REVIEW ARTICLE

The assessment of patient clinical outcome: A literature discussion

Mou'ath A. Hourani, Nidal M. Turab, Qusai Y. Shambour

Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan

Abstract

A patient's safety in clinical field is critical, important and complex. The patients are still suffering from preventable harms from diagnostic errors, procedure mistakes, teamwork failures, and the failure to deliver recommended therapies. Patient outcome is the status upon a patient's adherence to treatment. An assessment of patient's clinical outcome is one of the important aspects of patient safety, and requires the assessment of the benefits, harms and risks of therapeutic options and comparing between them. Very few methods are developed for the clinical field and there is still a need for more accurate methods for such assessment. To achieve the above objective, we have performed an integrative review of the literature using different online databases and search engines including PubMed, Scopus, Google, and Google Scholar to explore current issues regarding the assessment of patient clinical outcome. This paper presents:

1) an overview of the existing assessment methods for patient clinical outcome and their conceptual limitations; and
2) a discussion of the primitiveness of the current assessment methods.

Based on the literature research in this paper, researchers, clinicians and health care professionals working in the field of assessment of patient clinical outcome, will be able to

1) understand all the critical issues in this area, and
2) design and develop novel general methods for the assessment of patient clinical outcome that avoid the conceptual limitations of existing methods.

Key words: Assessment, clinical outcome, methods, patient

Introduction

Clinical Outcome Assessments (COAs) can be used to determine whether or not a drug has been demonstrated to provide desired treatment benefit. COA measures include: Patient-reported outcome (PRO) measures; Clinician-reported outcome (ClinRO) measures; and Observer-reported outcome (ObsRO) measures.

Patient-reported outcome (PRO) is a questioner or method for collecting responses from the patient directly or by interviewers.[1] A well-designed PRO questionnaire should consider single and multiple characteristics. These characteristics are known as constructs while the questionnaire that used to collect these characteristics is known as instrument or tool. Typically, PRO tools must be validated and tested extensively.[2,3] If questionnaires were designed to collect the characteristics of any disease population and cover a wide range of aspects it is known as generic;

Access this article online	
Quick Response Code:	Website: www.atmph.org
	DOI: ***

Cite this article as: We will update details while making issue online***

Correspondence:

Dr. Qusai Shambour, Al-Ahliyya Amman University, Amman, Jordan. E-mail: Q.Shambour@gmail.com

while it is known as condition-targeted if it designed to measure the characteristics of people with particular medical situations.[4] Unidimensional questionnaire measures a single characteristic; it has a scale of single score. A multi-dimensional questionnaire measures multiple characteristics and provides multiple scales; there is a separate report of each scale.[5]

Recently, there have been proposed quick, effective and understandable tools to observe patient clinical outcomes on a regular basis.[6] These tools allow patients to record clinical outcomes and experience in a semi-structured way and accordingly synoptic input data, while their physio-emotional sensitivity tracked automatically.[1] Modern advancements in psychometrics such as Item Response Theory (IRT)[7] and Computerized Adaptive Testing (CAT) are used to create reliable and validated measurement tools as part of the National Institute of Health's Roadmap Initiative.[8]

The need for an assessment of patient clinical outcome is compulsory, clinical outcome assessment methods in some clinical fields have been employed without sufficient understanding of their characteristics.[9] This insufficient understanding of the outcome assessment methods characteristics trigger the need for more research to better understand the existing patient clinical outcome assessment methods and their conceptual limitations. Consequently, researchers, clinicians and health care professionals will be able to design and develop novel general methods for the assessment of patient clinical outcome. Such methods will consider the conceptual merits of existing methods and avoid their conceptual limitations.

Accordingly, this paper is proposing a literature on the assessment of patient clinical outcome that will help researchers, clinicians and health care professionals to design novel general methods for the assessment of patient clinical outcome that avoid the conceptual limitations of existing methods. The literature is organized as follows: **[Section 2]** describe the literature search methodology used in this paper. **[Section 3]** presents an overview of the existing methods for patient outcome assessment and their conceptual limitations. A discussion of the primitiveness of the current assessment methods is shown at **[Section 4]**. Finally conclusion is given at **[Section 5]**.

Method of Literature Search

The information reported in this review was obtained from different online databases and search engines including PubMed, Scopus, Google, and Google Scholar. Keywords and expressions used for the search included Patient, Clinical Outcome, Healthcare, Assessment Method, Assessment of Clinical Outcome, Benefit-risk analysis, Quality-of-Life, Treatments, Cost-Effectiveness Analysis, and similar pre-identified terms were used in separate searches and in conjunction with each other to identify all related publications. Related publications published in English including Journal papers, books and reports were scrutinized to identify those that met a criteria of presenting information accredited to the purpose of this review. After elimination of studies that were not relevant to the subject matter, a total of 180 articles were reviewed and cited in this study.

Overview of the existing methods for patient clinical outcome assessment and their conceptual limitations

There are a number of methods to assess the benefits, harms, risk and patient outcome from different views; a summary of those methods are:

Categorizing of the Severity of Adverse Event and Disease States

In this approach, the severity of disease state is classified into specific degrees. One of the tries in this approach is the categorizing of severity of disease state into seven degrees ranging from mild disease or condition with symptoms, which are not progressive and which only cause a mild degree of discomfort or incapacity to life-threatening condition. Severity degrees of adverse events and diseases were assigned by group of physicians, converted to numerical scale of severity, and combined with data on frequency of benefits and adverse events to make quantitative assessment of drug benefits and harms.[10] Another try was repeated for the assessment of severity categories of adverse events by collecting family doctors opinions using five point category scales instead of seven.[11]

Recently, again, a six-category scale of severity for adverse events was used. Every degree of severity is assigned a score ranged from zero to one. The severity scores were then combined with the scores derived from an ADR causality algorithm by taking the average of both scores.[12]

Conceptual Limitations of Categorizing of the Severity of Adverse Event and Disease States

Adverse event severity grading do not have international acceptance.[13] It is intuitive, subjective assessment of safety[10,11] and based upon personal opinions.[14] Scoring available evidence is not definitive[15] and

grades definitions are not satisfactory.[16] Grades are set according to general rather than solid estimation without paying attention to different characteristics of the adverse effect.[17] They are biased and don't provide robust, consistent and valid clinical decision making.[17] Moreover, those methods don't define elements of benefit and risks and don't allow for tradeoffs between multiple elements.[17]

Evidence-based benefit and risk model

In Evidence-based Benefit and Risk Model,[18] the benefit is estimated by efficacy, responder rate, and data evidence; while risk is estimated by adverse drug reaction seriousness, adverse drug reaction frequency, and data evidence.

Conceptual limitations of evidence-based benefit and risk model

Evidence-based Benefit and Risk Model lacks a conceptual framework for trading-off the benefits and harms.[19] The criteria are not comprehensive and not well defined, and cannot be expressed in equivalent units.[19] In addition, patient health preferences are not considered.

Principle of Threes

This method is calculating the benefit score as the product of disease cure rate times disease seriousness times disease duration, and adverse event score as the product of adverse event incidence times adverse event seriousness times adverse event duration; each parameter in benefit and risk is rated as (low=1, medium=2, high=3).[20]

Transparent uniform risk benefit overview (TURBO)

The TURBO model is a quantitative and graphical method for benefit risk analysis; the risk factor R is calculated as the sum of two risks: the risk associated with the most serious adverse effect (severity score from 1 to 5), and the risk associated with the second most serious adverse effect or the most frequent adverse effect (severity score from 1 to 2). The Benefit factor B is calculated as the sum of the primary benefit, which is the change(s) in health status and social capabilities (score from 1 to 5), and the ancillary benefit (score from 1 to 2). The R factor and B factor are then represented in a diagram.

