatment of diabetes, hypertension and dyslipidaemia: a review. Int J Clin Pract, 62(1): 76-87.

4.   Agh T, Meszaros A. Adherence to therapy in chronic obstructive pulmonary disease, In: Ong K-C. Chronic obstructive pulmonary disease: current concepts and practice. InTech, 2012: 275-290.

5.   Agh T, et al. (2015) Relationship between medication adherence and health-related quality of life in subjects with COPD: a systematic review. Respir Care, 60(2): 297-303.

6.   Breekveldt-Postma NS, et al. (2008) Effect of persistent use of antihypertensives on blood pressure goal attainment. Curr Med Res Opin, 24(4): 1025-1031.

7.   Blackburn DF, et al. (2005) Cardiovascular morbidity associated with nonadherence to statin therapy. Pharmacotherapy, 25(8): 1035-1043.

8.   ABC Project Team. Ascertaining barriers for compliance: policies for safe, effective and cost-effective use of medicines in Europe. 2012. http://www.abcproject.eu/ [Accessed online: 22th October, 2015]

9.   Hughes D, et al. (2007) Methods for integrating medication compliance and persistence in pharmacoeconomic evaluations. Value Health, 10(6):498-509

10.   Petrilla AA, Benner JS. (2003) Critical evaluation of interventions to enhance patient compliance with chronic medications. Value Health, 6: 200.

11.   Capgemini Consulting. Patient Adherence: the next Frontier in Patient Care, 9th edition. 2011.

5. Preference (utility) measurement

Need for utility measurement

Limited resources necessitate prioritization among different health technologies. Therefore payers need to make choices between the available therapies and treatment alternatives. Explicit decision-making criteria reduce the opportunity cost of inappropriate decisions compared to implicit decision-making. The evaluation of only health benefits is not sufficient to justify reimbursement decisions, since it may result in running out of budget funds. On the other hand, evaluation of costs only may violate a patient's safety or jeopardize health outcomes, and therefore may result in the reimbursement of ineffective technologies or a decision not to reimburse an effective technology. In terms of health outcomes, three questions need to be considered by decision-makers [1]:

- Does the new therapy provide health gain compared with a placebo?
- Does the new therapy provide more health gain than the current standard care? (Comparison with other health technologies)
- Is the new therapy of good value for the money? (Comparing incremental cost and health gain)

The first two questions are in the focus of this module. In conclusion, to inform health policy decision-making, a full health economic evaluation is necessary which considers 1.) both costs (inputs) and consequences (outputs) and 2.) compares two or more alternatives [2]. To capture the whole spectrum of health technologies with various health outcomes, a universal construct is necessary. Therefore:

- The unit of consequences should be able to capture health gain in the entire range of health technologies (e.g. cholesterol reducing drugs, dialysis, hip prosthesis, PET CT etc.);
- The unit of consequences should be able to compare different health technologies (to ensure comparability) (e.g. reduction of blood pressure in Hgmm, higher resolution of imaging diagnostics etc.).

In cost utility analysis, the QALYs are applied as health outcome measure. This type of analysis therefore is suitable to compare not only health technologies with identical health outcomes (as in a cost minimization analysis), or health technologies where health outcomes can be measured in the same natural unit (e.g. blood pressure in Hgmm, as in a cost effectiveness analysis), but also health technologies with different natural units. Therefore cost utility analyses can improve allocative efficiency in healthcare [3, 4].

Economic background of measuring utility

Utility, or usefulness, is the ability of something to satisfy needs or wants. It represents satisfaction experienced by the consumer of a good. Not coincidentally, a good is something that satisfies human wants and provides utility, and satisfying needs improves utility. The term 'Utility' is based on the Neuman Morgenstern (NM) utility theory, which applies cardinal utility (measured in interval scale), under uncertainty [5]. Key premises of the NM utility theory are:

- Any individual whose preferences satisfied four axioms (completeness, transitivity, independence, continuity) has a utility function,
- Such an individual's preferences can be represented on an interval scale and
- The individual will always prefer actions that maximize the expected utility.

According to NM utility theory, only those preferences are utilities, which consider the subject's attitude to risk (uncertainty). Preferences without considering the subjects attitude to risk are values. However, in the wider medical literature, values are often considered as utilities [6]. Health state measures - beside the psychometric criteria discussed elsewhere – also need to fulfill the following economic criteria [7]:

- Comparability across disease,
- Interval scale for measurement,
- Individual preference-based scoring.

The economic and psychometric criteria are in trade-off in different health status measures: The specific measures have favorable psychometric attributes (e.g. sensitivity) but with only limited comparability. The survival-as-health-outcomes measure would have favorable economic attributes, but without considering the quality of life. As discussed in the introduction of this module, this wouldn't be in line with the holistic approach applied recently in healthcare. Considering only life expectancy would mean that living 1 year in perfect health would be equivalent to living 1 year in a coma. The optimal health outcome measures for healthcare decision-making would be the preference based generic index type measures: these have favorable economic attributes, but with some psychometric limitation compared to the specific measures [7].

What are the options for valuing health states? Asking for an expert's opinion or an expert panel's opinion would be cheap and feasible, but experts have different perspectives as (potential) patients. Another option would be to use the existing, and constantly increasing published data. However, transferability of results from the published literature to the target population could be an issue. Also, the quality of the publication needs to be carefully assessed. The third option is measuring the study population directly, which is probably the most accurate, but the most resource intensive option as well. Collective priority-setting requires comparing the benefits of different

kinds of healthcare techniques systematically. Therefore an extremely versatile benefit measure with an interval scale measurement property is needed to compare the size of differences in levels of health benefit between the treatments. Any measure that fails to fulfill these criteria is inadequate in principle as an aid to priority-setting.

Utility

According to the ISPOR Book of the terms, 'Utility is a quantitative expression of an individual's preference for, or ability of, a particular state of health under conditions of uncertainty'. As discussed in the Introduction, the conventional utility scale has two fix reference points: Utility of 0.0 for death and utility of 1.0 for perfect health. States being worse than death can have negative utilities [8]. Utilities may be aligned using direct preference measurement tools (such as time trade-off (TTO), standard gamble (SG), or rating scale (RS)), or indirectly, by using utility weighted index type generic QoL measures (such the EQ-5D index, the Short Form (SF)-6D or Health Utility Index 2 and 3) [9].

Direct preference (utility) measures

Rating scale (RS)

The RS method applies a VAS, which can have numbers (e.g. 0-100) or just a 10 cm line on a page with clearly defined end points. The preference of chronic health states can be measured on RS. The subject is provided a batch of health states including healthy and death, which are used as reference states. The subject is asked to select the best health state of the batch (probably normal healthy life) and the worst health state, which may or may not be death. Then the subject is instructed to locate the remaining states on the scale by concentrating on intervals and spacing (e.g. relative distance) and comparison of one interval to another rather than purely on scores (numbers). Intervals and spacing therefore should correspond to the differences in preference of health states of a subject. Death is assigned to 0.0 and healthy is assigned to 1.0, therefore the preference of health states are given by the formula (x-d)/(1-d), where x is the scale placement of the given health state and d is the placement of death [10].

Time trade-off (TTO)

Time trade-off was designed specifically to be applied in healthcare. When measuring chronic health state i, the subject is offered two alternatives:

- State i for time t (life expectancy of an individual with chronic condition) followed by death,
- Healthy for time x < t followed by death (Figure 5.1).

Time X is varied until the respondent is indifferent between the two alternatives, at which preference score is: x/t.
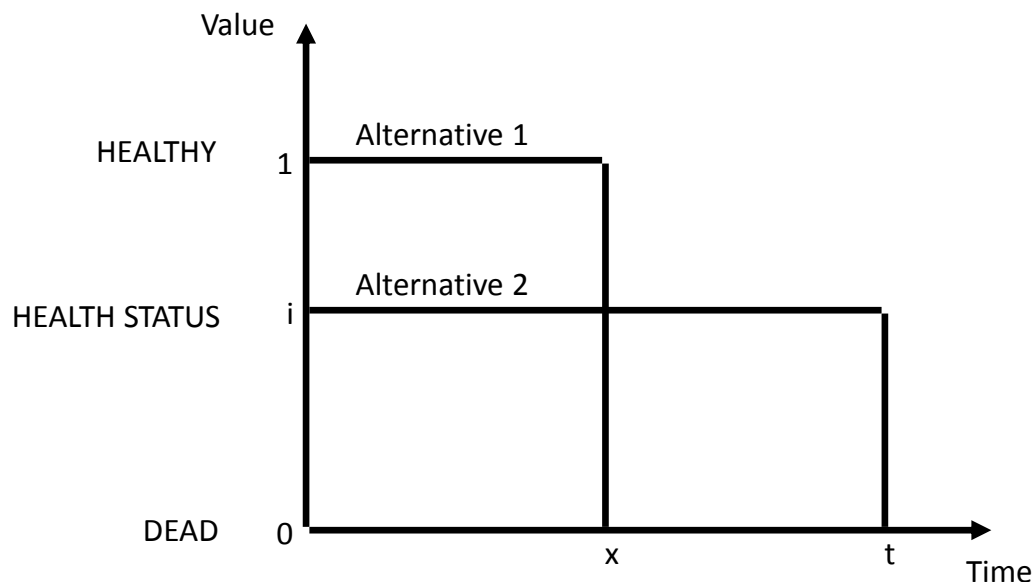
Figure 5.1. Time trade-off method [4]

There is no uncertainty among the two alternatives. Time frame of the TTO may be either:

- T time fixed (e.g. 10 years) or
- T time variable (e.g. life expectancy standardized to age and gender).

Also, the TTO interview might be conducted by using the so-called 'ping-pong' or 'iterative' question framing. Whichever is used, it should be applied consistently during the entire study. If a subject is indifferent between 40 years of life expectancy in a chronic health state and 30 years in a perfect health, his preference score is x/t=30/40= 0.75 [9, 11].

Standard Gamble (SG)

When measuring chronic health states i with standard gamble, the subject is offered two alternatives:

- A theoretical treatment with 2 possible outcomes where,
    i. Subject immediately returns to perfect health for an additional t years with a probability (p),
    ii. Subject immediately dies with a probability of (1-p);
- A certain outcome of staying in chronic state i for t years (Figure 5.2).

Probability p is varied until respondent is indifferent between the two alternatives. Then the preference score for state i for time t is p.
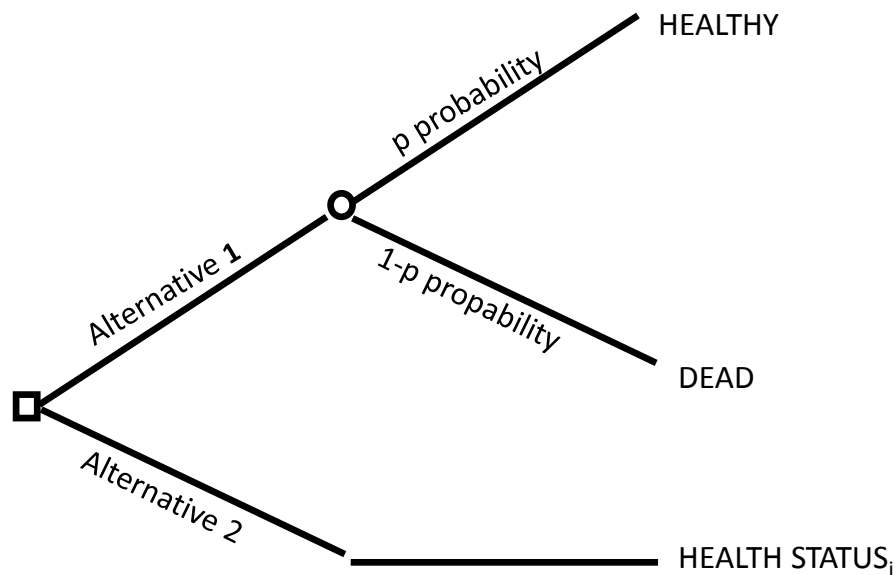
Figure 5.2. Standard gamble method [4]

In the SG method, one of the alternatives includes uncertainty (see p probability). If the respondent is indifferent between the two alternatives at 25% risk of mortality (1-p), his/her preference score for health state i is 0.75 [9, 12].

Comparing direct health state valuation methods

Many use 'utility', 'value' and 'preference' interchangeably. Preference is the umbrella term for the overall concept. Value is derived from a question framed under certainty by comparing outcomes and to choose between them (using TTO) or to scale them (using

RS/VAS). Utility is derived from a question framed under uncertainty by comparing two outcomes where at least one contained uncertainty i.e. there is a p probability to capture subjects risk attitude (SG) (Table 5.1) [6].

Response method	Question framing	
	Certainty (values)	Uncertainty (utilities)
Scaling	Rating scale (RS) Visual analogue scale (VAS)	-
Choice	Time trade-off (TTO) Person trade-off (PTO)	Standard gamble (SG)

Table 5.1. Classification of direct preference measurement tools [4]

RS does not involve a decision situation with a significant potential health loss (e.g. trading-off life expectancy (in TTO) or risking immediate death (in SG)). Many consider SG as a 'gold standard' because of handling uncertainty. SG provides a higher preference score compared to TTO for the same health state in the same subjects as:

- Subjects are risk avoidant (risk of immediate death, therefore SG overestimates preference scores),
- Subjects have positive time preference (less value is attributed for the lifetime just before death, therefore TTO underestimates the preference score).

Multi-attribute health status classification systems with preference scores

These measures define a finite number of health states. Health states are valued by population with direct valuation methods. These tools are the links between HRQoL and utility:

- Rosser-Kind matrix,
- EQ-5D,
- HUI, HUI2, HUI3,
- SF-6D.

EQ-5D 3L

The EQ-5D 3L index is a widely used instrument to assess general QoL, focusing on 5 dimensions ('5D'): mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 1 item, and each item provides 3 levels ('3L')

with level 1 denoting no problems and level 3 denoting extreme problems. Utility values derived from a UK population survey using the TTO method can be assigned to the 35=243 theoretically possible outcomes (and death and unconsciousness). The EQ-5D VAS (also known as EQ-5D Thermometer) is a visual analogue scale, calibrated from 0 (worst imaginable health state) to 100 (best imaginable health state).

The measure is simple, easy-to-understand, allows comparability of different diseases, and even population reference values exist for healthy population. The measure has been widely used in several countries of the world. However, as a generic measure, it has limited sensitivity to detect small changes [13, 14]. EQ-5D also has a five level version (EQ-5D 5L), with the same dimensions, but with 5 options, and a version developed for children (EQ-5D Y – 'Youth').

Whom should we ask about the relative utility of each health status?

If we ask patients, they are familiar with the health state they live in (e.g. there is no need for a detailed description of the given health state), however, they tend to rate their health state as well as the value of the therapy higher, since chronic patients are able to adapt even to devastating health states [15]. If we ask healthcare professionals, they are also familiar with the different health states, and that being the case, they also know all the potential negative outcomes of a certain health state, so they tend to underrate health states. Also, when arguing for a larger budget, healthcare professionals tend to overrate the value of their own profession. Manufacturers of health technologies would be even more biased towards health states in their own interest. Another option would be to ask the preference of the general population. As opposed to patients and healthcare professionals, representatives of the general population are unfamiliar with a particular health state; therefore they require a detailed description to imagine a particular condition. However, they are the sustainers of the healthcare system (as taxpayers) and also the potential future users of health technologies (as future patients). Therefore the most widely used approach is to ask patients to fill out the questionnaires, for which the preference scores of the general population have been assigned [16].

References

1.   Inotai A, Mészáros Á. (2009) Economic evaluation of NSAID strategies in rheumatoid arthritis. Cost-effectiveness of celecoxib in Hungary. Int J Technol Assess Health Care, 25: 190-195.

2.   Drummond MF, et al. Basic types of economic evaluation. Methods for economic evaluation of health care programs. 2nd edn. Oxford University Press, New York, 1997: 10.

3.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Cost-Utility Analysis. ISPOR, USA, 2003: 45-48.

4.   Drummond MF, et al. Cost utility analysis. Methods for economic evaluation of health care programs. 2nd edn. Oxford University Press, New York, 1997: 139-204.

5.   Neumann J, Morgenstern O. Theory of Games and Economic Behavior. Princeton, NJ, Princeton University Press, 1944.

6.   Drummond MF, et al. Cost utility analysis. Methods for economic evaluation of health care programs. 2nd edn. Oxford University Press, New York, 1997: 146.

7.   Fox-Rushby J, Cairns J. Approaches to measuring health and life. In Fox-Rushby J, Cairns J (eds), Economic Evaluation. Open University Press, McGraw Hill, 2005: 85-100.

8.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Utility. ISPOR, USA, 2003: 241-242.

9.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Utility measurement. ISPOR, USA, 2003: 242-244.

10.   Drummond MF, et al. Rating scale. Methods for economic evaluation of health care programs. 2nd edn. Oxford University Press, New York, 1997: 151-152.

11.   Drummond MF, et al. Time trade-off. Methods for economic evaluation of health care programs. 2nd edn. Oxford University Press, New York, 1997: 154-156.

12.   Drummond MF, et al. Standard Gamble. Methods for economic evaluation of health care programs. 2nd edn. Oxford University Press, New York, 1997: 153-154.

13.   The EuroQol Group. (1990) EuroQol: a new facility for the measurement of health-related quality of life. Health Policy, 16: 199–208.

14.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. EuroQoL (EQ-5D). ISPOR, USA, 2003: 104-106.

15.   Lyttkens CH. (2003) Time to disable DALYs? On the use of disability-adjusted life-years in health policy. European Journal of Health Economics, 4(3): 195-202.

16.   Williams A. (1996) QALYs and ethics: a health economist's perspective. Soc Sci Med, 43: 1795-1804.

6. QALY, DALY and critique

Quality Adjusted Life Years (QALY)

In 1968, Herbert Klarman suggested that kidney transplantation provides life years gain and also improved QoL compared to dialysis. This is considered to be one of the first references to QALY. QALY is a universal health outcome measure, designed to be applicable for all individuals, conditions and health technologies. QALY combines both quality of life (morbidity) and quantity of life (mortality) in a single construct [1]. In the QALY concept, life years are adjusted by a preference-based quality weight (utility) [2]. The construct is applied as a key health outcome measure in cost utility analyses (CUA) [3]. CUAs are used to estimate the incremental cost of a new therapy which has to be paid for one QALY gain compared to an appropriate comparator. As CUA can compare a wide spectrum of health technologies in terms of their health outcomes, it is used to support priority setting/decision-making in healthcare.

QALY= (0.5x1)+(0.5x0.25)+(0.5x0.5)+(1x0.75)+(0.5x0.5)=1.875
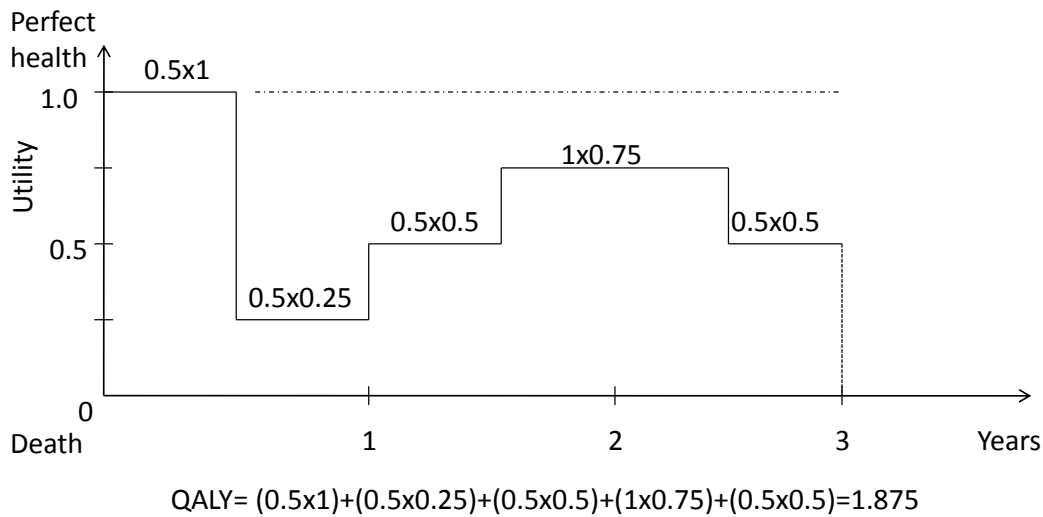
Figure 6.1. QALY calculation based on area-under-the- curve

QALYs can be calculated by estimating area-under-the-curve (Figure 6.1). As the next figure indicates, only considering life years would underestimate health gain in some cases (Figure 6.2).
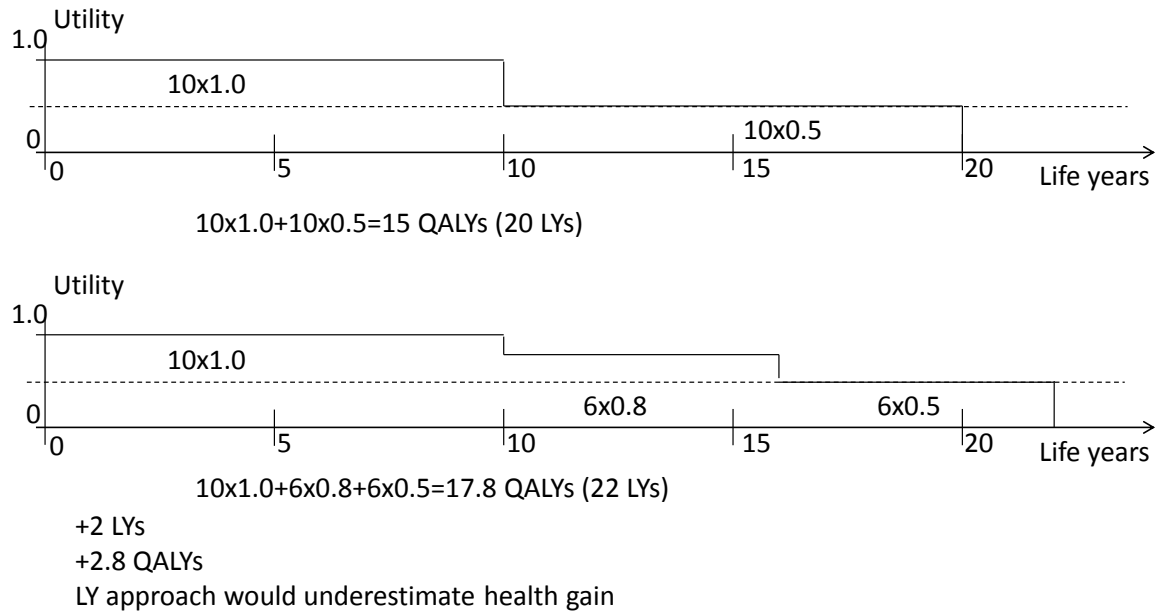
Figure 6.2. QALY and LYs in health outcome measurement

QALY is suitable to aggregate different dimensions of health outcomes even related to a single intervention. For example, in the case of complex oncology treatment there is short-term QoL deterioration (due to adverse events) and long-term life year gain and QoL improvement due to the benefits of the treatment. The QALY concept is suitable to estimate the magnitude and sign (+/-) of the aggregate health gain. By the use of QALY it is theoretically possible to rank different health technologies based on objective criteria ('QALY league table').

QALY in healthcare decision-making

One of the main objectives of health policy is to maximize the health gain of the population from the available healthcare budget. If we accept that health capital of a subject is the total number of QALYs accumulated by him- or herself, and health capital of the population is the sum of the health capital of the individual subjects, the following assumption is valid:

20 QALYs for 1 subject = 2 QALYs for 10 subjects = 0.2 QALY for 100 subjects

CUAs compare two or more alternatives in terms of both their costs and outcomes, where outcomes are measured in units of utility or preference multiplied by the duration of the particular health state [3].

Certainly, the QALY approach has some limitations: limitations of the utility construct, limitations of the QALY concept, and limitations of decision-making based on the QALY concept. The use of QALY has been criticized based on ethical, conceptual and methodological grounds.

Disability Adjusted Life Years (DALY)

Full economic evaluations are used to improve allocative efficiency in a healthcare system with limited resources by prioritizing among different interventions. However, prioritizing among different disease areas is not possible with economic evaluations. BoD studies are suitable for identifying unmet needs in healthcare. QALYs are suitable for estimating health capital, however they are not practical for estimating health loss in global burden of disease studies due to the high amount of input data needed. Therefore the WHO and the World Bank apply the DALYs to estimate health loss in global burden of disease studies. The DALY has two key components [4]:

DALY= Years of life loss (YLL) + Years lived with disability (YLD)

The impact of a particular disease on mortality is estimated by calculating the difference between life expectancy and the age at which death occurred and is expressed as years of life lost (YLL). To facilitate comparability across countries, the method is standardized by using the average life expectancy of Japanese women (82.5 yrs.) for women and an arbitrary value of 80 years is used for men [5]. Therefore

YLL = average life expectancy – age at death

Years lived with disability (YLD) measures the impact of morbidity in the DALY concept, considering the following factors:

- The extent of disability associated with non-fatal conditions (disability weights), which applies an endpoint of 0.0 for perfect health and 1.0 for death.
- The relative importance of healthy life at different ages which are weighted according to productivity (age weights, which are the highest in the middle-aged group, and lower in the elderly and the young).
- The time preference for health (the value of health gained now as compared to the value of health gained in the future) (per protocol discounting with 3%).

Disability weights are determined by experts by using person-trade-off method (PTO).

The link between QALYs and DALYs

Let's imagine a deaf patient to illustrate the relationship between QALY and DALY [6]. Before going deaf, this man lived in perfect health with a utility of 1. (For the simplicity of the example, we won't consider the impact of aging and short-term conditions on his QoL). At the age of 22 his utility will decrease by 30% to 0.7. From 22 to 50, he lives deaf. His accumulated QALYs can be summarized by calculating the area under the curve. Let's suppose that this man dies at the age of 50 from a stroke. The life years he lost compared to his potential life expectancy is summarized by the YLL. The health capital he lost due to deafness is summarized by YLD. His DALYs can be summarized as YLL+YLD. It is not possible then for QALY and DALY, utility weights and disability weights to be converted into the other measure (Figure 6.3).
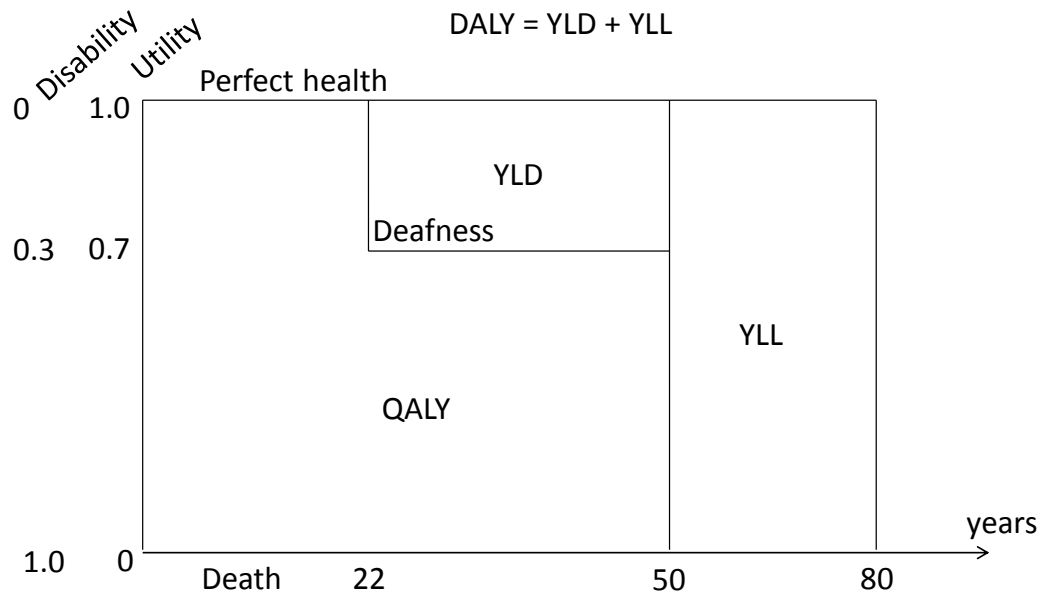
Figure 6.3. Relationship between QALY and DALY

DALYs can be used in burden of disease studies (WHO GBoD) and cost effectiveness studies (cost/avoided DALYs) to support evidence-based health policy. However, the applicability of DALY has been heavily criticized, especially when used for prioritization among health technologies [7].

Critique of DALY

The first area of DALY that receives criticism is its applicability. According to the opponents of DALY, the construct measures not only the BoD but the level of underdevelopment as well. When the standardized life expectancies from developed

countries are used, the assumption is made that health interventions alone are capable of achieving the same higher levels of increase in life expectancy in lower income countries as well. However, many non-health circumstances need to be improved for life expectancy to rise to the level of high income countries used as a reference in the DALY calculations.

Also, BoD is comprised of two core components: the economic burden of a disease and health loss. According to the opponents, DALYs focus only on disability and mortality, and do not take account of healthcare costs [7].

The second key area of criticism by opponents is the discrimination of people. According to these remarks, if DALY is used for priority-setting among different patients, it becomes more valuable to save the life of a young person rather than someone older (to avoid more DALYs), and similarly, DALY gives priority to save someone's life who is healthy compared to the one who is disabled. Therefore the use of DALY enhances inequality in the society. Also, the age weights of DALY discriminate between the young and elderly compared to those middle-aged. Similarly, discounting required by the DALY concept is against preventive health technologies, and would justify today's environmental degradation where the present generation benefits at the expense of future generations [6, 8].

The third group of concerns stem from the methodological debate. The DALY concept uses experts to value health states, which do not reflect preferences of the patients or the population. Also, the person-trade-off method has been heavily criticized since it incorporates both the valuation of states of health and people's views about distributional issues as well.

Critique of QALY

Allen Williams, the father of the QALY approach summarized the key critiques received from objectors of QALY in a review in 1996 [9]. According to this review, the QALY concept may be criticized on ethical grounds and on methodological grounds. Opponents may also criticize decision- making based on QALYs, and the cost of implementing the QALY approach.

Ethical based critiques

As stated in the introduction of the term Utility, there are scarce resources for any healthcare system even in very high income countries. Consequently, resources devoted to one person will be denied some other person who might have better benefitted from them (opportunity cost). One may therefore say that the cost represents

sacrifices made by other potential patients who did not get treated. The key objective of improving allocative efficiency in distributing limited resources, therefore, would be that benefits gained by those to whom treatments are offered would be greater than benefits sacrificed by those who were denied treatment. The aim of healthcare would be to do as much good (i.e. generate as much health gain) as possible with our limited resources. Therefore extending the healthcare budget wouldn't be an option to avoid the necessity of selecting between treatments and patients since it recreates the dilemma of scarce resources at a higher level of spending. Some opponents of the QALY concept don't accept limited resources in a healthcare budget. They state that where lives are at stake, a fundamental reappraisal of priorities would be necessary and the national budget should be reconsidered to allow for budget reallocations from non-lifesaving expense coverage. Thus some opponents accept scarce resources only if the expense headings of the national budget are assigned to even more important aims for the society than rescuing citizens in mortal danger [9].

Those who reject all collective priority setting as unethical, typically assert that it is immoral to sit in judgment over the worth of other people's lives. As they come to terms with the scarcity of resources, they have to acknowledge that some people must be denied the benefits of healthcare. However, they want this to happen in a manner free of interpersonal judgments of the relative worth of someone's life. What needs be kept in mind is that ultimately, someone has to make a conscious decision on how best to discriminate between people when confronted with scarcity. Would implicit decision-making be associated with lower opportunity cost of inappropriate decisions [9]?

Other opponents accept the need for priority-setting but believe that it is contrary to medical ethics. According to an extreme opinion, the doctors should do everything possible for the patient no matter what the costs are. For example in rare diseases, orphan drugs with very high costs are used, but often only with marginal health gain. So if we accept that costs represent the sacrifices of those who are denied treatment, this naturally implies that ignoring the costs translates into ignoring the sacrifices of those who are denied treatment. However, medical ethics do not require doing everything possible irrespective of the consequences it may have for other potential patients (opportunity cost) [9].

In a democratic society the views of all affected parties should count. The general public is considered to be the most appropriate reference group. However, the QALY approach only requires us to be explicit about what the values are that are being applied, and where they came from. Many clinicians believe that it is unethical for them to replace the values of each individual patient with some collective set of values. However, collective priority-setting requires a collective view, therefore some method of aggregation has to be adopted. In the real world, only in a purely private market (with no charity and no insurance) have doctors been in a position where they could do whatever

the patient demanded. In all other circumstances doctors have been constrained by somebody else's willingness to pay [9].

The simplest and most common use of QALY calculations at present is based on the assumption that a year of healthy life expectancy should be regarded with equal value by everybody. A strong egalitarian case could be made for that assumption, since it implies that it does not matter at all who the beneficiary is. However in the real world there are many implicit preferences in a society: the young are preferred to the elderly, non-smokers to smokers, the employed to the unemployed, etc. when allocating limited healthcare resources. Some other opponents argue that distributional concerns should not focus primarily on health gains, but on the level of health itself. This would imply that the difference between the health capital of subjects is minimized. These opponents suggest not devoting resources to improve the health of those who have already had a long and healthy life when those resources could be used to improve the health of someone who, otherwise, will have a shorter and/or unhealthier life [9]. This would discriminate between old and healthy people compared to the young who are in a poor state of health.

Methodological concerns

Some opponents are against the QALY approach based on methodological grounds. As was first mentioned when we discussed utilities, different valuation methods provide different utilities/values even for the same health state [10]. QALY therefore represents a comparable unit of health outcomes only if the same valuation methods are applied. The NICE in England therefore considers EQ-5D with a TTO-based value set as the gold standard. Also, test-retest reliability (re-valuation of the same health state after a while) often generates poor results in real world settings. There are some difficulties in valuing health states for QALYs with children, and in some conditions (blindness (with RS), deafness (with SG), mental/psychiatric disorders).

Concerns with QALY based decision-making

Counting QALYs does not differentiate between lifesaving and improving quality of life. On the contrary, some argue that saving lives should be the priority. However, the maximization of QALYs requires that the young are preferred to the elderly (i.e. more life years to be saved) or the healthy to the disabled in life saving interventions (i.e. more QALYs to be saved) ('Double jeopardy') [11]. Therefore maximization of the QALYs should not be the only ultimate aim of healthcare. According to recent findings, the society is willing to trade maximization of health gain for equity. This is a key reason why some orphan drugs have received reimbursement. In conclusion, other aspects

should be considered besides maximizing the QALYs, and these additional criteria might be captured by a multiple criteria decision analysis (MCDA).

Cost generated by the QALY concept

Opponents also cite the additional cost of implementing the QALY concept. Certainly, measuring health outcomes, developing, adapting, validating health state valuation methods, mapping studies, ensuring the transferability of health outcomes, generating EQ-5D value sets is quite resource intensive. Also, capacity building, establishing and implementing HTAs, collecting cost vectors for economic evaluations, conducting and evaluating cost utility analyses require significant resources. However, this cost should be weighed in comparison to the opportunity cost of inappropriate decisions that were made because explicit decision-making (cost utility analysis) was avoided.

