VALUE IN HEALTH 17 (2014) 275–279

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Quality-of-Life–Adjusted Hazard of Death: A Formulation of the Quality-Adjusted Life-Years Model of Use in Benefit-Risk Assessment

*Alberto Garcia-Hernandez, MSc**

Astellas Pharma Global Development Europe, Sylviusweg 62, Leiden, The Netherlands

ABSTRACT

Background: Although the quality-adjusted life-years (QALY) model is standard in health technology assessment, quantitative methods are less frequent but increasingly used for benefit-risk assessment (BRA) at earlier stages of drug development. A frequent challenge when implementing metrics for BRA is to weigh the importance of effects on a chronic condition against the risk of severe events during the trial. The lifetime component of the QALY model has a counterpart in the BRA context, namely, the risk of dying during the study. **Methods:** A new concept is presented, the hazard of death function that a subject is willing to accept instead of the baseline hazard to improve his or her chronic health status, which we have called the quality-of-life–adjusted hazard of death. **Results:** It has been proven that if assumptions of the linear QALY model hold, the excess mortality rate tolerated by a subject for a chronic health improvement is inversely proportional to the mean residual life. **Conclusions:** This result leads to a new representation of the linear QALY model in terms of hazard rate functions and allows utilities obtained by using standard methods involving trade-offs of life duration to be translated into thresholds of tolerated mortality risk during a short period of time, thereby avoiding direct trade-offs using small probabilities of events during the study, which is known to lead to bias and variability.

Keywords: benefit-risk assessment, hazard function, mean residual life, quality-adjusted life-years, tolerated risk.

Introduction

The concept of quality-adjusted life-years (QALY) has been routinely used to guide health care policymaking since its inception some 30 years ago. QALYs provide a very intuitive way of combining the two main components of health, namely, life duration and quality of life (QOL), into a single index. In its simplest form, which is linear with respect to time, the QALY model is formulated as QALY $(T,Q) = T \times g(Q)$, where T is the life duration and g is a utility function over the health states.

More recently, regulatory agencies, pharmaceutical companies, and other groups have begun to discuss and explore methods to improve and standardize the benefit-risk assessment (BRA) performed throughout all drug development phases and during the assessment of a marketing authorization application. In this context, three detailed reviews of quantitative methods with potential use during BRA have been published [1–3]. Some of the models and metrics proposed for BRA, such as relative value-adjusted number needed to treat, global benefit risk, and multi-criteria decision analysis, share the idea of combining risks and benefits into a single index by incorporating patients' or decision makers' preferences [4–6]. These models do not necessarily decompose the subject's outcomes into the two dimensions of the QALY model but are often based on a set of clinical trial end points selected for each evaluation.

Although the Work Package 2 report of the European Medicines Agency Benefit Risk Methodology Project concluded that regulators still find QALY insufficiently comprehensive for drug-related BRA, the existence of aspects common to QALY and other BRA models that deserve further research was acknowledged [2]. Herein, we concentrate on the inverse relationship between the life duration component of the QALY model and the risk of death during a clinical trial and use the hazard function, which is a key concept in time-to-event models, to establish a line of communication between the two frameworks.

Utilities in the QALY Model

In the QALY model, a chronic health state Q_0 is quantified on a 0 to 1 scale, with 1 representing perfect health and 0 representing death. Two common ways to elicit the utility of health status Q_0 are standard gamble (SG) and time trade-off (TTO).

SG is the most classical method for quantifying preferences in decision theory. With the SG probability equivalent (PE) method, decision makers are asked to fix a probability p such that they are

* *Address correspondence to:* Alberto Garcia-Hernandez, Global Data Science, Astellas Pharma Europe B.V., Sylviusweg 62, 2333 BE Leiden, The Netherlands.

E-mail: alberto.garciahernandez@astellas.com.

1098-3015/$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2013.11.013

indifferent between a certain consequence of living the rest of their lives in health state Q_0 and a lottery in which they will live with a better health state Q_1 with a certain probability p or will die immediately with the opposite probability $(1 - p)$. Although Q_1 normally represents perfect health, a more general use that compares two arbitrary states Q_0 and Q_1, the latter of which is preferred to Q_0, is assumed throughout this article. Assuming that the interviewee is a von Neumann–Morgenstern rational agent, the obtained probability p is taken as the utility of Q_0 on the 0 to 1 scale, where 0 is death and 1 is health state Q_1. SG utilities are believed to be upward biased due to loss aversion and probability weighting effects [7,8].

The TTO method asks for the duration (T_1) that yields indifference between living T_0 years in Q_0 status and living T_1 years in another health status Q_1 that is preferred to Q_0. When utility of life duration is assumed to be linear, the utility of health state Q_0 with respect to Q_1 is calculated as $u = T_1/T_0$. Some authors have warned of a possible downward bias of TTO utilities, mainly due to life duration curvature, whereas others have suggested a possible upward bias caused by loss-aversion and scale-compatibility effects [7,9]. It has also been hypothesized that both downward and upward biases might cancel out, thereby possibly explaining empirical evidence suggesting the greater accuracy of TTO when it comes to reflecting preferences over other methods such as SG [7–12].