Conceptual limitations of principle of threes and turbo model

The categorization of disease seriousness in both models is subjective, and lacks a conceptual framework for trading-off the benefits and harms.[21] Both models have limited number,[17,22] and not comprehensive benefit and risk criteria.[17,22,23] Moreover, patient health preferences are not considered. Both models are too simplistic for even moderately complex cases.[19]

Benefit-less-risk analysis (BLRA)

In this method, risk is represented by five different body functions selected to be of concern to the situation under consideration; next, each body functioning is assigned an intensity grade; then, the importance weight of each body functioning to the others are set by patient and reflect patient's overall well-being; after that, the intensity grades for different body functioning are combined using the importance weight of each body functioning. Benefit is typically measured on a smaller number of endpoints. Then, risk weighted score is multiplied by a conversion factor, the result is subtracted from the benefit score.[24,25]

Conceptual limitations of benefit-less-risk analysis (BLRA)

Benefit and risk in Benefit-Less-Risk Analysis (BLRA) method are not clearly defined,[17,25,26] interpretation is very complex,[24,25] and requires extensive quantification.[17,24] Also, setting of the conversion factor requires some critical thinking.[24] The assigned weights are subjective and subject to bias, and may affect validity.[24,25] Finally, it does not account for multiple adverse events.[17] Seeking more transparent method is necessary.[24]

*OMERACT 3*3*

In this method, benefits and harms are estimated by three categories of benefit and three categories of harm, creating a three by three table; None/minimal, major, and (near) remission for benefit, and none/minimal, major, and (near) death for harm.[27]

*Conceptual limitations of omeract 3*3*

Categorizing and weighting of benefits and harms in this method are subjective.[27] Three categories only simplifies the matters and does not replace deeper analyses.[27] The patient health preferences are also not considered.

Number needed to treat (NNT)

NNT Concept is defined as the inverse of the Absolute Risk Reduction (ARR).[28] The number needed to treat is the average number of patients needed to be treated to prevent an adverse outcome in one additional patient compared to a control treatment group;[29-32] in other words, number needed to treat is the number of people who need to be treated over a defined time to achieve the required outcome in one of them.[33] Physicians are widely using the number needed to treat because it is

usually reported as an integer and easy to understand;[34] it's understanding is relatively straightforward.[35]

$$NNT = \frac{1}{P1 - P2}$$

Where P1 is the proportion of the disease of interest in the control group, and P2 the proportion of disease in the treatment group.[28]

Absolute risk reduction (ARR)

ARR is the arithmetic difference between the incidence of harm condition of concern in the treatment group and the incidence of harm condition of concern in the control group.[32,33,36,37] It is suggested to use preferably more than NNT for both theoretical and practical reasons.[38]

Conceptual limitations of number needed to treat and absolute risk reduction-based methods

NNT and ARR-based methods does not consider multiple benefits and harms,[22,25,27,39,40] and does not account for utilities[24,25,39,40] and time dimension[39-41] of outcomes. Those limitations are exceeded only by The Adjusted Number Needed To Treat method,[39] and account for utilities only is considered by Relative-Value Adjusted Number-Needed-To-Treat method.[26]

NNT and ARR-based methods compare only two therapeutic options at a time; the situation will be more complex when comparing more than two therapeutic options.[35,40] NNT and ARR-based methods does not consider successful outcomes that are associated or not associated with treatment-related adverse events;[35,40] this limitation is exceeded by NNT_{US} and NNH_{UF} methods.[40]

NNT do not have good statistical properties;[34,42] among them is that when the denominator (ARR) is zero, in which the result of NNT and NNT-based methods become not interpretable and biased.[26,38] Also, the sums of different NNTs, can give meaningless results.[43] It is concluded that the nature of NNT scale is biased.[33,44,45]

Another weakness of those methods is that the severity importance of the adverse event relative to the benefit is not considered.[22,26,33,35] Moreover, NNT and ARR-based methods are dependent on baseline risk; this will lead to limited generalizability. It is inappropriate to extrapolate the results of NNT and ARR-based methods from one population to another population with a different baseline risk, and results are only applicable in similar settings for both populations.[22,32,33] Those methods also have significant limitations in terms of comprehensiveness and comparability,[46] they are not suitable for making value judgments,[21] and are not helpful for communicating harms.[45] The combining of two rates in one statistical measure was poorly expressed.[38]

Minimal necessary efficacy of the treatments

It is also named as Minimum Clinical Efficacy (MCE).[26,35] MCE calculates weighting of the benefit and the risk of a specific treatment. It aims to find the minimal therapeutic benefit at which a treatment is still worth administering.[47] It includes the calculations of benefits and risks for the new, old and no treatments to achieve this goal.[48] Risks of treatment of the disease and multiple adverse event profiles outcome are estimated by mortality and morbidity. The benefit is presented in terms of number needed to treat, relative risk reduction, and outcome utilities. Outcome utilities are expressed by length of life, absence of pain, cost, and the strength of individual patient preference for an outcome. Patient preference for an outcome is represented by the probability of patients who will be free from consequences of the disease or toxicity of the treatment.[47] MCE considers the natural characteristics of the disease in the general population.[35]

Conceptual limitations of minimal necessary efficacy of the treatments

Minimal Necessary Efficacy is difficult to explain to patients and stakeholders.[27] It does not include uncertainty in the benefit or risk measurements.[25] Values of the method need subjective and study-specific judgments.[49] While it is based on NNT, it inherits NNT limitations.

Disability-adjusted life years (DALYs)

The disability-adjusted life years (DALYs) had been developed in order to calculate the loss associated with premature mortality and disability. Severity of disability is classified into six categories, ranging from class 1 which is "limited ability to perform at least one activity in one of the areas of recreation, education, procreation or occupation", to class 6 which is "needs assistance with activities of daily living such as eating, personal hygiene or toilet use".[50] Classification of disabilities into the six classes and severity weighting of each class is set by expert panel.[51,52] The weights of class severity are between zero and one. The unit of measure for the burden of disease is time (in Daly);[50] DALYs are calculated by the sum of years lost from premature mortality and the years lost from disability.[50,53] For example, a woman with disability for ten years and disability weight of 0.4 and then she died ten years prematurely; her loss in health would be 14

DALYs; that is, the sum of the 10 years of lost life plus the four-year loss (10 x 0.4) from the disability.[51] Cost-effectiveness then could be directly calculated for each therapeutic option.

Quality-adjusted life years (QALYs)

The QALYs method is measuring both the quality and the quantity of life lived.[54,55] QALYs are the product of life expectancy (in years) and its quality (utility) over that time (estimated in QALY units);[55] each year of life considered is given a coefficient between 0 and 1;[46,56,57] 0 represents the value or utility score for death and 1 represents normal full health.[46,57,58] Thus, ten years of life expectancy at a utility of 0.5 is equivalent to five years with full health.[58] The patients estimate subjectively their own lived years-quality with handicap or serious discomfort by different methods like time trade off method, standard gamble, or from generic health-state questionnaires.[46] Time trade off and standard gamble methods are discussed later in this chapter. Cost-effectiveness/efficacy then could be directly calculated for each therapeutic option.[55]

Conceptual limitations of quality-adjusted life years (QALYs) and disability-adjusted life years (DALYs)

QALYs and DALYs are complex, abstract, and controversial, especially with the estimation of value preferences. Since preferences are totally subjective, results may open serious objections.[56] There are some concerns about the validity, and the application of those methods for decision-making.[25] The concerns about the sensitivity of such methods for measuring differences in health outcomes are fortified.[46] QALYs life-years unit and utility unit are not the same, joining them directly compromises QALYs arithmetic concept; to obtain coherent results, both scales should be expressed in the same measurement unit.[59] The same conclusion could be applied on DALYs.