References

1.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Quality Adjusted Life Year (QALY). ISPOR, USA, 2003: 195-197.

2.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Utility. ISPOR, USA, 2003: 241-242.

3.   Drummond MF, et al. Cost utility analysis. Methods for economic evaluation of health care programs. 2nd edn. Oxford University Press, New York, 1997: 139-204.

4.   Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Disability Adjusted Life Years (DALY). ISPOR, USA, 2003: 69-71.

5.   Murray CLJ. (1994) Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bulletin of the World Health Organisation, 72 (3): 429-445.

6.   Arnesen T, Nord E. (1999) The value of DALY life: problems with ethics and validity of disability adjusted life years. BMJ, 319(7222): 1423-5.

7.   Lyttkens CH. (2003) Time to disable DALYs? On the use of disability-adjusted life-years in health policy. European Journal of Health Economics, 4(3): 195-202.

8.   Anand S, Hanson, K. (1997) Disability-adjusted life years: a critical review. J Health Economics, 16: 685-702.

9.   Williams A. (1996) QALYs and ethics: a health economist's perspective. Soc Sci Med, 43: 1795-1804

10.   Nord E, et al. (2009) QALYs: Some Challenges. Value health, 12(Suppl1): S10-5.

11. Harris J. (1987) QALYfying the value of life. J Med Ethics, 13: 117-123

Slide decks

7. Principles of PRO measurement

7.1. Background of PROs

Background of PROs

"Financed from the financial support ELTE won from the Higher Education Restructuring Fund of the Hungarian Government"

Learning objectives

- Students should be aware of
  - The recent paradigm shift in healthcare resulting in the increased use of PROs
  - The difference between the perspective of a patient and a physician
  - The different methods of patient observation
  - The typology of patient reported outcome measures
  - The definition of quality of life and heath-related quality of life
  - The classification of QoL measures
  - The PROs and CONs of generic and specific measures
  - The advantages and disadvantages of measuring PROs with interviews and questionnaires

Content

- Paradigm shift in medicine
- Patient reported outcomes (PRO)
- Quality of life (QoL)
- Health-related quality of life (HRQoL)
- Measurement of PROs

PARADIGM SHIFT IN MEDICINE

"We should set the highest value not on living but on living well"

Socrates

What is health?

'Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.'

WHO 1948

Paradigm shift in medicine

- The perspective and perception of the physician and patient are different
  - These two perspectives complement each other

Scope	Patient	Physician
Observed phenomenon/ entity	Subjective perception	Objective diagnosis/ signs
Instruments	PRO measures/ instruments	Physical or instrumental examination
Outcomes	PROs	Physiological parameters

Changing disease structure

- Lower incidence of acute high mortality diseases (infectious diseases)
- Higher prevalence of chronic conditions with disability
- "Live a good life. In the end it is not the years in a life, but the life in the years." (Lincoln)
- Improving quality of life is becoming increasingly important besides improving life longevity
- Valuing health states based on PROs is relatively new in the field of Medicine
- Psychometrics: Discipline in psychology to measure PROs – new application in clinical trials

Paradigm shift

- Patient's perspective
  - Role of patient becomes more important
  - Knowledge of patients is improving
  - Patients receive drugs according to their own desire
  - Patients act as informed customers
- Physician's perspective
  - High workload
  - Limited time for patients
  - Limited time to keep up-to-date with recent scientific findings and best practices
  - Physicians are encountering informed consumer
- Patient-physician relationship is becoming more balanced

Objective and subjective perception of disease

- Perspective of patient and physician is different and complementary
- Objective perception
  - Observations
  - Physiological parameters
- Subjective perception
  - Symptoms
  - Patient satisfaction
  - Adherence (also as a reaction of perception)
  - Health-related quality of life

Perception of disease is different between patients and physicians
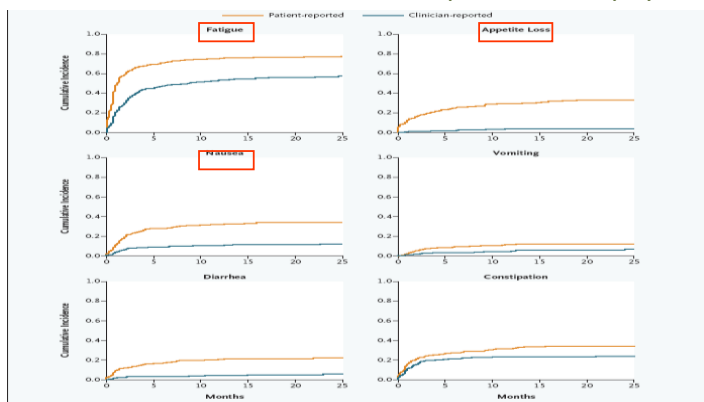
Cumulative Incidence of Adverse Symptom Events over Time as Reported by Patients versus Clinicians at Successive Office Visits. Patient-reported symptoms were collected directly from 467 persons with breast, lung, genitourinary, or gynecologic malignant conditions at a total of 4034 clinic visits at Memorial Sloan-Kettering Cancer Center, New York. Clinician-reported symptoms were recorded by physicians and nurses treating those patients at the same visits as a part of standard institutional documentation. Both patients and clinicians reported symptoms according to the National Cancer Institute's Common Terminology Criteria for Adverse

Ref: Basch E. (2010) N Engl J Med, 362: 865-869

Evaluation of antihypertensive therapies from different perspectives

- ACE inhibitors
- Effective drugs for hypertonia
- AE: dry cough

Hypertonia	Physician	Caregiver	Patient
Improved	75	1	36
Deteriorated	0	74	7
Unchanged	0	0	32
Total	75	75	75

Jachuck SJ et al. (1982) J Roy Coll Gen Pract, 32: 103–105.

Clinical Outcomes Assessment (COAs)
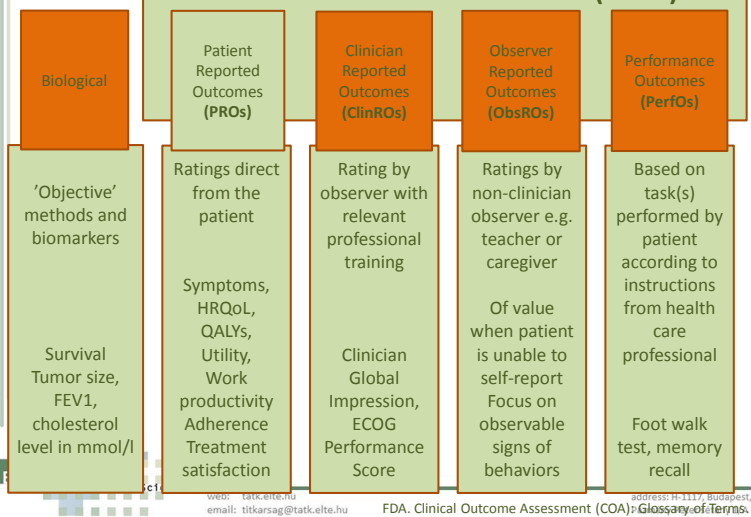
Biological	Patient Reported Outcomes (PROs)	Clinician Reported Outcomes (ClinROs)	Observer Reported Outcomes (ObsROs)	Performance Outcomes (PerfOs)
'Objective' methods and biomarkers	Ratings direct from the patient	Rating by observer with relevant professional training	Ratings by non-clinician observer e.g. teacher or caregiver	Based on task(s) performed by patient according to instructions from health care professional
Survival Tumor size, FEV1, cholesterol level in mmol/l	Symptoms, HRQoL, QALYs, Utility, Work productivity Adherence Treatment satisfaction	Clinician Global Impression, ECOG Performance Score	Of value when patient is unable to self-report Focus on observable signs of behaviors	Foot walk test, memory recall

FDA. Clinical Outcome Assessment (COA): Glossary of Terms.

PATIENT REPORTED OUTCOMES (PRO)

What is PRO (Patient Reported Outcome)?

- A PRO is any report on the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else.
- The outcome can be measured in absolute terms (e.g., severity of a symptom, sign, or state of a disease) or as a change from a previous measure.
- In clinical trials, a PRO instrument can be used to measure the effect of a medical intervention on one or more concepts (i.e., the thing being measured, such as a symptom or group of symptoms, effects on a particular function or group of functions, or a group of symptoms or functions shown to measure the severity of a health condition).

FDA. (2009) Guidance for Industry; Patient-Reported Outcome Measures
http://www.ispor.org/workpaper/FDA%20PRO%20Guidance.pdf

PROs

- Subjective perception of the patient – unique, valuable information
- Measurement with scientifically rigorous methodology (psychometrics)
- Similar regulation of measurement as for any other health outcomes
  - Informed consent of patient
  - Approval of Ethics Committee
  - Standardized methods for data collection and analysis

QUALITY OF LIFE (QOL)

Determinants of QoL

- Health status
- Society
- Economy
- Quality of life
- Culture
- Environment
- Other
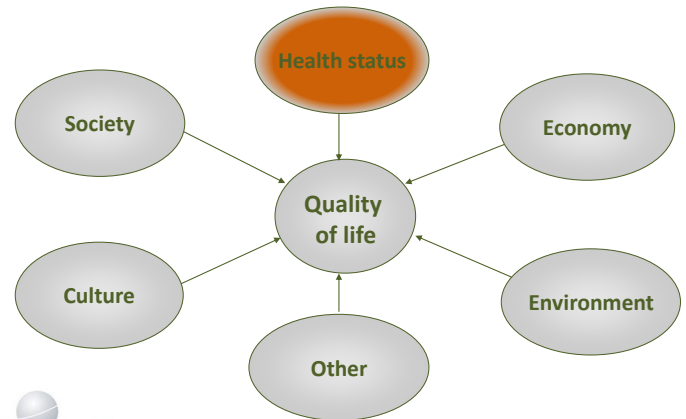
HEALTH-RELATED QUALITY OF LIFE (HRQOL)

Health-related quality of life

'Health-related quality of life (HRQoL) is a broad theoretical construct developed to explain and organize measures concerned with the evaluation of health status, attitudes, values and a perceived level of satisfaction and general well-being with respect to either specific health conditions or life as a whole from the individuals perspective'

Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Health Related Quality of Life (HRQoL). ISPOR, USA, 2003: 129-131.

'Health-related quality of life (HRQoL) measures extend patient outcome assessment beyond survival, adverse effects, and clinical efficacy, and reflect the patient's perspective on the impact of disease and its treatment on functioning and well-being.'

Revicki, DA. (2002) Value in Health. 5(4), 295-296.

Need for measuring HRQoL

- If no objective marker exists (e.g. physiological parameter)
- In chronic disorders
- In clinical decision-making
  - In clinical trials(->)
- In payer's decision-making
  - Cost utility analyses (->)
  - Burden of disease studies (->)

Other PROs: Patient satisfaction

- One of the most important quality assessment tools
- Multifaceted and a very challenging outcome to define
- Multiple determinants
- Various measures
  - Qualitative and quantitative questionnaires
  - Open questionnaires
  - Interviews
  - Direct observation
  - Analysis of medical documentation
- Relevance: affects medication adherence and therefore the clinical outcomes of patients

Other PROs: Medication adherence

- **Adherence**

  "the extent to which a person's behavior - taking medication, following a diet, and/or executing lifestyle changes - corresponds with agreed recommendations from a healthcare provider"[1]

- **Medication adherence**

  "refers to the act of conforming to the recommendations made by the provider with respect of timing, dosage and frequency of medication taking"[2]

- **Persistence**

  "the duration of time from initiation to discontinuation of therapy" [2]

WHO (2003) Adherence to long-term therapies. Evidence for action.
Cramer, JA. et al. (2008) Value in Health;11(1): 44-47.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

MEASUREMENT OF PROS

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Interview - CONs

- Hard to standardize
  - subjectivity of interviewer
- Hard to quantify
  - subjectivity of interviewer, heterogeneity of questions
- Subjective, more focus on qualitative research/exploration
- More expensive
  - need for (trained) interviewer
- Hard to compare
  - Diseases
  - Patients

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Interview - PROs

- Subject is unable to fill in a questionnaire
  - blindness,
  - Infant/child,
  - cognitive damage
- Supplement of a questionnaire
  - further relevant questions
- Scientific reasons
  - direct method for health state valuation (->)

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Questionnaires/Forms

- No need for interviewer
  - - > Relatively cheap
- No subjectivity of interviewer
  - - > Standardized, validated (->)
- Standardized
  - - > Can be evaluated quantitatively by statistical methods
  - - > Minimized influence on responder
  - - > Comparability among patients

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Consider application of PRO measure if...

- No objective marker exists (e.g. pain)
- Treatment improves
  - Disease symptoms
  - Quality of life
  - Patient satisfaction
  - Medication adherence
- Treatment has limited effect on survival (where the treatment benefit is mainly not increased longevity e.g. hip prosthesis)
- Treatment has subjective adverse reactions
- Effect of treatment is based on patient's perception

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

How to select the appropriate PRO measure?

- Research question/hypotheses (need for utilities) (**->**)
- Patient population (infants, blind, deaf)
- Adverse events (specific measure to capture effect of AEs)
- Quality (psychometric criteria: reliable, valid, practical, responsive) (**->**)
- Accessibility of appropriate measure for a particular disease
- Copyright, cost of license
- Availability of measure in local language (**->**)
- Quality of adapted version (psychometric criteria) (**->**)
- Length of measure (time to fill in – cognitive burden)
- Type of questions (irrelevant or offensive questions)

ELTE
Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Classification of measures 1: Index vs. profile

	Profile	Index
Generic measures	NHP, SF-36, SF-12	EQ-5D, SF-6D, HUI2, HUI3, QWB
Specific measures	Kidney disease questionnaire	RAQoL, SGRQ

- Profile: the measure presents different dimensions of health separately (SF-36: Physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, mental health)
- Index: provide a single summary index score
- Preference based scores: index scores with reference points within the instrument to death and perfect health (**->**) may be used to combine changes in quality and quantity of life (**->**) due to the possibility of linking, comparing and trading off these different aspects

ELTE
Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Classification of measures 2: Generic vs. specific

- Generic measures are designed
  - to be broadly applicable across types and severities of disease, different medical treatments or health interventions and demographic and cultural subgroups
  - to summarize a spectrum of core concepts of health and quality of life that apply to many different diseases, impairment, conditions, patients and populations
- Disease-specific measures
  - Focus on changes in severity of symptoms related to specific diseases or conditions
  - Differ not only by disease but also within disease groups

ELTE
Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Generic vs. Specific measures – PROs and CONs

	Generic measure	Specific measure
	Any conditions	One specific disease area
PROs	• Greater experience • Easier validation • Comparability (treatments and conditions) • May capture and compare several different domains within one condition • Population reference values are available	• Higher sensitivity • More suitable for clinical trial in a particular condition
CONs	• Less sensitive (may not detect small changes as clinically significant) • May over/ underemphasize particular dimensions	• Risk of overestimating minor health gain • Population and condition limits applicability • Not suitable for comparing different conditions • Captures only a limited number of aspects of a condition to reduce the burden of the responder • Limited applicability for healthcare decision-making (economic analyses)

Attributes of health state classification system

Torrance GW (1986) Journal of Health Economics

ELTE
Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A
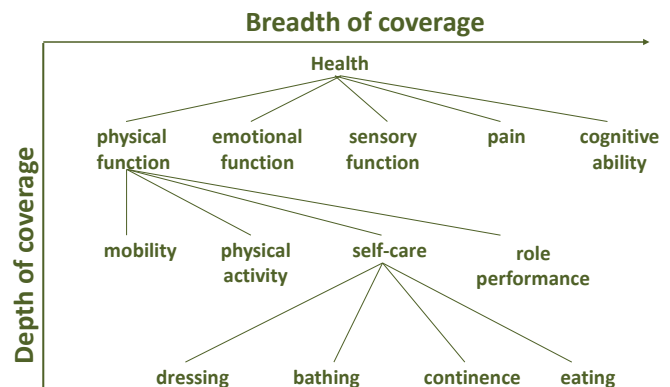
Measuring health-related quality of life

- Decision
  - To develop a new measure
  - To apply an existing measure
- If the decision is to develop a new measure: validation (**->**)
- If it is a non-preference measure: generation of values for health states described by the measure
- If it is an international multicenter trial: adaptation of measures to different languages (**->**)
- Adapted versions of the measure must be validated

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Measuring health-related quality of life

- Using consecutive responders to prevent selection bias
- Capturing demographic and clinical data
- Giving instructions to responders
  - The are no good or bad answers/expected answers
  - Results will be used for important decisions (be accurate)
- Checking of the forms
- Ensuring data protection (PROs are sensitive data)

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Take home messages

- The perspectives of patient and physician are different
- The relationship between patient and physician is in transition
- The importance of patient reported outcomes is becoming more and more important
- Methods for patient observation:
  - Patient reported outcomes
  - Caregiver reported outcomes
  - Clinician reported outcomes
  - Physiological data
- PROs
  - Adherence
  - Patient satisfaction
  - QoL
    - HRQoL

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Take home messages (2)

- A PRO is any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else
- Measurement with scientifically rigorous methodology (psychometrics)
  - Interview
  - Questionnaire
- Classification of measures
  - Generic/Specific
  - Profile/Index (preference based)

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Self-check questions

- What causes have led to the increased relevance and use of PROs?
- What are the advantages and disadvantages of generic and specific measures?
- What are the different methods of patient observation?
- When should we use interviews and when should we use questionnaires to measure PROs?
- How can patient reported outcome measures be classified?
- What should be considered when selecting PRO measures?

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Suggested reading

- FDA Guidance on PROs. 4. FDA: Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims, http://www.fda.gov/downloads/Drugs/.../Guidances/UCM193282.pdf

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

7.2. Concept of PRO measurement

Concept of PRO measurement

"Financed from the financial support ELTE won from the Higher Education Restructuring Fund of the Hungarian Government"

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Learning objectives

Students should:
- Understand the measurement methods of PRO
- Understand the types of data & measurement scales
- Be aware of how to select the right PRO instrument for a study

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Content

- Background: What are PROs?
- Measurement methods of PROs
- Types of data & measurement scales
- Selection of a PRO instrument

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Background: What are PROs?

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

What are PROs?

- Definition: A PRO is "any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else".
- PROs help us to understand how the patients feel they are doing.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

What are PROs?

- PRO is an umbrella term:
  - Quality of life (QoL)
  - Symptoms
  - Pain
  - Function
  - Treatment satisfaction
  - Adherence to medications or other therapies
  - ....

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Measurement methods of PROs

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Interview vs. questionnaire (1)

- Interview:
  - Involves direct questioning of participants
  - Can obtain more private feedback on aspects
  - Structured interview: All participants are asked the same questions from interviewers
  - Semi-structured interview: A fairly open framework where not all questions are phrased beforehand
  - Questions: Open questions (no given answers), or questions with multiple-choice answers
  - Bias: Interviewer (may interpret the behaviors of some participants as meaning one thing when it means something else)

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Interview vs. questionnaire (2)

- Questionnaire:
  - Used to collect information from a large sample of people
  - Data can be gathered quickly and relatively cheaply
  - Questionnaires collect more reliable information for research since they have fewer variables which can alter the information or cause false data to be collected
    - The collected information is not dependent on the mood of the interviewer

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Instrument vs. questionnaire (3)

- PRO instrument:
  - A means to capture data (i.e., a questionnaire) plus all the information and documentation that supports its use. Generally, it includes clearly defined methods and instructions for administration or responding, a standard format of collection, and well-documented methods, analysis, and interpretation of results in the target population.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Measuring PROs (1)

- Types of PRO instruments:
  - As per the measurement of the interested concept (horizontal-vertical coverage)
    - Generic instrument
    - Specific instrument

Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Classification of PROs

Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Generic PRO instruments

- Designed to measure very broad aspects of health
- Suitable for a wide range of patient groups
- Comparison across conditions is possible
- Can be used to generate normative data for the general population
- May not be sensitive enough

Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Specific instruments (1)

- Specific instruments:
  - Disease specific
    - Developed to measure the patient's perceptions of a specific disease or health problem (e.g., St George's Respiratory Questionnaire in COPD)
  - Population specific
    - Designed to be appropriate for particular demographic groups, such as children or elderly people (e.g., Child Health and Illness Profile-Child Edition)

Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Specific instruments (2)

- Specific instruments:
  - Function, symptom, or problem specific
    - Developed to assess one particular aspect of health status (e.g., Beck Depression Inventory)

Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Disease specific instruments

- Address a particular disease
- Responsive to clinically important changes in health
- Clinically sensitive
- Cannot be used in samples who do not have the relevant health problem
- Health status scores cannot be compared with those for the general population
- Do not allow for cross-condition comparisons

Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Generic vs. disease specific instruments

	Pros	Cons
Generic instrument	Can be used with healthy populations to generate normative data Comparison across conditions is possible	May not focus adequately on area of interest May not be responsive
Disease specific instrument	Clinically sensitive May be responsive	Does not allow for cross-condition comparison

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Measuring PROs (2)

- Types of PRO instruments:
  - Result presentation:
    - Profile
    - Index

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Index vs. profile type PRO instruments

PRO instruments	
Index (e.g., EQ-5D)	Express health states as a single index score
Profile (e.g., Nottingham Health Profile)	Provides a profile score for different dimensions of health state Can be used to detect differential effects on different aspects of health status

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Measuring PROs (3)

- Types of PRO instruments:
  - What does the instrument measure?
    - Discriminative instrument
    - Evaluative instrument
    - Predictive instrument

Fayers P, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. 2nd edition. Wiley, 2007.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Discriminative instruments

- PRO instruments which are intended to differentiate between people:
  - E.g., better QoL <-> worse QoL

- Less important to include symptoms that are common to all patients and unlikely to differ between various treatment groups

Fayers P, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. 2nd edition. Wiley, 2007.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Evaluative instruments

- PRO instruments which are intended to measure changes:
  - E.g., how much does QoL change over time?

Fayers P, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. 2nd edition. Wiley, 2007.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Predictive instruments

- PRO instruments are used to categorize individuals based on predetermined criteria
  - E.g., estimating mortality based on QoL in oncology patients

Fayers P, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. 2nd edition. Wiley, 2007.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Measuring PROs (4)

- Types of PRO instruments:
  - Complexity of the instrument:
    - Single-, or multi-item
    - Single-, or multi-domain

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
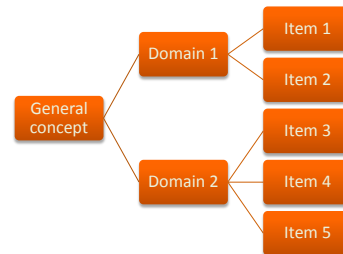address: H-1117, Budapest, Pázmány Péter sétány 1/A

Definitions of item & domain

- Item:
  - An individual question, statement, or task (and its standardized response options) that is evaluated by the patient to address a particular concept.
- Domain:
  - A sub-concept represented by a score of an instrument that measures a larger concept comprised of multiple domains. Domains are subdivided into items.

FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Conceptual framework

- Conceptual framework: an explicit description or diagram of the relationships between the questionnaire or items in a PRO instrument and the concepts measured.

General concept → Domain 1 → Item 1, Item 2; Domain 2 → Item 3, Item 4, Item 5

FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Example for conceptual framework

Sleep disturbance (concept):
- Falling asleep (domain)
  - Item 1: How difficult was it to fall asleep?
  - Item 2: How difficult was it to get comfortable?
- Staying asleep (domain)
  - Item 3: How difficult was it to stay asleep?
  - Item 4: How restless was your sleep?
- Impact (domain)
  - Item 5: How rested were you when you woke up?
  - Item 6: How difficult was it to start your day?

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Single-item PRO instruments

- Simple concepts can be assessed with a single-item scale on a particular concept of interest (the thing that is measured by the PRO), whose score is estimated by a single response to a single question
  - E.g., pain is often expressed as a single-item

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Multi-item PRO instruments (1)

- Multi-item scales are scales formed by more than one item and are used chiefly for concepts considered complex to measure
  - E.g., physical function
    - Adequate assessment of physical function requires a number of items to ensure that all of its relevant aspects are captured for the targeted patient population

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Multi-domain PRO instruments (2)

- Where multiple items are needed and they are referring to different aspects of the same concept or attribute, they form multiple domains
  - E.g., SF-36 (see next slide)

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Multi-domain PRO instruments (3)

http://www.sf-36.org/tools/sf36.shtml

Latent vs. manifest variables

- Measurements of PROs often involve unobserved constructs or concepts that are also referred to as latent (hidden) variables
  - Many psychological aspects are not directly observable and are hence considered latent variables

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Types of administration

- Self-administered - hospital
- Interview - hospital
- Self-administered - postal survey
- Telephone interview
- Web-based

Frequency of administration

- Patient-physician visit - once
- Patient-physician visit - several times
- Between patient-physician visit
- Defined intervals

Types of data & measurement scales

– Qualitative data
  • Nominal scale
  • Ordinal scale
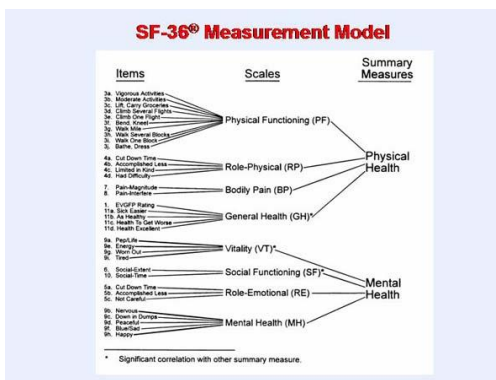
– Quantitative data
  • Interval scale
  • Ratio scale

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Types of data & measurement scales

Nominal scale

• Nominal scales are used for labelling variables, without any quantitative value (qualitative scale).
• A sub-type of nominal scale with only two categories is called "dichotomous"
  – E.g., what is your gender?
    • Male
    • Female

Ordinal scale

• With ordinal scales, it is the order of the values is what's important, but the differences between the values is not really known
• E.g., How do you feel today?
    • Unhappy
    • OK
    • Happy
• For measuring central tendency on a set of ordinal data, only the median and the mode can be used

Interval scale

• Interval scales are numeric scales (quantitative scale) in which we know not only the order, but also the exact differences between the values.
• Interval scales don't have a "true zero" point
  – E.g., Celsius temperature scale

Ratio scale

• Ratio scales tell us about the order, the exact value between units, and they also have an absolute zero point
• Absolute zero point: a point where none of the quality being measured exists
  – E.g., how many hours a day do you spend on a computer?

Special scales

- Special scales commonly used in PRO questionnaires:
  - Likert scale
  - Visual analog scale (VAS)

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Likert scale (1)

- The Likert scale is generally a five (or seven) point scale
- Likert scales use fixed choice response formats and are designed to measure attitudes
- It is important that these scales focus on "symmetry" and "balance"

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Likert scale (2)

- "Symmetry"
  - They contain equal numbers of positive and negative positions whose respective distances apart are bilaterally symmetric about the "neutral" value.
- "Balance"
  - The distance between each candidate value is the same
    - Allowing for quantitative comparisons across items containing more than two candidate values.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Likert scale (3)

- Examples:
  - Agreement:
    - Strongly agree / agree / undecided / disagree / strongly disagree
  - Frequency:
    - Very frequently / frequently / occasionally / rarely / never
  - Likelihood:
    - Almost always true / usually true / occasionally true / usually not true / almost never true

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Visual analog scale

- Visual analog scale (VAS) is a psychometric response scale
  - Respondents specify their level of agreement to a statement by indicating a position along a continuous line between two end-points
- The continuous aspect of the VAS differentiates it from discrete scales such as the Likert scale
- The simplest VAS is a straight horizontal line of fixed length, usually 100 mm
  - The ends are defined as the extreme limits of the parameter to be measured (e.g., symptom, pain) orientated from the left (worst) to the right (best)

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Selection of a PRO instrument

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Selection of a PRO instrument (1)

- Documentation
  - Is there formal written documentation, publications or a user manual?
- Development
  - How rigorous was the development procedure?
- Validity
  - Is there sufficient evidence that the scale is measuring what it is intended to measure?

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Selection of a PRO instrument (2)

- Reliability
  - Is there sufficient evidence that the scale is precise in accurately measuring its scores?
- Feasibility
  - Does the scale have questions that are easy to understand and a convenient mode of administration?
- Target population
  - Is the scale suitable for the target population?

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Selection of a PRO instrument (3)

- Language and cultures
  - Are there validated translations of the questionnaire?
- Scoring
  - How is the scoring procedure defined?
- Interpretation
  - Are there guidelines for interpreting scale scores?

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Take home messages

- PROs can be measured with interviews or questionnaires
- Types of PRO instruments as per the measurement of the interested concept: generic and specific instruments
- Types of measurement scales: nominal, ordinal, interval and ratio scales
- Critical points when selecting a PRO instrument: documentation, development, validity, reliability, feasibility, target population, language and cultures, scoring and interpretation

Self-check questions

- What are the pros and cons of interviews and questionnaires in measuring PROs?
- What is the difference between a PRO questionnaire and an instrument?
- What are the differences between profile and index type questionnaires?
- What are the pros and cons of generic- and disease-specific questionnaires?
- What is the difference between an item and a domain?
- What types of measurement scales do you know?
- What are the key characteristics of a Likert scale?
- What are the critical points when selecting a PRO instrument?

Suggested reading

- Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.
- Fayers P, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. 2nd edition. Wiley, 2007.
- FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009.

7.3. Development of PRO instruments

Development of PRO instruments

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Learning objectives

Students should:
- Understand the key steps of PRO development
- Be aware of types of validity and reliability

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Content

- Background: PRO instrument, PRO concept
- Developing a new PRO instrument
- Validity and reliability

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Background: PRO instrument, PRO concept

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

PRO instrument

- PRO instrument:
  - A means to collect data about a PRO concept
  - A questionnaire plus the information and documentation that support its use
  - Includes clearly defined methods and instructions for administration or responding, a standard format of collection, and well-documented methods, analysis, and interpretation of results

FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

PRO concept

- PRO concept:
  - The thing that is to be measured by a PRO instrument
  - Represents aspects of how patients function or feel related to a health condition or its treatment

FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Developing a new PRO instrument

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Developing a new PRO instrument

- Rigor in the development of a PRO instrument is essential
  - To ensure that the concept of interest is measured accurately
  - To capture the issues of most relevance to the patient
  - To subscribe to a language that allows patients to understand and respond without confusion

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

US-FDA: steps of PRO development

i. **Hypothesize Conceptual Framework**
- Outline hypothesized concepts and potential claims
- Determine intended population
- Determine intended application/characteristics (type of scores, mode and frequency of administration)
- Perform literature/expert review
- Develop hypothesized conceptual framework
- Place PROs within preliminary endpoint model
- Document preliminary instrument development

v. **Modify Instrument**
- Change wording of items, populations, response options, recall period, or mode/method of administration/data collection
- Translate and culturally adapt to other languages
- Evaluate modifications as appropriate
- Document all changes

ii. **Adjust Conceptual Framework and Draft Instrument**
- Obtain patient input
- Generate new items
- Select recall period, response options and format
- Select mode/method of administration/data collection
- Conduct patient cognitive interviewing
- Pilot test draft instrument
- Document content validity

iv. **Collect, Analyze, and Interpret Data**
- Prepare protocol and statistical analysis plan (final endpoint model and responder definition)
- Collect and analyze data
- Evaluate treatment response using cumulative distribution and responder definition
- Document interpretation of treatment benefit in relation to claim

iii. **Confirm Conceptual Framework and Assess Other Measurement Properties**
- Confirm conceptual framework with scoring rule
- Assess score reliability, construct validity, and ability to detect change
- Finalize instrument content, formats, scoring, procedures and training materials
- Document measurement development

PRO ↕ Claim

FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Key steps of developing a new PRO instrument

- Key steps:
  - Determine goals
  - Determine the context of use
    - Understand the disease or condition
    - Consider the target population
  - Item generation
    - Develop the conceptual framework
  - Item reduction, wording
  - Conduct cognitive interviews for evaluation of content validity

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Determine the context of use

- Understanding of the disease or condition in the target population
  - Pathophysiology of the disease
  - Risk factors
  - Diagnosis
  - Signs and symptoms
  - Possible therapies
    - Drug mechanism of action
    - Effects and side-effects of the treatment

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Disease model, psoriasis

Patrick DL, et al. (2011) Value Health, 14(8): 967-977.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A
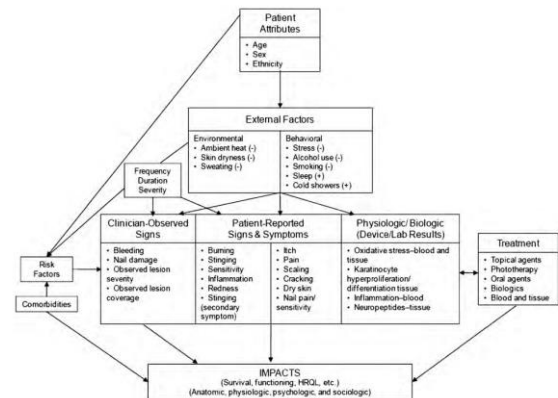
Consider the target population

- Characteristics of the target population
  - Language
  - Culture
  - Age groups: e.g., children, adolescents
  - Cognitive functions
  - ....

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Item generation

- To generate the content of a PRO
  - Individual interviews
  - Group discussion with patients (typically called focus groups)
- Generally, a mixture of the two approaches is beneficial

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Interviews vs focus groups

	Focus groups	Interviews
Pros	• Rich source of data • Allows the use of ideas from others as cues to express their own views • Participants can compare their experiences with others • Able to reach many participants at once	• Get more in-depth and detailed information about an individual's experience • Can be useful for sensitive topics • Data can be easier to analyze • Scheduling can be easier
Cons	• Data can be tough to analyze because the discussion will also contain reactions to the comments of other group members • Moderators need to be highly trained and able to lead the group • One strong group member can sway the tone of the entire group	• It may take longer to collect data • Limited to one participant's view at a time • Interviewers need to be trained with excellent communication skills • May be more costly

Patrick DL, et al. (2011) Value Health, 14(8): 967-977.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Conducting interviews/focus groups

- Requires a research protocol
  - Inclusion/exclusion criteria, No. of subjects, information about the questions to be asked, interview guide
- Obtain ethical approval
- Audio record or videotape interviews/focus groups to allow transcription and analysis

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Analysis of qualitative data (1)

- Developing a coding scheme:
  - Similar terms are given a code name
  - Quotes from subjects related to these specific terms are added under these codes
  - Allows us to see how how many different terms were used and how many times
  - Codes are usually developed iteratively that can be changed during data analysis
  - Codes will then be grouped into concepts and from these concepts a theory about the data is developed

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Analysis of qualitative data (2)

- Once the codes are agreed upon
  - Reviewers of the transcripts identify and agree on the common themes emerging from the data
  - This generates a list of patient statements per code, which allows assessment of how frequently the concept was discussed and by how many subjects
  - Reviewers can then determine whether a concept or category should form part of the new measure

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Saturation

- When interviewing patients, we reach the saturation point when no new relevant or important information emerges and collecting additional data will not add to the understanding of how patients perceive the concept of interest and the items in a questionnaire

FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009.

Process of determining saturation

- Example:
  - Step 1: Transcripts of the first 10 interviews are analyzed to identify consistency in the pattern of the responses to the concepts presented during the interview.
  - Step 2: The second set of 10 interviews will then be analyzed to determine if any new concepts have been identified. If they have not, then no further interviews are required.