Weights for Benefit-Risk Metrics

To compare both frameworks, in this section we define a simple weighted-sum multicriteria decision analysis model for BRA over two clinical trial end points, namely, the chronic health state achieved at steady state Q and the risk of death during the study D. If the variable Q has two possible values, Q_0 and Q_1, preferred to Q_0, the model can be expressed by

$$\text{Score} = w\,P(Q_1) - P(D) \tag{1}$$

where $P(Q_1)$ is the probability of achieving health state Q_1 and $P(D)$ is the probability of death during the study. This simple BRA metric requires the fixing of a weight w that represents the importance of a swing from the chronic health state Q_0 to Q_1 with respect to a swing from Q_0 to death during the study period. By using a PE gamble, decision makers might be asked to fix a probability p such that they are indifferent between a certain consequence of a chronic health state Q_0 at steady state in the clinical trial and a lottery in which the patient will achieve health state Q_1 with a certain probability p or will die during that period of time with the opposite probability $(q = 1 - p)$. If they accept a probability of death $q = 0.03$ (3%), then a swing from 0% to 3% in the percentage of subjects dying during the study will be concluded to have the same importance as a swing in the chronic health status from Q_0 to Q_1, leading to a weight $w = 0.03$.

One aspect that deserves attention is the fact that the quantification of preferences for the decision models proposed for BRA have often required trade-offs using small probabilities [13–15]. Individuals, however, are not perfect von Neumann–Morgenstern agents, and PE gambles that handle probabilities close to 0 are known to provide biased and variable quantifications [7,8,16–18]. Herein, we propose to avoid trade-offs that use low probabilities of events during the study and to focus on trade-offs of life duration. This article proposes a procedure to translate these TTO utilities into weights that could be useful for BRA metrics.

The Hazard Function and the Mean Residual Life

The hazard function $h(t)$, also known as failure rate or hazard rate, is a key concept in time-to-event statistics that represents the instantaneous risk of suffering the event of interest at time t for a subject who has survived to that moment in time. The hazard rate quantities are not probabilities but range from 0 to infinite and depend on both the strength of the risk and the time units used. For example, both a mortality hazard rate of 0.015 deaths per subject-month and 18 deaths per 100 subject-years represent the same instantaneous risk at a given time t but expressed in different units. The hazard function $h(t)$ of a random variable T is formally defined as

$$h_T(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \le T < t + \Delta t / T \ge t)}{\Delta t} \tag{2}$$

Another function that deserves special attention here is the mean residual life (MRL), also called expected remaining lifetime. The MRL provides the expected value for the lifetime remaining at any time t, given that the subject is known to have survived to t. Although two different random variables T and T' may share the same expected value $E(T) = E(T')$, the complete MRL(t) function over time uniquely defines the probability distribution of a random variable T. The MRL is given by

$$\text{MRL}_T(t) = \int_t^{\infty} \frac{S_T(x)\,dx}{S_T(t)} \tag{3}$$

where $S_T(t)$ is the survival function, that is, defined as $S_T(t) = P(T > t)$.

In expected utility, a subject is said to be risk-neutral with respect to the remaining life duration if he or she is indifferent over any lotteries with the same expected value. If lotteries over life duration are formulated in terms of hazard functions, risk neutrality over life durations is given by a subject being indifferent to any two hazard functions h_0 and h_1 provided the expected remaining lifetime is the same.

A risk-neutral 30-year-old subject would be indifferent between the two lotteries of lifetime durations L_0 and L_1 associated with the two hazard functions of death h_0 and h_1, illustrated in Figure 1, because both have the same expected value (52.3 years) when the subject is 30 years old. The same two hazard functions can be used to represent not only the lotteries at an initial time $t = 0$ but also the subsequent lotteries $L_0(t)$ and $L_1(t)$ for subjects who survive to later times. If the same risk-neutral subject survives to the age of 70 years and is again asked to choose between these two hazards, he or she will prefer the lottery represented in Figure 1 by the solid line because this function provides a higher MRL at this later age.

The Quality-of-Life–Adjusted Hazard of Death

In survival analysis, the term *baseline hazard* is used to represent the risk of death estimated for a subject or group of similar subjects over time and, as such, related to our knowledge of their diseases and demographic characteristics. The baseline hazard function h_0 is associated with a probability distribution of lifetimes L_0.

A quality-of-life–adjusted hazard of death (QAHD) function is defined as a hazard function of death, $h_1(t)$, that a person is willing to accept at any time t, instead of his or her baseline hazard of death $h_0(t)$, to improve his or her health status from any given level Q_0 to any other level Q_1.

In other words, $h_1(t)$ is the QAHD for an improvement from Q_0 to Q_1 if that person is indifferent between being in health state Q_0, with its baseline lifetime probability distribution L_0, for the rest of his or her life and being in health state Q_1, with distribution L_1, where $h_0(t)$ and $h_1(t)$ are the hazard functions associated with probability distributions L_0 and L_1, respectively. We intentionally restrict the definition of QAHD to lifetime distributions L_1 that meet the condition whereby if a subject survives to an interim