The role of both methods in benefit-risk assessment remains unclear.[17,60] The conceptual basis for the both metrics is flawed.[60,61] They are weak in acknowledging uncertainty of the outcomes.[19] They also discriminate against patients having limited treatment potential.[62] An example is if two patients are suffering the same health condition, but one has another disabling-health condition, and the other does not have; QALYs method will discriminate and give the priority for the non-disabled patient for the treatment of the shared health condition because his/her quality of life will improve more than the quality of life of the disabled patient.[53] This discrimination causes deep and unresolved difficulty for use of cost-effectiveness analysis with QALYs to prioritize health care.[53]

QALYs method is inappropriate to use as a basis for the comparison of different health outcomes, and health technologies.[55] QALYs don't perfectly measure the quality of life of various health conditions nor perfectly measure health states using interval scale.[55] It ignores salient societal concerns regarding resource allocation,[63] and health care priorities based on it cause public discomfort.[64] QALYs bias against palliative treatments that do not shorten premature death or improve the quality of life of patients with long life expectancies.[65] QALYs can be measured in different ways yielding different unreliable results[55,66] and should not be used for decision making.[55] While QALYs is using utility health methods, QALYs inherits their limitations.

DALYs is rough measure,[51] and not sensitive enough to capture the patient outcomes;[67] they consider all disabilities regardless their type or severity to be equal, which restricts its ability to rank various interventions.[51] The DALYs cannot replace traditional methods to assess disease, and treatment in clinical practice.[68] The validity of DALYs depends largely on the validity of the DALYs panel's composition and the underlying assumptions.[51] DALYs devalue the lives of women because they do not consider for social differences and how their lives are lived.[51] Critics argue that lives of the patient with disabilities worth less than the patient without disabilities by counting a year lived with a disability as less than a full year.[60,69] Because of that, it will drive away resources from disabled patients.[51,69]

DALYs' disability-concept does not accord with that in WHO's *International Classification of Functioning, Disability, and Health (ICF)*. DALYs approach ignores equity or acts directly counter to it.[60] It does not reflect life with a disability as experienced by disabled people.[51] Also, it does not take in consideration needs that patients with different functional statuses might have.[51] Moreover, it does not value interventions that enhance the lives of disabled patients.[51]

Disability-free life expectancy

Disability-free life expectancy deducts the disability years from life expectancy regardless the severity of different disabilities, and no weighting is used to differentiate between them.[51]

Conceptual limitations of disability-free life expectancy

It has DALYs' limitations. Moreover, it is a very rough estimate and do not consider the nature and severity of different disability conditions at all.

Time without symptoms of disease and toxic effects (TWIST), and quality adjusted time without symptoms of disease and toxic effects (Q-TWIST)

Time without symptoms of disease and toxicity of treatment (TWIST) is set to provide a single metric of length and quality of survival. Time with subjective adverse effects of treatment and time with unpleasant symptoms of disease are subtracted from overall survival time to calculate TWIST for each patient.[70] Health state preferences are estimated by assigning subjective weights.[17]

Quality Adjusted Time without Symptoms of Disease and Toxic Effects (Q-TWIST) uses a quality of life index (utility) to estimate the survival in which a day with low quality will not considered as a whole day, but a fraction of a day; the fraction is estimated by the utility weighting applied to that day.[24,71]

Conceptual limitations of time without symptoms of disease and toxic effects (TWIST), and quality adjusted time without symptoms of disease and toxic effects (Q-TWIST)

Both TWIST and Q-TWIST are intuitive.[24] Their utility weightings also are subjective, challenging,[24] and not explicit.[17] There are some concerns about the validity, and the application of those methods for decision-making.[25] For Q-TWIST, using different utility techniques yield different results,[25] and while Q-TWIST is using utility health methods, Q-TWIST inherit their limitations.

Incremental net health benefit (INHB)

Net Health Benefit (NHB) is the difference between the sum of the weighted benefits and the sum of the weighted risks of a treatment.[72] Benefits and adverse events of a treatment are quantified using available clinical trial or post-marketing surveillance data.[72] Importance weights of each outcome are usually incorporated using QALYs.[72,73] The difference between the NHB of a treatment and the NHB of an alternative treatment or standard of care represents the Incremental Net Health Benefit (INHB).[19,72,74] A positive INHB means that the net benefits of the treatment are positive relative to the comparator.[72]

Conceptual limitations of incremental net health benefit (INHB)

INHB is a decision aid not a replacement for expert judgment.[75] It is difficult to explain to patients and stakeholders.[27] While it is using QALYs method, INHB inherits its limitations.[25]

The clinical utility index (CUI)

Clinical utility is defined as the net benefit of therapy as perceived by the physician or expertise. A CUI gives value for each attribute in the product profile using physician preference data, or by relying on expert opinion to provide a single metric for multiple dimensions of benefit and risk.[76-79] The CUI can be expressed mathematically by the following equation:[76]

$$CUI = \sum_{i=1}^{n} W_i U_i$$

Where W is the weight of the attribute and U represents the utility function. The weight of the attribute represents the relative importance of each attribute to the others. Utility represents the clinically meaningful differences of an attribute. Each utility can be assigned value between zero (worst outcome) and one (best outcome) for both efficacy and safety measures. Another approach assumes that efficacy attributes are expressed with positive values (0 to 1) and safety attributes are expressed with negative values (-1 to 0).[78] Because of the subjectivity in deciding the weights and clinically meaningful differences for each attribute, a sensitivity analysis is usually performed.[76,78] The use of external data to minimize subjectivity, weighting and definition of clinical cut off are all important.[78]

Multiple-criteria decision-making techniques

Multiple-criteria decision-making techniques include decision trees, value trees, and other techniques. Decision trees are usually used for solving problems of choice under uncertainty with multiple objectives as usually takes place in different clinical decisions.[80] Decision trees aids modeling the logical flow of clinical problems and determines the best choice among multiple options by calculating probabilities of events and inserting the valuations of possible outcomes.[80] They can handle the data of different types.[81] Decision trees are a sequential probabilistic branches from an initial state of health or medical intervention[82] that branches from left to right.[19]

The first step of building a decision tree is to set a list for different therapy options with their consequences, every option is branched into one or more branches, every branch represents a consequence of the therapy option with probability of occurrence. Every sub branch could be also divided into more branches and so on until reaching the end outcomes in the most right level. Then, every end outcome is assigned a utility. After that, the utility of every end outcome is multiplied by the probabilities of all of its above branches, and the sum of all end outcomes scores is calculated for every therapeutic option. The therapy with the highest value represents the best option. Finally, sensitivity analysis is performed.[80] Sensitivity analysis includes changing repeatedly the probabilities and weights of the criteria

across a plausible range and testing the impact of this change on the overall decision.[80,83-85]

Value trees are almost sharing the same structure with decision trees. It handles problems of choice with multiple objectives under uncertainty.[17] An example of implementing the value trees in the clinical field is the Multi Criteria Decision Analysis (MCDA) approach, which is used to evaluate the benefit risk ratio of medicines.[21,83] In this approach, complex problems are partitioned into more manageable parts, which can be studied using data and judgment. The partitions are then recombined after scoring and weighting using computer software to provide a consistent overall picture for the decision makers.[83] The approach helps thinking for decision-makers to be more explicit, consistent and transparent in their discussions, but does not replace their judgment or take decisions.[83,86]

The first step in this approach is to detect a list of benefit and risk criteria, which form benefit-risk profile, and detection of the options to be evaluated. Then, for every option, all criteria are modeled graphically using value tree. After that, MCDA sub classifies every benefit and risk criterion in the value tree. Then, every criterion for an option in the most right level of the tree is scored using a scale and every criterion is assigned a weight to reflect its relative importance to the others. Scoring weights are set by experts. After that, the product of every criterion score for an option time's criterion's weight is calculated, and the sum of the products is calculated for both benefits and risks. Then, the total scores of benefits and risks are compared for every option.[21,22,46,83,87] The final output will be a single risk-adjusted benefit resulted from collapsed multiple dimensions.[88] The last step is running sensitivity analysis to detect the importance of each criterion and its impact on the result.[21,22,83,89]

Examples of the benefit criteria in this model are efficacy versus comparator and its clinical relevance, statistical adequacy of the trial, statistical significance of the efficacy results. Examples of the Risk criteria are overall incidence of adverse effects, Overall incidence of serious adverse effects, discontinuation rate due to adverse effects, incidence, seriousness, duration and reversibility of specific adverse effects, safety in subgroups, drugs and food interactions with other.[21]