Cappelleri JC. et al. (2014) Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series

Development of the conceptual framework

- Conceptual framework:
  - An explicit description or diagram of the relationships between the questions or items in a PRO instrument and the concepts measured

FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009.

Definitions of item & domain

- Item:
  - An individual question, statement, or task (and its standardized response options) that is evaluated by the patient to address a particular concept.
- Domain:
  - A sub-concept represented by a score of an instrument that measures a larger concept comprised of multiple domains. Domains are sub-divided into items.
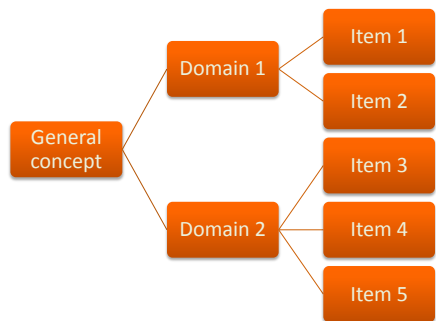
FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009.

Conceptual framework

ELTE
Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Item reduction

- Items are generated based on the language used by subjects from the individual interviews or focus groups.
  - Numerous items are formulated per concept/domain, often with significant overlap in wording
- It is not always clear which terminology is most appropriate
  - A further study can be conducted to help choose between the various options

ELTE
Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Item wording

- Items should be worded carefully and clearly
- Avoid double negations
- Character size: do not use too small characters
- Use effects for highlighting words (e.g. **Bold**, *Italic*, <u>Underline</u>)

ELTE
Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Cognitive interviews (1)

- Cognitive interview results for evaluation of content validity
  - Content validity: The degree to which the content of a measurement instrument is an adequate reflection of the construct to be measured (please find more detailed information on content validity in the section "Validity and reliability")

ELTE
Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Cognitive interviews (2)

- Patients are asked to complete the questionnaire, and while doing so, they are instructed to share what they are thinking and to explain how they are interpreting the content of the measure.
  - One-to-one interviews are recommended rather than focus groups

ELTE
Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Cognitive interviews (3)

- If there are issues with any part of the PRO, the interviewer can ask how they would reword something to make it clearer
- Once the content of the questionnaire has been confirmed via the cognitive interviews, the draft is ready for psychometric testing

ELTE
Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Validity and reliability

Validity

- Validity assesses the extent to which an instrument measures what it is meant to measure

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Types of validity (1)

Content validity (1)

- Evidence from qualitative research demonstrating that the instrument measures the concept of interest including evidence that the items and domains of an instrument are appropriate and comprehensive relative to its intended measurement concept, population, and use.

FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009.
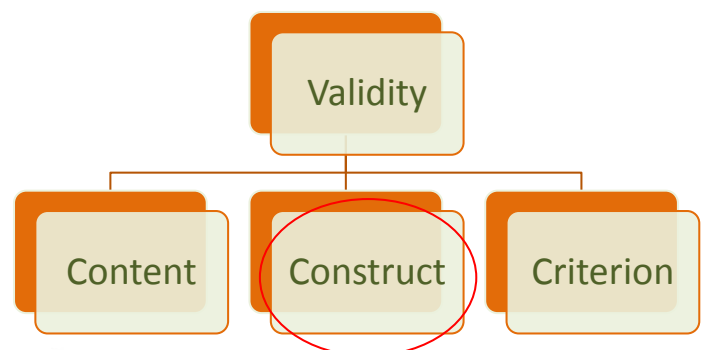
Content validity (2)

- A component of content validity is face validity
  - The degree to which a measurement instrument, indeed, looks as though it is an adequate reflection of the construct to be measured
  - Concerns whether items in an instrument appear on the face of it to cover the intended topics clearly and unambiguously

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.
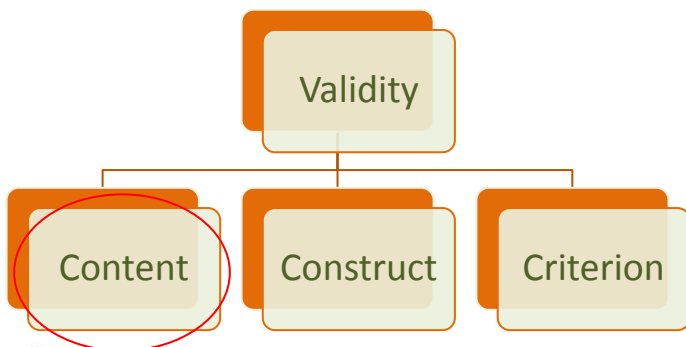

Types of validity (2)

Construct validity (1)

- Evidence that relationships among items, domains, and concepts conform to a priori hypothesis concerning logical relationships that should exist with other measures or characteristics of patients and patient groups.
- If there is a mismatch between the targeted PRO scale and its intended construct, then the problem could be that
  - The scale is good but the theory is wrong
  - The theory is good but the scale is not
  - Both the theory and the scale are useless or misplaced

FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009.

Construct validity (2)

- Mainly, assessments make use of correlations, changes over time, and differences between groups of patients
- Types of construct validity:
  - Convergent validity
  - Divergent validity
  - Known-group validity

Construct validity (3)

- Convergent validity:
  - Regards how much the target scale relates to other variables or measures to which it is expected to be related, according to the theory postulated
  - A correlation between 0.4 and 0.8 would seem to be reasonable
  - E.g., patients with higher levels of pain might be expected to also have higher levels of physical impairment

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Construct validity (4)

- Divergent validity:
 - Regards how much the target scale relates to other variables or measures to which it is expected to have a weak or no relation
  - E.g., little or no correlation might be expected between pain and intelligence scores

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Construct validity (5)

- Known-group validity:
  - The measurement scale of interest should be sensitive to differences between specific groups of subjects known to be different in a relevant way
  - The scale is expected to show differences, in the predicted direction, between these known groups
  - E.g., if a PRO instrument is intended to measure functional impairment, mean scores on it should be able to differentiate sufficiently between subjects with different levels of functional impairment (mild, moderate, severe)

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Types of validity(3)

Criterion validity (1)

- Criterion validity is the extent to which the scores of a PRO instrument are related to a known gold standard measure of the same concept.
- Types of criterion validity:
  - Concurrent validity
  - Predictive validity

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Criterion validity (2)

- Concurrent validity:
  - Involves an assessment of scores from the targeted PRO measure with the scores from a gold standard PRO measure administered at the same time
  - E.g., the result of a new single-item, disease-specific global quality of life questionnaire in patients with cystic fibrosis has to correlate with the result of the Cystic Fibrosis Quality of Life Questionnaire (gold standard)
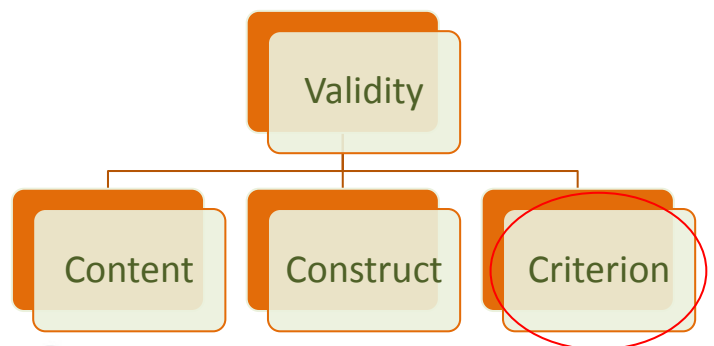
Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.
Johannes AM, et al. (2011) Health Qual Life Outcomes, 9: 105.

Criterion validity (3)

- Predictive validity:
  - Regards how well the target measure predicts the results of a gold standard measure in the future
  - E.g., in general health-related quality of life of oncology patients predicts mortality risk; therefore, the result of a new health-related quality of life instrument has to be a strong predictor of mortality in oncology patient

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.
Kaplan MS, et al. (2007) Qual Life Res, 16(9): 1539-1546.

Reliability (1)

- Reliability is the ability of a PRO instrument to yield consistent, reproducible estimates of true treatment effect.
  - It assesses how precise or stable the instrument measures what it is supposed to measure and is typically discussed in terms of reproducibility

FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009.

Reliability (2)

- Internal reliability:
  - Is based on item-to-item correlations and the numbers of items in the questionnaire

- Repeatability reliability:
  - Is based on the analysis of variances between repeated measurements on the same subjects

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Reliability (3)

- Repeatability reliability:
  - *Test-retest reliability*
    - If a patient is in a stable condition, an instrument should yield reproducible results when it is repeated on that patient
    - Selecting the right period between test and retest, or between times or occasions, is crucial
      - a period too long would increase the chance of a true change in the status
      - a period too short would allow subjects to recall their responses

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

Reliability (4)

- Repeatability reliability:
  - *Equivalent-forms reliability*
    - Involve different variants of the same attribute or construct
    - Absolute/relative agreement between scores from two or more instruments that are designed to measure the same attribute
      - Not equivalent to convergent validity which addresses how much the target scale correlates with another similar measure to which it is expected to be related

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Reliability (5)

- Cronbach's alpha coefficient:
  - Is the most widely used method to assess internal consistency reliability
  - Presumes that the multi-item scale reflects a single concept and is therefore unidimensional

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Reliability (6)

- Cronbach's alpha coefficient:

$$\text{Cronbach's alpha} = \frac{n}{n-1}\left(1 - \frac{\text{Sum of item variances}}{\text{Sum of variances and covariances}}\right)$$

  - Values are between 0 and 1
    - Good: 0.9-0.95
    - Acceptable: 0.75-0.95

Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Take home messages

- Key steps of PRO development: (1) determine goals, (2) determine the context of use, (3) item generation, (4) item reduction and wording, and (5) cognitive interviews for evaluation of content validity
- Validity assesses the extent to which an instrument measures what it is meant to measure. The main types of validity are: (1) content validity, (2) construct validity, and (3) criterion validity.
- Reliability assesses how precise or stable the instrument measures what it is supposed to measure and is typically discussed in terms of reproducibility. The main types of reliability are: (1) internal reliability, and (2) repeatability reliability.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Self-check questions

- What is the difference between a PRO instrument and PRO concept?
- What are the key steps of PRO development?
- What are the pros and cons of focus groups and interviews?
- What does saturation mean in PRO development?
- How can you define conceptual framework?
- What are the types of validity? What do these assess?
- What are the types of reliability? What do these assess?

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Suggested reading (1)

- FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009.
- Patrick DL, et al. (2011) Content validity - Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 1 - Eliciting concepts for a new PRO instrument. Value Health, 14(8): 967-977.
- Patrick DL, et al. (2011) Content validity - Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 2 - Assessing respondent understanding. Value Health, 14(8): 978-988.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Suggested reading (2)

- Cappelleri JC. et al. Patient-reported outcomes: measurement, implementation and interpretation. Chapman&Hall/CRC Biostatistics Series, 2014.
- Example on PRO development and validation:
  - Osborne et al. (2015) Improving the assessment of quality of life in the clinical care of myeloma patients: the development and validation of the Myeloma Patient Outcome Scale (MyPOS). BMC Cancer, 15:280

ELTE
Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

7.4. Translation and cultural adaptation of PRO instruments

Translation and cultural adaptation of PRO instruments

"Financed from the financial support ELTE won from the Higher Education Restructuring Fund of the Hungarian Government"

Learning objectives

- Students should:
  - Know the rationale behind the need for PRO instrument translation/adaptation
  - Know the levels of equivalence
  - Know the substantial difference between translation and cultural adaptation
  - Know the main steps of instrument cultural adaptation
  - Know the main differences between forward-backward translation and dual panel approach in cultural adaptation

Content

- Need for translation
- Levels of equivalence and adaptation
- Recommendations and good practices for instrument cultural adaptation
- Examples for forward-backward translation and dual panel approaches

Need for translation

Background

- International researches require data
  - to be compared
  - to be pooled
- Instrument development is
  - complex
  - expensive
  - a time-consuming process
- Specific costs of an instrument development for languages that are spoken by few people are even much higher

Need for translation

Aim

- To produce
  - totally equivalent PRO results
  - regardless of different languages/cultural background

Methodology

⇒ Conducting a special kind of translation process that results in a translated measurement version that is fully adapted to the target language and cultural environment

Levels of equivalence

Conceptual equivalence

- Construct exists and is relevant and acceptable in both cultures
- Items represent the definition of the construct well

Semantic equivalence

- Items mean the same thing to the people in both population
- The same expression exists in the target culture
- Language technical features (complexity, syntax, grammar) are equivalent

Operational equivalence

- Standardized methods of survey administration (e.g. questionnaire format, instructions, respondent burden) are appropriate for target culture

Stewart AL, Napoles-Springer A. (2000) Med Care, 38(9): 102-124.

Levels of equivalence

Psychometric equivalence

- There are comparable psychometric properties (e.g., construct validity, test-retest reliability, internal consistency)
- There are comparable effect sizes

Item equivalence

- Items are not more difficult in target language
- There are comparable item weights among items
- There is similar meaning of (and distance between) response categories

Criterion equivalence

- There is the same interpretation of scores
- There is the same relationship with independent criterion/criteria

Stewart AL, Napoles-Springer A. (2000) Med Care, 38(9): 102-124.

Levels of adaptation

Literal translation

- In the simplest way possible
- Its application is limited, and probably causes several types of misinterpretation

Linguistic adaptation

- Expressions that are uninterpretable word-for-word ("feeling blue", "heart burn", "loose track")
- Need to find the most appropriate target language idioms (even with more words) to cover the same concept that the original version did

Inotai A, Lovas K, Kaló Z. Az egészségnyereség mérése a betegek értékelése alapján [Patient reported outcomes]. Springmed, 2014.

Levels of adaptation

Experimental adaptation

- When the translation is adequate from a conceptual point of view
- But solutions should be found to adapt concepts that are irrelevant in the target cultural environment (eating with a "fork" vs. "chopstick"; being hampered in "driving a car" vs. "getting on a bus")

Conceptual adaptation

- When the same phrase has a distinct meaning ("symptoms", "rubeola", "brother", "girl friend")

Inotai A, Lovas K, Kaló Z. Az egészségnyereség mérése a betegek értékelése alapján [Patient reported outcomes]. Springmed, 2014.

Comparing the phases of original development and translation/adaptation

Original development	Phases	Translation/adaptation
• Exploration of the field of interest (disease, health state, population) • Eliciting the relevant domains • Coding • Item generation • Response option selection • First version of the measure	Preparing the final draft version	• Translation • Reviewing the first target language version • Comparing it against the original version • Completion of the final draft version
• Psychometric validation in the target population (field test/cognitive interview) • Completion of the final version of the original measure	Validation process	• Psychometric validation in the target population • Completion of the final target language version (field test/cognitive interview)
SIMILARITIES, DIFFERENCES		

General adaptation steps in different guidelines

Recruitment criteria forward translation

- Forward translators (e.g.): bilingual / target language native speaker / certified translators / expert in psychology, psychiatry, psychometry / panel members/participants
- No. of forward translators: 2 / minimum 2 / 8-12 from the panel members / 1 from the USA and 1 from the target country / not specified → most typically 2

⇒ first draft target language version

Synthesis

- Done by (e.g.): the translators / the coordinator / third translators / a focus group / bilingual panel / not specified / no synthesis

⇒ reconciled first draft target language version

Acquadro C, et al. (2008) Value Health, 11(3): 509-521.

General adaptation steps in different guidelines

Back-translation

- No. of backward translators (e.g.): 1 / 2 / 2 of whom 1 is a native English speaker / 20 raters involved / not specified / not recommended
- Working conditions: ± working independently / having no knowledge of the other translators involved in the back translation

⇒ Back-translated version to should then be harmonized to the original one

Review

- By whom (e.g.): health professionals / developer (team) / project coordinator / experts / no review

⇒ Final draft target language version to be pre-tested

Acquadro C, et al. (2008) Value Health, 11(3): 509-521.

General adaptation steps in different guidelines

Pretesting

- Involving: subjects from the target population / interview with individuals / subjects with low level of education / unspecified lay panel / involving both clinicians and respondents / bilingual subjects

⇒ Final target language version

Acquadro C, et al. (2008) Value Health, 11(3): 509-521.

ISPOR Adaptation Good Practice

" Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation"
(International Society For Pharmacoeconomics and Outcomes Research , 2005)

Summary

- Gather the major guidelines through a literature review
- Provide detailed and well-structured description of adaptation methodology, labels and key-actors and explanation of several PROs

Wild D, et al. (2005) Value Health, 8(2): 94-104.

ISPOR Adaptation Good Practice

Part I: Definitions

- For concepts used in the adaptation process
- For key actors in the adaptation process

Part II: Translation and cultural adaptation

- 10 steps from preparation to reporting
- listing critical components within the steps
- specifying the rationale, the responsible persons and the risks <u>if not doing for each critical component</u>

Wild D, et al. (2005) Value Health, 8(2): 94-104.

ISPOR Adaptation Good Practice

Part II – Step 1: Preparation

- Obtaining permission (legal aspects)
- Inviting instrument developer to be involved
- Development of explanation of concept
- Recruiting key in-country persons to the project

Part II – Step 2: Forward translation

- At least two independent forward translations
- Provision of translation explanations (underlying the importance of conceptual equivalence)

Part II – Step 3: Reconciliation

- Making a single forward translation out of the various translations
- Resolving discrepancies, seeking agreement

Wild D, et al. (2005) Value Health, 8(2): 94-104.

ISPOR Adaptation Good Practice

Part II – Step 4: Back translation
- Providing evidences that the original and the back-translation versions have the same meaning
- (some concepts require more literal rather than conceptual translation)

Part II – Step 5: Back-translation review
- Against the source language version to ensure conceptual equivalence

Part II – Step 6: Harmonization
- Of all new translations with each other and the source version to detect and deal with any translation discrepancies

Wild D, et al. (2005) Value Health, 8(2): 94-104.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

ISPOR Adaptation Good Practice

Part II – Step 7 – Cognitive debriefing
- Usually with patients drawn from the target population
- To assess the comprehensibility, to highlight any inappropriate items, to test translation alternatives

Part II – Step 8 – Review of cognitive debriefing results and finalization
- Reviewing and completing the translation, incorporating any findings of the cognitive debriefing

Part II – Step 9 – Proofreading
- Detecting any minor errors

Part II – Step 10 – Final Report
- Clearly explaining the whole process
- Helping future translations of the same measure to be harmonized

Wild D, et al. (2005) Value Health, 8(2): 94-104.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Dual panel methodology

Basic concepts supporting dual panel methodology
- "It is better to produce quality in the translation, rather than checking it through back-translation"
- "Translation is only the start of the adaptation process"

Key actors in the process
- Expert translator panel
- Coordinator (from the developer team)
- Lay panel
- Sample of patients from the target population

Differential characteristics
- No back-translation
- Lay actors play an essential role in the wording of the final draft version (before field testing)

Lovas K, et al. (2003) Health Policy, 63: 49-61.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Dual panel methodology

Main steps of cultural adaptation
- Recruit 5-7 translators with varied profiles to work as a team in a group meeting
- Inform the group of the concepts underlying the questionnaire
- Inform them of the translation's requirements
- Have them work under the supervision of an experienced coordinator
- Have the agreed translation assessed by a lay panel working as a focus group
- Pilot testing by means of face-to-face (cognitive) interviews with several (15–20) representatives of the target population

Lovas K, et al. (2003) Health Policy, 63: 49-61.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Example for a forward-backward approach

Algorithm of linguistic validation of the PedsQL questionnaires
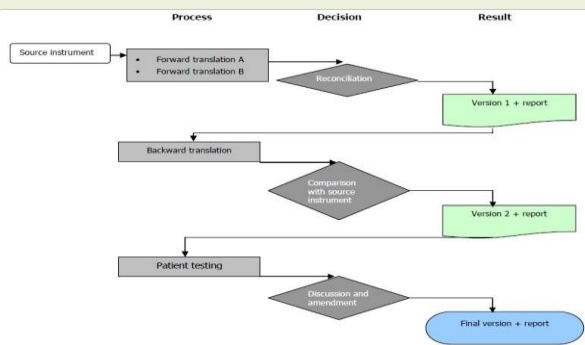
Mapi Research Institute. Linguistic validation of the PedsQL - a Quality of Life Questionnaire. 2002.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Take home messages

- PRO measures need to be translated and culturally adapted because
  - International researches require data to be compared and pooled
  - Instrument development is an expensive and time-consuming process
- There are several levels of equivalence such as conceptual, semantic, and criterion equivalence, etc. between the original and translated versions
- There are similarities between the development of the original and the adaptation of the translated versions
- The main steps of instrument adaptation are preparation, forward translation, reconciliation, back-translation, reviewing, harmonization, cognitive debriefing, reviewing cognitive debriefing, proofreading and process reporting
- A dual panel approach excludes the back-translation step and involves lay actors in the earlier phase of the adaptation

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Self-check questions

- What are the most important levels of equivalence between two versions of PRO instruments?

- What are the main general steps of PRO instrument cultural adaptations?

- What are the similarities between original instrument development and translation/adaptation processes?

- What are the steps of the dual panel approach?

Suggested reading

- Acquadro C, et al. (2008) Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. Value Health, 11(3): 509-521.

- Wild D, et al. (2005) Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health, 8(2): 94-104.

- Lovas K, et al. (2003) Establishing a standard for patient-completed instrument adaptations in Eastern Europe: experience with the Nottingham Health Profile in Hungary Health Policy, 63: 49-61.

7.5. International regulations on the use of PRO measures

International regulations on the use of PRO measures

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Learning objectives

- Students should:
  - Have a general knowledge of the process of PRO instrument development
  - Know the ethical principles relevant in PRO instrument development and application
  - Know the definitions and concepts covered by EMA Reflection Paper and FDA Guidance
  - Know the main steps of the whole development process according to ISPOR Good Practices

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Content

- The professional and ethical framework for the use of PRO measures
- Authority guidelines and requirements
- Recommendations and good practices of institutions internationally

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

The whole process of PRO instrument development

Creating a measure 1 – Conceptual framework
- disease/pathology
- patient population
- HRQoL domains to be assessed

Creating a measure 2 – cognitive interview
- item generation/modification
- validity/reliability testing

Measuring process
- involving patients from the target population
- in interventional/non-interventional trials

Data analysis
- with appropriate methodology

Publication of the results
- usually in a peer-reviewed journal

Using the results in drug application
- in drug authority applications

Domain of ethical regulations
- WMA Declaration of Helsinki
- national ethical regulations

Domain of authority regulations
- FDA Guides
- EMA Reflection paper

Domain of professional requirements
- ISOQOL Minimum Standard
- ISPOR Good Research Practices
- others (e.g. PROMIS Protocol, translation guides)
- ....

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

WMA Helsinki declaration

"WMA Declaration of Helsinki - Ethical principles for medical research involving human subjects"
(64th World Medical Association [WMA], General Assembly, Fortaleza, Brazil, October 2013)

Summary

- Is a set of ethical principles regarding human experiments
- Is the most important cornerstone of research ethics
- Was amended several times since its first version in 1964 (9th currently)
- Originated from the Nuremberg Code and Declaration of Geneva

WMA. Declaration of Helsinki - Ethical principles for medical research involving human subjects. 2013.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

WMA Helsinki declaration

It declares certain ethical principles in the following areas:

- Benefits weighed against risk and burden
- Specific protection of vulnerable groups
- Scientific requirements of research involving human subjects
- Mandatory involvement of an ethical committee
- Privacy protection
- Involving subjects based on informed consent
- The use of placebo treatment
- Research registration and publication

WMA. Declaration of Helsinki - Ethical principles for medical research involving human subjects. 2013.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

EMA reflection paper

"Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products" (European Medicines Agency [EMA], 2005)

Summary

- Defines itself as a "reflection paper" rather than a guideline on what EMA thinks about HRQoL as a specific type of instrument under the "umbrella" term of PRO
- Traces its definition of HRQoL back to the WHO health definition (1948)
- Distinguishes different PROs assessed by the patients themselves (such as disease symptoms)

EMA. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. 2005.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

EMA reflection paper

HRQoL in the drug evaluation process

- EMA considers HRQoL as an endpoint similar to other efficacy or safety endpoints and requires that "HRQoL improvement" as an endpoint/claim be supported by a validated instrument
- EMA accepts both generic and specific instruments
- EMA accepts improvement in only specific domains but recommends that the changes of all domains are presented (even if they remained unchanged or worsened)
- EMA evaluates the HRQoL results-based drug claim according to
  - HRQoL instrument assessment
  - Instrument validation results
  - Objectives and the justification of instrument choices
  - Adequacy of data analysis and the relevance of changes

EMA. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. 2005.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

EMA reflection paper

Specific study design for HRQoL assessment

- The instrument planned to be used should be validated by the time it is applied in the drug trial
- An HRQoL instrument may be applied in two ways
  - Simultaneously to the efficacy/safety assessment in the placebo or active drug- controlled trial
  - After its efficacy and/or safety has been shown and an additional trial has been conducted to show the HRQoL improvement in an active drug-controlled trial

Statistical analysis

- An adequate analysis plan should be provided
- In general the methodology is similar to the analysis of efficacy and safety results
- Statistical "overpower" for HRQoL results should be avoided
- A multi-dimensional instead of one-score type analysis is recommended to avoid "fading" the change of one domain into the others

EMA. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. 2005.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

EMA reflection paper

Aspects of special interest

- In some diseases (e.g. tumors) HRQoL has specific significance
- In decisions between two drugs with similar efficacy and safety, HRQoL may have importance

EMA. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. 2005.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

EMA reflection paper for oncology studies

"Reflection paper on the use of patient reported outcome (PRO) measures in oncology studies (Draft)"
(European Medicines Agency, 2014)

Summary

- PROs have special significance in malignancies and anti-tumor therapies (often symptomatic rather than definitive treatments)
- PROs frequently have add-on values over efficacy results
- Provides a more comprehensive description of PRO instrument use in drug labeling (than the "Reflection paper" from 2009)

EMA. Reflection paper on the use of patient reported outcome (PRO) measures in oncology studies (Draft) 2014.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

EMA reflection paper for oncology studies

Reasons why PRO measures should be included in clinical trial programs

- They provide a patient focused assessment by
  - Helping to understand how patient functioning changes
  - Identifying symptoms that need additional supportive care
- They provide complementary data to efficacy and safety results
- They help to differentiate two treatment options which have similar efficacy

Clinical trial design aspects

- Appropriate frequency and duration of assessment (optimal: when changes are expected)
- Patients' general health/disease status should be kept in mind (length of questionnaire, mode of data collection)
- Data loss should be minimized with appropriate stat. methods (high mortality rate -> high drop-out rate)

EMA. Reflection paper on the use of patient reported outcome (PRO) measures in oncology studies (Draft) 2014.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

EMA reflection paper for oncology studies

Other requirements related to instrument use

- Valid, reliable and culturally adapted measures should be used
- Specific instruments are preferred
- The role of proxy-reporting is very limited

EMA. Reflection paper on the use of patient reported outcome (PRO) measures in oncology studies (Draft) 2014.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

FDA guidance

"Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims"
(Food and Drug Administration [FDA], 2009)

Summary

- A set criteria through which PROs are evaluated
- A framework of instrument development
- Additional information for specific populations
- A good summary of the most important aspects in clinical trials that are relevant in the use of PRO measures
- A useful glossary of PRO-related conceptions

Limitation

- It restricts itself to cover only those instruments that are used in labeling applications ("labeling claim")

FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
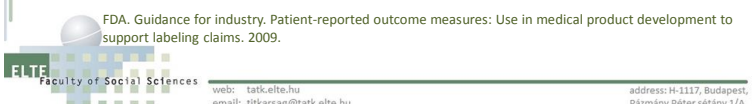address: H-1117, Budapest, Pázmány Péter sétány 1/A

FDA guidance

Definition of PRO that is:

"... any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else"

Use of PRO instruments

their use is advised when "... measuring a concept best known by the patient or best measured from the patient perspective" (e.g. symptoms, signs, functioning related to a disease or status)

FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

FDA guidance

Aspects that FDA takes into consideration in a PRO instrument evaluation

- Concepts being measured
- Conceptual framework of the instrument
- The medical condition for intended use
- The population for intended use
- Number of items
- Data collection method
- Administration mode
- Recall period
- Response options
- Scoring
- Weighting of items or domains
- Format
- Respondent burden
- Translation or cultural adaptation availability

FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

FDA guidance

Most important aspects 1. – Endpoint model

- It focuses on the PRO derived endpoint (e.g. eating out with mates [as social functioning] related to avoidance of allergenic foods)
- The PRO endpoint should be defined beforehand (by the clinical trial sponsor/instrument developer)
- Understanding the role of the PRO endpoint (its relationship to the efficacy of the investigated drug) is essential so that the endpoint of interest can be evaluated
- The endpoint can be primary or secondary

FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

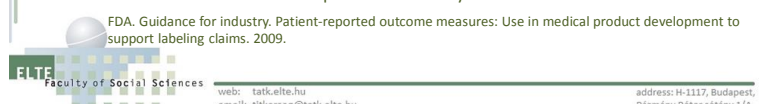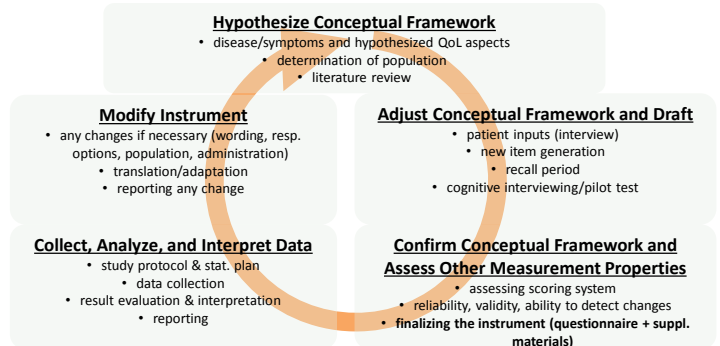ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

FDA guidance

Most important aspects 2. – Development as an iterative process

Hypothesize Conceptual Framework
- disease/symptoms and hypothesized QoL aspects
- determination of population
- literature review

Adjust Conceptual Framework and Draft
- patient inputs (interview)
- new item generation
- recall period
- cognitive interviewing/pilot test

Confirm Conceptual Framework and Assess Other Measurement Properties
- assessing scoring system
- reliability, validity, ability to detect changes
- **finalizing the instrument (questionnaire + suppl. materials)**

Collect, Analyze, and Interpret Data
- study protocol & stat. plan
- data collection
- result evaluation & interpretation
- reporting

Modify Instrument
- any changes if necessary (wording, resp. options, population, administration)
- translation/adaptation
- reporting any change

FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

FDA guidance

Most important aspects 3. – Concept - domain - item structure

Item 1, Item 2, Item 3, Item 4 → Domain 1

Item 1, Item 2 → Domain 2

Domain 1, Domain 2 → GENERAL CONCEPT

CONCEPTUAL FRAMEWORK

FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.
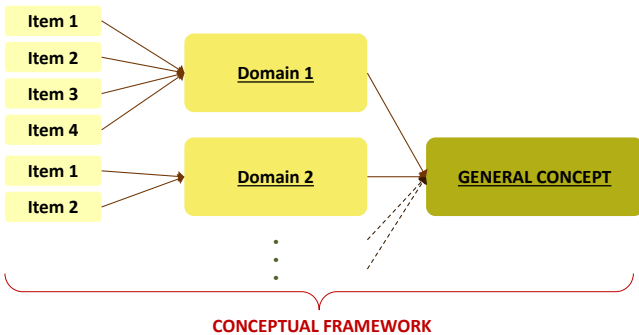
ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

FDA guidance

Most important aspects 4. – Defined instrument properties

- Validity indicators (content, construct, responsiveness)
- Reliability indicators (e.g. test-retest, intra- and inter-interviewer, inter-item correlations)
- Descriptions and criteria by which the indicators of an instrument are evaluated

FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

FDA guidance

Most important aspects 5. – Types of response options

- Examples of response options
  - Visual analog scale (VAS)
  - Likert scale
  - Rating scale
  - Recording of events as they occur
  - Pictorial scale
  - Checklist (Y/N)
- Evaluated in the context of the items:
  - How do responses correspond to the specific item
  - How do responses correspond to the characteristics of the population intended to be involved (age, severity of the disease, mental capacity etc.)
  - How clear are the distinctions between the options
  - Avoid any bias

FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

FDA guidance

Other important aspects in the evaluation of PRO instruments:

- Information provided to the user (instructions, training)
- Assessment of subject understanding
- Item scoring and scoring algorithm
- Respondent and administrator burden
- Application in specific populations (e.g. children, patients with cognitive impairment)

FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

FDA guidance

Other questions covered by the FDA Guidance
- Clinical trail design
  - How PRO endpoints are added into the trial protocol
  - How the PRO instrument is incorporated into the documents collecting patient data
  - How general clinical trial features (e.g. blinding, random patient allocation) and conducting (e.g. quality control) appear in the trial design
- Specific concerns when an electronic PRO instrument is used

Analysis of data

Glossary
- Useful definition of widely used (misused??) concepts

FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

ISOQOL minimum standards

"ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research"
(International Society for Quality of Life Research [ISOQOL], 2013)

Summary
- Provides a literature review for guidelines and expert opinions
- Provides the result of a survey of ISOQOL membership opinion about PRO minimum standards
- Introduces two concepts: "patient-centered outcome research" and "comparative effectiveness research" – combined to integrate patients' perspectives about their health and clinical data to assess the efficacy and safety of a medical intervention

Reeve BB, et al. (2013) Qual Life Res, 22(8):1889-1905.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

ISOQOL minimum standards

ISOQOL members were asked

To value requirement statements regarding (e.g.):
- Measurement model
- Reliability
- Different types of validity
- Score interpretability
- Translation
- etc.

whether the given statement is
- "Required as a minimum standard" or
- "Desirable but not required" or
- "Not required" or
- "Not sure" or
- "No opinion"

designated as ''recommended'' if > 50 % valued "required as .... "

Reeve BB, et al. (2013) Qual Life Res, 22(8):1889-1905.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

ISOQOL minimum standards

Some examples:

Requirement statements	Examples for valuation rates
A PRO measure should have documentation defining and describing the concept(s) included and the intended population(s) for use	Required as a minimum standard: 90 %
Reliability for a multi-item unidimensional scale should include an assessment of internal consistency	Required as a minimum standard: 79 %
Reliability for a multi-item unidimensional scale should include an assessment of test–retest reliability	Desirable but not required: 51 %
A PRO measure should have evidence supporting its content validity	Required as a minimum standard: 78 %
Documentation of background and experience of the persons involved in the translation	Not required: 8 %

Reeve BB, et al. (2013) Qual Life Res, 22(8):1889-1905.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

ISPOR good research practice 1

"Content Validity—Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 1 – Eliciting concepts for a new PRO instrument"
(International Society For Pharmacoeconomics and Outcomes Research [ISPOR], 2011)

Summary
- Provides a very detailed methodology for developing PRO instruments
- Provides interpretation and explanation of several PRO-related concepts and expressions
- Covers the whole process of development from understanding the disease to fully elaborating on the conceptual framework
- Focuses on the development of medical product-related PROs
- Summarizes the recommendation in 5 "good practices"

Patrick DL, et al. (2011) Value Health, 14(8): 967–977.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

