Concept and Scope of Benefit–Risk Evaluation of Medicines

1.1 HISTORICAL BACKGROUND

The regulation of modern medicines started after the thalidomide tragedy, which was undoubtedly the most significant adverse event in pharmaceutical history. It was a tragedy because the toxic effects of the medicine were expressed through the damage to the unborn foetus between the fortieth and fiftieth days of gestation after the mother had taken therapeutic doses of the medicine as a sedative or hypnotic during the pregnancy. The baby was born with characteristic reduction deformities of the limbs with shortening or complete absence of long bones, the hands and feet being attached as 'flippers' or absent altogether (Burley, 1988). The plight of these children, about 1000 who were born in the United Kingdom and several thousands in West Germany, caused widespread horror and emotional reactions and calls for all medicines to be the subject of governmental control and regulation in order to avoid a repetition of such an event.

A previous disaster in 1938, when 107 deaths were caused in the United States by the consumption of an 'elixir' of sulfanilamide containing a toxic solvent, had led to the setting up of new legislation in the United States whereby manufacturers of medicines had to be registered, had to carry out safety tests and were liable to receive factory inspections and seizure of products in violation (Burley, 1988). The thalidomide tragedy led to the strengthening of the regulatory process, particularly in respect of prescription medicines and the introduction of the requirement for pre-marketing submission to the regulatory authorities in the United States as well as in Europe.

Already in 1962, after the thalidomide tragedy had become known, MacGregor and Perry (1962) made the following very wise comments with regard to the impact of thalidomide on regulations worldwide:

> Such hazards can become apparent only on prolonged clinical use of drugs in human patients. Tragedies will no doubt continue to occur with new remedies, and this is part of the price to be paid for therapeutic progress, because we cannot, as a community, ask for new drugs without being prepared to accept such risks. We cannot legislate ourselves out of this dilemma . . .

Since 1962, a significant number of breakthrough medicines have been put on the market in a vast number of therapeutic areas such as hypertension, heart failure, hypercholesterolaemia, schizophrenia, cancer, human immunodeficiency virus (HIV), etc. The regulatory authorities have assessed all these products with regard to their quality, safety and efficacy, but the emphasis was usually put on the safety. Whenever Food and Drug Administration (FDA) Commissioners were called to testify before Congress, the charge was usually 'Why did you put this drug on the market which later turned out to be toxic?' It was not until the acquired immune deficiency syndrome (AIDS) epidemic surfaced that Congress began to ask the FDA why certain medicines were not on the market (Lasagna, 1998). The statement by the FDA (1999b) provides a very good and balanced perspective with regard to the benefits (efficacy) and risks (safety) of medicines: 'Although medical products are required to be safe, safety does not mean zero risk. A safe product is one that has reasonable risks, given the magnitude of the benefit expected and the alternatives available. All participants in the medical product development and delivery system have a role to play in maintaining this benefit–risk balance by making sure that products are developed, tested, manufactured, labelled, prescribed, dispensed, and used in a way that maximizes benefit and minimizes risk.'

The role of regulatory authorities is well defined in this respect, and regulatory authorities approve a new medicine or confirm approval of a marketed medicine when they judge that the benefits of using a medicine outweigh the risks for the intended population and use, and they ensure that the medicine is truthfully and adequately labelled for the population and use (Food and Drug Administration, 1999b). The key question is however *how* the regulatory authorities judge whether the benefits outweigh the risks, or in other words, how is the benefit–risk balance of a medicine established? This basic question triggered this book, also because it is recognized that balancing the benefits and risks is probably one of the most difficult tasks for anyone involved either in the development of new medicines or in the post-approval re-assessment of marketed medicines, and that basic research is still needed in this area (Edwards and Hugman, 1997). The next question is then whether methods, models or other tools have been developed or can be used to aid the regulatory authorities and others in determining the overall benefit–risk balance of

medicines. Since the regulatory authorities are not the only stakeholder with regard to medicines, these questions should be considered in the context of how the other stakeholders (pharmaceutical companies, prescribers, and patients) judge whether the benefits of a medicine outweigh the risks.

An overview on benefit–risk assessment is provided in the next sections of this chapter. Following a review of the regulatory systems for assessing medicines, definitions, views and perceptions of benefits and risks will be outlined. Next, concepts as well as current practices in benefit–risk assessment will be discussed, and finally an overview on the use of methods and models for benefit–risk assessment will be provided.

1.2   THE REGULATORY SYSTEMS FOR ASSESSING MEDICINES

As outlined below, the concept of benefit versus risk is well captured in the EU and US legislation, which provide the legal framework for benefit–risk assessments by the regulatory authorities.