Another example which is similar to the MCDA technique[90] is The Benefit-Risk Assessment Model (BRAM).[17] BRAM includes evaluative judgments with relevant data to provide a platform for trading off multiple benefit and risk components in a transparent and consistent manner.[17] In this model, a branched hierarchy of benefits and risks are presented instead of the value tree.[17] Benefit here includes efficacy, life effects, and convenience, and risk includes safety, tolerability, and improper use of drugs.[17]

Other example of using value tree is the BRAT Framework, which is developed by The Benefit Risk Action Team, and it is a set of processes and tools for selecting, organizing, summarizing, communicating, and interpreting data for benefit-risk assessments. It provides a standardized and flexible platform for incorporating outcomes and preference weights for decision-making.[91,92]

This method first defines the decision context, which includes drug, dose, formulation, indication, patient population, comparator(s), and time horizon for outcomes. Second, it identifies all-important outcomes and creates the initial value tree. Then, it assesses outcome importance by applying any ranking or weighting of outcome importance to decision makers or other stakeholders.[91,92] The BRAT framework does not apply any particular method for weighting, and does not require the use of weights.[91] The value tree could be updated as new data or more precise definitions of the outcome end points or measures become available.[91]

Examples of benefits are reducing pain, reduction in functional disability and other specific case-related benefits. Example of risk is different individual case-related risks.[91,92]

Previous models have some merits. They combine judgments numerically in a transparent way.[17,22,86] The balance of benefits and risks can be evaluated for multi-therapy and against placebo, or against active control.[22] The models consider a comprehensive benefit and risk criteria of potential relevance, and one or more additional benefit and risk criteria could be added in flexible way.[21,22,93] They enable a discussion and trade-offs about the subjective weights.[21,22] They also consider potential uncertainty in the case of incompleteness of the evidence.[21] They demonstrate the value of the social effect,[83] and they are applicable to all kinds of medicines and medical devices[21] because they can handle data of mixed type.[81]

Conceptual limitations of multiple-criteria decision-making techniques

Multiple-Criteria Decision-Making Techniques can't generate decisions; instead, they serve as a aid to thinking and decision-making.[83] Setting up the model is time consuming, burdensome, and require building up a complex model for every situation, therapeutic

area or even product or indication.[19,22,25,73,94] There are no constant benefit and risk safety criteria[17,95] and utilities with their limitations are usually used for this evaluation.[80] The results of the models are uncertain,[96] and a danger may take place if a decision over relies on them.[22] Additionally, optimal weights of the criteria are not guaranteed because weights are subjective, and require tradeoff between models' criteria.[17] Decision makers also may not achieve a consensus about the weights.[97] These models do not support the calculation of relative benefit, harm and risk of therapeutic options in a health system; instead, they estimate the benefits and risks for every clinical case alone.[17,98] The models are complex to explain to patients and nontechnical stakeholders.[27] Models' results are snapshots in time.[17] Finally, some of those models do not apply any weights, which is a limitation by itself.

Discussion of the primitiveness of the current methods

The process of assessment of benefits, harms, risks, and patient clinical outcome is still primitive,[27] wanting[99] and rudimentary,[100] primarily undefined,[26] or ill-defined,[46] not well-developed,[101,102] biased[103] and limited to make informed decisions,[104] in its infancy and early stage,[78,105-107] and is largely uncharted.[108] It relies on subjective judgment of experts[91,92] and rarely done in a quantitative fashion.[78] A consistent quantitative assessment structure is lacking,[109] and there is very little or no well-established and recognized approach on how to do it.[24,54,110,111] It is not typically performed, not presented in a consistent, systematic, analytical and integrated framework using single scale,[25,73,112] and not standardized.[113,114] Currently, there is no general universal solution available for the assessment process[73,91,115] and this process is still subject to continuous development.[114,116]

The estimation of the ratio between the benefits of drugs and their harms is not obvious or easy[117] and not applicable in a straightforward manner.[21,112,118] Although benefits and harms assessment is the core of the drug development, standardized and validated quantitative conceptual models, which measures patient outcome are lacking.[25,119] This also leads to lack of consistency in the comparison of pharmaceuticals[73,120] and health programs.[121]

Using the term benefit-risk ratio without explaining its meaning is common in the literature.[122] There is no single, clear definition of "benefits" and "risks".[123] Rarely, any quantitative analysis is used or attempted to synthesize in the articles where benefit, and risk words are mentioned in the title,[26] and even in the highest impact medical journals, benefits and harms evidence is not consistently presented to make direct interpretation much easier.[45] Some researchers are not willing to adopt a quantitative benefit-risk assessment because it does not accurately represent the whole picture to patient.[24]

It is frustrating[19,21] that there is no generally agreed metric or methodology, and no standard with widely acknowledged definitions for benefit-risk assessment.[19,21,123] Benefits and harms are not usefully combined into one scale.[96] There are no standards in which comparisons against might be made[26] nor clearly showing the benefits or harms of treatments in a clinically useful way.[33] The common practice of providing separate lists of benefits and adverse effects cannot be justified as a scientific analysis and the decision made relying on them will be subjective.[26,124] Also, there is no method for measuring the quality of the decision made.[120] This process is necessary to track how safety is being monitored and assessed.[125]

The current US system for assessing the risks of therapeutics is outdated and inadequate.[126] Regulatory authorities in EU, US and Japan did not issue criteria for benefit and risk assessment.[54] Neither the US Food and Drug Administration (FDA) nor the CHMP have released methods for benefit-risk analyses. The available methods of analyses are limited to non-regulatory situations,[46] and research domain.[54] The public's health techniques for the detection, verification, and quantification of safety issues are also scattered and disappointing but could be improved.[127]

Benefits, harms, and risk assessment is more than the subjective opinion of a group of experts.[105] The benefit-risk assessment differs between countries, and regulatory authorities differ in the threshold for taking action and for handling of therapeutic risk management plans.[128] The definition of benefit can be quite varied.[116] Decisions are made on a relatively informal and irrational basis.[129] Because of the subjective judgments, evaluation process is varied between different assessors and assessments.[130] Different regulatory authorities and countries have different decisions and actions using the same data inconsistently.[21,107,128]

There is an increasing importance of cost-effectiveness analysis.[82,131] Assessment of therapy effectiveness is primary driver of cost-effectiveness analysis and economic modeling.[132] Any economic evaluation should be based on a representation of the effectiveness data.[133] There is a lack of consensus on evaluation criteria and standards for cost-effectiveness analysis and economic modelling,[134] and how to weigh those

criteria.[120] Economic models can have political, discriminatory, or arbitrary biases[135] and there are many shortcomings in the existing cost-effectiveness models, which consequently affect the legitimacy of their recommendations.[132] In addition, they fail to identify existing misallocation of resources.[136] The procedures to assure transparency for many of these models are also unclear.[132] Unfortunately, the majority of decisions about the cost-effectiveness of interventions are also made on an uncertain information.[137,138]

Accordingly, communication of clinical benefits and harms is also currently infant,[139] modest[104] limited[140] and in a sorry state.[141] It needs more attention to the theory and practice.[142] There is a lack of transparent method of communicating these information,[45,112] and there is a need for better effective clear ways of communicating risk information to patients and healthcare practitioners.[112,143,144] There is a remained space for development in this area.[145,146] More effective ways should be developed for clinicians to understand and interpret clinical data and to assess patient perceptions for the harms and benefits of the drugs, devices and biologics that they use.[126] Current ways of how information presented can alter receiver's decisions.[112,147-149] Consequently, the lack of shared communication and understanding can increase safety problems.[150]

The Institute of Medicine's committee on the Assessment of the US Drug Safety System recommends that the new Office of Drug Safety Policy and Communication should develop a cohesive risk communication plan to review all risk communication activities of the center, and evaluation of communication tools.[125]