ISPOR good research practice 1

Good practice 1: Determine the context of use
- Understanding the disease, identifying the concepts
- Developing the specific endpoint model
- Determining the optimal assessment frequency and recall period
- Considering the special characteristics of the target population (e.g. age, disabilities, subjects from different countries etc.)
- Considering the theoretical and qualitative methodological approach

⇒Developing a hypothesized conceptual framework

Patrick DL, et al. (2011) Value Health, 14(8): 967–977.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

ISPOR good research practice 1

Good practice 2: Develop the research protocol for qualitative concept elicitation and analysis

- Defining the target sample characteristics so that the sample population mirrors the target population (patient segments according to age, sex, ethnicity, disease severity, duration, course of disease etc.)
- Selecting the data collection method - focus groups vs. individual interviews vs. both
- Determining the setting and location for data collection

⇒Developing the interview guide

Patrick DL, et al. (2011) Value Health, 14(8): 967–977.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

ISPOR good research practice 1

Good practice 3: Conduct the concept elicitation interviews and focus groups

- Recruiting and training sites, selecting and training interviewers
- Recruiting participants
- Conducting and recording the interviews (both audio and video signs)
- Transcribing (immediately after the interview was performed)

⇒Collected data

Patrick DL, et al. (2011) Value Health, 14(8): 967–977.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

ISPOR good research practice 1

Good practice 4: Analyze the qualitative data

- Analyzing qualitative data (no quantitative metrics)
- Establishing a preliminary coding framework; update as data are coded
- Establishing coding procedures and train coders
- Assessing saturation

⇒Results interpretation

Good practice 5: Document concept development and elicitation methodology and results

- Summarizing all development steps described above from hypothesizing to results
- Providing the endpoint model, conceptual framework, protocol, interviews, data collection, analysis and interpretation

⇒Final report

Patrick et al. (2011) Value Health, 14: 967–977.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

ISPOR good research practice 2

"Content Validity—Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 2 – Assessing respondent understanding" (ISPOR, 2011)
Summary

- A very detailed methodology for developing PRO instruments
- A very detailed description of cognitive interviewing ("debriefing"): preparation, conducting, documenting
- The aim is to answer the questions:
  - "What do respondents believe the question is asking?"
  - "What do specific words and phrases mean to respondents?"
- Summarize the recommendation in 5 "good practices"

Patrick DL, et al. (2011) Value Health, 14(8): 978-988.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

ISPOR good research practice 2

Good practice 1: Item development based on concepts elicited

- Select items and their wording for cognitive interviews
- Select recall period (in chronic/episodic disease, frequent or rare symptom)
- Select modes of administration (patient diary/visit questionnaire, paper/electronic; types of data to be collected)
- Match each new item to response scale
- Format the instrument (item order)

⇒Format the actual instrument for cognitive interviewing

Patrick DL, et al. (2011) Value Health, 14(8): 978-988.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

ISPOR good research practice 2

Good practice 2: Design cognitive interview process

- Select study sample (similar characteristic to those of the target population)
- Determine sample size (more complexity, more patients)

⇒Final protocol for the cognitive interview

Good practice 3: Conducting cognitive interview

- Train interviewers (knowledgeable, focused, sensitive, friendly)
- Train subject to think aloud ("Tell me what you think this item is asking you about?")
- Conduct new interview rounds for each instrument revision
- Record and transcribing
- Prepare result summaries

⇒Results of cognitive interviews

Patrick DL, et al. (2011) Value Health, 14(8): 978-988.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

ISPOR good research practice 2

Good practice 4: Making decisions to revise the instrument
- The decision to make/not make revisions is not always straightforward
- Reduce ambiguity in item language
- Assess saturation
- Use an iterative process to reach a point when no further revision is needed

⇒ Make decisions about the final version

Good practice 5: Document cognitive interview results
- Complete Item tracking matrix including
  - Final item
  - Final response scale
  - Description of intent of item, patient quotes supporting item intent

⇒ Reporting the whole cognitive interview process

Patrick DL, et al. (2011) Value Health, 14(8): 978-988.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

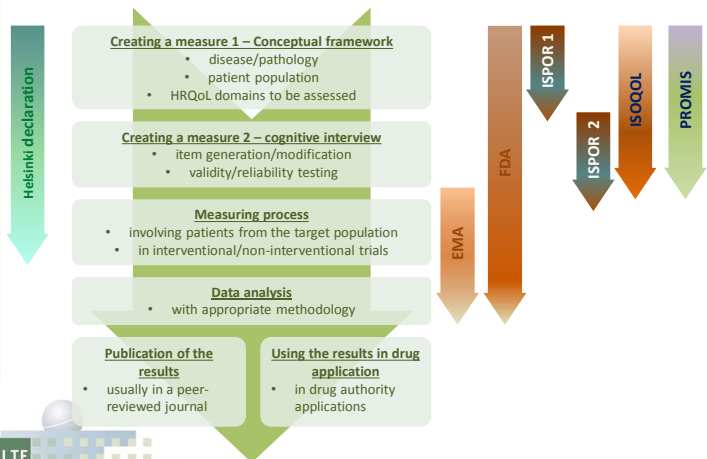## PROMIS scientific standards

"PROMIS® – Instrument Development and Validation Scientific Standards Version 2.0 (revised May 2013)"
(Patient Reported Outcomes Measurement Information System [PROMIS])

Summary
- PROMIS is a system of highly reliable, valid, flexible, precise, and responsive assessment tools that measure patient reported health status.
- A very detailed description of cognitive interviewing ("debriefing"): preparation, conducting, documenting
- The aim is to answer the questions:
  - "What do respondents believe the question is asking?"
  - "What do specific words and phrases mean to respondents?"
- Summarize the recommendation in 5 "good practices"

PROMIS. Instrument development and validation scientific standards version 2.0 (revised May 2013)

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Regulation of use in PRO measuring instruments

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Example for concept elicitation

Examples for concepts elicited during qualitative development of a symptom measuring PRO instrument (for patients with diarrhea-predominant irritable bowel syndrome)

Concept elicited	Spontaneously elicited in focus groups total (N=4)	Decision and rationale
Diarrhea	29	Included; saturated; bothersome
Immediate need (urgency)	18	Included; saturated; bothersome
Frequency of BMs	16	Included; saturated; bothersome
Cramps	16	Included; saturated; bothersome
Tired/weakness	10	Excluded; not specific to IBS-D
Nausea	9	Excluded; upper GI symptom
Completely emptied bowels/incomplete evacuation	9	Included; saturated; confirmed as core concept and bothersome; included as yes/no on event log

Marquis P, et al. (2014) Clin Transl Gastroenterol, 5: e59

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Example for conducting cognitive interview

A cognitive debriefing interview in the adapting process of an instrument that assesses HRQoL of patients with adulthood growth hormone deficiency
- Disease: Growth hormone deficiency in adulthood does result in dwarfism but also fat deposition, somatic hypertension, depression, anxiety, low level of energy, decreased libido, deteriorating memory etc. as psychological signs and symptoms -> all can impact quality of life
- Task: To conduct a cognitive interview with patients with adulthood growth hormone deficiency to assess the translated and adapted version of an HRQoL questionnaire
- Previous steps: The translation/cultural adaptation of the original instrument by means of a dual panel method

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Example for conducting a cognitive interview

Defining the purpose
- To assess the relevance, acceptability, comprehensiveness and understandability of questionnaire items;

Preparations
- Defining the patients who are eligible for the interview
- Defining the interview setting: patient's home or investigators' office that is quiet and where the patients is not disturbed by other people
- Besides the patient and the interviewer, no other person can be present

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Example for conducting a cognitive interview

During the interview

- As an initial step, the interviewer should
  - Explain the purpose of the interview i.e. To test the instrument itself, not the patient
  - Assure the patient of the confidential nature of the interview
  - Collect demographic information
  - Ask the patient to fill in the questionnaire and urge him/her to make comments on the questionnaire at any time
- While completing the questionnaire, the interviewer should
  - Note whether the respondent reads the instructions before starting
  - Note whether the respondent makes any general comments
  - Note whether any questions take a long time to answer
  - Record the finish time

Example for conducting a cognitive interview

During the interview

After completing the questionnaire, the interviewer should

- Ask whether the questions were relevant and understandable
- Ask whether any important aspect has been omitted
- Ask the patient to choose the most appropriate phrase if the lay panel could not decide
- Thank the patient for his/her participation

Example for conducting a cognitive interview

After the interview

- All information recorded during the interviews should be summarized in a report
- The report should state any changes in the wording of items that were made and indicating the reasons for these changes and their relation to the original English version

Take home messages

WMA Helsinki declaration is a set of ethical principles stating that (e.g.) benefits must be weighed against risk and burden; scientific rationale is required in researches involving human subjects; an ethical committee must be involved in the approval process, informed consent must be obtained from all subjects participating in scientific research

FDA guidance gives a well-referenced definition for PRO and outlines several aspects of instrument development and application such as: endpoint model, iterative development framework, conceptual framework, psychometric properties and response options

ISPOR good research practices outlines how ISPOR thinks about instrument development through the steps of concept elicitation, conducting interviews with subjects/focus groups, item generation and conducting cognitive interview

Self-check questions

- What are the similarities and differences between EMA Reflection papers and FDA Guidance?
- What do the following concepts mean: endpoint model, conceptual framework, iterative development process, content, construct validity, recall period, cognitive interview, coding answers?
- Give the definition of PRO by FDA Guidance?
- Give examples for questionnaire response options?
- What are the main steps of instrument development from understanding the disease to finalizing the instrument according to ISPOR Good Practice?

Suggested reading

- WMA. Declaration of Helsinki - Ethical principles for medical research involving human subjects. 64th World Medical Association, General Assembly, Fortaleza, Brazil, October 2013.
- EMA. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. EMA, 2005.
- FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. FDA, 2009.
- Patrick DL, et al. (2011) Content validity - Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 1 - Eliciting concepts for a new PRO instrument. Value Health, 14(8): 967-977.

Suggested reading

- Patrick DL, et al. (2011) Content validity - Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 2 - Assessing respondent understanding. Value Health, 14(8): 978-988.

7.6. PRO measuring in childhood

Patient reported outcome measuring in childhood

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Learning objectives

- Students should:
  - Have a knowledge of factors that facilitated the measurement of PROs in children and adolescents
  - Have a knowledge of age-specific domains that should be assessed in childhood and adolescence
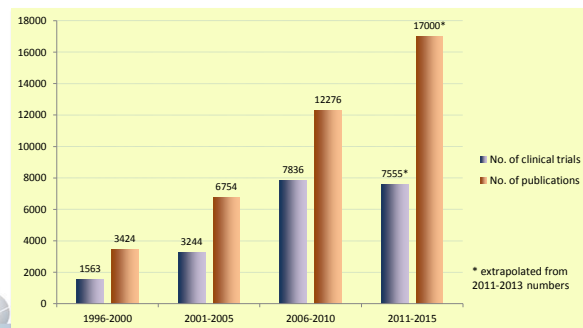 - Have a knowledge of specificities, advantages and disadvantages of parent/proxy-reporting

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Content

- Development and regulatory aspects of PRO measures used with children/adolescents
- Special aspects 1: Domain adapted to this age-range and specific domains
- Special aspects 2: Instrument formats and features adapted to this age-range
- Utility measures
- Proxy (parent) reporting instead of or parallel with child reporting

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Development of PRO measures for children

Dramatic increasing in number of

- clinical trials including children (www.clinicaltrial.gov)
- publications with QoL results in children (PubMed)

Period	No. of clinical trials	No. of publications
1996-2000	1563	3424
2001-2005	3244	6754
2006-2010	7836	12276
2011-2015	7555*	17000*

* extrapolated from 2011-2013 numbers

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Development of PRO measures for children

- Factors that facilitated measuring PROs in childhood
  - Changing in approach:
    – Taking into account child's well-being to greater extent
  - Changing medical point of view:
    – Considerably improvement in survival rates (in oncology, in autoimmune disorders, after transplantation, in neonatology etc.)
    – More definitive complications
    – Innovations in medicines and drug formulation technology
    – Development in home health care

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Development of PRO measures for children

- Factors that facilitated PRO measure development for childhood
  - Financing point of view:
    – Expensive health technologies appeared
    – Patient-centered endpoint can be differential endpoint if hard clinical endpoints are similar between two (or more) health technologies
    – Authority requirement

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Regulatory aspects of measuring PROs in children

- Regulatory guidance for measuring PROs
  - FDA Guidance (USA):
    - Has a short section that discusses "specific populations" including the pediatric population
    - Special aspects taken into account:
      – Age-related vocabulary
      – Language comprehension
      – Comprehension of the health concept measured
      – Duration of recall
  - EMA Guidance (Europe):
    - Has no specific instructions for the pediatric population

FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.
EMA. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. 2005.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Regulatory aspects of measuring PROs in children

- Ethical requirements for using PRO measures
  - Legal regulations of medical research
  - Ethical requirements of medical research

FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.
EMA. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. 2005.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

SPECIAL ASPECTS OF USING PRO MEASURES FOR CHILDREN

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Special aspects of using PRO measures for children

Main differences from PRO measurements for adulthood
- Differences in
  - Aethiology, course, complications, treatment, time horizon of the diseases
  - Social interactions of a child
  - Functions to be measured in the context of PRO
    ➡ age-specific domains needed
- Differences in approach
  - Intellectual, emotional, social development
    ➡ age-specific approach needed

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Special aspects of using PRO measures for children

Adaptation of domains used in adulthood

- Physical function
  - Sport, play (instead of employment, housework, ability to self-care)
- Psychological/emotional function
  - Referring to concrete situations, persons (instead of abstract concepts of mood, attitudes etc.)
- Social
  - Relationship with friends, peers, parents, boyfriend/girlfriend (instead of spouse, fellow-workers)

Eiser C, et al. (2001) Health Technology Assessment. 5(4): 1-157.

Special aspects of using PRO measures for children

Additional domains

- Cognitive function
  - Children are more vulnerable to cognitive impairment
  - The younger the child, the more severe the impairment
  - Lower ability to learn, to understand, to remember, lower motivation, and consequently lower school performance
  (e.g. cognitive impairment from birth hypoxia)
- Body image
  - Transition or permanent physical disabilities as a consequence of diseases, complications, treatments (body height, baldness, skin appearance)
  - Become extremely important in adolescence
  (e.g. embarrassment about a topic skin/eczema)

Eiser C, et al. (2006) Acta Paediatr. 96: 963-968.
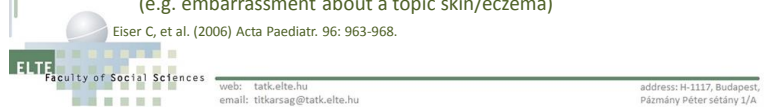
Special aspects of using PRO measures for children

Additional domains

- Autonomy
  - Attempts to establish autonomy and independence in adolescents
  - Socialization to autonomy is compromised
  - Dependence:
    - From the disease itself and its consequences;
    - From parents' care
  (e.g. continuous care because of dialysis)
- Peer socialization
  - Interdependent with autonomy domain
  - Integration into a peer group
  - Acceptance or rejection or even bullying by peers
  (e.g. outcasting because of cerebral palsy or motor dysfunction)

Eiser C, et al. (2006) Acta Paediatr. 96: 963-968.

Special aspects of using PRO measures for children

Additional domains

- Intimacy
  - Being important in adolescence
  - Sexual socialization (that does not necessarily mean sexual liaison)
  - Visible body image changes, feelings and desires
  (e.g. body disability or permanent "presence" parents may interfere with intimal relationship)
- Family relationship
  - The most important relationship
  - A model for the child: In life-style, coping strategies, treatment adherence
  (e.g. treatment and diet adherence of a diabetic teenager which depends on the intensity of family bonds)

Eiser C, et al. (2006) Acta Paediatr. 96: 963-968.
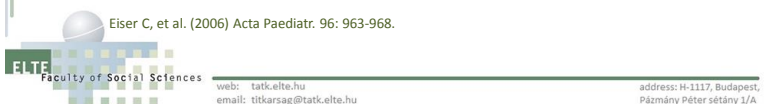
Special aspects of using PRO measures for children

Questionnaire format

- General aspects that should be taken into consideration
  - Language skills and literacy development
  - Development in abstract thoughts, understanding causality
  - Recall period (less defined)
  - Decision strategy (rather "yes/no", more concrete)

Matza LS, et al. (2004) Value Health, 7(1): 79-92.

Special aspects of using PRO measures for children

Questionnaire format

- Length
  - The younger the child, the shorter the questionnaire needs to be
- Appearance
  - Simple and clear-cut
  - Can be read easily
  - Pictorial support and use of computers to make the instrument more attractive
- Help in response
  - The aim is self-report without any help
  - But, interviewer is needed below a certain age limit

Matza L S, et al. (2004) Value Health, 7(1): 79-92.
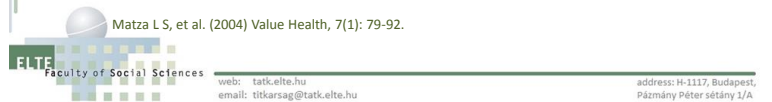
Special aspects of using PRO measures for children

Questionnaire format

E.g. for age-specific wording in PedsQL Generic Measure v4.0

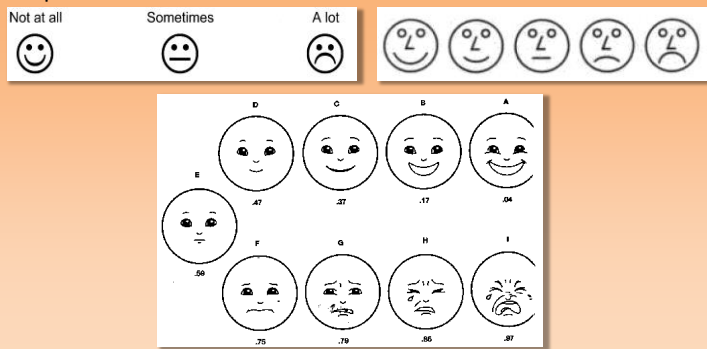| Items for ages 5-7 (an interviewer asks)	Items for ages 8-12
"Is it hard for you to walk?"	"It is hard for me to walk more than *one block* "
"*Do you ever feel too tired to play?*"	"*I have low energy* "
"Do you feel *mad*?"	"I feel *angry* "
"Can other kids do things that you cannot do "	"I cannot do things that other kids *my age* can do"

Special aspects of using PRO measures for children

Questionnaire format

- Response options – Likert-scale / Rating scales
  - 5-7-point Likert scale can be used from age 8
  - Young children tend to reduce to "yes/no" responses (choosing always two extreme options)
  - "Graphic Likert-scales" for ages 5-8

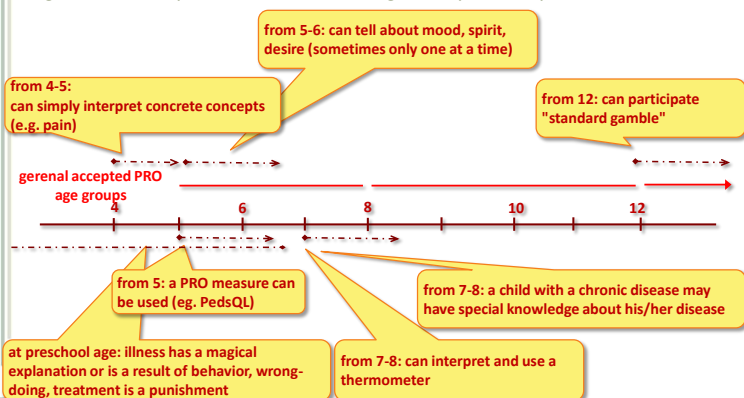## Special aspects of using PRO measures for children

Graphic Likert-scales

Christie MJ, et al. (1993) Psychosom Med, 55: 541-548.
French D, et al. (1994) Qual Life Res, 3: 215-224.

Special aspects of using PRO measures for children

Graphic Likert-scales/rating scales

Rebok G. (2001) Qual Life Res, 10: 59-70.
Ernst M, et al. (2000) J Am Acad Child & Adol Psych, 39(1): 94-99.

Special aspects of using PRO measures for children

Cognitive development – understanding concepts and questions

from 5-6: can tell about mood, spirit, desire (sometimes only one at a time)

from 4-5: can simply interpret concrete concepts (e.g. pain)

from 12: can participate "standard gamble"

gerenal accepted PRO age groups

4   6   8   10   12

from 5: a PRO measure can be used (eg. PedsQL)

from 7-8: a child with a chronic disease may have special knowledge about his/her disease

at preschool age: illness has a magical explanation or is a result of behavior, wrong-doing, treatment is a punishment

from 7-8: can interpret and use a thermometer

Special aspects of using PRO measures for children

Proxy as a responder

- Who can be a responder :
  - Primarily a mother or father
    - Mother much more often
    - Always the same person if the test is repeated
  - Sometimes: Medical staff or teacher or other caregiver
- Main issue: To what extent is the parent rate valid?
- The proxy cannot be a responder in some domains
  - E.g. pain, school performance, intimacy

Eiser C, et al . (2001) Health Technology Assessment, 5(4): 1-162.
FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

Special aspects of using PRO measures for children

Proxy as a responder
- Use of parent (proxy) as a responder is inevitable
  - Below a certain age limit
  - In serious mental or physical disabilities
  - In a serious acute disease
- Advantage of parent (proxy) responding
  - Increases the amount of data (and point of view)
- Disadvantages of parent (proxy) responding
  - It is contrary to the principles of PRO
  - May decrease validity
  - A parent may be a source of information on health state rather than HRQoL
  - Parents rates can be influenced by their own QoL

Eiser C, et al . (2001) Health Technology Assessment, 5(4): 1-162.
FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Special aspects of using PRO measures for children

Parallel rating
- Advantages
  - Can make the "point of view" wider (providing a fuller informational picture)
  - The results of comparing parents' vs. children's responses can help later when only parent responses are available
- Difficulties
  - May require a separate development
  - May require a more complex study design, that is more expensive

Matza LS, et al. (2004) Value Health, 7(1): 79-92.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Special aspects of using PRO measures for children

Parallel rating
- Agreement and differences
  - Factors that can influence the results
    – QoL domain
    – Age
    – Item interpretation
    – Points that are considered
- The more similarities there are
  - The younger the child is
  - The more external the domain is

Matza LS, et al. (2004) Value Health, 7(1): 79-92.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Special aspects of using PRO measures for children

Where QoL measures can be utilized with children
- In clinical trials:
  - Primary or secondary end-points
- In epidemiology surveys
- In the evaluation of health technologies
- In the assessment of chronic patient care
- In health policy and resource allocation decisions

Examples:
How long acting are the present prevention treatments for asthmatic attacks?
What kind of impact does an insulin pump have on independent living/social function
How satisfied are patients with the new use of insulin pens?

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

MEASURING UTILITY IN CHILDREN

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Measuring utility in children

EQ-5D-(Youth)
- Structure
  - 5 domains / 5 item
  - 3 levels to rate the domains 35 =243+2 potential outcomes
- Age ranges and responder versions
  - For age 0-7: only a proxy reporting version exists
  - For age 8-11: EQ-5D-Y version can be used
  - For age 12-15: EQ-5D or EQ-5D -Y version can be used
  - Over age 16: EQ-5D
- Disadvantages
  - There are no validated value sets by Euroqol

Euroqol. EQ-5D-Y user guide. 2004.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Measuring utility in children

Child Health Utility 9D (CHU-9D)
- Characteristics
  - 9 dimensions: worried, sad, annoyed, tired, pain, sleep, daily routine , work, able to join in activities
  - 5 levels for each domain : 1.9 million outcomes
  - Age range: 7-17
  - General population preferences by standard gamble

Stevens K. (2012) Pharmacoeconomics, 30(8): 729-747.

Take home messages

- Patient-reported outcome (quality of life) measures for children were delayed some 20 years after QoL measures were introduced for adults
- The most important domains for childhood instruments are the same as those used in adulthood instruments but should be extensively adapted
- There are age-specific domains that should be built in to cover functions and late aspects that are relevant and important for children/adolescents
- Almost all questionnaire format elements and features should be adapted to this age-range
- Utility measures are quite limited in childhood, especially in early life

Self-check questions

- What are the factors that facilitate the measurement of PROs in children and adolescents?
- How should we adapt domains that are used in instruments developed for adults?
- What are those domains that have special importance in childhood / adolescence?
- What is the role of a proxy when reporting PROs instead of the patients themselves?

Suggested reading

- Eiser C, et al. (2001) Quality-of-life measures in chronic diseases of childhood. Health Technology Assessment. 5(4): 1-157.
- Frisén A. (2006) Measuring health-related quality of life in adolescence. Acta Paediatr. 96: 963–968
- Euroqol. EQ-5D-Y User Guide. 2004.
- Matza LS, et al. (2004) Assessment of health-related quality of life in children: A review of conceptual, methodological, and regulatory issues. Value Health. 7(1): 79-92.
- FDA. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

7.7. Patient adherence

Patient adherence

"Financed from the financial support ELTE won from the Higher Education Restructuring Fund of the Hungarian Government"

Learning objectives

- Students should:
  - Be familiar with the key definitions and measurement and calculation methods of adherence
  - Be familiar with the determinants and consequences of non-adherence
  - Be familiar with the key methods of integrating adherence in pharmacoeconomic evaluations
  - Be familiar with adherence-enhancing interventions

Content

- Terminology
- Methods for measuring and calculating adherence
- Prevalence of non-adherence
- Consequences of non-adherence
- Determinants of adherence
- Methods for integrating adherence in pharmacoeconomic evaluations
- Adherence-enhancing interventions

TERMINOLOGY

Terms for describing the extent to which a patient undertakes the recommendations of healthcare providers

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
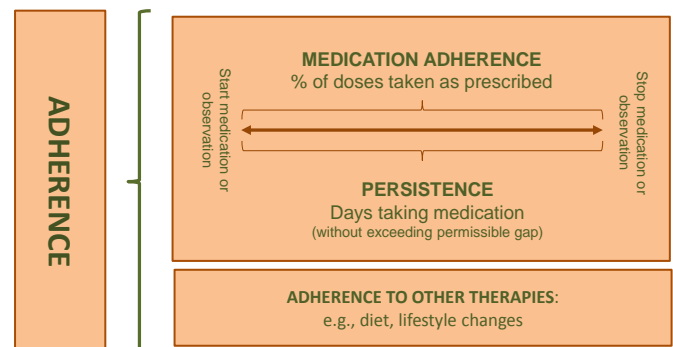address: H-1117, Budapest, Pázmány Péter sétány 1/A

Key definitions (1)

- **Adherence:**

  "The extent to which a person's behavior - taking medication, following a diet, and/or executing lifestyle changes - corresponds with agreed recommendations from a healthcare provider"

- **Medication adherence:**

  "Refers to the act of conforming to the recommendations made by the provider with respect of timing, dosage and frequency of medication taking"

World Health Organization. Adherence to long-term therapies: evidence for action. WHO, 2003.
Cramer JA, et al. (2008) Value Health, 11(1): 44-47.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Key definitions (2)

- **Persistence**

  "The duration of time from initiation to discontinuation of therapy"

Cramer JA, et al. (2008) Value Health, 11(1): 44-47.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Key definitions (3)

Cramer JA, et al. (2008) Value Health, 11(1): 44-47.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Adherence vs. compliance

- Adherence and compliance are often used interchangeably
- Adherence has become the preferred term instead of compliance
- Primary difference:
  - Compliance connotes a paternalistic relationship between the healthcare provider and the patient
  - Adherence represents the patient as an equal partner with the healthcare provider

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Primary and secondary adherence

- Primary adherence to medication: adherence to fill prescription

- Secondary adherence to medication: adherence to take the medication as prescribed

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

METHODS FOR MEASURING MEDICATION ADHERENCE

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Measurement methods (1)

- Indirect methods:
  - Patient self-report: adherence questionnaire, patient diary
  - Pharmacy dispensing data records
  - Pill counting, canister weighing
  - Electronic adherence monitoring
  - Therapeutic outcome monitoring

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Measurement methods (2)

- Direct methods:
  - Pills with microchip
  - Direct observation
  - Biological assay

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Measurement methods (3)

	Pros	Cons
Patient self-report: adherence questionnaire, patient diary	Easy to obtain	May overestimate adherence
Pharmacy dispensing data records	Rapid Inexpensive	Pharmacy database can be incomplete No evidence on medication intake
Pill counting, canister weighing	Easy to obtain	No evidence on medication intake
Electronic adherence monitoring	Accurate measure of dosing history	Expensive No evidence on medication intake
Therapeutic outcome	Easy to obtain	Clinical outcomes may be affected by other factors as well
Pills with microchip	Confirm medication intake	Expensive Require cooperation from the patient Can be used only with tablets/capsules/pills
Direct observation of the medication intake	Confirm medication intake	Unpleasant for the patient Require large human resources
Biological assay	Confirm medication intake	Expensive Unpleasant for the patient Limited information regarding use over time Insensitive to inhaled drugs

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Adherence questionnaires (1)

- Morisky Medication Adherence Scale 4-item (MMAS-4)
  - Consists of 4 items (see next slide)
  - Scoring scheme: 'Yes=0', 'No=1'
  - Summary score: 0 – 4, higher scores indicate greater adherence

http://dmorisky.bol.ucla.edu/

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Adherence questionnaires (2)

- Morisky Medication Adherence Scale 4-item (MMAS-4)
  - Questions:
    - Do you ever forget to take your medicine?
    - Do you ever have problems remembering to take your medication?
    - When you feel better, do you sometimes stop taking your medicine?
    - Sometimes if you feel worse when you take your medicine, do you stop taking it?

http://dmorisky.bol.ucla.edu/

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Adherence questionnaires (3)

- Morisky Medication Adherence Scale 8-item (MMAS-8):
  - Consists of 8 items
  - Scoring scheme:
    - Questions 1 - 7: 'Yes=0', 'No=1'
    - Question 8: Likert scale
      » 0=0, 1=0.25, 2=0,5, 3=0,75, 4=1
    - Summary score: 0 – 8, higher scores indicate greater adherence

http://dmorisky.bol.ucla.edu/

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Pills with microchip

- The first digital medicine was approved by FDA in 2012
- How does it work?
  - Patients take pills which have been modified to contain an edible microchip
    - Size of the microchip is ~0.5 mm
  - After pill is swallowed, chip is activated by stomach fluids, sending a signal to a patch on patient's arm
  - Patch contains a receiver which decodes data about drug
  - Receiver transmits information to patient's cell phone, telling them when their next dose is due and provides other health data

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Mono-pharmacotherapy

- Several measures for calculating adherence to mono-pharmacotherapy from pharmacy dispensing records have been proposed in the literature
  - **Medication Possession Ratio (MPR)**
  - Proportion of Days Covered (PDC)
  - ….

CALCULATING ADHERENCE FROM DISPENSING DATA RECORDS

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Medication Possession Ratio (1)

- Observation period:
  - Fixed interval

$$MPR = \frac{Total\ day's\ supply\ dispensed\ over\ the\ observation\ period}{Number\ of\ days\ of\ the\ observation\ period}$$

  - Start with the first refill and end with the last refill

$$MPR = \frac{Total\ day's\ supply\ dispensed\ over\ the\ observation\ period}{Last\ refill\ date - First\ refill\ date + Days'\ supply\ dispensed\ at\ last\ refill}$$

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Medication Possession Ratio (2)

- Continuous variable
  - Ranges from 0 (no medication dispensed) to 1 (maximal adherence)
    - Oversupply: >1 MPR –> Non-adherence
  - Can also be expressed as a percentage

- Dichotomous variable (Adherent vs. Non-adherent)
  - Adherent: patients with adherence 0.8 -1.0 MPR
  - Disease and therapy specific cut-off value

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Medication Possession Ratio (3)

Example:

Refill date	Number of days in the interval	Drug A Quantity dispensed
01.09.2014		30
12.10.2014	42	30
20.11.2014	39	30

- Fixed interval: 01.09-20.11.2014 (91 days)

MPR=(30+30+30)/91=0.99 (99%)

- Start with the first refill and end with the last refill

MPR=(30+30+30)/((42+39)+30)=0.81 (81%)

Days' supply dispensed at last refill

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Calculating persistence (1)

- Persistence
  - The duration of time from the initiation (or at chronic disease from an optional date) to the discontinuation of drug therapy
  - Percentage of individuals remaining on therapy (persistent) until a specified time interval

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Calculating persistence (2)

- Permissible gap
  - Is reported as the maximum allowable period of the refill interval without discontinuation of the therapy
  - Should be defined in a disease and therapy specific way
    - Most evaluations use a 60 day permissible gap

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Calculating persistence (3)

Example:

Patient A			Patient B		
Refill date	Gaps in the interval	Drug A Quantity dispensed	Refill date	Gaps in the interval	Drug A Quantity dispensed
06.09.2014		30	15.09.2014		30
15.11.2014	-40=70-30	30	20.10.2014	-5=35-30	30
20.12.2014	-5=35-30	30	21.01.2015	-63=93-30	30
25.01.2015	-6=36-30	30			

- Permissible gap: 60 days
- Observation period: 01.09.2014 - 31.01.2015
- Persistent at the end of the observation period:
  - Patient A: YES; Patient B: NO

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Calculating persistence (4)

Example:

Patient A			Patient B		
Refill date	Gaps in the interval	Drug A Quantity dispensed	Refill date	Gaps in the interval	Drug A Quantity dispensed
06.09.2014		30	15.09.2014		30
15.11.2014	-40=70-30	30	20.10.2014	-5=35-30	30
20.12.2014	-5=35-30	30	21.01.2015	-63=93-30	30
25.01.2015	-6=36-30	30			

- Permissible gap: 60 days
- Observation period: 01.09.2014 - 31.01.2015
- Patients were persistent with the therapy:
  - Patient A: 06.09-31.01.2015 = 148 days
  - Patient B: 15.09-19.10.2014 + 30 days = 65 days

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Poly-pharmacotherapy

- Methods developed for mono-pharmacotherapy tend to over/under-estimate adherence in patients with treatment regimens consisting of multiple medications
- Some new approaches have been developed in recent years to calculate adherence to poly-pharmacotherapy:
  - The multiple-Medications Prescribing Ratio (mMPrR), the multiple-Medications Possession Ratio (mMPR) and the Prescription and Medication Possession Graph (PMPG)
  - The daily polypharmacy possession ratio (DPPR)

Ágh T, et al. A novel method for calculating medication adherence to poly-pharmacotherapy by linking general practice prescribing data and pharmacy dispensing records. ISPOR 18th Annual European Congress (Milan, Italy)
Arnet I, et al. (2014) Int J Clin Pharm, 36(1):192-201.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

PREVALENCE OF NON-ADHERENCE

RCTs vs. real-world studies

- Efficacy ⟷ Effectiveness

  Non-adherence

- In RCTs, conditions are highly controlled and the stringent follow-up protocol limits the occurrence of medication non-adherence; therefore non-adherence rates derived from RCTs do not reflect an objective picture on medication adherence
  - Adherence to COPD medication
    - RCTs: 70-90% / Real-world studies: 20-60%

Ágh T, Mészáros Á. In: Ong K-C, editor. Chronic Obstructive Pulmonary Disease - Current Concepts and Practice. InTech, 2012: 275-290.

