The Assessment of Patient Clinical Outcome: Challenges, Methods, and Conceptual Limitations

Mou’ath Hourani and Qusai Shambour*

Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan
Corresponding Email: Q.Shambour@ammanu.edu.jo

ABSTRACT

The patient clinical outcome assessment is a very important factor in order to guarantee the patient safety. It is the assessment of patient outcome in respect of mental status, symptoms, or impact of ill health on how the patient functions. This assessment aims to improve the quality of health care in general by predicting the performance of any therapeutic option in terms of benefits or harms. Due to the importance of this field, an in-depth review of current issues regarding the assessment of patient clinical outcome should be carried out. Accordingly, this paper go over three main points: 1) the challenges involved in the assessment of patient clinical outcome; 2) a number of the existing methods for patient clinical outcome assessment; and 3) a discussion of the general conceptual limitations and difficulties of patient outcome quantitative assessment. This paper will advance the understanding of the assessment of patient clinical outcome field in regards to its challenges, methods and conceptual limitations.

Key-words: Assessment; Patient; Clinical; Outcome; Methods.

INTRODUCTION

The need for quantitative assessment of patient clinical outcome is compelling, urgent [1], important [2], vital [3] and necessary [4] with high research priority [5]. It has considerable public attention [6-8]. The optimum solution is to develop a model which generates data in a consistent format and could be universally applied [9]. The global approach is important to compare the benefits of different patient groups across different health care interventions [10, 11] and to prioritize health interventions [12].

There is a need for developing outcome measures, which are appropriate to therapists and patients across different specialties and which are simple to use in daily routine practice [13]. If the model is plausible, it will make a behavioral change [14] and it will lead to a greater shared evidence for what really works in clinical practice [15]. The evidence-based results of randomized clinical trials cannot be assessed if applied in everyday care without a routine measurement of outcomes in daily practice too [16]. Measuring and reporting on the health of populations or individuals requires a valid, reliable and comparable way to measure health status and perform the clinical assessment [17-19] which is considered as the challenge of future [20-22].

Academia should play a central role diligently in the creation of new and not biased instrument for key health indicators [23-25]. There is an emphasis on the development of theoretical models, which incorporate more quantitative and comprehensive assessment of outcomes [26, 27]. Stakeholders and research are needed to construct a global, standardized method with considerable effort [10, 28, 29].

Accordingly, this paper proposes an in-depth review of the literature to explore different issues regarding the assessment of patient clinical outcome. This paper presents: 1) an overview of the challenges involved in the assessment of patient clinical outcome; 2) an overview of the existing assessment methods for patient clinical outcome; and 3) a discussion of the general conceptual limitations and difficulties of patient outcome quantitative assessment.
1. Overview of the Challenges Involved in the Assessment of Patient Clinical Outcome

The assessment of benefits, harms, risk, and patient clinical outcome and their communication is a complex scientific challenge [30-32], and formidable problem [33]. It is remote to achieve [34], and difficult to quantify on a common scale [4, 35-38]. It becomes more difficult when comparing drugs with other options like surgery with different measurements and scales [4]. When sometimes tens of adverse effects recorded for some drugs, combining them into a useful benefit-harm assessment seems impossible, and making it challengeable to determine the best therapy [39, 40]. Comparing different therapy regimens is still challenging [41]. The definition of clinically important components and their weighting scheme across different team members may be difficult to achieve [32]. The assignment of value and weighting to every component is still challengeable [42, 43]. Additionally, There is considerable difficulty in weighing up benefits and harms over the short and long term [44]. Also weighing and incorporating patients’ preferences in the assessment is one of the tremendous fundamental intellectual challenges [21, 30, 45, 46], and still a key question in the field [9].

Other main challenges include how quantitatively representing a complex drug benefit-risk profile, quantitatively characterizing drug benefit-risk for individuals who are different in physiology and preferences, updating benefit-risk assessments with new information through the drug life cycle, addressing the uncertainty in benefit-risk assessment, addressing the cost of adopting a quantitative framework, and effectively presenting and communicating quantitative information [30]. This process is a prominent challenge for all sectors of health care [47], and considered as the most difficult step in the approval process [48, 49].