There are also theoretical and practical problems for estimating patient health preferences[55,151] such as the proper source of preferences weights.[151] Significantly, different scores may be yielded after the same intervention for the same patients.[152] In spite that current patient wants to contribute actively in his/her treatment[153] and there is an agreement to include patient perception in the outcome, it is not easy to achieve.[56,154] There is no federal agency, which has a formal method for weighing preference variation, and no consensus in the literature how to do it.[112,155] It is obvious that both patient outcomes and preferences are often inadequately measured,[156] and consumers and patients are often not sure how to weigh risks and benefits for different options.[157] There is a need for systematic method, which incorporates patients' preferences and values into clinical decisions.[158]

Patient clinical outcome assessment is based primarily on the definition of health and health status.[159] A shared definition of health is needed for valid assessment method[160] and will enhance the quality in health care. The factors of health outcome and what considered as dimensions of health and their relative contributions are unspecified, variable[161] and questionable.[162] The attempts to define health are futile and lack operational value.[162,163] Health services administrators lack a good working definition of health, and no universally-accepted instruments for measuring it.[14] Quantization the components of health is also a complex task.[164]

WHO defines health as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity."[165] It is not obvious how this definition supports clinical and public health practice or how it can be measured or operationalized.[14,166] Terms in the definition like "complete", "social well-being", and "disease and infirmity" all are not clear and need to be defined.[163] The definition is too abstract and oversimplified. It expresses the final goal in health rather than a method for solid action. It does not illustrate the relative importance of its components, and does not include mortality.[14] It does not distinguish health from happiness, which is not intended to be measured in the health domain for many reasons.[167]

Another common term, which is related to patient clinical outcome assessment, is the Health Related Quality of Life (HRQOL) term. Quality of life has a vague and difficult concept to define.[168] The concept is abstract and complex, and has no definition consensus,[56,152,169] which reflects the lack of theoretical conceptualization of the term.[152] There are serious methodological and logical troubles in the construction of HRQOL measurement, and it is recommended that HRQOL measurement be neglected.[170] The components of quality of life are a personal issue, which leads to a philosophical rather than a scientific approach.

A good definition could be operationalized and operational definitions of quality of life term are woefully inadequate.[56] The use of the term quality of life to reflect the values and perceptions of patients has confusion, and misunderstanding among health professionals, and patients because of unclear conceptual definition.[56] Evaluation of quality of life is based on arguments rather than on rational debate, and so, comparing instruments on scientific grounds is difficult.[56]

No HRQOL instrument is universally recommended, and no gold standard is available.[171] The differences between quality of life measurement methods highlight the difficulties of a standard definition of the concept.[56] There is not a single instrument, which stands out above the rest,[172] and there are no "worst" or "best" instruments.[173] Accordingly, it is difficult to reflect a decision maker's preference. Lack of standard instrument forms obvious difficulty in the validation of health-related quality of life measures.[174] It is difficult to progress in the field if there is no consensus over concept-definition,[56] and those critical scientific, and logistic obstacles in this field need to be overcome.[175]

Well-being is also a widely used term, which is related to the outcome assessment. Well-being is even more ambiguous.[168] The definitions of the term are diverse[176] and inconsistent, and its vague concept will hamper the development of knowledge and theory in research.[176] Well-being is a complex and many-sided construct, which still eluding researchers to define and measure.[177]

Quality of life and Well-being have a complex construct with variable meaning.[178] They are both elusive concepts having problems in measurement and definition.[179] General accepted definitions for both are still lacking[168,180] and they are most often used interchangeably.[180]

Conclusion

An assessment of patient clinical outcome is a very important facet of patient safety, and involves the measurement of the therapeutic options in terms of their benefits, harms and risks. Limited research has been devoted for the development of assessment methods for the clinical field. Consequently, clinical outcome assessment methods in some clinical fields have been employed without sufficient understanding of their characteristics. This triggers the need for more research to better understand the existing patient clinical outcome assessment methods and their conceptual limitations. Accordingly, this paper proposes an informative review on the assessment of patient clinical outcome that will help researchers, clinicians and health care professionals to design novel general methods for the assessment of patient clinical outcome that avoid the conceptual limitations of existing methods.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Deshpande PR, Rajan S, Sudeepthi BL, Abdul Nazir CP. Patient-reported outcomes: a new era in clinical research Perspect Clin Res 2011;2:137-44.
2. Scott NW, Fayers PM, Aaronson NK. The relationship between overall quality of life and its subdimensions was influenced by culture: analysis of an international database J Clin Epidemiolo 2008;61:788-95.
3. Marquis P, Caron M, Emery MP. The role of health-related quality of life data in the drug approval processes in the US and Europe Pharmaceutical Medicine 2011;25:3147-60.
4. Cambria E, Benson T, Eckl C. "Sentic PROMs: Application of sentic computing to the development of a novel unified framework for measuring health-care quality Expert Systems with Applications 2012;39;12.
5. Scott NW, Fayers PM, Aaronson NK. Differential item functioning (DIF) analyses of health-related quality of life instruments using logistic regression Health Qual Life Outcomes 2010;8;81.
6. Benson T, Sizmur S, Whatling J, Arikan S, McDonald D, Ingram D. Evaluation of a new short generic measure of health status: howRu. Inform Prim Care 2010;18;89-101.
7. Yu, C. H. A Simple Guide to the Item Response Theory (IRT) and Rasch Modeling. Retrieved from www.creative-wisdom.com/computer/sas/IRT.pdf 2011;1-30.
8. Walton MK, Powers JH, Hobart J. Clinical Outcome Assessments: Conceptual Foundation-Report of the ISPOR Clinical Outcomes Assessment-Emerging Good Practices for Outcomes Research Task Force Value Health 2015;18;741-52.
9. Spuls PI, Lecluse LL, Poulsen MLN How good are clinical severity and outcome measures for psoriasis?: quantitative evaluation in a systematic review J Invest Dermatol 2010;130;933-43.
10. Tallarida RJ, Murray RB, Eiben C. A scale for assessing the severity of diseases and adverse drug reactions Application to drug benefit and risk Clin Pharmacol Ther 1979;25;381-90.
11. Kind P, Dolan P, Martínez C. Adjusting outcomes for quality: scaling the severity of adverse drug reactions. Presented at the 13th Annual Meeting of the International Society Technology Assessment in Health Care, Barcelona, Spain, May 1997.
12. Koh Y, Yap CW, Li SC. Development of a combined system for identification and classification of adverse drug reactions: Alerts Based on ADR causality and severity (ABACUS)," J Am Med Inform Assoc 2010;17;720-2.
13. Edwards IR, Aronson JK. Adverse drug reactions: Definitions, diagnosis, and management Lancet 2000;356;1255-9.
14. J S. Larson. The weighting of an international health status index *Social Indicators Research* 1994;31;3:265-75.
15. Brown MJ. A rational basis for selection among drugs of the same class Heart 2003;89;687-94.
16. Aronson JK, Ferner RE. Clarification of terminology in drug safety Drug Saf 2005;28:851-70.
17. Felli JC, Noel RA, Cavazzoni PA. A multiattribute model for evaluating the benefit-risk profiles of treatment alternatives Med Decis Making 2009;29:104-15.
18. Beckmann J. "Basic aspects of risk-benefit analysis Semin Thromb Hemost 1999;25:89-95.
19. Phillips LD, Fasolo B, Zafiropoulos N, Beyer A.Benefit-risk methodology project work package 2 report: Applicability of current tools and processes for regulatory benefit-risk assessment. London: European Medicines Agency; 31 Aug 2010. Report No.: EMA/549682/2010.
20. Edwards IR, Wiholm BE, Martinez C. Concepts in risk-benefit assessment. A simple merit analysis of a medicine? Drug Saf 1996;15:1-7.