Unfilled prescriptions

- Kennedy et al.
  - 4.4% of patients failed to fill or refill 1 or more prescriptions (N=14,500)
    - Failure-to-fill rates
      - Psychiatric conditions: 8.0%
      - Arthritis: 5.2%
      - Cardiovascular disease: 5.2%
      - Emphysema, asthma, or chronic obstructive pulmonary disease: 6.6%

Kennedy J, et al. (2008) J Manag Care Pharm, 14: 553–560.

Long-term therapies

- WHO: adherence to long-term therapies averages only 50%
  - Hypertension, dyslipidemia and diabetes
    - Adherence: 67-76% / Persistence: 63%
  - Asthma, COPD
    - Adherence: 20-60% / Persistence: 7-16%

Cramer JA, et al. (2008) Int J Clin Pract, 62: 76-87.
World Health Organization. Adherence to long-term therapies: evidence for action. WHO, 2003.
Ágh T, Mészáros Á. In: Kian-Chung Ong. Chronic Obstructive Pulmonary Disease - Current Concepts and Practice. InTech, 2012: 275-290.

Short-term treatment

- Non-adherence to short-term treatment is also a significant problem

- Kardas et al.
  - Non-adherence to antibiotic regimens in 5 EU countries: 20.8%*

* MMAS-4; adherent = 4 scores

ABC Project Team. Ascertaining barriers for compliance: policies for safe, effective and cost-effective use of medicines in Europe. 2012.

Cross-country differences

- There are cross-country differences in prevalence of non-adherence

Country	Non-adherence in patients with hypertension*
Hungary	70,3%
United Kingdom	41,5%
Poland	57,6%
Austria	33,7%

* MMAS-4; adherent = 4 scores

ABC Project Team. Ascertaining barriers for compliance: policies for safe, effective and cost-effective use medicines in Europe. 2012.

CONSEQUENCES OF NON-ADHERENCE

Consequences of non-adherence

- Clinical consequences

- Cost consequences

- Impact on quality of life

Clinical consequences of non-adherence

- Adherence ↑ -> clinical outcomes ↑
  - Persistent anti-hypertensive therapy is associated with a 40% increased chance of BP goal attainment
  - Non-adherence to statin therapy in CV patients doubled the incidence of myocardial infarction
- Drug overuse may also have negative clinical consequences (e.g., NSAID overuse -> gastric ulcer, nephropathy)

Breekveldt-Postma NS, et al. (2008) Curr Med Res Opin, 24: 1025-1031.
Blackburn DF, et al. (2005) Pharmacotherapy, 25: 1035-1043.

Cost of non-adherence (1)

- The effect of non-adherence on medical costs works in two ways:
  - Immediate and direct impact on drug costs
  - Less immediate and indirect impact on health service utilization / healthcare costs
- In general, non-adherence is likely to reduce drug costs, but increase subsequent overall health service utilization / healthcare costs
- Highly dependent on
  - Condition
  - Therapy
  - Time

Cost of non-adherence (2)

- The impact of medication adherence on drug costs:
  - Is determined mostly by the extent of non-adherence
  - Non-adherence does not always result in decreased drug costs!
    - Patients continuing to dispense prescriptions but stockpiling their medicines
    - Medication overuse

Cost of non-adherence (3)

- Durg cost
  - The UK's NICE produced guidelines for patient adherence in which it estimated that around £4 billion of medicines supplied on prescription through the NHS are not used correctly
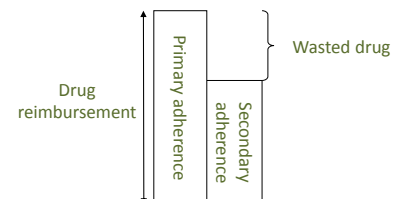
NICE. Medicines adherence: Involving patients in decisions about prescribed medicines and supporting adherence. 2009.

Cost of non-adherence (4)

- The impact of non-adherence on overall healthcare utilization is determined primarily by clinical effectiveness of the medicine
  - If health service use is highly associated with the extent of the management of the condition and the medication has a key role in the management of the condition -> the impact of medication non-adherence on resource use is large
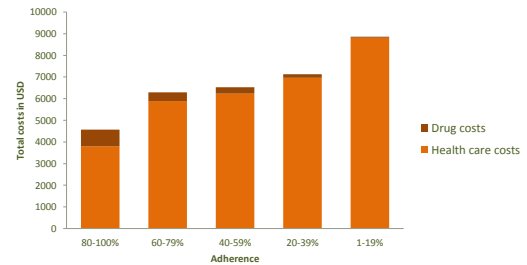
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Cost of non-adherence (5)

- The impact of medication adherence on healthcare cost for diabetes

Sokol MC, et al. (2005) Med Care, 43: 521-530.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Cost of non-adherence (6)

- In the US, the overall cost of poor adherence, measured in otherwise avoidable medical spending, is close to $310 billion annually, representing approximately 14% of total healthcare expenditures
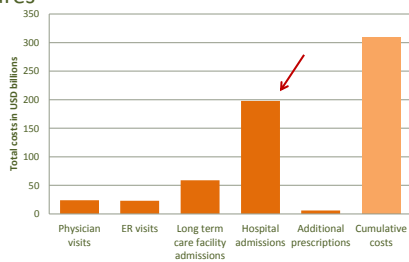
Capgemini Consulting. Patient Adherence: the next Frontier in Patient Care, 9th edition. 2011.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Impact of non-adherence on quality of life (1)

- Association between medication adherence and health-related quality of life (HRQoL) is dual
  - Adherence -> HRQoL
  - HRQoL -> adherence
- Psychiatric comorbidities (e.g., depression) may influence the relationship between medication adherence and HRQoL
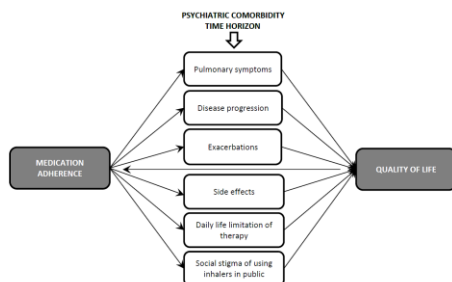
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Impact of non-adherence on quality of life (2)

- A framework for understanding the relationship between medication adherence and quality of life in COPD

Ágh T, et al. (2015) Respir Care, 60(2): 297-303.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Impact of non-adherence on quality of life (3)

- The effect of adherence on HRQoL may be a consequence of the effectiveness of therapy and the negative effects that it can generate (i.e., side-effects, daily life limitation of therapy, social stigma)

- Dynamics between adherence and HRQoL may differ over time
  - The negative effects of medication non-adherence may become more and more dominant in the long-term

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Impact of quality of life on adherence (1)

- A patient's decision to adhere and to what extent is a personal trade-off between the benefits and the negative effects of the therapy at any given time
- The dynamics between adherence and HRQOL might change over time

Impact of quality of life on adherence (2)

- Example:
  - The initiation of a pharmacological therapy in newly diagnosed COPD patients might significantly improve their HRQoL
  - Later, this HRQoL improvement due to medical treatment might be smaller and could be detected only in the long-term
  - Therefore, if patients have previously been treated for longer durations, the benefits from medication non-adherence might temporarily outweigh the effects of the disease deterioration in the short-term

Ágh T, et al. (2015) Respir Care, 60(2): 297-303.

DETERMINANTS OF ADHERENCE
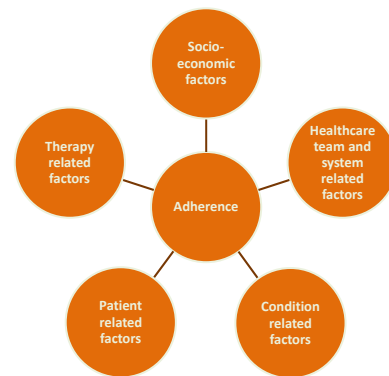
Determinants of adherence (1)

Final report of the ABC Project - 2012; www.abcproject.eu

Determinats of adherence (2)

Socio-economic factors	Family/social support (emotional, financial), social stigma, co-payment, income, empoyement status, etc.
Healthcare team and system-related factors	Barrires to healthcare, prescription by a specialist, healthcare provider-patient communication and relationship, etc.
Condition-related factors	Disease severity, persence of symptoms, psychiatric condition, clinical improvement, etc.
Therapy-related factors	Adveres effects, number of drugs/daily doses, duration of the treatment, etc.
Patient-related factors	Age, forgetfulness, marital status, education, etc.

ABC Project Team. Ascertaining barriers for compliance: policies for safe, effective and cost-effective use of medicines in Europe. 2012.

METHODS FOR INTEGRATING ADHERENCE IN PHARMACOECONOMIC EVALUATIONS

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Adherence in health-economic models

- To evaluate the impact of non-adherence/non-persistence on both health outcomes and costs requires the use of health-economics models
- The choice of the economic model is dependent on:
  - The condition being treated (e.g., acute vs. chronic)
  - Data availability (individual vs. aggregated data)
  - Type of adherence data (adherence data vs. persistence data)

Hughes D, et al. (2007) Value Health, 10(6):498-509.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Decision-analytic model

- A decision-analytic model presents individuals' possible prognoses following some sort of intervention by a series of pathways
- These models may be appropriate particularly for acute conditions
- In most conditions, a decision-analytic model can be developed from existing published sources
- Branches of the decision tree may be used to represent different levels of adherence (adherent/non-adherent)

Hughes D, et al. (2007) Value Health, 10(6):498-509.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Decision-analytic model vs. Markov model

- When there are numerous health states, including the possibility of transitions from one health state to another and back again, the decision tree may become far too complex to handle the problem efficiently.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Markov model (1)

- The Markov model places patients into discrete "health states", and time is partitioned into discrete periods, known as "cycles", during which patients are assumed to stay in the same health state.
- In each cycle, a patient's health state may change from his or her current health state to another health state. The probabilities of moving from one health state to another (are called "transition probabilities") .
- The costs and utility associated with each health state are combined with the time that patients spend in the state to estimate the overall costs and utility expected for different adherence/persistence rates.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Markov model (2)

- A sample hypothetical example for integrating persistence in health economic evaluations by use of a Markov model
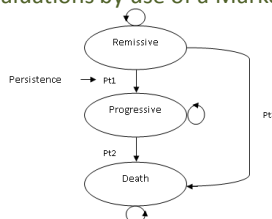
  - For those patients who discontinue treatment, Pt1 transition probability is assumed to increase and disease progression is accelerated in comparison with patients remaining under treatment

Hughes D, et al. (2007) Value Health, 10(6):498-509.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Discrete event simulation (1)

- In a discrete event simulation (DES), the experience of individuals is modeled over time in terms of the events that occur and the consequences of those events
- DES specifies patients as entities and treatment discontinuation as events
- DES facilitates interactions between adherence and time, as well as individual characteristics (e.g., adherence to drugs for asthma may be highly correlated with the severity of symptoms)
- A reasonably informative DES requires more detailed data than a typical Markov cohort model

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

ADHERENCE-ENHANCING INTERVENTIONS

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Adherence-enhancing interventions (1)

- Leading influencers of adherence:
  - Health policy
  - Pharma industry
  - Healthcare providers
  - Social environment
  - Patient

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Adherence-enhancing interventions (2)

- Interventions:
  - **Clinical innovation**: simplified regimen, long action medication, combination drug
  - **Patient education**: print materials, online communication, CD-ROMs
  - **Patient reminders**: tele-calling, e-mails, text messages, apps
  - **Cost-related approaches**: reducing co-payment, discounts, vouchers,
  - **Others**: nurse education, pharmacist programs, patient organizations, self-monitoring

Capgemini Consulting. Patient Adherence: the next Frontier in Patient Care, 9th edition. 2011.
Petrilla AA, Benner JS. (2003) Value Health, 6: 200.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Adherence-enhancing interventions (3)

- For the analysis of the cost–effectiveness of adherence-enhancing interventions, it is important to look at both costs of the intervention and outcomes, not only in terms of adherence, but also in terms of the subjective value of the clinical outcome for the patient.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Take home messages

- Medication adherence "refers to the degree or extent of conformity to the recommendations for day-to-day treatment by the provider with respect to the timing, dosage, and frequency".
- Medication non-adherence is common and poses a significant barrier to optimal disease management, since it can result in poor health outcomes and increased healthcare costs.
- Non-adherence has a number of causes, including socio-economic, patient-related, condition-related, therapy-related, and healthcare team and system-related factors.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Self-check questions

- What is the definition of adherence/medication adherence/persistence?
- What kind of direct and indirect adherence measurement methods exist? What are their pros and cons?
- How can you calculate MPR and persistence?
- What are the clinical/cost/quality of life consequences of non-adherence?
- What are the determinants of adherence?
- How can you integrate adherence in health-economic evaluations?
- What kind of adherence-enhancing interventions do you know?

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Suggested reading

- World Health Organization. Adherence to long-term therapies: evidence for action. WHO, 2003. (http://www.who.int/chp/knowledge/publications/adherence_report/en/)
- Cramer JA, et al. (2008) Medication compliance and persistence: terminology and definitions. Value Health, 11(1): 44-47.
- Hughes D, et al. (2007) Methods for integrating medication compliance and pesisitence in pharmacoeconomic evaluations. Value Health, 10(6): 498-509.
- Ascertaining barriers for compliance: policies for safe, effective and cost-effective use of medicines in Europe. Final report. 2012. (http://abcproject.eu/img/ABC%20Final.pdf)

7.8. Patient satisfaction

Patient satisfaction

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Learning objectives

- Student should:
  - Be familiar with the definition of patient satisfaction
  - Be familiar with the measurement methods of patient satisfaction

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Content

- Definition
- Methods for measuring patient satisfaction
- Benefits of patient satisfaction measurement

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

What is patient satisfaction?

- Patient satisfaction is a measure of the extent to which a patient is content with the healthcare which they received from their healthcare provider

- Patient satisfaction is one of the most important quality assessment tools

- Patient satisfaction is multifaceted and a very challenging outcome to define

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Determinants of patient satisfaction

- Determinants:
  - Patient expectations of care
  - Health status
  - Current emotional state
  - Socioeconomic status
  - Physical environment
  - Other factors: e.g. opinion of friends

- Determinants of satisfaction and dissatisfaction are not necessarily the same

A theoretical model of patient satisfaction
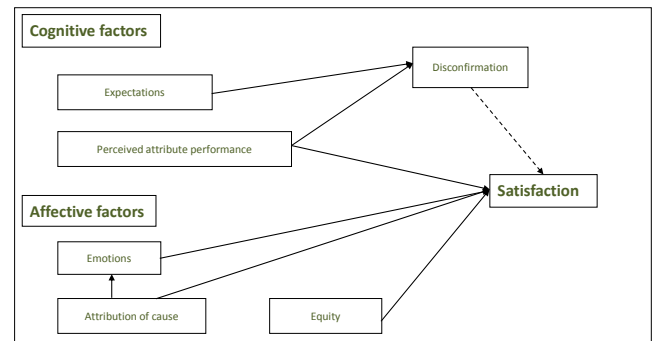
Oliver RL (1993). J Cons Res, 20: 418.

Measure of patient satisfaction

- Problem-oriented survey: focus on a given problem

- Patient-oriented survey: focus on the patient, allows for a more complex evaluation

Measurement methods of patient satisfaction

- Methods:
  - Qualitative and quantitative questionnaires
  - Open questionnaires
  - Interviews
  - Direct observation
  - Analysis of medical documentation

Questionnaires

- Source: previously developed or self-developed questionnaires
- Criteria:
  - Validity: the degree to which a questionnaire reflects reality.
  - Reliability: the degree to which a questionnaire will produce the same result if administered again, or the "test-retest" concept. It is also a measure of the degree to which a questionnaire can reflect a true change.

Real benefits of a patient satisfaction measurement

- Macro-level:
  - system performance management
  - benchmarking
  - competition/contestability through markets
- Micro-level:
  - feedback to professionals and managers
  - acceptability of processes / social model of health
- Patient satisfaction affects medication adherence and therefore the clinical outcomes of patients

Thompson A. What is patient satisfaction? VII Meeting of INGID, 2006

Take home messages

- Patient satisfaction is a measure of the extent to which a patient is content with the healthcare which they received from their healthcare provider
- Patient satisfaction can be measured with qualitative and quantitative questionnaires, open questionnaires, interviews, direct observation, and an analysis of medical documentation

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Self-check questions

- What is patient satisfaction?
- What are the determinants of patient satisfaction?
- How can you measure patient satisfaction?
- What are the benefits of measuring patient satisfaction?

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Suggested reading

- Oliver RL (1993) Cognitive, affective, and attribute bases of the satisfaction response. J Cons Res, 20: 418.
- Farley H, et al. (2014) Patient satisfaction surveys and quality of care: an information paper. Ann Emerg Med, 64: 351-7.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

8. Applicability of PROs in economic evaluation

8.1. Utility measurement

Utility measurement

"Financed from the financial support ELTE won from the Higher Education Restructuring Fund of the Hungarian Government"

Learning objectives

- Students should:
  - Be aware of the necessity of a universal construct to measure QoL
  - Be familiar with the core characteristics of 'utility' construct
  - Be familiar with measuring utility using direct and indirect measures
  - Be able to conduct a time trade-off interview
  - Be able to calculate utility with EQ-5D 3L
  - Be familiar with the usefulness and policy relevance of measuring utility

Content

- Background
- Health outcome measurement in health economic evaluation
- Utility
- Measurement of utility - 1. direct measures
- Measurement of utility - 2. indirect measures

BACKGROUND

Background

- Limited resources in all countries (even in high income countries)
- Need to make choices between available therapies
- Necessity to compare at least 2 alternatives
- Decision
  - Implicit (occurrent decisions, higher opportunity cost)
  - Explicit (full economic evaluations, MCDA)

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Need for full economic evaluation

- An evaluation of health benefits only is not sufficient to justify reimbursement decisions
  - Running out of budget
- Evaluation of costs may violate patient safety or jeopardize outcomes
  - Reimbursement of ineffective technologies or the risk of not reimbursing cost effective technologies
- A cost evaluation and/or budget impact analysis without economic evaluation may result in false conclusions

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Key questions for decision-makers

- Does the new therapy provide health gain?
  - Comparison with placebo
- Does the new therapy provide more health gain than the current standard care?
  - Comparison with other technologies
- Is it of good value for the money?
  - Comparing incremental cost and health gain

Inotai A. et al. (2009) Int J Technol Assess Health Care, 25: 190-195.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Distinguishing characteristics of healthcare evaluations

1) Comparison of two or more alternatives?
2) Both costs (inputs) and consequences (outputs) of the alternatives examined?

	(2) No		(2) Yes
	Examines only consequences	Examines only costs	
(1) No	1A Partial evaluation 1B outcome description	cost description	2 Partial evaluation cost - outcome description
(1) Yes	3A Partial evaluation 3B efficacy or effectiveness evaluation	cost analysis	4 Full economic evaluation cost-minimzation analysis cost-effectiveness analysis cost-utility analysis cost-benefit analysis

Drummond et. al. Methods for economic evaluation of health care programs.
Oxford University Press, 1997

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

HEALTH OUTCOME MEASUREMENT IN HEALTH ECONOMIC EVALUATION

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Requirements for measuring consequences in full health economic evaluations

- The unit of consequences should be able to capture health gain of the entire range of health technologies
  - Cholesterol reducing drugs, dialysis, hip prosthesis, PET CT

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Natural units as health outcomes

Natural units

- Hgmm blood pressure reduction
- mmol/L cholesterol reduction
- disease (ulcer) free days
- symptom free days
- avoided amputation
- life years gained
- ...

Full economic evaluation

Type of analysis	Abbr.	Unit of inputs (costs)	Unit of outputs (consequences)
cost-minimization	CMA	money	identical
Cost effectiveness	CEA	money	natural unit (e.g. Hgmm, life years)

Drummond et. al. Methods for economic evaluation of health care programs.
Oxford University Press, 1997

How to interpret and compare the clinical benefits of health technologies?

- 50% reduction of epilepsy attacks
- 50% reduction of tumor size
- 15% improvement in lung functions
- 4 points improvement in ADAS-Cog score
- 20% less patients need symptomatic therapy
- 2 points improvement in the 10-point pain VAS scale

Which outcome is more important?

Requirements for measuring consequences in full health economic evaluations

- The unit of consequences should be able to capture health gain of the entire range of health technologies
  - Cholesterol reducing drugs, dialysis, hip prosthesis, PET CT
- The unit of consequences should be able to compare different health technologies (comparability)
  - (reduction of blood pressure in Hgmm, higher resolution of imaging diagnostics)

Full economic evaluation

Type of analysis	Abbr.	Unit of inputs (costs)	Unit of outputs (consequences)
cost minimization	CMA	money	identical
cost effectiveness	CEA	money	natural unit (e.g. Hgmm, life years)
cost-utility	CUA	money	quality adjusted life years (e.g. QALY)
cost-benefit	CBA	money	money

Drummond et. al. Methods for economic evaluation of health care programs.
Oxford University Press, 1997

Economic background of utility

- Utility, or usefulness, is the ability of something to satisfy needs or wants
- It represents satisfaction experienced by the consumer of a good. Not coincidentally, a good is something that satisfies human wants and provides utility, so satisfying needs improves utility
- Evolution of utility theories
  - Cardinal utility under certainty
  - Ordinal utility under certainty
    - Preference
  - Ordinal utility under uncertainty
  - Neumann-Morgenstern:
    - Cardinal utility (interval scale), under uncertainty
- Utility theories can be differentiated according to
  - Handling uncertainty
  - Scale of measurement (ordinal/cardinal)

Neumann Morgenstern utility theory

- Any individual whose preferences satisfied four axioms (completeness, transitivity, independence, continuity) has a *utility function*;
- such an individual's preferences can be represented on an *interval scale* and
- the individual will always prefer actions that *maximize expected utility*.

Utility, value, preference

- According to NM utility theory, only those *preferences* are *utilities*, which consider the subject's attitude toward risk (uncertainty)
- *Preferences* without considering the subject's attitude toward risk are *values*
- However, in the wider medical literature, *values* (indices) are often considered as *utilities*

Methods of measuring *preferences*

Response method	Question framing	
	Certainty (*values*)	Uncertainty (*utilities*)
Scaling	I. Rating scale (RS) Visual analogue scale (VAS)	II. –
Choice	III. Time trade-off (TTO) Person trade-off (PTO)	IV. Standard gamble (SG)

Drummond et. al. Methods for economic evaluation of health care programs. Oxford University Press, 1997

Economic criteria of health status measures

- Comparability across disease (**<-**)
- Interval scale (**<-**)
- Individual preference-based scoring?

Economic and psychometric criteria of health status measures

			Psychometric criteria			
			Reliable	Valid	Practical	Responsive
Mortality			Depends on surveillance system			No
Morbidity			Depends on surveillance system			No
Disease specific measures			Yes & No	Yes & No	Yes & No	Yes
Generic health profiles			Yes	Yes	Yes	More than preference-based indices, less than disease specific measures
Generic indices (non-preference)			Yes	Yes	Yes	More than preference-based indices, less than disease specific measures
Generic indices (preference based)			Yes (unless very many health states)	Yes	Yes	Often not

Fox-Rushby J, et al. Open University Press, McGraw Hill, 2005: 85-100.

Economic and psychometric criteria of health status measures

	Economic criteria			Psychometric criteria			
	Compara-bility across disease	Interval scale	Individual preference-based scoring?	Reliable	Valid	Practical	Responsive
Mortality	Yes	Yes	No	Depends on surveillance system			No
Morbidity	No	No	No	Depends on surveillance system			No
Disease specific measures	No	Yes & No	No	Yes & No	Yes & No	Yes & No	Yes
Generic health profiles	Yes	No	No	Yes	Yes	Yes	More than preference-based indices, less than disease specific measures
Generic indices (non-preference)	Yes	No	No	Yes	Yes	Yes	More than preference-based indices, less than disease specific measures
Generic indices (preference based)	Yes	Yes	Yes	Yes (unless very many health states)	Yes	Yes	Often not

Fox-Rushby J, et al. Open University Press, McGraw Hill, 2005: 85-100.

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Economic and psychometric criteria of health status measures

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Economic and psychometric criteria of health status measures

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Trade off: economic vs. psychometric criteria

- Specific measures: favorable psychometric attributes with limited comparability (**<-**)
- Survival: favorable economic attributes without considering quality of life vs. a holistic approach in healthcare (**<-**)
  - Considering only life expectancy would mean: living 1 year in perfect health = living 1 year in a coma
- Preference-based generic index type measures: favorable economic attributes (**<-**) with psychometric limitations

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Methods for valuing health states

- Expert opinion/expert panel
  - PROs: cheap and feasible
  - CONs: experts have different perspectives as (potential) patients
- Existing literature
  - PROs: expanding number of published data
  - Is the population of the identified paper comparable to the population to be valued (e.g. age, disease severity etc.)?
  - Was the measure and methodology applied in the paper accurate (e.g. validated)?
- Measuring
  - PROs: most accurate
  - CONs: resource intensive

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Collective priority-setting

- To systematically compare the benefits of different kinds of healthcare techniques
- An extremely versatile benefit measure with interval scale measurement properties to compare the size of differences in levels of benefit between treatments
- Any measure which fails to fulfill these criteria is inadequate in principle as an aid to priority-setting

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

UTILITY

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Definition of utility

- "Utility is a quantitative expression of an individual's preference for, or ability of, a particular state of health under conditions of uncertainty" (**<-**)

Berger ML et al. Health care cost, quality and outcomes – ISPOR book of terms. Utility. ISPOR, 2003: 241-242.

Utility

- Conventional utility scale
  - Fix (reference) points:
    - Utility of 0.0 for dead
    - Utility of 1.0 for perfect health
  - States worse than dead can have negative utilities
- Assignment of utilities
  - By direct measurement (TTO, SG, RS) (**->**)
  - Indirectly, by using utility weighted index (EQ-5D index, SF-6D) (**->**)

Describing a health state

- You have a life-threatening illness which is stable on treatment. You are receiving cycles of treatment which require you to go to the outpatient clinic.
- You have recently had a blood disorder which led to you being hospitalized for about 5 days with a fever and severe flu-like symptoms. You received intravenous antibiotics because this blood disorder could have caused you to die within a few days of onset. You are at risk of it happening again following your next cycle of treatment.
- Your appetite is reduced. You sometimes experience significant pain which can be treated with painkillers. You are able to visit family and friends but often have to cut visits short because you get tired.
- You are able to wash and dress yourself and do jobs around the home. Shopping and daily activities take more effort than usual. You were unable to do these things when you had the fever and flu-like symptoms.
- You feel less physically attractive than usual and your sexual drive is reduced.
- You feel quite anxious and depressed. You worry that your disease may progress in the future.

Lloyd A et al. (2006). Br J Cancer, 95: 683-690.

MEASUREMENT OF UTILITY
1. DIRECT MEASURES

Direct methods for health state valuation

- RS (rating scale)
- TTO (time trade-off)
- SG (standard gamble)

Visual analog scale

Best imaginable health state

Your own health state today
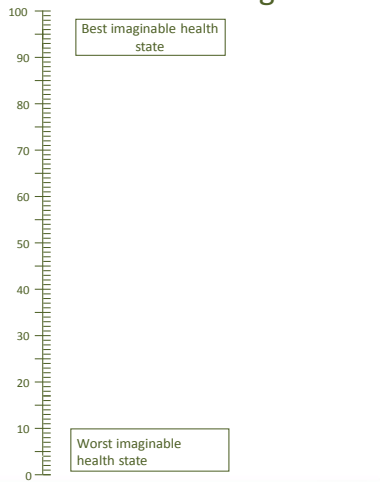
To help people to say how good or bad a health state is, we have drown a scale (rather like a thermometer) on which the best state you can imagine is marked by 100 and the worst state you can imagine is marked by 0.

We would like you to indicate in this scale how good or bad your health state is today, in your opinion. Please do it by drawing a line from the box abowe to whichever point on the scale indicates how good or bad your current health state is today.

Worst imaginable health state

Valuing health states with a rating scale

- A batch of health states are given to the respondent
- The respondent ranks health states from most preferred to least preferred, then locates them on the VAS
- Intervals and spacing correspond to the differences in preference of subjects
- The subject is instructed to concentrate on these intervals and spacing and comparison of one interval to another rather than on scores
- Ratios of scale values are meaningless in an interval scale
  - X 'Outcome A is twice as desirable as B so I will place it twice as high on the scale'
  - ✔ 'The difference in desirability between outcome A and B is twice as great as the difference between C and D hence I will make the interval (space) between A and B twice as large'

Measuring chronic states

- Preferences for chronic states can be measured on a RS
- Chronic state: irreversible, from age of onset until death (same for all health states)
- The batch of health states include:
  - normal healthy life (1.0)
  - death at age of onset (0.0)
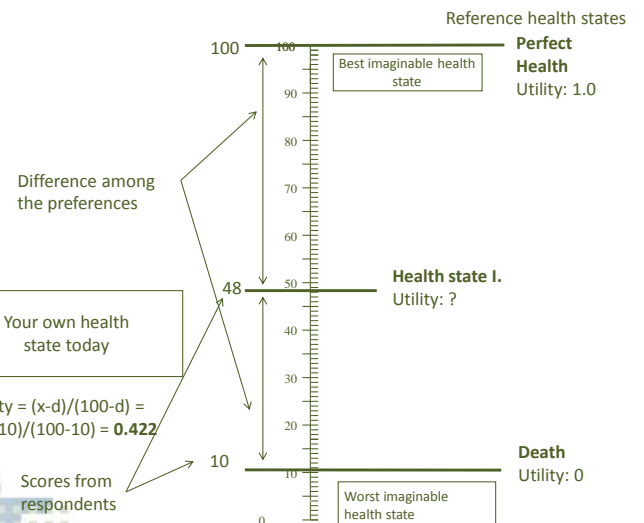- Locate the whole batch on the rating scale
- If death is not judged to be the worst health state at the lower end of the scale, utilities are calculated as (x-d)/(1-d) where:
  - d: score of death
  - x: score of health state to be measured

Reference health states

100 — Perfect Health — Utility: 1.0

Best imaginable health state

Difference among the preferences

Your own health state today

$$\text{Utility} = (x-d)/(100-d) = (48-10)/(100-10) = \mathbf{0.422}$$

48 — Health state I. — Utility: ?

Scores from respondents

10 — Death — Utility: 0

Worst imaginable health state

Exercise

Reference health states

100 — Perfect Health — Utility: 1.0

Best imaginable health state

Difference among the preferences

65 — Dialysis — Utility: 0,6

Your own health state today

50 — Health state I. — Utility: ?

Worst imaginable health state

Time Trade-Off

Value

HEALTHY — 1 — Alternative 1

HEALTH STATUS — i — Alternative 2

DEAD — 0

x   t   Time

Torrance GW. (1986) J Health Econ

Time Trade-Off

- The subject is offered two alternatives
  - state i for time t (life expectancy of an individual with chronic condition) followed by death
  - healthy for time x < t followed by death
- Time X is varied until the respondent is indifferent between the two alternatives, at which preference score is: x/t

Time Trade-Off

- Designed for use in healthcare
- No uncertainty (<-) among the two alternatives
- If subject is indifferent between 40 years of life expectancy in chronic health state and 30 years in perfect health, his preference score is x/t=30/40= 0.75
- Time frame:
  - t time fixed (e.g. 10, 20, 40 years)
  - t time variable (e.g. life expectancy standardized to age and gender)
- Question framing
  - 'Ping-pong' method
  - 'Iterative' method

'Ping-pong' method

Iterative method
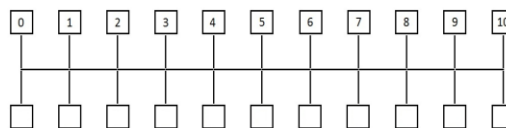
- t=fixed, for 10 years
- Would you trade 10 years in dialysis for t=5 years in perfect health?
- **If YES:** Would you trade 10 years in dialysis for t=4 years in perfect health?
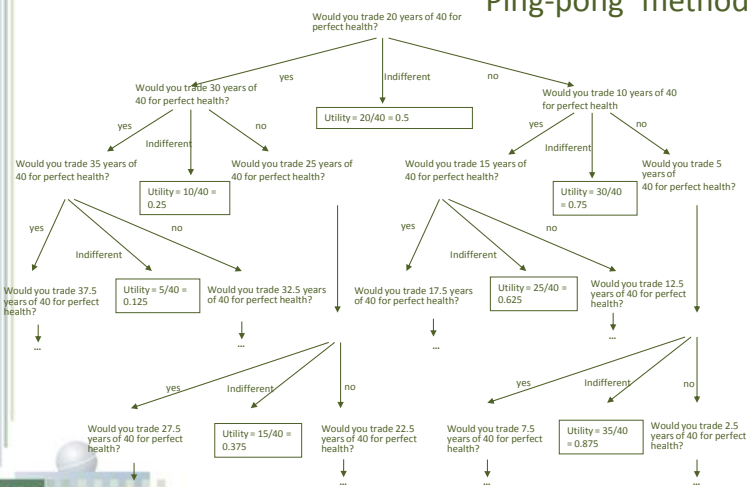 - If the evaluator is indifferent about two options: t/10=4/10=0.4
- **If NO:** Would you trade 10 years in dialysis for t=6 years in perfect health?
  - If the evaluator is indifferent about two options: t/10=6/10=0.6
- **If** the evaluator is **indifferent** about two options: t/10=5/10=0.5

Standard Gamble
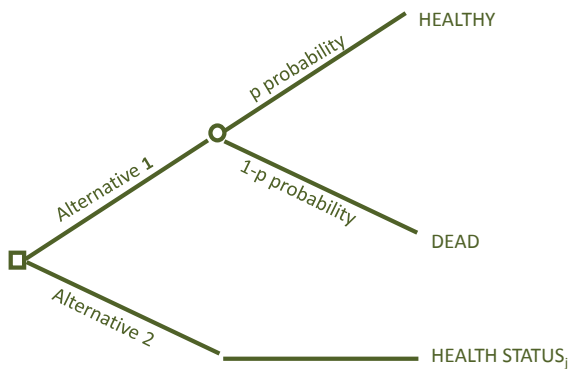
Ref. Torrance GW. (1986) J Health Econ

Standard Gamble

- The subject is offered two alternatives
  - Treatment with 2 possible outcomes:
    - Subject immediately returns to perfect health for an additional t years (p)
    - Subject immediately dies (1-p)
  - Certain outcomes of chronic state i for t years
- Probability p is varied until respondent is indifferent about the two alternatives
- The preference score for state i for time t is p

117

Standard Gamble

- One of the alternatives includes uncertainty (**<-**) (p probability)
- 25% risk of mortality (1-p)
- p=75%
- Preference score for health state i is 0.75
- To illustrate p for respondent: a probability wheel with gyroscope
- Possible outcomes indicated with different colors
  - Perfect health: light gray
  - Immediate death: dark gray
- Sector of the circle corresponds with p value and it's size can be changed according to the preference of the respondent

Probability wheel

Perfect health

Death

This side to interviewee

0%
10%
20%
30%

This side to interviewer
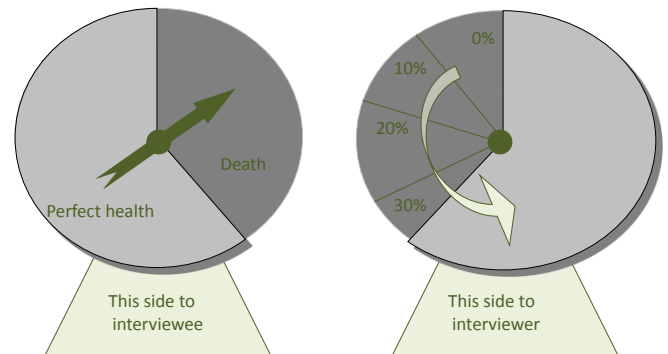
Utility, value, preference

- Many people use 'utility', 'value', 'preference' interchangeably
- Preference: umbrella term for the overall concept
- Value: derived from a question framed *under certainty* (**<-**) by *comparing 2 outcomes* where you need to choose between them or scale them (RS/VAS,TTO)
- Utility: derived from a question framed *under uncertainty* (**<-**) by *comparing 2 outcomes* where at least one contained uncertainty i.e. probability to capture subject's risk attitude (SG)

Categoring methods of measuring preferences

| Response method	Question framing	
	Certainty (values)	Uncertainty (utilities)
Scaling	I. Rating scale (RS) Visual analogue scale (VAS)	II. -
Choice	III. Time trade-off (TTO) Person trade-off (PTO)	IV. Standard gamble (SG)

Drummond et. al. Methods for economic evaluation of health care programs. Oxford University Press, 1997

Comparing direct health state valuation measures

- RS does not involve decision situation with potential health loss (e.g. trading-off life expectancy or risking immediate death)
- Many consider SG as the gold standard because of how it handles uncertainty (**<-**)
- SG provides higher preference scores compared to TTO for the same health state in the same subjects:
  - For subjects who are risk avoidant (risk of immediate death – SG will overestimate the preference score
  - For subjects who have positive time preferences (less value attributed for lifetime just before death – TTO will underestimate the preference score)

MEASUREMENT OF UTILITY
2. INDIRECT MEASURES

Indirect measurement of utility with HRQoL instruments

- Multi-attribute health status classification systems with preference scores
- Define finite numbers of health states
- These health states are valued by a population with direct valuation methods
- The link between HRQoL and utility
  - Rosser-Kind matrix
  - EQ-5D
  - HUI, HUI2, HUI3
  - SF-6D

Valuation matrix of Rosser and Kind

Disability	Distress			
	A	B	C	D
I	1.000	0.995	0.990	0.967
II	0.990	0.986	0.973	0.932
III	0.980	0.972	0.956	0.912
IV	0.964	0.956	0.942	0.870
V	0.946	0.935	0.900	0.700
VI	0.875	0.845	0.680	0.000
VII	0.677	0.564	0.000	-1.486
VIII	-1.028	-	-	-

FIXED POINTS: HEALTHY = 1   DEAD = 0

York Discussion paper 38.