Pessimistic thinking is observed in literature by using words like “impossible” [50, 51], “not possible” [8, 52] and “unreasonable” [53] to construct benefits and harms arithmetic ratio applicable for all cases. It is believed that it is not expected to replace expert judgment [54-56] for the foreseeable future [9]. Sometimes it is believed that we will ever have a perfect method which is free of limitations for comparing health states [57, 58].

The construction of proper tool for benefits, harms, and risk assessment is an outstanding challenge [59, 60]. It is considered one of the questionable and unresolved issues in the field [4, 28, 40, 53, 56, 61, 62]. It requires much critical thought [6, 37], and needs first-rate minds from the world of clinical medicine [59] with considerable skill and effort [63]. It needs organizing workshops with specialists from different fields or supporting specific research projects like post doctorate research [50, 54, 56, 62, 64].

2. Overview of the Existing Methods for the Assessment of Patient Clinical Outcome

2.1. Mortality, Morbidity Rates and Indexes of Other Parameters

There are a lot of previous attempts to estimate health status by different parameters like mortality, morbidity, life expectancy, mental hospital admissions, the prevalence of states of coma, unconsciousness in an institutional setting, death from suicide rate, infant mortality, or combinations of them [65]. An example is a scale composed from infant mortality, life expectancy, the literacy rate, death rate for persons aged 65 and older, and mental hospital admissions [65]. Many of these parameters are not directly related to health, others reflect only specific areas related to health [65]. Also, they don’t reflect the complex conception of health [65].

2.2. Number Needed to Treat (NNT)

NNT is defined as the average number of patients needed to be treated over a defined time to achieve the required outcome in one of them [66]. Physicians are widely using this method because it is relatively simple and easy to comprehend [67, 68].

\[
NNT = \frac{1}{P_1 - P_2}
\]

Where \(P_1\) is the proportion of the disease of interest in the control group, and \(P_2\) the proportion of disease in the treatment group [69].

2.3. Number Needed to Harm (NNH)

The number needed to harm (NNH) is the number of patients who need to receive an intervention to cause one additional adverse event; the NNH is the inverse of the absolute difference in adverse event rates between the experimental and control groups [70].

\[
NNH = \frac{1}{R_1 - R_2}
\]

Where \(R_1\) is the risk of an adverse event of interest in the untreated group, and \(R_2\) is the risk of an adverse event of interest in the treated group [70].
2.4. Unqualified Success NNT and Unmitigated Failure NNH (NNT_{US} and NNH_{UF})

NNT had been extended to include treatment with adverse effects. NNT for unqualified success (NNT_{US}) is the number of patients who must be observed on average to encounter one successfully treated patient who did not suffer adverse events due to the treatment [71, 72]. NNH for unmitigated failure (NNH_{UF}) is the number of patients on average will suffer an adverse, treatment-related effect without benefiting from the therapy [72].

\[
NNT_{US} = \frac{1}{(P_1-P_2)[1-(F_1-F_2)]}
\]

\[
NNH_{UF} = \frac{1}{P_2(F_1-F_2)}
\]

Where \(P_1\) is the proportion of the disease of interest in the control group, \(P_2\) the proportion of disease in the treatment group, \(F_1\) is the frequency of the adverse event in the treated group, and \(F_2\) is the frequency of this event in a control or untreated group [71, 72].

2.5. Disease Impact Number and Population Impact Number

NNT could be extended to calculate disease impact number which is defined as “the number of patients with the disease in question among whom one event will be prevented by the intervention” [73]; it’s formula is 1/(absolute risk reduction × proportion of people with the disease who are exposed to the intervention) [73].