21. Mussen F, Salek S, Walker S. A quantitative approach to benefit-risk assessment of medicines-part 1: The development of a new model using multi-criteria decision analysis Pharmacoepidemiol Drug Saf 2007;16:(Suppl 1): S2-S15.
22. CHMP. *Report of the CHMP working group on benefit-risk assessment models and methods*, European Medicines Agency, London, 2007.
23. Mussen F, Salek S, Walker S. *Benefit-risk appraisal of medicines: A systematic approach to decision-making,* UK: Wiley-Blackwell, 2009.
24. Chuang-Stein C, Entsuah R, Pritchett Y. "Measures for conducting comparative benefit:risk assessment," Drug Information Journal 2008;42;3:223-33.
25. Guo JJ, Pandey S, Doyle J. A review of quantitative risk-benefit methodologies for assessing drug safety and efficacy–report of the ISPOR risk-benefit management working group Value Health 2010;13:657-66.
26. Holden WL. Benefit-risk analysis: A brief review and proposed quantitative approaches Drug Saf 2003;26:853-62.
27. Maarten B, Peter B, James FF. A first step to assess harm and benefit in clinical trials in one scale J Clin Epidemiol 2010;63:627-32.
28. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment N Engl J Med 1998;318:1728-33.
29. Bender R. "Number Needed to Treat (NNT)," *Encyclopedia of biostatistics*, P. Armitage and T. Colton, eds.,Chichester John Wiley and Sons, Ltd. 2005;pp.3752-61.
30. Baglin T. "Communicating benefit and risk," British Journal of Haematology 2009;146:31-33.
31. Carneiro A. Relative Risk, absolute risk and Number Needed to Treat: Basic concepts Rev Port Cardio 2009;28:83-7.
32. A K. Akobeng. Understanding measures of treatment effect in clinical trials Arch Dis Child 2005;90:54-6.
33. Akobeng AK. Communicating the benefits and harms of treatments Arch Dis Child 2008;93:710-3.
34. Hutton JL. Number Needed to Treat: Properties and problems," *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2000;163:381-2.
35. Holden WL, Juhaeri J, Dai W. Benefit-risk analysis: A proposal using quantitative methods Pharmacoepidemiol Drug Saf 2003;12:611-16.
36. Cook RJ, Sackett DL. The Number Needed to Treat: A clinically useful measure of treatment effect BMJ 1995;310:452-4.
37. McHugh ML. Clinical statistics for primary care practitioners: Part II–Absolute Risk Reduction, Relative Risk, Relative Risk Reduction, and Number Needed to Treat J Spec Pediatr Nurs 2008;13:135-8.
38. Hutton JL. "Number Needed to Treat and Number Needed to Harm are not the best way to report and assess the results of randomised clinical trials," *British Journal of Haematology* 2009;146:27-30.
39. Riegelman R, Schroth WS. Adjusting the Number Needed to Treat: Incorporating adjustments for the utility and timing of benefits and harms Med Decis Making 1993;13:247-52.
40. Mancini GBJ, Schulzer M. Reporting risks and benefits of therapy by use of the concepts of unqualified success and unmitigated failure: Applications to highly cited trials in cardiovascular medicine Circulation 1999;99:377-83.
41. Christensen PM, Kristiansen IS. Number Needed to Treat (NNT)-Needs treatment with Care Basic Clin Pharmacol Toxicol 2006;99: 12-16.
42. Grieve AP. "The Number Needed to Treat: A useful clinical measure or a case of the Emperor's new clothes?," Pharmaceutical Statistics 2003;2:87-102.
43. Lesaffre E, Pledger G. A note on the Number Needed to Treat Control Clin Trials 1999;20:439-47.
44. Duncan BW, Olkin I. Bias of estimates of the Number Needed to Treat Stat Med 2005;24:1837-48.
45. Sedrakyan A, Shih C. "Improving depiction of benefits and harms: Analyses of studies of well-known therapeutics and review of high-Impact Medical Journals," *Medical Care* 2007;45:S23-S28.
46. Hughes DA, Bayoumi AM, Pirmohamed M. Current assessment of risk-benefit by regulators: Is it time to introduce decision analyses? Clin Pharmacol Ther 2007;82:1237-27.
47. Djulbegovic B, Hozo I, Fields KK. High-dose chemotherapy in the adjuvant treatment of breast cancer: Benefit/risk analysis Cancer Control 1998;5:394-405.
48. Holden WL, Juhaeri J, Dai W. Benefit-risk analysis: Examples using quantitative methods Pharmacoepidemiol Drug Saf 2003;12:693-97.
49. Terwee CB, Dekker FW, Wiersinga WM. On assessing responsiveness of health-related quality of life instruments: Guidelines for instrument evaluation Qual Life Res 2003;12:349-62.
50. Murray CJL. Quantifying the burden of disease: The technical basis for Disability-Adjusted Life Years Bull World Health Organ 1994;72:429-45.
51. Daniel M. Measuring health and disability Lancet 2007;369:1658-63.
52. Bognar G. Well-being and Health Health Care Anal 2008;16:97-113.
53. Edejer TTT, Baltussen R, Adam T Making choices in health: WHO guide to cost-effectiveness analysis, Geneva: World Health Organization, 2003.
54. CHMPReflection paper on benefit-risk assessment methods in the context of the evaluation of marketing authorisation applications of medicinal products for human use, European Medicines Agency, London, 2008.
55. McGregor M, Caro JJ. QALYs: Are they helpful to decision makers? Pharmacoeconomics 2006;24:947-52.
56. Leplege A, Hunt S. The problem of quality of life in medicine JAMA 1997;278:47-50.
57. Brazier J. Valuing health states for use in cost-effectiveness analysis Pharmacoeconomics 2008;26:769-79.
58. Garrison LP, Towse A, Bresnahan BW. Assessing a structured, quantitative health outcomes approach to drug risk-benefit analysis Health Aff (Millwood) 2007;26;684-95.
59. Prieto L, Sacristá J. n Problems and solutions in calculating Quality-Adjusted Life Years (QALYs) Health Qual Life Outcomes 2003;1;80.
60. Anand S, Hanson K. Disability-Adjusted Life Years: A critical review J Health Econ 1997;16:685-702.
61. Duru G, Auray JP, Béresniak A. Limitations of the methods used for calculating quality-adjusted life-year values Pharmacoeconomics 2002;20:463-73.
62. Gold MR, Stevenson D, Fryback DG. HALYs and QALYs and DALYs, OH MY: Similarities and differences in summary measures of population health Annu Rev Public Health 2002;23:115-34.
63. Nord E. Towards cost-value analysis in health care? Health Care Anal 1999;7:165-75.
64. Ubel PA, Richardson J, Prades JLP. Life-saving treatments and disabilities: Are all QALYs created equal? Int JTechnol Assess Health Care 1999;15:738-48.
65. Hughes J. Palliative care and the QALY problem Health Care Anal 2005;13:289-01.
66. Chapman RH, Stone PW, Sandberg EA. A comprehensive league table of cost-utility ratios and a sub-table of "panel-worthy" studies Med Decis Making 2000;20:451-58.
67. Lam CLK. "Subjective quality of life measures–General principles and concepts," Handbook of disease burdens and quality of life measures, V. R. Preedy and R. R. Watson, eds., USA: Springer, 2010:381-99.
68. Sundby J. Are women disfavoured in the estimation of Disability Adjusted Life Years and the global burden of disease? Scand J Public Health 1999;27:279-85.
69. Arnesen T, Nord E. The value of DALY life: Problems with ethics and validity of disability adjusted life years BMJ 1999;319:1423-25.
70. Gelber RD, Gelman RS, Goldhirsch A. A quality-of-life-oriented endpoint for comparing therapies Biometrics 1989;45:781-95.
71. Gelber RD, Cole BF, Gelber S "Comparing treatments using quality-adjusted survival: The Q-Twist Method," *The American Statistician* 1995;49:161-69.
72. Hauber AB, Johnson FR, Andrews E. "Risk-benefit analysis methods for pharmaceutical decision-making – Where are we now?," *International Society for Pharmacoeconomics and Outcomes Research* 2006;12:3-5.
73. Cross JT, Garrison LP. Challenges and opportunities for improving benefit-risk assessment of pharmaceuticals from an economic perspective, Office of Health Economics (OHE), London, 2008.