EQ-5D 3L questionnaire ('EuroQoL')

- Two components:
- EQ-5D VAS
  - Cannot be interpreted as utility (<-)
- EQ-5D index

EQ-5D 3L questionnaire

Mobility
I have no problems in walking about ☐
I have some problems in walking about ☐
I am confined to bed ☐

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

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

Pain/discomfort
I have no pain or discomfort ☐
I have moderate pain or discomfort ☐
I have extreme pain or discomfort ☐

Anxiety/depression
I am not anxious or depressed ☐
I am moderately anxious or depressed ☐
I am extremely anxious or depressed ☐

EuroQoL © EuroQol Group. EQ-5D™ is a trade mark of the EuroQol Group

EQ-5D 3L index

- $3\times3\times3\times3\times3=3^5=243$ possible health state
- \+ death + unconsciousness
- PROs:
  - Simple,
  - Easy to understand,
  - Widely used – international references,
    - Comparability among diseases (even with a healthy population, and among different countries
  - Several validated translations in different languages to conduct multinational trials
- CONs:
  - Limited sensitivity to detect small changes

EQ-5D tariffs (Dolan N3)

dimension	Coefficient
constant	0.081
mobility	
1.2. level	0.069
2.3. level	0.314
self-care	
1.2. level	0.104
2.3. level	0.214
Usual activities	
1.2. level	0.036
2.3. level	0.094
pain / discomfort	
1.2. level	0.123
2.3. level	0.386
anxiety / depression	
1.2. level	0.071
2.3. level	0.236
N3 (3. level in any dimension)	0.269
unconscious	-0.402

Health Status: 11223

calculated utility:
$$1.0 - 0.081 - 0.036 - 0.123 - 0.236 - 0.269 = \mathbf{0.255}$$

Drummond et. al. Methods for economic evaluation of health care programs. Oxford University Press, 1997

EQ-5D 5L

MOBILITY
- I have no problems in walking about ☐
- I have slight problems in walking about ☐
- I have moderate problems in walking about ☐
- I have severe problems in walking about ☐
- I am unable to walk about ☐

SELF-CARE
- I have no problems washing or dressing myself ☐
- I have slight problems washing or dressing myself ☐
- I have moderate problems washing or dressing myself ☐
- I have severe problems washing or dressing myself ☐
- I am unable to wash or dress myself ☐

USUAL ACTIVITIES (e.g. work, study, housework, family or leisure activities)
- I have no problems doing my usual activities ☐
- I have slight problems doing my usual activities ☐
- I have moderate problems doing my usual activities ☐
- I have severe problems doing my usual activities ☐
- I am unable to do my usual activities ☐

PAIN / DISCOMFORT
- I have no pain or discomfort ☐
- I have slight pain or discomfort ☐
- I have moderate pain or discomfort ☐
- I have severe pain or discomfort ☐
- I have extreme pain or discomfort ☐

ANXIETY / DEPRESSION
- I am not anxious or depressed ☐
- I am slightly anxious or depressed ☐
- I am moderately anxious or depressed ☐
- I am severely anxious or depressed ☐
- I am extremely anxious or depressed ☐

EuroQoL © EuroQol Group.
EQ-5D™ is a trade mark of the
EuroQol Group

HUI 2, HUI 3

- Indirect utility measures
- HUI2
  - 24 000 health states
  - 7 dimensions
  - 3-5 levels
- HUI3
  - 972 000 health states
  - 8 dimensions
  - 5 or 6 levels

HUI2 - scoring

Considering only one dimension

Level	Sensation	Mobility	Emotion	Cognition	Self-Care	Pain	Fertility
1	1	1	1	1	1	1	1
2	0.87	0.92	0.86	0.86	0.85	0.95	0.75
3	0.65	0.61	0.6	0.66	0.55	0.75	0
4	0	0.34	0.37	0	0	0.42	
5		0	0			0	

Considering multiple dimensions at the same time

Sensation	Mobility	Emotion	Cognition	Self-Care	Pain	Fertility
x1 b1	x2 b2	x3 b3	x4 b4	x5 b5	x6 b6	x7 b7
1 1.00	1 1.00	1 1.00	1 1.00	1 1.00	1 1.00	1 1.00
2 0.95	2 0.97	2 0.93	2 0.95	2 0.97	2 0.97	2 0.97
3 0.86	3 0.84	3 0.81	3 0.88	3 0.91	3 0.85	3 0.88
4 0.61	4 0.73	4 0.70	4 0.65	4 0.80	4 0.64	
	5 0.58	5 0.53			5 0.38	

$$u^* = 1.06 (b_1 * b_2 * b_3 * b_4 * b_5 * b_6 * b_7) - 0.06$$

Torrance et al. (1996) Med Care, 34: 702-722

HUI3 - scoring

Considering only one dimension

Level	Vision	Hearing	Speech	Ambulation	Dexterity	Emotion	Cognition	Pain
1	1	1	1	1	1	1	1	1
2	0.95	0.86	0.82	0.83	0.88	0.91	0.86	0.92
3	0.73	0.71	0.67	0.67	0.73	0.73	0.92	0.77
4	0.59	0.48	0.41	0.36	0.45	0.33	0.7	0.48
5	0.38	0.32	0	0.16	0.2	0	0.32	0
6	0	0		0	0		0	

Considering multiple dimensions at the same time

Vision	Hearing	Speech	Ambulation	Dexterity	Emotion	Cognition	Pain
x1 b1	x2 b2	x3 b3	x4 b4	x5 b5	x6 b6	x7 b7	x8 b8
1 1.00	1 1.00	1 1.00	1 1.00	1 1.00	1 1.00	1 1.00	1 1.00
2 0.98	2 0.95	2 0.94	2 0.93	2 0.95	2 0.95	2 0.92	2 0.96
3 0.89	3 0.89	3 0.89	3 0.86	3 0.88	3 0.85	3 0.95	3 0.90
4 0.84	4 0.80	4 0.81	4 0.73	4 0.76	4 0.64	4 0.83	4 0.77
5 0.75	5 0.74	5 0.68	5 0.65	5 0.65	5 0.46	5 0.60	5 0.55
6 0.61	6 0.61		6 0.58	6 0.56		6 0.42	

$$u^* = 1.371 (b_1 * b_2 * b_3 * b_4 * b_5 * b_6 * b_7 * b_8) - 0.371$$

$$Utility(21121213) = 1.371 (0.98 * 1.00 * 1.00 * 0.93 * 1.00 * 0.95 * 1.00 * 0.90) - 0.371 = 0.70$$

Furlong W, et al. CHEPA WP#98-11, App B

Whom should we ask about the relative utility of each health status?

- Patients?
  - X Overrate health state
  - X Overrate value of therapy
  - ✓ Familiar with the health state
- Medical professionals (e.g. physicians)?
  - ✓ Familiar with the health state
  - X Familiar with potential outcomes - underrate health state
  - X Overrate the value of their own profession
- General population?
  - X Unfamiliar with the health state
  - ✓ Sustainers of the healthcare system (taxpayers)
  - ✓ Potential future users of health technology

Mapping

- Generic QoL measures are not sensitive enough to detect small differences (**<-**)
- Specific measures are very often not validated for utility measurement, consequently their capacity to be used in healthcare decision-making is limited
- Many studies apply specific measures
  - PRO: more favorable psychometric characteristics (**<-**) (suggested by EMA) guidance
  - CON: unable to capture utility
- Let's map the specific measure to utility scale
- One cross-sectional study
  - Specific measure (not suitable for measuring utility)
  - A further generic measure (suitable for measuring utility)
  - Regression model

Policy relevance of mapping

- Application of generic method: increase the risk of detecting non-significant differences in QoL
- Payer may not accept QoL improvement as a value message
- Specific measures are often associated with higher sensitivity
  - Minimize the risk of detecting non-significant differences in QoL
  - Through mapping enable specific measures to capture utility

Take home messages

- Heterogeneity of health technologies and health outcomes necessitates a universal construct to measure QoL
- These are index scores with reference points within the instrument to death (0.0) and perfect health (1.0) and may be used to combine changes in quality and quantity of life because there is the possibility of linking, comparing and trading off these different aspects
- States considered worse than death can have negative weights

Take home messages (2)

- Key economic criteria of health status measures
  - Comparability across disease
  - Interval scale
  - Individual preference-based scoring
- Utilities are assigned to health states by direct (RS, TTO, SG) and indirect (EQ-5D, SF-6, HUI 2-3) utility measurement tools
- Policy relevance: utilities are necessary to calculate QALYs applied in a cost-utility analysis to inform paying decision-makers in healthcare

Self-check questions

- Why do we need a universal construct to measure QoL?
- What would be the key characteristics of such a measure?
- What are the core features of a utility scale?
- How can utility be measured?
- What are the pros and cons of different direct health state valuation tools?
- What is the difference between utility, value and preference?
- What are the critical steps of measuring utility with RS?
- How should a TTO interview be conducted?
- How can EQ-5D 3L be used to calculate utility?
- What is the policy relevance of measuring utility to inform healthcare decision-makers?

Suggested reading

- Torrance GW. (1986) Measurement of health state utilities for economic appraisal. J Health Econ, 5: 1-30.

- Berger ML, Bingefors K, Hedblom EC, Pashos CL, Torrance GW. Health care cost, quality and outcomes – ISPOR book of terms. ISPOR, USA, 2003
- Drummond MF, O'Brien BJ, Stoddart LG, Torrance GW. Cost utility analysis. Methods for economic evaluation of health care programs. 2nd edn. Oxford University Press, New York, 1997

8.2. QALY approach

QALY approach

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Learning objectives

- Students should:
  - Be able to describe the necessity of a universal measure to estimate health gain
  - Be able to introduce the key features of the QALY concept
  - Be familar with a cost-utility analysis
  - Be aware of the criticisms concerning QALY

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Content

- Concept of the QALY
- Decision-making

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

CONCEPT OF THE QALY

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Treatment outcomes of malignant disorders
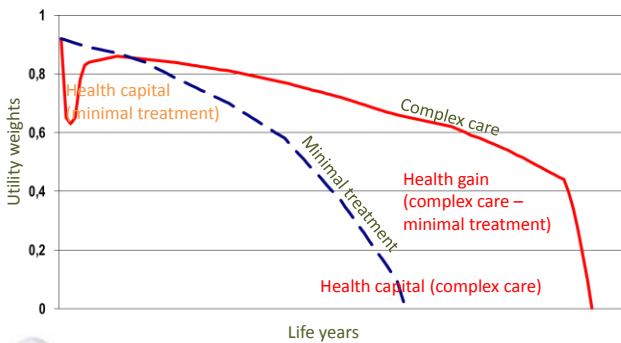
History of QALY

- Herbert Klarman, 1968: Kidney transplantation provides life years gain AND improved QoL compared to dialysis
- This is considered to be one of the first references to quality adjusted life years

QALY

Quality Adjusted Life Years (QALY)

- Is a universal health outcome measure which is applicable to all individuals and all diseases
- Enables comparisons to be made among all disease areas
- QALY combines both quality of life (morbidity) and quantity of life (mortality)
- Life years adjusted by preference-based quality weight (**<- utility**)
- Applied in CUAs

Berger ML et al. Health care cost, quality and outcomes – ISPOR book of terms. Quality Adjusted Life Year (QALY). ISPOR, 2003: 195-197.

QALY

- Typical mistake: commuting QALY and QoL
- Time preference – discounting
- QALY approach: designed to support priority-setting/decision-making in healthcare
- QALYs are used in cost-utility analyses (**->**)
- to estimate the incremental cost of a new therapy for one QALY gain compared to an appropriate comparator

QALY approach

QALY= (0.5x1)+(0.5x0.25)+(0.5x0.5)+(1x0.75)+(0.5x0.5)=1.875

QALY approach

10x1.0+10x0.5=15 QALYs (20 LYs)

10x1.0+6x0.8+6x0.5=17.8 QALYs (22 LYs)

+2 LYs
+2.8 QALYs
LY approach would underestimate health gain
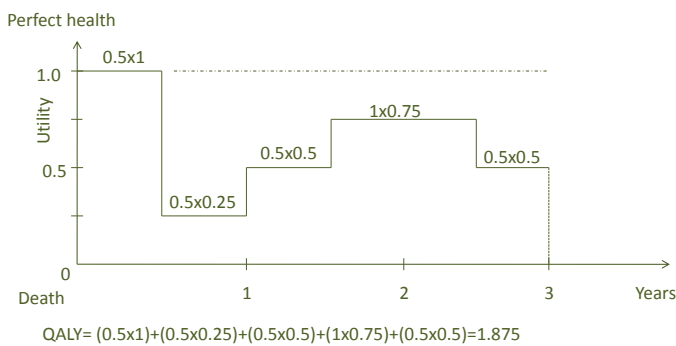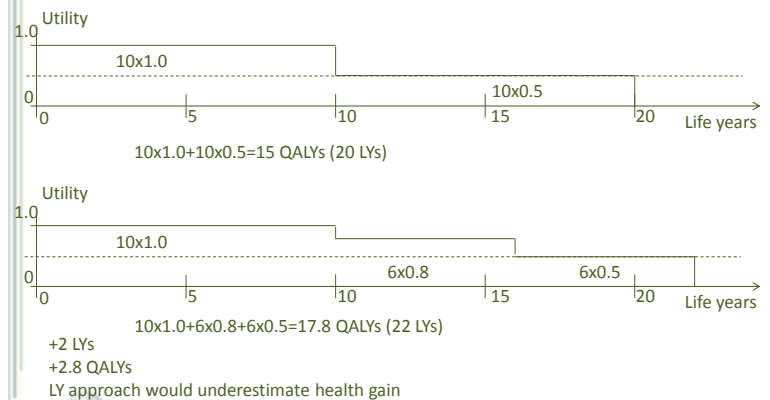
Relevance of QALY

- QALY is suitable to aggregate different dimensions of health outcomes (e.g. in the case of complex oncology treatment **<-**):
  - Long-term life years gain
  - Short-term QoL deterioration (due to AEs)
  - Long-term QoL improvement
- QALY is suitable to estimate the magnitude and sign of aggregate health gain
- Set up objective ranking between health technologies
  - QALY league table

Limitations of the QALY approach

- Limitations of the utility measurement (**<-**)
- Limitations of the QALY concept (**->**)
- Limitations of decision-making based on the QALY approach

DECISION-MAKING

Drummond et. al. Methods for economic evaluation of health care programs. Oxford University Press, 1997

QALY and decision-making

- One of the main objectives of health policy: to maximize the health capital of the population from the available healthcare budget
  - Health capital of a subject: total number of QALYs accumulated by him/herself
  - Health capital of the population: sum of the health capital of individual subjects
  - Assumption: 20 QALYs for 1 subject = 2 QALYs for 10 subjects = 0.2 QALY for 100 subjects
- CUAs compare 2 or more alternatives in terms of both cost and outcomes, where outcomes are measured in units of utility or preference multiplied by the duration, often as QALYs
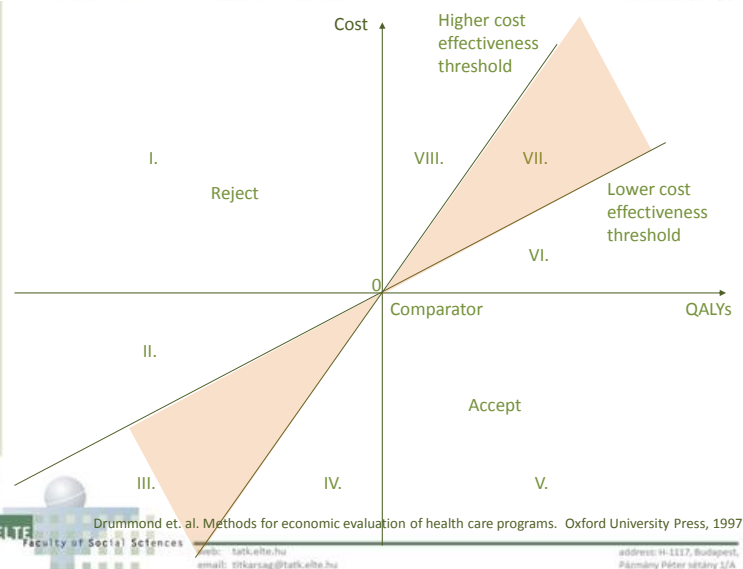
Cost utility analysis

- Unit of health outcomes: QALYs
- Designed to compare 2 alternative health technologies even if their health outcomes cannot be measured with the same natural unit
- Capable of ranking within the entire spectrum of health technologies (Oregon experiment)

Will decision-making in healthcare by applying the QALY concept be more transparent?

- YES, but
  - Not necessarily more rational – methodological concerns (->)
  - Not necessarily more fair – the concept privileges some groups over others
  - Not necessarily cheaper – measuring QoL costs money

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Take home messages

- QALY is a universal construct to measure health gain in the whole spectrum of health technologies
- Policy relevance: QALYs are applied as a health outcome measure to capture both quality and quantity of life in cost-utility analysis to inform paying decision-makers in healthcare
- QALY is suitable for aggregating different dimensions of health outcomes
- QALYs may be used to set up objective rankings between health technologies
- A wide range of critiques exist in the literature about the application of QALY

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Self-check questions

- Why the development and the use of QALY was necessary?
- What is the policy relevance of QALY in informing healthcare decision-making?
- What are the criticisms concerning the application of QALY?

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Suggested reading

- Berger ML, Bingefors K, Hedblom EC, Pashos CL, Torrance GW. Health care cost, quality and outcomes – ISPOR book of terms. ISPOR, USA, 2003
- Drummond MF, O'Brien BJ, Stoddart LG, Torrance GW. Cost utility analysis. Methods for economic evaluation of health care programs. 2nd edn. Oxford University Press, New York, 1997

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

8.3. The transferability of the QALY

The transferability of the QALY

"Financed from the financial support ELTE won from the Higher Education Restructuring Fund of the Hungarian Government"

Learning objectives

- Students should:
  - Be aware of the principles of transferability of health gain within different countries
  - Be aware of the heterogeneity of health state valuation methods
  - Be aware of the development of EQ-5D 3L value sets
  - Be aware of the valuation of EQ-5D Health States based on the MVH protocol
  - Be aware of the classification of EQ-5D 3L health states

Contents

- Introduction
- Transferability of EQ-5D
- Implementation

INTRODUCTION

Health Economic Analysis

Goal: Optimal use of scarce resources

-Comparison of 2 (or more) health technologies

TYPE	ABBREVIATION	MEASUREMENT of HRQoL
Cost-minimization analysis	CMA	Not measured; basic assumption: compared health technologies are non-inferior
Cost-effectiveness analysis	**CEA**	**Natural units (i.e. 1% change in HbA1c level)**
Cost-utility analysis	**CUA**	**QALYs**
Cost-benefit analysis	CBA	Monetary units

Most widely used

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Cost-effectiveness analysis vs. Cost-utility analysis

$$ICER = \frac{\Delta Cost}{\Delta Efficacy} = \frac{C_2 - C_1}{E_2 - E_1}$$

Cost-effectiveness analysis

The incremental cost of 1 unit of incremental health gain (measured in natural units)

Cost-utility analysis

The incremental cost of 1 incremental QALY gain

How can the relative level of the ICER be interpreted from a decision-maker perspective?

Is the level of ICER high or low?

5 000 EUR / 1 month PFS	20 000 EUR /QALY
1 500 EUR/avoided hospitalization	20 000 EUR / avoided bone fracture
60 000 EUR /QALY	18 000 EUR / LY

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Measuring Health Gains

- **HE analysis** → An indicator is necessary which measures HRQoL by a single index number
- **Cost-utility analysis (CUA)** → Measures health gains in QALYs
- **QALY** → Calculated from utility values (but is QALY really a „common denominator"?)

» **Measurement of HRQoL:**

  – Several questions:

  - Who's preferences are measured?
  - Which measurement method/instrument is used?
  - When are the preferences measured?
  - Where are the preferences measured?

Bombardier et al. 1982	Needs walking stick when walking	Needs supervision when walking	Needs two assistants for moving
Standard Gamble	0.85	0.64	0.38
Time Trade-Off	0.78	0.41	0.11
Rating Scale	0.65	0.29	0.08

Nord E et al. Value in Health 2009;12:S10-5

Different utility values can belong to the same health state!!

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Lack of standardization - consequences

Measure	A: Walking with stick, mild pain, unable to work	B: Difficulties in leaving home, discomfort, able to do some work	Difference
	values	values	
EQ-5D (UK TTO)	0.45	0.8	0.35
HUI 2	0.7	0.92	0.22
15-D	0.86	0.92	0.06

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

The assessment of cost effectiveness
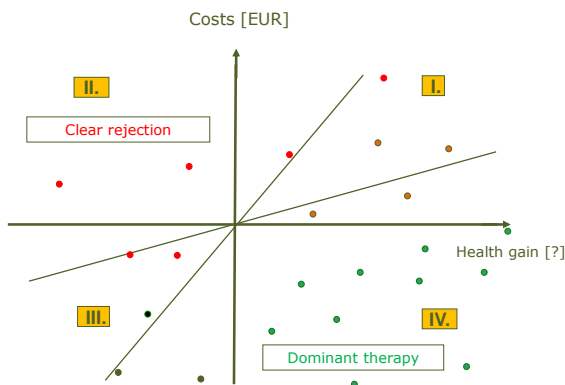
Costs [EUR]

II.

Clear rejection

I.

Health gain [?]

III.

IV.

Dominant therapy

Drummond et. al. Methods for economic evaluation of health care programs. Oxford University Press, 1997

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

The standardization of HRQoL measurement in HE analyses

- There are various HRQoL measures
- There is no "best" HRQoL measurement method (all methods have pros and cons)
- HrQoL measurement in HE analyses should be standardized

HTA in the UK (NICE)

Unit of Health gain/outcome	QALY
HRQoL measuring instrument	EQ-5D
Valuation method	TTO
Source of utility preference	Society (UK EQ-5D value set, sample size: 3 395)

NICE – Guide to the methods of technology appraisal

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

The calculation of the QALY

Stable disease — 0.76
Stable disease with AE — 0.71
Progressed disease — 0.68
Death — 0

The calculation of the QALY

Health State	Utility	Time spent in a Health state	
		„A" medicine	„B" medicine
Stable disease	0.76	5	4
Stable disease with AE	0.71	4	3
Progressed disease	0.68	3	3
Death	0		
QALY		**0.723**	**0.60**
Δ QALY		**0.12**	
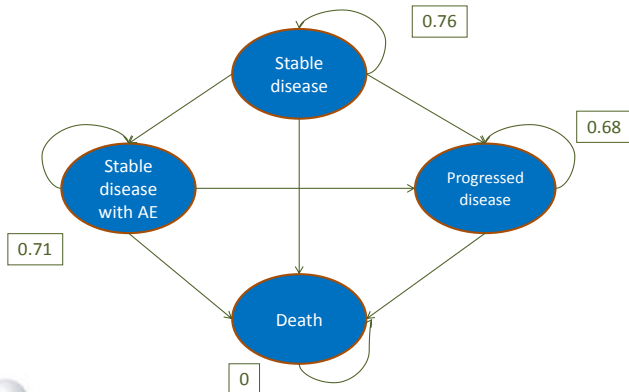
QALY = Utility value of a certain health state multiplied by *years* spent in a certain health state

	Therapy			
	"A"medicine		"B"medicine	
Month	Health State	Utility	Health State	Utility
1	Stable disease	0.76/12	Stable disease	0.76/12
2	Stable disease	0.76/12	Stable disease	0.76/12
3	Stable disease	0.76/12	Stable disease	0.76/12
4	Stable disease	0.76/12	Stable disease	0.76/12
5	Stable disease	0.76/12	Stable disease with AE	0.71/12
6	Stable disease with AE	0.71/12	Stable disease with AE	0.71/12
7	Stable disease with AE	0.71/12	Stable disease with AE	0.71/12
8	Stable disease with AE	0.71/12	Progressed disease	0.68/12
9	Stable disease with AE	0.71/12	Progressed disease	0.68/12
10	Progressed disease	0.68/12	Progressed disease	0.68/12
11	Progressed disease	0.68/12	Death	0
12	Progressed disease	0.68/12	Death	0
13	Death	0	Death	0
TOTAL QALYs		0.723		0.6

Determination of utilities

- Who's preferences should we capture?
  - Patients'?
  - Healthcare Professionals'?
  - Society's?
- When and where should we measure preferences?
- Which measurement technique should we use?
  - Direct methods: VAS, TTO, SG
  - Questionnaires:
    - General: EQ-5D, SF-6D, HUI, etc.
    - Specific: Disease, problem or function specific instruments

The consequence of using different HRQoL measurement methods

- Different HRQoL measurement techniques → different utilities for the same health states → different QALY gains for the same health technology → different ICERs → may affect decision
- Utilities measured with different methods can affect the cost effectiveness of a health technology

Health State	Utility of health state ("X" method)	Time spent in health state (month)	
		"A" medicine	"B" medicine
Stable disease	0,76	5	4
Stable disease with AE	0,71	4	3
Progressed disease	0,68	3	3
Death	0		
QALY		0,723	0,60
Δ QALY		0,123	

Health State	Utility of health state ("Y" method)	Time spent in health state (month)	
		"A" medicine	"B" medicine
Stable disease	0,61	5	4
Stable disease with AE	0,47	4	3
Progressed disease	0,4	3	3
Death	0		
QALY		0,51	0,42
Δ QALY		0,09	

Δ Cost = 2 945 EUR

$ICER_1$ = 23 943 EUR/QALY

$ICER_2$ = 32 722 EUR/QALY

In case threshold = 30000 EUR/QALY $ICER_1$ is cost effective $ICER_2$ not cost effective

Utility determination and Health Economic Guidelines

	Recommended determination of utility values used in HE analyses
England , Wales	EQ-5D
Austria	Not specified
Baltic countries	EQ-5D, HUI
Belgium	EQ-5D, SF-36,TTO, SG
Denmark	TTO, SG
Finland	Not specified
France	QWB, HUI, EuroQol
Netherland	SG, TTO, VAS
Ireland	Base case: EQ-5D or SF-6D, is some cases direct methods are permitted as well: TTO, SG
Poland	Indirect methods, EQ-5D
Hungary	EQ-5D, HUI, QWB, RS, TTO, SG
Germany	Not specified
Norway	General HrQoL instruments (pl.: EQ-5D, SF-6D)
Italy	TTO, SG
Portugal	Not specified
Scotland	EQ-5D
Spain	Direct and indirect methods
Sweden	Direct (SG, TTO) and indirect methods (EQ-5D)
Slovakia	TTO, SG

Huszti Z et al. (2012) Value in Health, 15: A318-A319.

Health outcome measures in Hungarian HTA dossiers (2004 - 2011)

Specific measure 12%
Different measures for differetn health states 3%
Not adequately specified 38%
SG 1%
TTO 6%
VAS (other) 1%
SF-36/SF-6D 9%
EQ-5D VAS 3%
EQ-5D TTO 27%
N=189
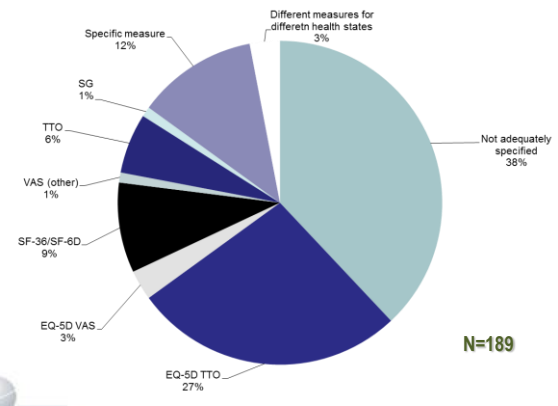
HTA Office, Borsi A (2012) IME 11(1): 30-33.

TRANSFERABILITY OF EQ-5D

The standardization of HRQoL measurement in HE analysis

No "best" method, but HrQoL measurement should be standardized

HTA in UK (NICE)

Expression of health outcomes	QALY
Preferred measure of HRQoL	EQ-5D
Valuation method	TTO
Source of preference	Public preferences elicited using a choice-based method in a representative sample of the UK population

NICE – Guide to the methods of technology appraisal

The EQ-5D method

- HrQoL measurement method developed by EuroQol Group
  - Measures 5 dimensions of HrQoL: mobility, self-care, usual activities, pain/discomfort, anxiety/depression
  - General questionnaire, applied already in more than 120 therapeutic areas
  - Adapted and validated for use in several countries
  - Index type questionnaire
  - Multiple versions (3L, 5L, Y)
  - 243+2 unique health states (3L version)

The EQ-5D questionnaire

By placing a tick in one box in each group, please indicate which statements best describe your health today.

Mobility
I have no problems in walking about
I have some problems in walking about
I am confined to bed

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

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

Pain/Discomfort
I have no pain or discomfort
I have moderate pain or discomfort
I have extreme pain or discomfort

Anxiety/Depression
I am not anxious or depressed
I am moderately anxious or depressed
I am extremely anxious or depressed

EQ-5D User guide

The EQ-5D - 3L

- **245 unique health states**
  - 3^5 = 243 different health states +
  - 2 further health states (death and unconsciousness)
- **A utility score belongs to each health state**
  - Best possible health state (full health) = 1.0 ("11111")
  - Death = 0
  - Health states can have negative utility values
- **HrQoL measurement**
  - **1. Determination of health state:**
    Patients fill out the EQ-5D questionnaire
  - **2. The valuing of determined health state:**
    Calculate utility values for health states determined by patients

Determination of EQ-5D health states

- **"12212" health state:**

  - **Mobility (1)**
    - I have no problems in walking about
  - **Self-Care (2)**
    - I have some problems washing or dressing myself
  - **Usual activities (2)**
    - I have some problems performing my usual activities
  - **Pain/Discomfort (1)**
    - I have no pain or discomfort
  - **Anxiety/Depression (2)**
    - I am moderately anxious or depressed

Determining utility values and calculating QALYs for EQ-5D health states

- **11111 = 1.0**
- **Death = 0**
- **11223 = ?**
- **21232 = ?**
- **12231 =?**
- **VAS, TTO?**

= ? QALY

EQ-5D health states -> QALY

1. Scoring EQ-5D health states
   - $utility^{11223} = 1.0 - 0.081 - 0.036 - 0.123 - 0.236 - 0.269 = \mathbf{0.255}$

2. Calculation of QALYs
   - $QALY = 0.255 * 0.5 = 0.1275$

Utility   Time spent in certain health state (in years)

Valuing EQ-5D health states

- Utility represents (should represent) the preferences of the general population
- Valuation of health states by the members of society
  - With RS/VAS?
  - With TTO?

Utility represents preferences of the general population

- **11111 = 1.0**
- **Death = 0**
- **12212 = ????**
- **21221 = ????**

Even the same preference measurement method can result in different values according to the preference of the society

UK values	Thai values
0.812	0.707
0.691	0.546

Consequence

Year	Health State	UK value	Thai value
1st year	12212	0.812	0.707
2nd year	12212	0.812	0.707
3rd year	21221	0.691	0.546
QALY		**2.315**	**1.96**

EQ-5D value sets

Country	Sample Size	Valuation method
Belgium	722	EQ-5D VAS
Denmark	1686	EQ-5D VAS
Denmark	1332	TTO
Europe	8709	EQ-5D VAS
Finland	1634	EQ-5D VAS
France	452	TTO
Germany	339	EQ-5D VAS
Germany	339	TTO
Japan	621	TTO
New Zealand	1360	EQ-5D VAS
Netherlands	309	TTO
Slovenia	733	EQ-5D VAS
Spain	300	EQ-5D VAS
Spain	1000	TTO
United Kingdom	3395	EQ-5D VAS
United Kingdom	3395	TTO
USA	4048	TTO
Zimbabwe	2440	TTO

http://www.euroqol.org/

Valuation of EQ-5D Health States based on the MVH* protocol

1. Selection of health states to evaluate
2. Interviews & Evaluation

*Measurement and Valuation of Health study; EuroQol Group

Selection and classification of health states

- **243 EQ-5D health states**
  - <u>Mild:</u> Contains no level 3, and maximum three of level 2s
  - <u>Moderate:</u> Health states which are considered neither mild nor severe
  - <u>Severe:</u> no level 1, and minimum two of level 3s
- Further classification of health states
  - <u>Mild:</u> "Distance Group" 1-3
  - <u>Moderate:</u> "Distance Group" 4-6
  - <u>Severe:</u> "Distance Group" 7-9
- Same amount of mild, moderate and severe states should be evaluated by every interviewee AND
- Every interviewee should evaluate "11111" and "33333" health states

Distance Group

Based on the distance from "11111" (full health) health state:

- "12112" health state = 0+1+0+0+1 = 2: Distance group 2 – Mild health state
- "31111" health state= 2+0+0+0+0 = 2: Distance group 2 – based on distance groups this would be a mild health state but because it contains a level 3 →moderate state
- "23333" = 1+2+2+2+2=9: Distance Group 9 – severe health state

The number of evaluable health states

Measurement and Valuation of Health (MVH) study:

- Individuals were required to assess 13 health states (15 states including 11111 and dead) using a TTO procedure
- In case of 11 evaluated health states, the length of evaluation is around 1 hour.