2.6. Relative-Value Adjusted Number-Needed-To-Harm (RV-NNH)

In RV-NNH method, harms value could be estimated by utility which is a numeric representation of patients’ preferences for health states or desirability for specific outcomes [74-76] by adding the relative utility value (RV) into the NNH calculation [74]; RV can be calculated as:

\[
RV = \frac{1-\text{utility of AE}}{1-\text{utility of disease of interest}}
\]

NNH adjusted for relative value can then be calculated [74] as:

\[
RV\text{-NNH} = \frac{1}{(F_2-F_1) \times RV}
\]

RV-NNH also is accommodated to include the harms of multiple adverse events [74]. Utility could be measured using some utility methods (standard-gamble method or the time-trade-off) [74, 77] which is discussed later in this chapter. However, the NNH method has some merits for clinicians because it is simple to use [54]. This method considers both the benefit and harm of the therapeutic intervention.

2.7. Likelihood of Being Helped vs. Harmed (LHH)

This method combines NNT and NNH into a single ratio in the following formula:

\[
LHH = \frac{1}{NNT} : \frac{1}{NNH}
\]

The resulting number is no of times more likely to benefit from treatment option than to be harmed [70].

2.8. The Adjusted Number Needed To Treat

This method incorporates qualities and timings of the potential outcomes of the therapeutic option(s) to NNT Method [78]. First, NNT and NNH are combined to calculate a number needed to treat that include the probabilities of both benefits and harms resulting from the two options as:

\[
\text{Adjusted NNT} = \frac{1}{(B_2-B_1)-(H_2-H_1)}
\]
Where $B_2$ is the probability of benefit under option 2, $B_1$ is the probability of benefit under option 1, $H_2$ is the probability of harm under option 2, and $H_1$ is the probability of harm under option 1. Then, qualities and timings are included in the formula as follows:

\[
1 = \frac{\left( \text{LEB}_2 \times B_2 \times \text{UB}_2 \right) - \left( \text{LEB}_1 \times B_1 \times \text{UB}_1 \right)}{\left( \text{LEH}_2 \times H_2 \times \text{UB}_2 \right) - \left( \text{LEH}_1 \times H_1 \times \text{UB}_1 \right)}
\]

Where: $\text{LEB}_2$ is the time-discounted life expectancy of the average individual receiving the benefit under option 2. $B_2$ is the probability of benefit under option 2. $\text{UB}_2$ is the average utility change for an individual receiving the benefit under option 2. $\text{LEB}_1$ is the time-discounted life expectancy of the average individual receiving the benefit under option 1. $B_1$ is the probability of benefit under option 1. $\text{UB}_1$ is the average utility change for an individual receiving the benefit under option 1. $\text{LEH}_2$ is the time-discounted life expectancy of the average individual receiving the harm under option 2. $H_2$ is the probability of harm under option 2. $\text{UB}_2$ is the average utility change for an individual receiving the harm under option 2. $\text{LEH}_1$ is the time-discounted life expectancy of the average individual receiving the harm under option 1. $H_1$ is the probability of harm under option 1. $\text{UB}_1$ is the average utility change for an individual receiving the harm under option 1. Life expectancy represents timing, and utility represents quality of life. The concept of QALYs is used to estimate utility. A sensitivity analysis could be used to assess the impact of uncertainty on the decision, and to detect the variables which are more relevant to it [78].

2.9. **Survey-Based Health-Related Quality of life (HRQOL) Instruments**

Health-Related Quality of life (HRQOL) is the individual’s subjective perception of the impact on health status [79], their perceived need for health care, and their preferences for therapy and outcomes [80]. It cannot be measured directly; instead of that, it can be measured indirectly by asking a series of questions known as items to a respondent [81, 82]; respondent ranks the items to give their weights using techniques such as visual analog, standard gamble, and/or time trade off [82-84]. Aggregating the scored items is performed to derive scale scores [81, 82, 85] using some mathematical formula, additive or multiplicative [85]. Validation of the scale is then performed in a large number of patients [82]. Those items are set in a questionnaire and administered by an interviewer, face-to-face or completed by the patient himself [82, 86, 87].

Instrument’s items cover multidimensional outcomes which may include global well-being, symptoms, economic welfare, characteristics of community and environment, physical functioning, social functioning, cognitive functioning, mental health, general health perceptions, vitality, and pain [63, 79, 82, 84, 88-93]. Respondent are patients, clinicians, or family members, and some instruments have more than one respondent [63].