74. Craig BA, Black MA. Incremental cost-effectiveness ratio and incremental net-health benefit: Two sides of the same coin Expert Rev Pharmacoecon Outcomes Res 2001;1:37-46.
75. Lynd LD, Najafzadeh M, Colley L Using the incremental net benefit framework for quantitative benefit-risk analysis in regulatory decision-making- A case study of Alosetron in irritable bowel syndrome Value Health 2009;13:411-17.
76. Khan AA, Perlstein I, Krishna R. The use of clinical utility assessments in early clinical development AAPS J 2009;11:33-8.
77. Korsan B, Dykstra K, Pullman W. Transparent trade-offs: A Clinical Utility Index (CUI) openly evaluates a product's attributes and chance of success, Pharmaceutical Executive, 2005.
78. Ouellet D. Benefit-risk assessment: the use of Clinical Utility Index Expert Opin Drug Saf 2010;9:289-300.
79. Ouellet D, Werth J, Parekh N The use of a Clinical Utility Index to compare insomnia compounds: A quantitative basis for benefit-risk assessment Clin Pharmacol Ther 2008;85:277-282.
80. Lee A, Joynt G, Ho A. "Tips for Teachers of Evidence-based Medicine: Making Sense of Decision Analysis Using a Decision Tree," J Gen Intern Med 2009;24:642-48.
81. Blower PE, Cross KP. Decision tree methods in pharmaceutical research Curr Top Med Chem 2006;6:31-39.
82. Hay J. Evaluation and review of pharmacoeconomic models Expert Opin Pharmacother 2004;5:1867-80.
83. Walker S, Philips L, Cone M. Benefit-risk assessment model for medicines: Developing a structured approach to decision making, CMR International Institute for Regulatory Science, Washington D.C., 2006.
84. Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: The role of sensitivity analysis Health Econ 1994;3:95-104.
85. Brandeau ML. Modeling complex medical decision problems with the Archimedes model Ann Intern Med 2005;143:303-4.
86. Towse A. Net clinical benefit: The art and science of jointly estimating benefits and risks of medical treatment Value Health 2010;13:S30-S32.
87. Baltussen R, Niessen L. Priority setting of health interventions: The need for multi-criteria decision analysis Cost Eff Resour Alloc 2006;4;14.
88. Phillips L, Bana e Costa C. "Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing," Annals of Operations Research 2007;154:51-68.
89. Mussen F, Salek S, Walker S. A quantitative approach to benefit-risk assessment of medicines-part 1: The development of a new model using multi-criteria decision analysis; part 2: The practical application of a new model Pharmacoepidemiol Drug Saf 2007;16;S42-6.
90. Tervonen T, Van Valkenhoef G, Buskens E. A stochastic multicriteria model for evidence-based decision making in drug benefit-risk analysis Stat Med 2011;30:1419-28.
91. Coplan PM, Noel RA, Levitan BS. Development of a framework for enhancing the transparency, reproducibility and communication of the benefit-risk balance of medicines Clin Pharmacol Ther 2011;89:312-5.
92. Levitan B, Andrews E, Gilsenan A. Application of the BRAT framework to case studies: Observations and Insights Clin Pharmacol Ther 2011;89:217-24.
93. Brass EP, Lofstedt R, Renn O. Improving the decision-making process for nonprescription drugs: A framework for benefit-risk assessment Clin Pharmacol Ther 2011;90:791-3.
94. Hleb Babič Š, Kokol P, Podgorelec V. The art of building decision trees J Med Syst 2000;24;43-52.
95. Zorman M, Štiglic MM, Kokol P. The limitations of decision trees and automatic learning in real world medical decision making J Med Syst 1997;21:403-15.
96. Simon LS, Strand CV, Boers M. How to ascertain drug safety in the context of benefit. Controversies and concerns Journal Rheumatol 2009;36:2114-21.
97. Tervonen T, Figueira JR. A survey on stochastic multicriteria acceptability analysis methods," *Journal of Multi-Criteria Decision Analysis* 2008;15:1-14.
98. Baltussen R, Youngkong S, Paolucci F. Multi-criteria decision analysis to prioritize health interventions: Capitalizing on first experiences Health Policy 2010;96:262-4.
99. Smith DM, Brown SL, Ubel PA. Are subjective well-being measures any better than decision utility measures? Health Econ Policy Law 2008;3:85-91.
100. Col NF. Challenges in translating research into practice J Women's Health 2005;14:87-95.
101. O'Neill RT. "A perspective on characterizing benefits and risks derived from clinical trials: Can we do more?," Drug Information Journal 2008;42:235-45.
102. Waller PC, Evans SJW. A model for the future conduct of pharmacovigilance Pharmacoepidemiol Drug Saf 2003;12:17-29.
103. Green C. On the societal value of health care: What do we know about the Person Trade-Off technique? Health Econ 2001;10:233-43.
104. Califf RM. CERTs Benefit Assessment Workshop Participants. Benefit assessment of therapeutic products: The centers for education and research on therapeutics Pharmacoepidemiol Drug Saf 2007;16:5-16.
105. Edwards IR, Biriell C. "WHO programme–Global monitoring," Pharmacovigilance, R. D. Mann and E. B. Andrews, eds., West Sussex John Wiley and Sons Ltd. 2007;pp. 151-66.
106. Sederer LI, Dickey B, Eisen SV. "Assessing outcomes in clinical practice," *Psychiatric Quarterly* 1997;68:311-25.
107. Walker S, McAuslane N, Liberti L Measuring benefit and balancing risk: Strategies for the benefit-risk assessment of new medicines in a risk-averse environment Clin Pharmacol Ther 2009;85:241-6.
108. WHO. Patient safety: Rapid assessment methods for estimating hazards, World Health Organization, Geneva, 2003.
109. Quartey G, Wang J. Statistical aspects in comparative benefit-risk assessment: Challenges and opportunities for pharmaceutical statisticians Pharm Stat 2012;11:82-5.
110. Honig P. Benefit and risk assessment in drug development and utilization: A role for clinical pharmacology Clin Pharmacol Ther 2007;82:109-12.
111. Garattini S. Evaluation of benefit-risk Pharmacoeconomics 2010;28:981-6.
112. IOM. Understanding the benefits risks of pharmaceuticals: Workshop summary Institute of Medicine of The National Academies Washington D.C, 2007.
113. Brinsmead R, Hill S. Use of pharmacoeconomics in prescribing research. Part 4: Is cost-utility analysis a useful tool? J Clin Pharm Ther 2003;28:339-46.
114. Luteijn JM, White BC, Gunnlaugsdóttir H. State of the art in benefit–risk analysis: Medicines Food Chem Toxicol 2012;50:26-32.
115. MHRA. Forum on benefit: risk decision analysis: Summary of discussions and recommendations, Ministerial Industry Strategy Group (MISG), 2008.
116. Stang PE, Pham SV, Kinchen K. The Identification of benefit in medical intervention: An overview and suggestions for process Am J Ther 2008;15:495-03.
117. McNamee D. Communicating drug-safety information Lancet 1997;350:1646-46.
118. Schosser R. Risk/benefit evaluation of drugs: The role of the pharmaceutical industry in Germany Eur Surg Res 2002;34:203-7.
119. Evans S. "Special section: Benefit: risk evaluation in clinical trials," *Drug Information Journal* 2008;42:221-22.
120. Clouse J, Gagnon JP, Boyer G Panel 5: Application of healthcare intervention economic evaluations in healthcare decision-making Value Health 1999;2:92-98.
121. Lopert R, Lang DL, Hill SR. Use of pharmacoeconomics in prescribing research. Part 3: Cost-effectiveness analysis–A technique for decision-making at the margin J Clin Pharm Ther 2003;28:243-49.
122. Ernst E, Resch KL. Risk-benefit ratio or risk-benefit nonsense? J Clin Epidemiol 1996;49;1203-4.
123. Brizmohun N. Standardising benefit:risk assessment: Heads DIA EuroMeeting news, RAJ Pharma, Monaco, 2010.
124. Schiller LR, Johnson DA. Balancing drug risk and benefit: Toward refining the process of FDA decisions affecting patient care Am J Gastroenterol 2008;103:815-19.