IMPLEMENTATION

The process of health state evaluation

1. **Background information** (age, sex, education, income, marital status, etc.)

2. **Evaluation of their own health state with a**
   - » EQ-5D questionnaire
   - » VAS scale

3. **Evaluation of selected health states**
   - » Ranking exercise
   - » Evaluation of selected health states on VAS scale
   - » *Classification of health states*
     - States better than dead
     - States worse than dead

4. **TTO exercise**

EQ-5D health state "cards"

TTO exercise – State better than dead

- Reference state: "11111"
- Life A means you live for *t years* in full health (11111) and then die.
- Life B means you live for *10 years* in the health state to be evaluated and then die.
- Base case: t = 5
- TTO question: *Would you prefer Life A or Life B or are they the same?*

0	1	2	3	4	5	6	7	8	9	10

X										
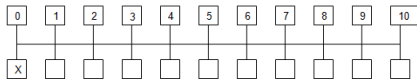
- Calculation of utility: t/10: i.e. t = 4 -> 0.4 (t = years of full health)

TTO exercise – State worse than dead

- Life A means you would live for *t years* in the health state to be evaluated followed by *10-t years* in full health (11111).
- Life B is *immediate* death.
- TTO question: Would you prefer Life A or Life B or are they the same?
- Base case: t=5

0	1	2	3	4	5	6	7	8	9	10

X										
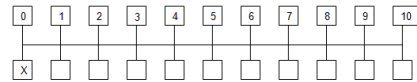
- Calculation of utility: - ((10-t)/(t)); i.e. t = 4 -> -1.5

Next steps

- Identifying inconsistent evaluations
  - i.e. interviewee evaluates "21332" health state (dist. group 6) better than "23212" health state (dist. group 5)
- Selection of regression method to determine utility values for non-evaluated health states
  - Dolan (1997),
  - Dolan & Roberts (2002),
  - Shaw (2005)

Opened questions regarding valuation of EQ-5D health states

- Should it be a 3L or 5L version?
- How about the sample size or representativeness?
  - A TTO exercise is complex and time consuming
- How many health states should be evaluated?
- What is the number of health states an interviewee should evaluate?
- Should every interviewee evaluate the same health states or different ones?
- Etc.

Take home messages

- Different utility measures may result in different utilities – even when applied on the same patients with the same health status
- This might have considerable policy implications and affect cost effectiveness and therefore reimbursement decisions
- Only QALYs derived from the same utility measure and measured in the same countries are comparable
- NICE considers the EQ-5D 3L instrument with TTO valuation methodology as the gold standard to be applied in cost utility analyses
- The development of local EQ-5D value sets can contribute to more relevant reimbursement decisions locally

Self-check questions

- Why can transferability of health gain be an issue?
- What is the policy relevance of different HRQoL measurement methods?
- What are the principles of transferability of health gain within different countries?
- How were EQ-5D 3L value sets developed?
- How were EQ-5D 3L Health States valued based on the MVH protocol?
- How can EQ-5D 3L health states be classified?
- How was a TTO interview conducted in the MVH protocol?

Suggested reading

- Paul Kind. A revised protocol for the valuation of health states defined by the EQ-5D-3L classification system. Learning the lessons from the MVH study

ELTE
Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

8.4. Burden of disease studies

Burden of disease studies

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Learning objectives

- Students should:
  - Be aware of the concept and policy relevance of burden of disease studies
  - Be familiar with PROs and CONs of applying different constructs of measuring health capital (DALY and QALY) in burden of disease studies

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Content

- Burden of disease studies
- Estimating health loss by QALY concept - A Case study

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

BURDEN OF DISEASE STUDIES

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Role of burden of disease studies

- Economic evaluations: improve allocative efficiency in a healthcare system with limited resources (prioritizing among different interventions)
- Economic evaluations are not able to prioritize among different disease areas
- Burden of disease studies are appropriate for identifying unmet needs in healthcare

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Burden of disease

- Economic burden (Cost-of-illness)
  - Direct healthcare cost
  - Direct non-healthcare cost
  - Indirect cost
- Health loss
  - Health capital loss due to early mortality (<-)
  - Health capital loss due to impairment in quality of life (<-)

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Aim of BoD studies

- To aid in setting health service (both curative and preventive) priorities;
- To aid in setting health research priorities;
- To aid in identifying disadvantaged groups and targeting of health interventions;
- To provide a comparable measure of output for intervention, program and sector evaluation and planning.

Murray CLJ. (1994) Bulletin of the World Health Organisation, 72 (3): 429-445.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Cost categories (Cost of Illness - CoI)

	Healthcare cost	Non-healthcare cost
Direct cost	• Cost of drugs • Cost of hospitalization • Outpatient cost	• Travelling cost • Sickness benefit
Indirect cost	• Healthcare cost due to life years gained	• Productivity loss • Societal cost of premature mortality • Income loss of relatives due to medical attendance

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Calculation of health loss (simplified model)
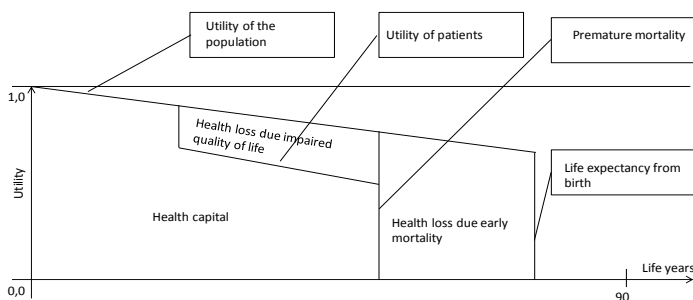
Limitations of the model

- Assumption: Avoiding a chronic condition will result in perfect health (in the real world generally another condition occurs)
- A high number of inputs are needed to feed the model

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

135

Factors influencing magnitude of health loss

- Characteristics of the disease (acute or chronic)
- Disease-specific mortality
- Impact of the disease on utility
- Number of affected patients (prevalence)

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Acute vs. chronic conditions

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

ESTIMATING HEALTH LOSS BY QALY CONCEPT - A CASE STUDY

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Case study: calculation of annual health loss

- In three chronic conditions
  - RA
  - COPD
  - Asthma
- Two components of health loss
  - Health loss due to utility decrement
  - Health loss due to early mortality

Inotai A, et al. (2012) International Journal of Person Centered Medicine, 2: 505-510.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Annual health loss due to utility decrement

Annual health loss due to utility decrement =

Number of population in the age cohort x Prevalence x Utility decrement x 1yr

QALY loss per patient

Number of patients

Mean EQ-5D index of population − mean EQ-5D index of patients

Input:
- Number of population in each age cohort
- Age- and gender-specific prevalence,
- Age- and gender-specific utility (both for patients and population)

Inotai A, et al. (2012) International Journal of Person Centered Medicine, 2: 505-510.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Annual health loss due to early mortality

Annual health loss due to premature mortality =

Population x Mortality x Remaining life expectancy x EQ-5D index of patients

Affected population

QALY loss per patient

Input:
- Number of population in each age cohort,
- Age- and gender-specific prevalence,
- Age- and gender-specific mortality,
- Age- and gender-specific life expectancy,
- Age- and gender-specific utility

Inotai A, et al. (2012) International Journal of Person Centered Medicine, 2: 505-510.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Pros and Cons of applying QALY in the calculation of health loss

- An accurate method for calculation, but with very data-intensive calculations (disease-, age- and gender-specific utility, prevalence and mortality)
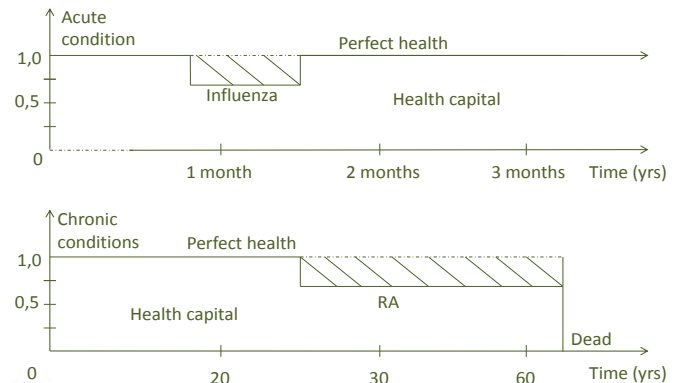- Generalizability (does the sample represent the population?) and comparability (values from different methods) of utilities (<-)
- Assumption: Avoiding a chronic condition will result in perfect health (in the real world another condition occurs)

Take home messages

- Burden of disease studies are appropriate for identifying unmet needs in healthcare
- The two core parts of BoD are health loss and the economic burden of a disease
- The most widely used construct to estimate BoD is DALY (<-)

Self-check questions

- How can burden of disease studies be used to inform healthcare decision-makers?
- What are the core components of disease burden?
- How would you measure health capital/health loss?

Suggested reading

- Inotai A, Ágh T, Mészáros Á. (2012) Quality of life, utility and health burden in asthma, chronic obstructive pulmonary disease and rheumatoid arthritis. International Journal of Person Centered Medicine, 2: 505-510.

8.5. DALY

DALY

Learning objectives

- Students should:
  - Be familiar with the key characteristics of a DALY construct
  - Be familiar with the similarities and differences of QALY and DALY
  - Be familiar with the critique of DALY
  - Be familiar with the advantages and disadvantages of DALY compared to QALY

Contents

- The DALY concept
- Critique of the DALY concept
- Comparison of QALY and DALY

THE DALY CONCEPT

DALY

- DALY – Disability Adjusted Life Years
- A unit of measurement of the impact of disease in terms of both time lost due to premature death (mortality) and time lived with disability (morbidity) (<-)
- Origin: World Bank/WHO
- DALY=YLL+YLD
    - YLL: years of life lost
    - YLD: years lived with disability

Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Disability Adjusted Life Years (DALY). ISPOR, 2003: 69-71.

YLL

- The impact of a particular disease on mortality is estimated by using the difference between life expectancy and the age at which death occurred and is expressed as years of life lost (YLL)
- YLL = average life expectancy – age at death
- The method is standardized by using
    - The average life expectancy of Japanese women (82.5 yrs.) for women
    - Arbitrary value of 80 yrs. for men

Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Disability Adjusted Life Years (DALY). ISPOR, 2003: 69-71.

YLD

- YLD measures the impact of morbidity on DALYs, considering the following factors:
    - The extent of disability associated with non-fatal conditions (disability weights). Endpoints:
        - 0.0 – perfect health
        - 1.0 - death
    - The relative importance of a healthy life at different ages (age weights)
    - The time preference for health (the value of health gained now as compared to the value of health gained in the future) (discounting with 3%)
- Multiplied by the duration of the disability

Berger ML, et al. Health care cost, quality and outcomes – ISPOR book of terms. Disability Adjusted Life Years (DALY). ISPOR, 2003: 69-71.

DALY Disability weights

Category	Original description	weight	
1	Limited ability to perform at least one activity in one of the following areas: recreation, education, procreation or occupation.	0.096	Better health
2	Limited ability to perform most activities in one of the following areas: recreation, education, procreation or occupation.	0.220	
3	Limited ability to perform activities in two or more of the following areas: recreation, education, procreation or occupation.	0.400	
4	Limited ability to perform most activities in all of the following areas: recreation, education, procreation or occupation	0.600	
5	Needs assistance with instrumental activities of daily living such as meal preparation, shopping or housework.	0.810	Greater disability
6	Needs assistance with activities of daily living such as eating, personal hygiene or toilet use.	0.920	

Murray CLJ. (1994) Bulletin of the World Health Organisation, 72 (3): 429-445.

Relative value of one year at different age (DALY age weights – indicative figure)
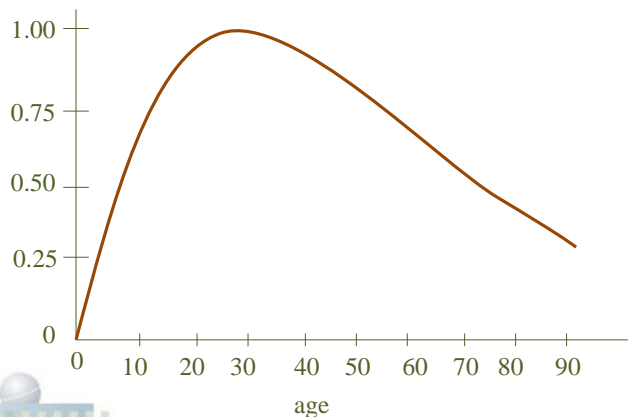
QALY vs. DALY
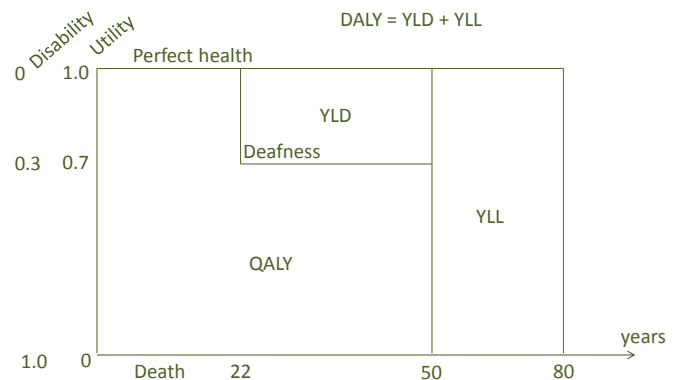
Arnesen et al. (1999) BMJ, 319, 1423-25

WHO GBD
European DALYs (2004)

Cause	EUROPE
TOTAL DALYs (1000.0)	255 410
I. Communicable diseases, maternal and perinatal conditions and nutritional deficiencies	29 090
Infectious and parasitic diseases	10 661
Respiratory infections	6 148
Maternal conditions	1 019
Perinatal conditions (e)	8 793
Nutritional deficiencies	2 470
II. Noncommunicable conditions	192 393
Malignant neoplasms	33 299
Mouth and oropharynx cancers	1 096
Oesophagus cancer	862
Stomach cancer	2 600
Colon/rectum cancer	3 751
Liver cancer	1 030
Pancreas cancer	1 476
Trachea/bronchus/lung cancers	6 385
Melanoma and other skin cancers	547
Breast cancer	3 379
Cervix uteri cancer	882
Corpus uteri cancer	651
Ovary cancer	962
Prostate cancer	1 113
Bladder cancer	861
Lymphomas, multiple myeloma	1 405
Leukaemia	1 441
Other neoplasms	579
Diabetes mellitus	4 520
Nutritional/endocrine disorders	2 194
Neuropsychiatric disorders	33 318
Sense organ disorders	13 236
Cardiovascular diseases	67 548
Respiratory diseases	10 239
Digestive diseases	11 304
Diseases of the genitourinary system	2 266
Skin diseases	435
Musculoskeletal diseases	8 465
Congenital abnormalities	3 850
Oral diseases	1 140
III. Injuries	33 927
Unintentional injuries	24 630
Intentional injuries	9 297

WHO estimates the European BoD to be 255 million DALYs annually. 33.3 million DALYs are attributable to malignancies.

3.38 million DALYs are attributable to BRC

http://www.who.int/healthinfo/global_burden_disease/estimates_regional/en/index.html
Results are presented in 1000 DALYs

Relevance of DALY

- Applicability of DALY
  - Burden of disease studies (WHO GBD) (->)
  - Cost-effectiveness studies (cost/avoided DALYs)
  - Support evidence-based health policy decision-making
- Applicability of DALY has been heavily criticized
- According to these critiques, DALY cannot be used for priority-setting among healthcare interventions

---

CRITIQUE OF THE DALY CONCEPT

Critique of the DALY approach: applicability

- DALY measures the burden of disease (BoD) *and* underdevelopment
  - Using the standardized life expectancies assumes that health interventions alone are capable of achieving an increase in life expectancy to these higher levels
 - However, many non-health circumstances needs to be changed for life expectancy to rise to the level used in the DALY calculations
- DALY is not suitable for measuring BoD
  - BoD consists of health burden and economic burden (->) DALYs focuses on disability and does not take account of healthcare costs.

Lyttkens CH. (2003) European Journal of Health Economics, 4(3): 195-202.

Critique of the DALY approach: discrimination

- DALY discriminates among subjects (maximizing avoided DALYs)
  - It becomes more valuable to save the life of a young person rather than someone older (more DALYs to avoid)
  - Similarly, there is a priority to save someone's life who is healthy compared to disabled – enhances inequality ('Double jeopardy')
- DALY discriminates among subjects (age weights)
  - DALY discriminates against the young and elderly compared to the middle-aged
- DALY discriminates among programs (discounting)
  - DALY is discriminative against preventive health technologies, and discounting would justify environmental degradation today where the present generation benefits at the expense of future generations

Lyttkens CH. (2003) European Journal of Health Economics, 4(3): 195-202.

Critique of the DALY approach: methodological concerns

- Group of respondents to estimate disability weights
  - DALY uses experts to value health states
- Disability weights
  - The most severe disability group by Murray (Class 6): '*Needs assistance with activities of daily living such as eating, personal hygiene or toilet use*'. Infants are not capable of feeding themselves – does this imply that they are disabled?
  - PTO incorporates both the valuation of states of health and people's views about distributional issues

Lyttkens CH. (2003) European Journal of Health Economics, 4(3): 195-202.

COMPARISON OF QALY AND DALY

QALY vs. DALY

DALY	Attribute	QALY
Part of the concept - Arbitrary value (Life expectancy of Japanese women/80 yrs. for men)	Life expectancy	Not required for the concept
Disability	Name of weights in the concept	Utility/value/preference
Modified person trade-off	Estimation method for the weights	Direct health state valuation methods, general indices with preference scores
2 endpoints (0.0 for perfect health, 1.0 for dead)	Calibration of scale	2 fixed points (0.0 for dead, 1.0 for perfect health)
Experts	Source of weights	General population
yes	Applies age weights?	no
yes, 3% discounting per protocol	Applies discounting?	Only in economic evaluations. Not part of the concept, discount rates are country-specific

Advantages of QALY compared to DALY

- No data on life expectancy is required for a QALY calculation
- Valuation method is less problematic compared to PTO, and methods are much more accepted
- QALY weights are widely available
- Use of QALY in CUAs (priority-setting among different health technologies) is far more accepted

Disadvantages of QALY compared to DALY

- QALY is not suitable for studies such as WHO GBD (extremely high amount of input)
- No worldwide data on BoD exists with the QALY concept
- DALY bears constant bias due to the homogeneity of the applied methodology to estimate disability weights. The magnitude of bias is heterogenic and unknown because of different valuation methods, value sets etc.

Conclusion

- DALY has been heavily criticized mainly due to discounting, age weights and methods to estimate disability weights
- In its present form DALY cannot be used for measuring burden of disease (DALY does not capture cost) or priority setting
- DALY measures are suitable for measuring health loss
- The use of QALYs in priority-setting is far from unproblematic, however, QALY is much more accepted for prioritizing among health technologies

Take home messages

- DALY is a unit of measurement of the impact of disease in terms of both time lost due to premature death (mortality, years of life lost) and time lived with a disability (morbidity, years lived with disability)
- DALY has been criticized as a tool to estimate BoD on ethical and methodological grounds and has been heavily criticised as a tool to perform cost utility analyses
- DALY is the most widely used construct to estimate BoD

Self-check questions

- What are the key characteristics of a DALY construct?
- What are the similarities and differences of QALY and DALY?
- What are the main conceptual, methodological and ethical critiques of DALY?
- What are the advantages and disadvantages of DALY compared to QALY?

Suggested reading

- Murray CLJ. (1994) Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bulletin of the World Health Organisation, 72 (3): 429-445.
- Lyttkens CH. (2003) Time to disable DALYs? On the use of disability-adjusted life-years in health policy. European Journal of Health Economics, 4(3): 195-202.

8.6. Critique of the QALY approach

Critique of the QALY approach

Based on the paper of Alan Williams (1996)

"Financed from the financial support ELTE won from the Higher Education Restructuring Fund of the Hungarian Government"

Learning objectives

- Students should
  - have a balanced view on the use of QALY and DALY
  - Have knowledge on advantages and disadvantages of QALY and DALY
  - Be able to argue for and against the use of QALY and DALY
  - Be familiar with methodological, conceptual and ethical critiques of QALY and DALY

Content

- QALY – Ethical concerns
- QALY – methodological concerns
- Concerns with QALY-based decision-making

Objectors on ethical grounds to the use of QALYs in priority-setting in public healthcare systems

1. Those who reject all collective priority-setting as unethical should be asked – **Are implicit decisions better? How can we reduce corruption related to implicit decisions?**
2. Those who accept the need for collective priority-setting but believe that it is contrary to medical ethics should be asked – **OK, but how should we handle the scarcity of health care resources?**
3. Those who accept the need for collective priority-setting and do not believe that it is contrary to medical ethics, but reject the role of QALYs in it on other ethical grounds should be asked – **How should we ensure optimal allocation based on individual preferences from public resources, and deal with discrimination and equity?**
4. Those who accept the need for collective priority-setting in principle, but are unwilling to specify how it should be done in practice should be asked – **OK, what else is better?**

Williams A (1996) Soc Sci Med, 43(12): 1795-1804

QALY – ETHICAL CONCERNS

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Those who reject all collective priority-setting as unethical

- There are scarce resources in healthcare even in very high income countries
- Resources devoted to one person will be denied some other person who might have better benefitted from them.
- Cost represents sacrifices made by other potential patients who did not get treated
- Benefits gained by those to whom treatments are offered should be greater than benefits sacrificed by those who are denied treatment – allocative efficiency
- We should do as much good as possible with our limited resources
- Extending the healthcare budget: Recreates the dilemma of scarce resources at a higher level of spending

Williams A (1996) Soc Sci Med, 43(12): 1795-1804

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Those who reject all collective priority-setting as unethical

- Harris does not accept limited resources in a *healthcare budget* where lives are at stake, in which cases a fundamental reappraisal of priorities are necessary and the *national budget* should be reconsidered.
- Harris accepts scarce resources only if in the national budget headings of expenditure have more important aims for the society than rescuing citizens in mortal danger (investment to infrastructure?)
- Those who reject all collective priority-setting as unethical typically assert that it is immoral to sit in judgement over the worth of other people's lives
- As they come to recognise the reality of scarcity, they acknowledge that some people must be denied the benefits of healthcare
- They want to do this in a manner which is free of interpersonal judgements concerning the relative worth of someone's life
- Ultimately, someone has to make a conscious decision on how best to discriminate between people when confronted with scarcity

Harris (1987), J Med Ethics, 13: 117-123.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Those who accept the need for collective priority-setting but believe that it is contrary to medical ethics

- They accept the need of priority setting but believe it is contrary to medical ethics
- Extreme opinion: The doctor's duty to do everything possible for the patient no matter what the costs are
- Rare diseases – orphan drugs with very high costs and marginal health gains
- If cost represents a sacrifice, ignoring costs mean ignoring sacrifices of those who are denied treatment
- Medical ethics does not require everything possible to be done for one patient no matter what the consequences are effecting any others

Williams A (1996) Soc Sci Med, 43(12): 1795-1804

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Those who accept the need for collective priority-setting and do not believe that it is contrary to medical ethics, but reject the role of QALYs in it on other ethical grounds

They would ask:

- Whose values should be counted? (**<-**)
- How should we move from individual values to collective values?
- Shouldn't we be concerned with the distribution of the benefits of healthcare across different people groups, as well as with the total amount of such benefits? (**<-**)
- Are there any other benefits from healthcare that QALYs do not pick up?

Williams A (1996) Soc Sci Med, 43(12): 1795-1804

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Is the QALY approach unacceptable because it uses the wrong people's values?

- ED case study
  - Males assigned a 0.26 utility decrement to ED (in other words: respondents would trade 26% of their life expectancy to avoid ED)
  - Women reported 0.02 utility decrement
- In a democratic society the views of all affected parties should count
- The general public is considered as the most appropriate reference group
- The QALY approach requires us to be explicit about
  - what the values are that are being used,
  - and where they came from

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Is the QALY approach unacceptable because of the way it moves from individual to group values?

- Collective priority-setting requires a collective view, so some method of aggregation has to be adopted
- Median: It gives less weight to extreme views than would the taking of a simple average
- Many clinicians believe that it is unethical for them to replace the values of each individual patient with some collective set of values
- Only in a purely private market (with no charity and no insurance) have doctors been in a position where they could do whatever the patient demanded. In all other circumstances doctors have been constrained by somebody else's willingness to pay

Williams A (1996) Soc Sci Med, 43(12): 1795-1804

Is the QALY approach unacceptable because it ignores the interpersonal distribution of health gains?

- The simplest and most common use of QALY calculations at present is based on the assumption that a year of healthy life expectancy is to be regarded by everybody as having equal value
- A strong egalitarian case could be made for that assumption, since it implies that it does not matter at all who the beneficiary is

Is the QALY approach unacceptable because it ignores the interpersonal distribution of health gains?

Preferred	Parameter	Non preferred
Young	Age	Elderly
Child	Age	Neonate
People with young children	Household	Childless
Non-smoker	Lifestyle	Smoker
Non-drinker	Lifestyle	Alcohol dependent
University professor	Profession	Physical worker
Employed	Employment status	Unemployed

Is the QALY approach unacceptable because it ignores the interpersonal distribution of health gains?

- A 0.2 improvement in utility: from 0.1 to 0.3 or from 0.7 to 0.9
- Some argue that our distributional concerns should not focus primarily on *health gains*, but on the *level of health* itself.
  - The aim would be: to minimize the difference between the *health capital* of subjects
- They suggest not devoting resources to improving the health of those who have already had a long and healthy life when those resources could be used to improve the health of someone who, otherwise, will have a shorter and/or more unhealthy life
- Discrimination of old and healthy people vs. young and with a poor state of health

QALY – METHODOLOGICAL CONCERNS

Methodological concerns

- Different valuation methods provide different utilities/values even for the same health state
- QALY represents comparable unit of heath outcomes only if the same valuation methods were applied (**-> transferability of QALY**)
- NICE gold standard: EQ-5D with TTO-based value set

	Needs walking stick when moving about	Needs supervision when walking	Needs two assistants for moving about
SG	0.85	0.64	0.38
TTO	0.78	0.41	0.11
RS	0.65	0.29	0.08

Nord E, et al.: (2009) Value health, 12(Suppl1): S10-5.

Difficulties of health state valuation

- Test-retest reliability (re-valuation of the same health state) (**<-**)
- Feasibility concerns
  - Under a certain age (e.g. 6 years)
  - In some conditions (blindness (RS), deafness (TTO, SG), mental/psychiatric disorders)

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Life saving vs. improving QoL

- Counting QALYs does not differentiate between life-saving and improving quality of life
- However, some argue that saving lives should be the priority

ED
(ICER: 3639
GBP/QALY)

Hip prosthesis

Dialysis

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

‚Double jeopardy'

- Patients with paraplegia have a lower number of QALYs to be saved compared to a patient of similar age and gender but in perfect health
- Double jeopardy: QALY maximization would give us no choice but to bring an even graver misfortune upon an individual who is a victim of disaster and becomes a paraphlegic
- The first disaster leaves her with poor quality of life while the second rules her out as a candidate for life-saving treatment
- Similarly, saving the lives of younger people with all other factors being equal, are always likely to produce more QALYs than saving older people.

Harris (1987), J Med Ethics, 13: 117-123.

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

CONCERNS WITH QALY-BASED DECISION-MAKING

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Preference of the individual

- Chemotherapies

Alternative	Life expectancy	Utility	QALYs accumulated
Palliative care	3 years	0.8	2.4 QALYs
Aggressive treatment	10 years	0.6	6.0 QALYs

- Let's suppose that the aggressive treatment is cost-effective compared to palliative care
- What if someone considers life as pointless if their QoL is less than 80%? Can he/she receive the less cost-effective reimbursed treatment for a higher QoL from public reimbursement?

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

Maximising QALY

- Maximization of the QALYs should not be the only ultimate aim of healthcare
- The society is willing to trade maximization of health gain for equity
- Other aspects should be considered (MCDA)
  - Equity: sildenafil in ED vs. orphan drugs

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

146

MCDA

#	Item	Max. score
1	Healthcare priority	20
2	Severity of disease	15
3	Equity	15
4	Cost-effectiveness and Quality of life	30
5	Budget Impact	10
6	Level and type of International and Hungarian Professional Evidence	10
	Total score	**100**

Harris J vs. Williams A

Harris	Williams
1. Healthcare priorities should not be influenced by any other consideration than keeping people alive;	1. Healthcare priorities should be influenced by our capacity both to increase life expectation and to improve people's quality of life;
2. Everyone has an equal right to be kept alive if that is what they wish, irrespective of how poor their prognosis is, and no matter what sacrifices others have to bear as a consequence;	2. A particular improvement in health should be regarded with equal value, no matter who gets it, and should be provided unless it prevents a greater improvement from being offered to someone else;
3. When allocating healthcare resources, we must not discriminate between people, not even according to their differential capacity to benefit from treatment.	3. It is the responsibility of everyone to discriminate wherever necessary to ensure that our limited resources go where they will do the most good.

Williams A (1987) J Med Ethics, 13: 117-123.

Cost generated by the QALY concept

- Cost of measuring health outcome
  - Developing, adapting, validating health state valuation methods, mapping (**<-**)
  - Valuing health states (**<-**)
  - Transferability (generating EQ-5D value sets) (**<-**)
- Cost of HTA
  - Cost of cost vector collection
  - Cost of conducting CUAs
  - Cost of evaluating CUAs
- Opportunity cost – Value of information
  - Value of more accurate information vs. cost generating more accurate information

Conclusion

- 'A typical stance is to point out all the difficulties involved with some particular approach, and then to sit on the fence waiting for the next candidate to come by, and then do the same
- This would be fine if the implied ideal method were available to us, or if we could suspend all healthcare decision-making until it were. But there is no perfect system being offered, and we cannot wait
- If the same criteria as are used to criticise the QALY approach were used in an even-handed way to criticise current practice, or any feasible alternative to it, how would these alternatives make out? It is irresponsible to do less.'

Williams A (1996) Soc Sci Med, 43(12): 1795-1804

Take home messages

- The use of QALY has been far from unproblematic
- QALY has been criticized on ethical, methodological and conceptual grounds
- The use of QALY also generates costs, which should be compared with the opportunity cost of inappropriate decisions made without using the QALY approach
- Until the advent of the perfect methodology is here to replace QALY, day- to-day decisions still have to be made in healthcare

Self-check questions

- What are the key conceptual, methodological and ethical critiques of using QALY and DALY to inform decision-makers in healthcare?
- What are the advantages and disadvantages of QALY and DALY?

Suggested reading

- Lyttkens CH. (2003) Time to disable DALYs? On the use of disability-adjusted life-years in health policy. European Journal of Health Economics, 4(3): 195-202.

- Wiliams, A. (1996) QALYS and ethics: a health economist's perspective. Soc sci med, 43(12): 1795-1804.

- Harris J. (1987) QALYfying the value of life. J Med Ethics, 13: 117-123

ELTE Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

148

8.7. Mapping

Utility mapping

"Financed from the financial support ELTE won from the Higher Education Restructuring Fund of the Hungarian Government"

Learning objectives

- Students should be
  - Familiar with the concept and steps of utility mapping
  - Aware of the policy relevance of utility mapping

Content

- Exercise 1 – utility calculation
- Exercise 2 – utility mapping

EXERCISE 1

Exercise 1 – Utility decrement of secondary disease

- Average utility of BRC patients: 0.7
- Prevalence of malnutrition among BRC patients: 20%
- Utility decrement of malnutrition among BRC patients: 0.1
- Excess mortality of malnourished patients RR=1.25
  (well nourished patients RR=1.00)
- What is the utility of malnourished BRC patients?

Solution of the exercise

	Prevalence	Utility
Well nourished	100-20%=80%	X
BRC patients		0.7
Malnourished	20%	X-0.1

provided inputs are marked red

Malnourished   Well nourished

$$(X-0.1)*0.2+X*0.8=0.7$$
$$0.72=X$$

Utility of malnourished BRC patients = X-0.1 = 0.72-0.1 = **0.62**

Exercise 2 – Utility mapping

Task:

- to provide an estimate on the QALY gain of tiotropium vs. placebo based on the study of Tonnel et al. (International Journal of COPD 2008:3(2),301–31.)

EXERCISE 2

Background of the applied measures

- St. George's Respiratory Questionnaire (SGRQ): disease-specific index-type PRO measure
- 50 items relevant for respiratory diseases, summarized in 3 categories + one composite index score (SGRQ total score)
- SGRQ total score: 0-100, where:
  - 0 = best possible health status,
  - 100 = worst possible health status
- Let's compare it to utility measure (<-):
  - Calibration of the scale x
  - Direction of the scale x
  - Endpoints of the scale x

Solution: mapping of SGRQ

- SGRQ is not a utility measurement tool
- Solution: Let's map the SGRQ to the utility scale
- Utility mapping in a cross-sectional study:
  - One instrument validated to measure utility (e.g. EQ-5D)
  - One disease-specific measure (e.g. SGRQ)
  - Demographic data
- Aim: To map SGRQ to the utility scale by using a regression model

Mapping - theory

SGRQ total score

100

EQ-5D index

-0.594   0   0.5   1.0
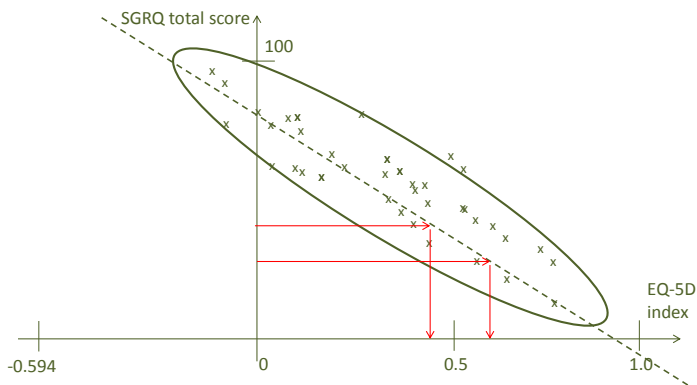
Inputs – baseline data

Table 1 Patient demographics and characteristics at pretreatment baseline

	Tiotropium	Placebo
Patients, n	266	288
Sex, M/F	231/35	246/42
Age, years	64.9 ± 9.7	63.5 ± 10.1
Body mass index, kg/m²	26.0 ± 4.8	25.8 ± 4.7
Duration of COPD, years	7.9 ± 7.6	8.0 ± 7.9
Smoking history, pack-years	44.4 ± 21.3	43.0 ± 22.5
Current smokers, n (%)	63 (23.7)	87 (30.2)
HRQoL[a]		
SGRQ total score	45.8 ± 17.7	48.9 ± 18.4
VSRQ total score[b]	47.3 ± 16.0	43.3 ± 16.9
Pulmonary function		
FEV₁, L	1.38 ± 0.44	1.35 ± 0.46
FEV₁, % predicted	47.49 ± 13.27	46.19 ± 12.40
FVC, L	2.50 ± 0.68	2.49 ± 0.75
FEV₁/FVC, %	55.30 ± 11.32	54.62 ± 11.27
IC, L	2.14 ± 0.69	2.09 ± 0.69
SVC, L	2.78 ± 0.74	2.70 ± 0.78
FIV₁, L[c]	2.02 ± 0.64	2.04 ± 0.71
Reversibility, n (%)[d]	108 (40.6)	116 (40.3)

Data are presented as mean ± SD unless otherwise stated.
[a]Lower SGRQ but higher VSRQ scores indicate an improvement.
[b]VSRQ total score was complete in 415 patients (tiotropium: n = 233; placebo: n = 237).
[c]FIV₁ was assessed only at selected sites (80.5% of patients in the full analysis population, tiotropium: n = 233; placebo: n = 237).
[d]Reversibility to a short-acting β₂-agonist.