Hundreds of different survey-based instruments for measuring HRQOL are available [83, 94-96]. No gold standard for HRQOL exists [86, 90]. No specific instrument can be used to measure all purposes, setting or population [63, 97, 98]. There is no ‘worst’ or ‘best’ performing instruments [82, 83]. The decision to choose an instrument over another, to choose a combination(s), or to choose the type of instrument is driven by the purpose of the measurement, the characteristics of the population and the environment [83].

Instruments are different in construction methodology, the questions asked to respondents, their intended focus, the robustness of the results [89, 96], their validity and reliability [89, 99], defining health, and clarifying the purposes of the instrument [99].

Instruments may enhance physician-patient communication, and facilitate important discussions by focusing on patients with health status impairment [100]. They also may increase efficiency by allocating most limited resources. Good tools can also help managing time spent in collecting needed information from patients by asking the “right” questions [100].

Examples of instrument are Dartmouth Primary care Cooperative Information Project (COOP) Charts, Duke Health Profile, EuroQol Instrument (EQ-5D), General Health Questionnaire, Health Assessment Questionnaire, Health Utility Index, Nottingham Health Profile (NHP), Quality of Well-Being (QWB) Scale, Short Form 36 (SF-36), Sickness Impact Profile (SIP) [90, 94, 101], the Asthma Quality of Life Questionnaire, the Oxford Hip Score, the Shoulder Disability Questionnaire, and McGill Pain Questionnaire [96].

2.10. **Stated Preferences or Preference-Based (Utility) Approach for Assessing Health-Related Quality of Life**

Health-related quality of life is relating to the health domain of the individual’s existence. Utility concept which is defined as “the numeric representation of patients’ preferences for specific outcomes” [74] could be used to measure HRQOL. The utility approach is measuring a single cardinal value, usually between zero and one, which reflects the
health-related quality of life of the individual at a specific point in time where zero is death and one represents perfect health state [79, 83, 86, 102-104]. The measured value represents the difference between the gains from the treatment and the burdens of side effects [86]. In this approach, there is no criteria for demarcation [59]. Patients usually evaluate their HRQOL using pair wise comparisons, rating scales, time trade off, and standard gamble measurement techniques [86, 103, 105-107].

In Pair wise Comparison, patients are asked to rate a series of pairs of health states whether one is worse and by how many times is worse, or whether both are equal in severity [106-108]. Internal consistency in this method can be calculated, and it can assess the quality of each respondent’s performance and the extent of agreement between individuals [108].

Rating scales directly determine a given health state for respondents on a scale [104] which can be a scale from 0 to 100 in which 0 represents “worse imaginable health state” and 100 represents “best imaginable health state” in Visual Analogue Scale (VAS) method, or a scale of nine degrees grading from “worst imaginable health state” through “best imaginable health state” [107-110]. Other expressions for health like magnificent, excellent, good, fair, poor, and terrible are also used for such scales [59, 108]. The patient is asked to rate his or her health using those scales.

Time trade-off (TTO) technique presumes that shorter life with a satisfactory state of health is more preferred by patients than longer life with a considerable handicap or serious discomfort [82, 89, 111]. It is performed by asking the patient to choose between two options; the first option is to live with specific health state for the rest of patient’s life (t) years, and the second option is to live in full health for (h) years, followed by death where (h < t). TTO value then calculated as h/t at the point where the patient has no distinction between the two options [104, 109, 110, 112, 113].

Standard gamble (SG) technique is performed by asking the patient to choose between two options. The first option is to live with specific health state for the rest of patient’s life (t) years, and the second option is to receive a therapy with a probability (p) of reviving to full health or immediate death with the probability (1-p). SG value is (p) at the point where the patient has no distinction between the two options [104, 108-110, 113].