125. Committee on the Assessment of the US Drug Safety System: Board on Population Health and Public Health Practice, The future of drug safety: Promoting and protecting the health of the public, Institute of Medicine of The National Academies, Washington, DC, 2007.
126. Strom BL. Risk assessment of drugs, biologics and therapeutic devices: Present and future issues Pharmacoepidemiol Drug Saf 2003;12:653-62.
127. Stricker BH, Psaty BM. Detection, verification, and quantification of adverse drug reactions BMJ 2004;329:44-47.
128. Hirst C, Cook S, Dai W. A call for international harmonization in therapeutic risk management Pharmacoepidemiol Drug Saf 2006;15:839-49.
129. Greenhalgh T, Kostopoulou O, Harries C. Making decisions about benefits and harms of medicines BMJ 2004;329:47-50.
130. Yuan Z, Levitan B, Berlin JA. Benefit-risk assessment: to quantify or not to quantify, that is the question Pharmacoepidemiol Drug Saf 2011;20:653-56.
131. Klepser DG. Pitfalls associated with commonly used methods for pharmacoeconomic analyses Pharmacotherapy 2002;22:35S-38S.
132. Berger ML, Teutsch S. Cost-effectiveness analysis: From science to application Med Care 2005;43;(7 Suppl). 49-53.
133. Drummond M, Sculpher M. Common methodological flaws in economic evaluations Med Care 2005;43:(7 Suppl). 5-14.
134. IOM. Valuing health for regulatory cost-effectiveness analysis, Institute of Medicine of The National Academies, Washington, D.C, 2006.
135. Warren JM. Rationing health care resources. Is the quality-adjusted life-year a helpful guide? Canadian Family Physician. 1994;40:123-28.
136. Murray CJL, Evans DB, Acharya A Development of WHO guidelines on generalized cost-effectiveness analysis Health Econ 2000;9:235-51.
137. Bojke L, Claxton K, Palmer S Defining and characterising structural uncertainty in decision analytic models, Centre for Health Economics, University of York, York, 2006.
138. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence (NICE) Lancet 2002;360:711-15.
139. Lofstedt RE. Academic analysis of the Institute of Medicine report: The future of drug safety Expert Rev Clin Pharmacol 2008;1:617-25.
140. Moore RA, Derry S, McQuay H What do we know about communicating risk? A brief review and suggestion for contextualising serious, but rare, risk, and the example of cox-2 selective and non-selective NSAIDs Arthritis Res Ther 2008;1;R20.
141. Avorn J, Shrank WH. Communicating drug benefits and risks effectively: There must be a better way Ann Intern Med 2009;150:563-4.
142. Edwards IR, Hugman B. The challenge of effectively communicating risk-benefit information Drug Saf 1997;17:216-27.
143. Kramer JM. Managing the risks of therapeutic products: Proceedings of a workshop Pharmacoepidemiol Drug Saf 2005;14:619-28.
144. Liberti L, McAuslane N, Walker SR. Progress on the development of a benefit/risk framework for evaluating medicines, Regulatory Affairs Professionals Society (RAPS), 2010.
145. Breckenridge A. For the good of the patient: Risks and benefits of medicines Pharmacoepidemiol Drug Saf 2003;12:145-50.
146. Hazel T. Communicating the benefits, harms and risks of medical interventions: In journals; to patients and public Int J Surg 2009;7:3-6.
147. Berry DC. Informing People about the Risks and Benefits of Medicines: Implications for the Safe and Effective Use of Medicinal Products Curr Drug Saf 2006;1:121-26.
148. Halvorsen PA, Selmer R, Kristiansen IS. Different ways to describe the benefits of risk-reducing treatments: A randomized trial Ann Intern Med 2007;146:848-56.
149. Peters E, Dieckmann N, Dixon A Less is more in presenting quality information to consumers Med Care Res Rev 2007;64:169-90.
150. Shortell SM, Singer SJ. Improving patient safety by taking systems seriously JAMA 2008;299:445-7.
151. Neumann PJ, Goldie SJ, Weinstein MC. Preference-based measures in economic evaluation in health care Ann Rev Public Health 2000;21:587-61.
152. Carr AJ, Higginson IJ. Are quality of life measures patient centred? BMJ 2001;322:1357-60.
153. Härmark L, van Grootheest A. Pharmacovigilance: Methods, recent developments and future perspectives Euro J Clin Pharmacol 2008;64:743-52.
154. Fitzpatrick R, Davey C, Buxton MJ. Evaluating patient-based outcome measures for use in clinical trials Health Technol Assess 1998;2:1-74.
155. Walker S, Liberti L, McAuslane N. Refining the benefit-risk framework for the assessment of medicines: Valuing and weighting benefit and risk parameters Clin Pharmacol Ther 2011;89:179-82.
156. Kerba M. "Assessing health-related quality of life," *University of Toronto Medical Journal* 2001;78:3196-99.
157. Peters E, Hibbard J, Slovic P Numeracy skill and the communication, comprehension, and use of risk-benefit information Health Aff (Millwood) 26:741-48.
158. Eichler HG, Abadie E, Raine JM. Safe drugs and the cost of good intentions N Engl J Med 2009;360:1378-80.
159. Larson JS. "Two scales for measuring international health status," Evaluation and the Health Professions 1991;14:422-37.
160. Lock K. Health impact assessment BMJ 2000;320:1395-98.
161. Larson JS. The conceptualization of health Med Care Res Rev 1999;56:123-36.
162. Jadad AR, O'Grady L. How should health be defined? BMJ 2008;337:1363-64.
163. Locker D, Gibson B. The concept of positive health: A review and commentary on its application in oral health research Community Dent Oral Epidemiol 2006;34:161-73.
164. Testa MA, Simonson DC. Assessment of quality-of-life outcomes N Engl J Med 1996;334:835-40.
165. WHO. Basic documents: Constitution of the World Health Organization. 45th Geneva; World Health Organisation 2006.
166. Yfantopoulos J. "Quality of life and QALYs in the measurement of health," *Archives of Hellenic Medicine* 2001;18:114-30.
167. Saracci R. The world health organisation needs to reconsider its definition of health BMJ 1997;314:1409-10.
168. Galloway S, Bell D, Hamilton C. Quality of life and well-being: Measuring the benefits of culture and sport: Literature review and thinkpiece, Scottish Executive Social Research, 2006.
169. Chassany O, Sagnier P, Marquis P "Patient-reported outcomes: The example of health-related quality of life–a European guidance document for the improved integration of health-related quality of life assessment in the drug regulatory process," *Drug Information Journal* 2002;36:1209-38.
170. Cummins RA, Lau ALD, Stokes M. HRQOL and subjective well-being: Noncomplementary forms of outcome measurement Expert Rev Pharmacoecon Outcomes Res 2004;4:413-20.
171. Ubel PA, Loewenstein G, Jepson C. Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public Qual Life Res 2003;12:599-07.
172. Kind P, Gudex CM. Measuring health status in the community: A comparison of methods J Epidemiol Community Health 1994;48:86-91.
173. Sajid MS, Tonsi A, Baig MK. "Health-related quality of life measurement," *International Journal of Health Care Quality Assurance* 2008;21;4:365-73.
174. Nease RFJ. Challenges in the validation of preference-based measures of health-related quality of life Med Care 2000;38;(9 Suppl):I155-9.
175. Meyboom R, Egberts A. Comparing therapeutic benefit and risk Therapie 1999;54:29-34.
176. Wang H, Shieh C. Concept analysis of well-being Kaohsiung J Med Sci 2001:17286-93.
177. Pollard EL, Lee PD. "Child well-being: A systematic review of the literature," *Social Indicators Research* 2003;61:59-78.
178. Ersser S. Reflections on wellbeing, quality of life and their significance Perspect Public Health 2010;130;256.
179. Kahn RL, Juster FT. "Well-being: Concepts and measures," *Journal of Social Issues* 2002;58;4:627-44.
180. Svensson O, Hallberg LR-M. HHunting for health, well-being, and quality of life. International Journal of Qualitative Studies on Health and Well-being. 2011;6: 10.3402/qhw.v6i2.7137.