Tonnel AB et al. (2008) International Journal of COPD, 3(2): 301–310.

Inputs - effectiveness

SGRQ total score decrease from baseline	Tiotropium	Placebo
0 month	0	0
3 months	- 6.90	- 3.55
6 months	- 7.70	- 4.32
9 months	- 8.50	- 4.32

a Tiotropium (n = 247)   Placebo (n = 245)
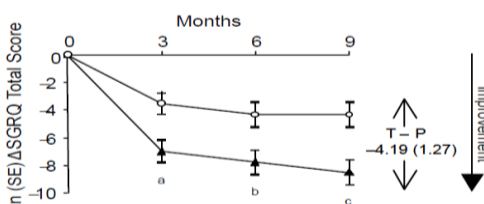
[a]p = 0.003, [b]p = 0.007, [c]p = 0.001 vs placebo

Tonnel AB et al. (2008) International Journal of COPD, 3(2): 301–310.

Utility mapping algorithm

EQ-5D = 0.9617 - 0.0013 X SGRQtotal - 0.0001 X SGRQtotal² + 0.0231 X %male

Starkie HJ et al. (2011) Value in Health, 14: 354–360.

Step 1 Transforming SGRQ total scores to utilities

Equation				
Const	Comp	Comp	Comp	Utility
0,9617	0,13	1	0	-0,167
		Inputs:		
SGRQ total sco	100			
Male (%)	5%			

Ref: Starkie et al, ViH, 2011, 14, 354-360

Step 1 Transforming SGRQ total scores to utilities

Time	SGRQ total score decrement		SGRQ total score (absolute values)		Utilities	
	Tiotropium	Placebo	Tiotropium	Placebo	Tiotropium	Placebo
0 month	0	0	45.8	48.9	0.712	0.679
3 months	- 6.90	- 3.55	38.9	45.35	0.780	0.717
6 months	- 7.70	- 4.32	38.1	44.58	0.787	0.725
9 months	- 8.50	- 4.32	37.3	44.58	0.794	0.725

Step 2 Use the QALY approach to generate QALYs from utilities
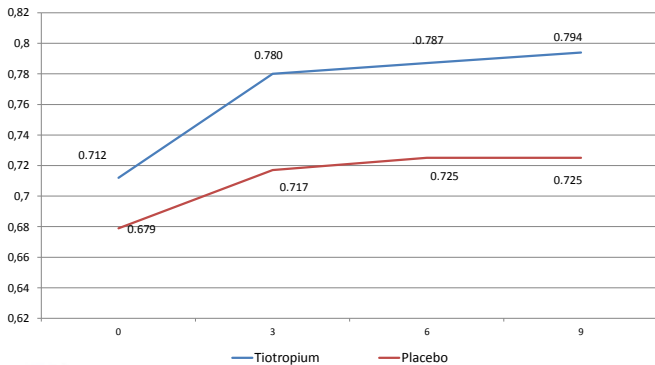
Step 2 Use the QALY approach to generate QALYs from utilities

Tiotropium utility	Placebo utility	Row difference	Tiotropium utility gain (amended with baseline difference)
0.712	0.679	0.033	0
0.780	0.717	0.063	0.0293
0.787	0.725	0.062	0.0286
0.794	0.725	0.069	0.0357

Step 2 Use the QALY approach to generate QALYs from utilities
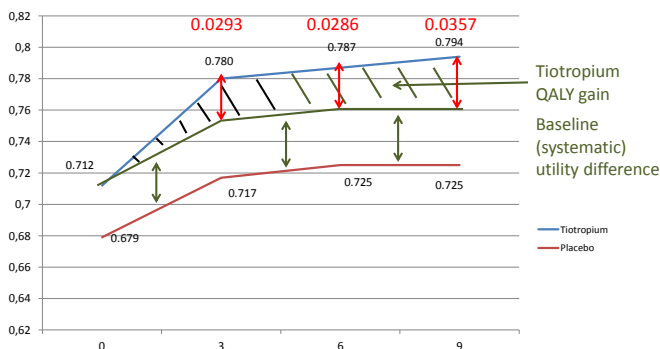
Step 3 Calculation of QALY gain

- Area under the curve
- Area of trapezoid: Area=(a+c)/2*h
- QALY gain=0.25*(0+0.0293)/2+0.25*(0.0293+0.0286)/2+0.25*(0.0286+0.0357)/2
- QALY gain=0.01895 /QALY/

Relevance of utility mapping

- Generic QoL measures are not sensitive enough to detect small differences
- Specific measures are very often not suitable to measure utility, consequently their capacity to be used in healthcare decision-making is limited
- There is a risk that a generic measure will not detect statistical significance, and consequently improvement in QoL may not be accepted by the payer

Take home messages

- Specific measures are very often not suitable for measuring utility, consequently their capacity to be used in healthcare decision-making is limited
- Mapping enables specific measures to estimate utility
- To map a specific measure, there is a need for individual patient level data from the same patient population of both the specific and the utility measure

Self-check questions

- What are the key steps of mapping?
- What is the policy relevance of mapping?

Suggested reading

- Tonnel AB, Perez T, Grosbois JM, Verkindre C, Bravo ML, Brun M. (2008) Effect of tiotropium on health-related quality of life as a primary efficacy endpoint in COPD. International Journal of COPD, 3: 301–310.
- Starkie HJ, Briggs AH, Chambers MG, Jones P. (2011) Predicting EQ-5D Values Using the SGRQ. Value in Health, 14: 354–360.

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

ELTE
Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest,
Pázmány Péter sétány 1/A

8.8. PROs in non-interventional clinical trials

PROs in non-interventional clinical trials

"Financed from the financial support ELTE won from the Higher Education Restructuring Fund of the Hungarian Government"

Learning objectives

- Students should:
  - Be aware of the relevance of non-interventional trials
  - Be familiar with key principles of regulating non-interventional trials
  - Be able to design and conduct non-interventional trials to collect PROs
  - Be familiar with key components of non-interventional trial surveys/questionnaires

Contents

- Regulation of non-interventional trials
- Non-interventional trial: Case study

REGULATION OF NON-INTERVENTIONAL TRIALS

Background

- Real word evidence is becoming central to decision-making when used in conjunction with clinical trials
- Non-interventional studies (NIS) are a key method in gathering this type of evidence
- There is a need for hypothesis generating

Non-interventional studies - Definition by the EC

A study where the medicinal products are prescribed in the **usual manner in accordance with** the terms of the **marketing authorization**.

The **assignment of the patient** to a particular therapeutic strategy is not decided in advance by a trial protocol, but <u>**falls within the current treatment practice**</u>, and the **prescription of the medicine is clearly separated from the decision to include** the patient in the study.

No additional diagnostic or monitoring procedures shall be applied to the patients and epidemiological methods shall be used for the analysis of the collected data.'

2001/20/EC

Benefits and limitations of NIS

Benefits	Limitations
• Ability to evaluate large number of patients/consumers • Lower cost • Ability to establish and increase dominant status in market in a therapeutic area • Ability to heighten disease awareness • Support of research and scientific inquiry	• Methodological challenges in attributing causality to outcomes selection bias • Challenges in conducting global NIS due to different stages of the life cycle in different countries • Limited ability to follow up

Checklist for NIS

- Is this a study of one or more medicinal products, which has/have a marketing authorization in the EU Member State concerned?
- Are the **products prescribed in the usual** manner in accordance with the terms of the **authorization**?
- **Does the assignment of any patient** involved in the study to a particular therapeutic strategy **fall within current clinical practice** and is not decided in advance by a clinical trial protocol?
- Is the decision to **prescribe a particular medicinal product clearly separated from the decision to include** the patient in the study?
- Will **no diagnostic or monitoring procedures** be applied to the patients included in the study, **other than** those which are applied in the course of **current clinical practice**? (The issue of diagnostic and monitoring procedures is often a topic of contention and really needs to be discussed with an ethics committee. However, it should be noted that Volume 9A provides clarification that **interviews, questionnaires and blood sampling may be considered as normal clinical practice**)
- Will **epidemiological methods be used for the analysis** of data arising from the study?

NIS studies must comply with the following criteria:

- There must be a written protocol and written contracts between health professionals and institutes at which the study will take place (that specify the nature of the services and the payment)
- Any remuneration must be reasonable and reflect the fair market of the work
- In countries where ethics committees are prepared to review such studies, the study protocol must be submitted to the committee for review
- Data protection legislation must be complied with

NIS studies must comply with the following criteria:

- The study must not constitute an inducement to prescribe, supply, administer or sell medicine
- The company's scientific service must approve the protocol and must supervise the conduct of the study
- Study results must be analyzed and summaries must be made available; reports should be sent to health professionals who participated in the study
- If study results are important to the assessment of benefit/risk, a summary should be immediately forwarded to the relevant authority
- Sales reps may only be involved in an administrative capacity, and such service must be under the supervision of the company's scientific service

EFPIA Code of practice for the Pharmaceutical Industry (2015)

Good Clinical Practice

- The **principles of GCP should generally apply to all clinical research** involving human subjects, and **not just research involving pharmaceutical** or other medical products
- Although some principles of GCP may not apply to all types of research on human subjects, **consideration of these principles is strongly encouraged** wherever applicable as a means of ensuring the ethical, methodologically sound and accurate conduct of human subjects research

ICH-GCP

Declaration of Helsinki

- (8.) While the primary purpose of medical research is to generate new knowledge, this goal can never take precedence over the **rights and interests of individual research subjects**.
- (19.) Some groups and individuals are particularly **vulnerable** and may have an increased likelihood of being wronged or of incurring additional harm. All vulnerable groups and individuals should receive specifically considered protection
- (22.) The design and performance of each research study involving human subjects must be clearly described and justified in a **research protocol**.
- (23.) The research protocol must be submitted for consideration, comment, guidance and approval to the concerned **research ethics committee** before the study begins.

Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects (2013)

Declaration of Helsinki

- (26.) In medical research involving human subjects capable of giving informed consent, each potential subject must be **adequately informed** of the
  - aims, methods, sources of funding,
  - any possible conflicts of interest,
  - institutional affiliations of the researcher,
  - the anticipated benefits and potential risks of the study and the discomfort it may entail,
  - post-study provisions and any other relevant aspects of the study.

Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects (2013)
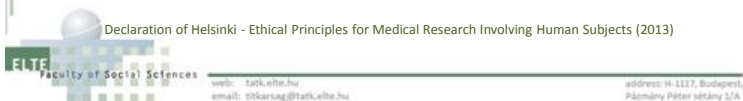
Declaration of Helsinki

- (26.) The potential subject must be informed of the **right to refuse** to participate in the study or to withdraw consent to participate at any time without reprisal. After ensuring that the potential subject has understood the information, the physician or other appropriately qualified individual must then seek the potential subject's **freely-given informed consent**, preferably in writing.
- (30.) Research involving subjects **who are physically or mentally incapable of giving consent**, for example, unconscious patients, may be done only if the physical or mental condition that prevents giving informed consent is a necessary characteristic of the research group.

Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects (2013)

Observational studies - legislation

- Eudralex Volume 9A states, "for non-interventional post-authorisation safety studies, the marketing authorisation holder and investigator should follow relevant national legislation in those member states where they exist."
- A patient's right to confidentiality is crucial

Informed consent of the patient

Informed consent is a decision to participate in the research, taken by a competent individual, who has **received the necessary information**, who has adequately **understood the information**, and who, after considering the information, has arrived at a decision without having been subjected to coercion, undue influence or inducement or intimidation.

Further definitions

- Freely given consent
  - Competent individuals are entitled to choose freely whether to participate in the research
- Voluntary participation in the study
  - Subjects who are junior or subordinate members of a hierarchical group require careful consideration, as their agreement to volunteer may be unduly influenced…by the expectation of preferential treatment if they agree or by fear of disapproval or retaliation if they refuse

CIOMS Ethical guidelines commentary

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Examples for non-interventional studies

- Purely observational **database review** or research
- **Retrospective review of patients** where all events of interest have already happened (case-control, cross-sectional and retrospective cohort studies)
- **Registries** in which the data collected is derived from routine care
- Studies which evaluate **patterns of the usage of medicines** (drug utilization and occurrence of health outcomes)

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Assembling questionnaire for observational studies

- Demographic questionnaire
- Classification of patients (according to the model structure)
- Generic QoL measures (EQ-5D) and direct health state valuation/utility measures (TTO, SG)
- Specific measures
- Resource utilization
- AEs

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

CASE STUDY

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Case study: Generating utilities from observational studies

- Malignant disease with rapid progression
- Treatment guidelines identify the following health states
  - Stable disease
  - Local metastasis
  - Distant metastasis
  - Dead
- New drug: slows down disease progression
  - Reduce development of metastasis
  - Reduce mortality
- Aim: to adapt a central model with local cost vectors and utilities

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

Markov model

Stable disease
Utility: ?

Local metastasis
Utility: ?
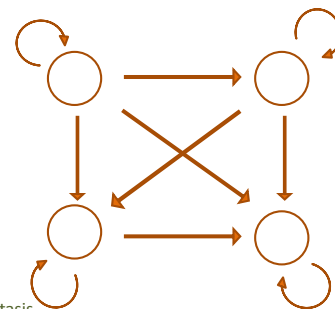
Distance metastasis
Utility: ?

Dead
Utility: 0

ELTE Faculty of Social Sciences
web: tatk.elte.hu
email: titkarsag@tatk.elte.hu
address: H-1117, Budapest, Pázmány Péter sétány 1/A

157

Design and setting of the clinical trial

- Prospective cross-sectional non-interventional study
- Consecutive patients (random selection)
- Inclusion criteria
  - Diagnosed disease
  - Aged 18+
  - Informed written consent for participation

Questionnaire

- Patient information form
- Consent form
- Demographic questionnaire (age, gender, disease duration)
- Disease state
  - Existence of local metastasis
  - Existence of distant metastasis
- Questionnaire on resource utilization
- Disease-specific measure (mapping) (<-)
- Utility measure (risk of statistically non-significant difference) (<-)

Results

	Number of patients	Utility (mean)	SD
Stable disease	58	0.75	0.30
Local metastasis	67	0.50	0.28
Distant metastasis	52	0.25	0.33

The Markov model populated with utilities

Stable disease
Utility: 0.75

Local metastasis
Utility: 0.50

Distant metastasis
Utility: 0.25

Dead
Utility: 0

The Markov-Model

Stable disease
Utility: 0.75

Local metastasis
Utility: 0.50

Distant metastasis
Utility: 0.25

Dead
Utility: 0

Cycle 1.
Cycle 2.
Cycle 3.
Cycle 4.
Cycle 5.
Cycle 6.
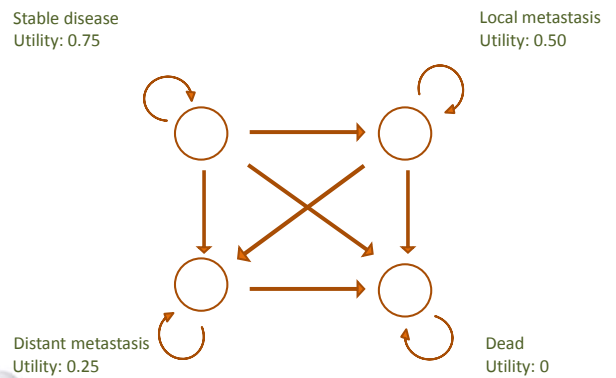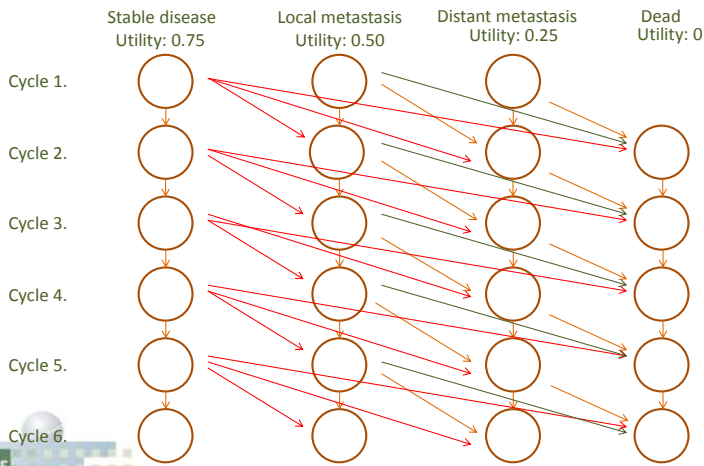
Calculation of QALYs - example

Stable disease
Utility: 0.75

Local metastasis
Utility: 0.50

Distant metastasis
Utility: 0.25

Dead
Utility: 0

Month 1-3
Month 4-6
Month 7-9
Month 10-12
Month 13-15
Month 16-18

158
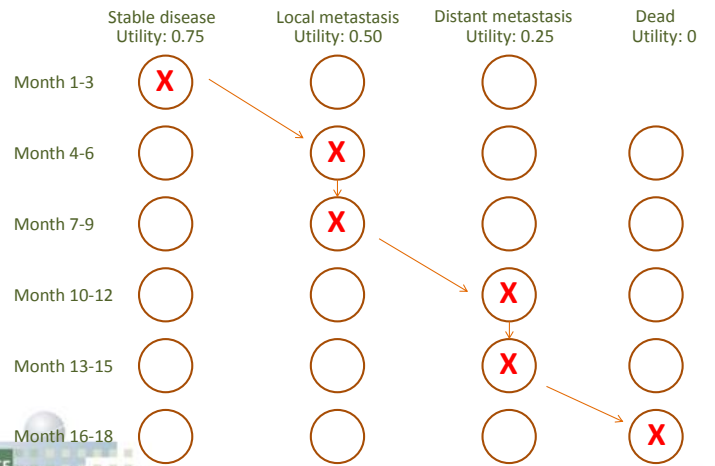
Calculation of QALYs – area under the curve

Total QALY = 0.75/4+0.5/4+0.5/4+0.25/4+0.25/4+0 = 0.5625 QALY

Take home messages

- PROs may be collected in non-interventional clinical trials beside RCTs
- The principles of GCP should generally apply to all clinical research involving human subjects, and not just research involving pharmaceutical or other medical products
- For non-interventional trials, the research protocol must be submitted for consideration, comment, guidance and approval to an independent research ethics committee before the study begins
- The potential subject must be informed of the right to refuse to participate in the study
- After ensuring that the potential subject has understood the information, the physician must then seek the potential subject's freely-given informed consent

Take home messages (2)

- Core elements of the questionnaire
  - Patient information form
  - Consent form
  - Demographic questionnaire (age, gender, disease duration)
  - Disease/medical state of patient
  - Generic and specific measures (utility measures)

Self-check questions

- What are the principles of the regulation of non-interventional trials?
- What are the principles of the Declaration of Helsinki?
- How would you select measures to be included in a non-interventional trial questionnaire?

Suggested reading

- Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects (2013)

9. Appendix

9.1. Valuation of EQ-5D 3L health states

Valuation of EQ-5D
Health States with RS

Based on the
Measurement and Valuation of Health
(MVH) protocol

"Financed from the financial support ELTE won from the
Higher Education Restructuring Fund of the Hungarian
Government"

ELTE
Faculty of Social Sciences

web: tatk.elte.hu
email: titkarsag@tatk.elte.hu

address: H-1117, Budapest,
Pázmány Péter sétány 1/A

‾‾‾‾‾‾‾‾‾

COMPARING STATES OF HEALTH

Different people have different states of health. Some people's health is better than others.

We are interested in comparing different states of health and measuring how good or bad they are.

We would like to start by asking you about your own health.

Then we will ask you to think about some different states of health and tell us how good or bad you think they are.

There are no right or wrong answers. We just want to know what you think.

First, we would like you to tell us about the state of your own health today. Which of the following statements best describe the state of your health today?

Please tick ONE box for each group of statements ☒

Mobility
- I have **no** problems in walking about ☐
- I have **some** problems in walking about ☐
- I am **confined to bed** ☐

Self-care
- I have **no** problems with self-care ☐
- I have **some** problems washing or dressing myself ☐
- I am **unable** to wash or dress myself ☐

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

Pain/Discomfort
- I have **no** pain or discomfort ☐
- I have **moderate** pain or discomfort ☐
- I have **extreme** pain or discomfort ☐

Anxiety/Depression
- I am **not** anxious or depressed ☐
- I am **moderately** anxious or depressed ☐
- I am **extremely** anxious or depressed ☐

To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked 100 and the worst state you can imagine is marked 0.

We would like you to indicate on this scale how good or bad your own health is today, in your opinion. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state is today.

Best imaginable
health state

100

90

80

70

60

50

40

30

20

10

0

Your own health
state today

Worst imaginable
health state
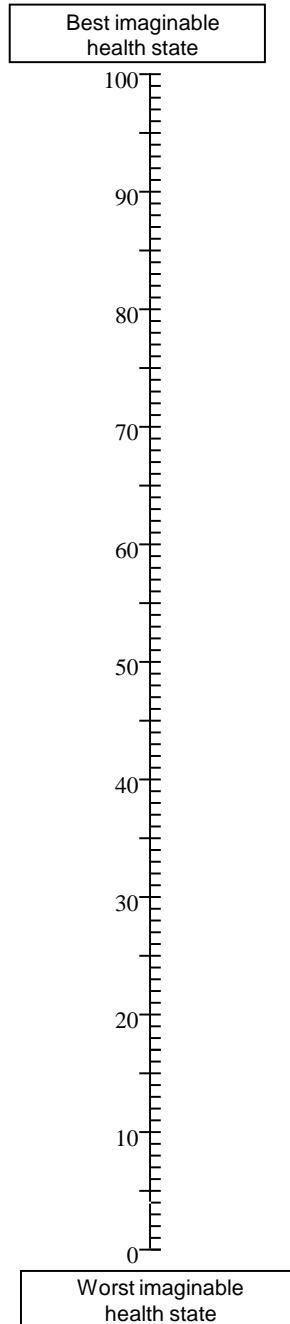
- We would now like you to think about some states of health you might not have experienced yourself.

- The following pages contain descriptions of different states of health. Each one is grouped in a box, labelled with a code letter.

- For each box, imagine that you have to live in that state of health for **one year.**

- How good or bad do you think each state of health is compared to the others?

- Please give each state of health a **score** between 0 and 100, where 0 = the worst health state you can imagine and 100 = the best state of health you can imagine.

- Write the **score** in the bottom right hand corner of each box.

Best imaginable
health state

100

A

- No problems in walking about,
- No problems with self care,
- Some problems perfoming usual activities (e.g. work, study, housework, family or leisure activities),
- No pain or discomfort, and
- Not anxious or depressed

SCORE= ☐

90

E

- No problems in walking about,
- No problems with self care,
- No problems perfoming usual activities (e.g. work, study, housework, family or leisure activities),
- Moderate discomfort, and
- Not anxious or depressed

SCORE= ☐

80

70

B

- No problems in walking about,
- No problems with self care,
- No problems perfoming usual activities (e.g. work, study, housework, family or leisure activities),
- No pain or discomfort, and
- Not anxious or depressed

SCORE= ☐

F

- Some problems in walking about,
- Some problems with self care,
- Some problems perfoming usual activities (e.g. work, study, housework, family or leisure activities),
- Extreme pain or discomfort, and
- Extremely anxious or depressed

SCORE= ☐

60

50

C

- Some problems with walking about,
- No problems with self care,
- Some problems perfoming usual activities (e.g. work, study, housework, family or leisure activities),
- Extreme pain or discomfort, and
- Moderately anxious or depressed

SCORE= ☐

40

G

- Confined to bed,
- Unable to wash or dress self
- Unable to perform usual activities (e.g. work, study, housework, family or leisure activities),
- Extreme pain or discomfort, and
- Extremely anxious or depressed

SCORE= ☐

30

20

D

- No problems with walking about,
- No problems with self care,
- No problems with perfoming usual activities (e.g. work, study, housework, family or leisure activities),
- Moderate pain or discomfort, and
- Moderatley anxious or depressed

SCORE= ☐

H

- Confined to bed,
- Unable to wash or dress self
- Unable to perform usual activities (e.g. work, study, housework, family or leisure activities),
- Moderate pain or discomfort, and
- Not anxious or depressed

SCORE= ☐

10

0

Worst imaginable
health state

• In your opinion, would living in any of the eight health states on the facing page for one year be **worse** than being **dead**?

Yes ☐   No ☐

• Please write the code letter of those health states in the space below. The code letter is in the bottom left-hand corner of each box

• On a scale from 0 to 100, what **score** would you give to the state of being **dead**? (where 0 = the worst health state you can imagine and 100 = the best health state you can imagine).

Score = ☐

Best imaginable
health state

100

—

—

90

—

—

80

—

70

—

—

60

—

—

50

—

—

40

—

—

—

30

—

—

20

—

—

10

—

—

0

Worst imaginable
health state

- Some problems in walking about,
- No problems with self care,
- No problems perfoming usual activities
  (e.g. work, study, housework, family or
  leisure activities),
- No pain or discomfort, and
- Not anxious or depressed

I   SCORE= ☐

- No problems in walking about,
- No problems with self care,
- No problems perfoming usual activities
  (e.g. work, study, housework, family or
  leisure activities),
- No pain or discomfort, and
- Not anxious or depressed

B   SCORE= ☐

- Unconscious

J   SCORE= ☐

- No problems in walking about,
- Some problems with self care,
- No problems perfoming usual activities
  (e.g. work, study, housework, family or
  leisure activities),
- No pain or discomfort, and
- Not anxious or depressed

K   SCORE= ☐

- No problems in walking about,
- No problems with self care,
- No problems perfoming usual activities
  (e.g. work, study, housework, family or
  leisure activities),
- No discomfort, and
- Moderately anxious or depressed

L   SCORE= ☐

- Confined to bed
- Some problems with self care,
- Some problems perfoming usual activities
  (e.g. work, study, housework, family or
  leisure activities),
- No pain or discomfort, and
- Not anxious or depressed

M   SCORE= ☐

- Confined to bed,
- Unable to wash or dress self
- Unable to perform usual activities (e.g.
  work, study, housework, family or leisure
  activities),
- Extreme pain or discomfort, and
- Extremely anxious or depressed

G   SCORE= ☐

- Some problems in walking about,
- Some to wash or dress self
- Unable to perform usual activities (e.g.
  work, study, housework, family or leisure
  activities),
- Moderate pain or discomfort, and
- Extremely anxious or depressed

N   SCORE= ☐

- In your opinion, would living in any of the eight health states on the facing page for one year be **worse** than being **dead**?
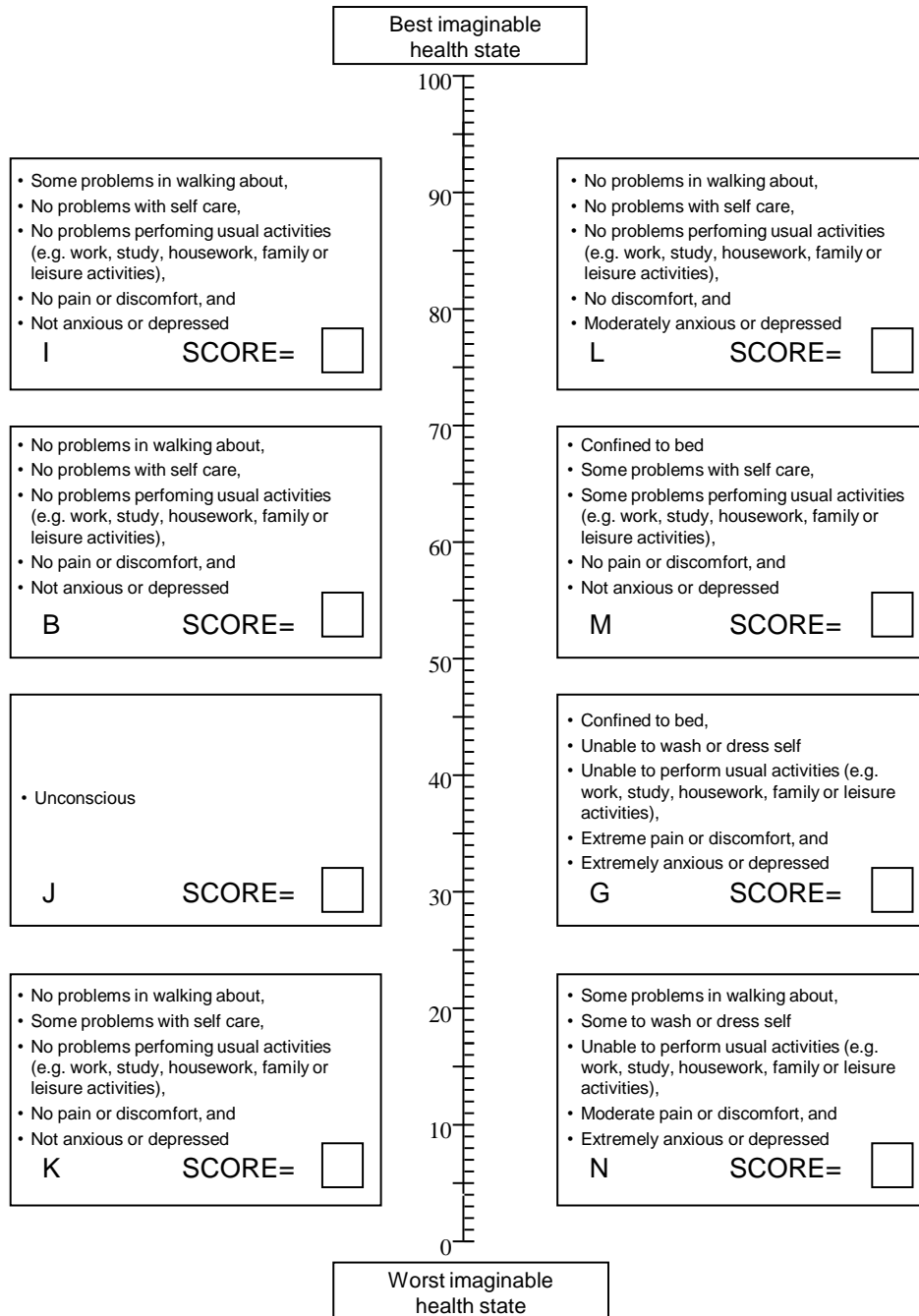
  Yes ☐   No ☐

- Please write the code letter of those health states in the space below. The code letter is in the bottom left-hand corner of each box

  ☐

- On a scale from 0 to 100, what **score** would you give to the state of being **dead**? (where 0 = the worst health state you can imagine and 100 = the best health state you can imagine).

  Score = ☐

List of abbreviations

AE	adverse event
BoD	burden of disease
BP	blood pressure
BRC	breast cancer
ClinRO	clinician reported outcome
CBA	cost-benefit analysis
CEA	cost-effectiveness analysis
CMA	cost-minimization analysis
COA	clinical outcome assessment
COPD	chronic obstructive pulmonary disease
CUA	cost utility analysis
CV	cardiovascular
DALY	disability adjusted life years
DES	discrete event simulation
EMA	European Medicines Agency
EQ-5D 3L/5L/Y	EuroQoL 5 dimensions 3 levels/5 levels/youth
EU	European Union
FDA	Food and Drug Administration
GCP	good clinical practice
GP	general practitioner
HRQoL	health-related quality of life
HTA	health technology assessment
HUI	Health Utility Index
ICER	incremental cost-effectiveness ratio
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ISOQOL	International Society for Quality of Life Research
LY	life year
MCDA	multiple-criteria decision analysis
MMAS-4/8	Morisky Medication Adherence Scale 4-item/8-item
MPR	medication possession ratio
NICE	National Institute for Health and Care Excellence
NM	Neumann Morgenstern
No	number
PDC	proportion of days covered
PRO	patient reported outcome
QALY	quality adjusted life year
QoL	quality of life

RA	rheumatoid arthritis
RCT	randomized controlled trial
RR	relative risk
SD	standard deviation
SGRQ	St. George's Respiratory Questionnaire
SF-6D/36	Short Form 6D/36
SG	standard gamble
PROMIS	Patient Reported Outcomes Measurement Information System
PTO	person trade-off
RS	rating scale
TTO	time trade-off
UK	United Kingdom
US	United States
VAS	visual analogue scale
WHO	World Health Organization
WMA	World Medical Association
YLD	years lived with disability
YLL	years of life loss

Tamás Ágh PhD graduated from the Semmelweis Medical University (MD, 2006), the Corvinus University of Budapest (MSc in Economics, 2010) and the Semmelweis University School of PhD Studies (PhD in Pharmacoeconomics, 2013). Dr. Ágh is a principal researcher at Syreon Research Institute. With more than 8 years of experience in health economics research, he developed expertise in the fields of patient reported outcomes, health technology assessment, pharmaceutical policy and systematic literature reviews. Dr. Ágh practices as a general practitioner since 2009. He is a part time university lecturer (2011-) and module leader (2015-) at Eötvös Loránd University (MSc in Health Policy, Finance and Analysis) in the field of Patient Reported Outcomes. Dr. Ágh is member of the leadership group of the ISPOR Multiple Medication Adherence Measurement Working Group (2015-). He is author and co-author of 1 e-book, 2 book chapters and 16 peer-reviewed scientific publications in the area of health economics.

András Inotai PhD is a graduate of Semmelweis Medical University in Budapest (PharmD, 2007), the Eötvös Loránd Science University in Budapest (Diploma in pharmaceutical policy and pharmacoeconomics, 2008) and the Semmelweis University School of PhD Studies in Budapest (PhD in Pharmacoeconomics, 2011). Dr. Inotai is the Head of Pharmaceutical Policy Department at Syreon Research Institute. He has 8 years of research experience in pharmaceutical policy, pharmacoeconomics, health technology assessment, patient reported outcomes, drug utilization and burden of disease studies. He is a part time university lecturer and module leader at Eötvös Loránd University (MSc in Health Policy, Finance and Analysis) in the field of Patient Reported Outcomes since 2011. Dr. Inotai was the president of ISPOR Hungary Chapter between 2014-15. He is a co-editor of the ISPOR CEE Network Newsletter. He is editor of 3 books, author and co-author of 1 e-book, 3 book chapters and 29 peer-reviewed scientific publications in the area of health economics.