2.11. Stochastic Multi Objective Acceptability Analysis (SMAA)

Stochastic multi criteria acceptability analysis (SMAA) is a method, which handles problems with inaccurate, uncertain, or lack of preference information [114-116]. It can be used also in the case of lack of measures of the criteria, and the presence of several decision makers [114]. In general, the decision makers prefer methods, which describe the potential decisions and their consequences in an appropriate form rather than methods, which require them to express their preferences explicitly [114, 117]. In SMAA, the decision makers do not need to express their preferences implicitly or explicitly [116, 117] because the technique can be used with or without preference information [114].

SMAA is a multi criteria decision support method, which explore what are the criteria valuations that will make each alternative the preferred one [115, 117]. The method produces an acceptability index for each alternative to support specific alternative to be the preferred one [114, 115, 117]. SMAA technique produces weight combinations for criteria that support the preference of specific alternative [114, 115, 117]. In addition, the method measures a confidence factor for each alternative, which indicates whether the input data is accurate adequately for making an informed decision [115, 117]. The core of the technique is the using of multidimensional integrals, which is impossible to compute analytically. So the computations can be implemented using Monte Carlo simulation [115, 116]. The results and ranking of alternatives are then presented to the decision makers for their final evaluation [117].

3. Discussion of the General Conceptual Limitations and Difficulties of Patient Outcome Quantitative Assessment

When constructing a measure, value judgments assign weights to outcome criteria [4, 37, 118]; such judgments suffering different types of bias [119], and can be highly subjective and different between health professional and patients [4, 120]. Different methods produce different health state values and cannot always be compared to each other’s [5, 20, 43, 121-123]. Consequently, the choice of method for health state assessment directly affects the estimated cost-effectiveness of interventions [98, 124] and the lack of a standardized approach limits the interpretability of cost-effectiveness analysis [104, 125]. Methods do not reflect the intellectual process of balancing harms and benefits [54], and allow a very crude assessment [126]. They begin with inherited imbalance assessment of benefits and harms simply from the metric used [7]. Many tools ignore patients’ choices [127, 128].
A convincing method of patient outcome quantitative assessment that considers a drug’s benefits, adverse events, patient health preferences and the natural history of the condition is not presented yet [74]. Models do not consider many of the previous benefits, harms and risk criteria already identified [126]. Many criteria in the models are not well defined [126, 129] while real complex variables are affecting outcome [63].

There is no agreement exists about what is meant by health and health status [59]. Also there is a confusion between health status and quality of life which created technical, conceptual, and ethical problems [89]. There is a difficulty in operationalizing outcome’s measurement [63]. Some current methods are complex [130]. Some methods can’t handle multiple benefit, harm and risk outcomes of different severity [6]. Quantitative methods have a margin of error, and this uncertainty should be kept in mind [56]. The clinical methods derived from biomedical, psychosocial, or mathematical methods are inadequate [59]. Current methods aid in the process of decision-making, and cannot substitute the existing decision making process [53, 77, 131-133].

There is no consensus how to define a clear gold standard for patient outcome assessment [77]. Current methods are not universally accepted nor systematically adopted by the pharmaceutical industry or regulatory agencies [6, 36, 77, 130, 134]. There is also a difficulty appears when the same clinical situation needs different assessment methods for different purposes and medical settings [59]. Finally, it is observed from literature that the term balancing benefits and risks means a pure rational judgment whether or not the harms outweigh the benefits [35]. It is also observed that a lot of studies that record information about benefits and harms do not use the same metric for recording [2].

**CONCLUSION**

The patient clinical outcome assessment is about measuring various aspects of the patient’s health status by recording clinical outcomes resultant from medical treatments and interventions to observe their effectiveness [135,136]. A number of patient clinical outcome assessment methods have been developed to measure the effectiveness of therapeutics and drugs in terms of their benefits and harms. Throughout this literature, challenges involved in the assessment of patient clinical outcome, existing assessment methods for patient clinical outcome, and general conceptual limitations and difficulties of patient outcome quantitative assessment have been reported.

**REFERENCES**


[34] Wade G. Outcome and health status measurement. Physiotherapy. 1999;85:231-.