Europe

In the European Union, medicines must meet the three exclusive criteria laid down in the Community law which are quality, safety and efficacy, for a marketing authorization to be granted (Brunet, 1999). In Directive 2001/83/EC on the Community code relating to medicinal products for human use it is stipulated that a marketing authorization shall be refused if: (a) the risk–benefit balance is not considered to be favourable, or (b) its therapeutic efficacy is insufficiently substantiated by the applicant, or (c) its qualitative and quantitative composition is not as declared (Official Journal of the European Communities, 2004). Furthermore, in the preambles to this Directive, the following is stated:

> The concepts of harmfulness and therapeutic efficacy can only be examined in relation to each other and have only a relative significance depending on the progress of scientific knowledge and the use for which the medicinal product is intended. The particulars and documents which must accompany an application for marketing authorization for a medicinal product demonstrate that potential risks are outweighed by the therapeutic efficacy of the product.

As an example of the application of the EU regulations, in the United Kingdom the Committee on Safety of Medicines (CSM) looks for evidence of the medicine's safety, efficacy and quality when considering an application for a product licence. The CSM also uses benefit–risk analyses for medicines already licensed, and they

have the right to change a license to minimize the risks from taking a medicine whilst allowing specific patients continuing benefit (Risk:benefit analysis of drugs in practice, 1995).

United States

The US drug law first embraced the idea of risk versus benefit in 1962. Providing evidence of safety before marketing was first required by the Federal Food, Drug and Cosmetic Act in 1938. However, it was not until the Kefauver–Harris Drug Amendments of 1962 that firms had to show a drug's effectiveness before marketing (FDA, 1999a). Today, the FDA decides before any medicine is marketed whether the studies submitted by the drug's sponsor show it to be safe and effective for its intended use. This decision comes down to two questions: (1) do the results of well-controlled studies provide substantial evidence of effectiveness? (2) do the results show that the product is safe under the condition of use in the proposed labelling (safe, in this context, means that the benefits of the drug appear to outweigh its risks) (FDA, 1999a)?

Japan

Similar to the EU and the United States, in Japan the quality, efficacy and safety of a medicine are assessed by the Ministry of Health and Welfare (Japan Pharmaceutical Manufacturers Association, 2003), and approval is an official confirmation that the medicine is both safe and effective (Danjoh, 1999).

The YUYOSEÏ instrument, which is required in Japanese clinical trials and which is a global assessment given by the physician to reflect an individual's overall response to a treatment, will be discussed later in this chapter.

1.3   BENEFIT–RISK ASSESSMENT: DEFINITIONS

Several authors have defined the words 'benefit' and 'risk', and some have proposed using other terms because in their view the term 'benefit–risk assessment' is incorrect.

According to Webster's Dictionary (1991), 'benefit' originates from the Latin words '*bene factum*' and means 'something that promotes well-being, advantage'. 'Risk' means 'possibility of loss or injury, peril' or 'a dangerous element or factor'. The FDA (2002) defined, in its information to consumers, benefits of medicines as 'helpful effects you get when you use medicines, such as lowering blood pressure, curing infection or relieving pain', and risks as 'the chances that something

unwanted or unexpected could happen to you when you use medicines'. The WHO Collaborating Centre for International Drug Monitoring used similar definitions for benefit ('the proven therapeutic good of a product; should also include the patient's subjective assessment of its effects') and risk ('the probability of harm being caused; the probability (chance, odds) of an occurrence') (The Uppsala Monitoring Centre, 2002a).

An important difference between benefit and risk is that 'benefit' is a quantity (in preventive medicine, the magnitude of the loss or harm averted) whereas 'risk' is a probability (the likelihood of the occurrence of an unfavourable event) (Cheung and Kumana, 2001). In this respect, Cheung and Kumana argued that a minimal benefit can never be attractive, even if there is a 99% chance of gaining it. On the other hand, a tiny risk, say 1%, cannot always be ignored, especially if the penalty is something unpleasant. Similarly, Herxheimer (2001) stated that benefits and risks have completely different dimensions (a benefit is a material or experiential good thing, while a risk is a probability, the chance that something bad will happen), and that therefore 'benefit' should be weighed against 'harm', and the probability of benefit against the probability of harm. Veatch (1993) also mentioned that the proper contrasting terms are 'benefit' and 'harm', not benefit and risk. Edwards *et al*. (1996) went one step further and argued that the phrase 'benefit–risk' is clumsy and has been used in too many different contexts. The preferred term was therefore 'merit assessment' since it unambiguously indicates the determination of the worth of a medicine in a given context.

The term 'benefit–risk ratio' is used very often in the literature, but several authors have argued that this term is very often inaccurate. Herxheimer (2001) and Ernst and Resch (1996) stated that, because the benefit and the risk are not of the same nature, no one can really weigh them. Similarly, in the CIOMS IV report it was mentioned that for many benefit–risk assessments, neither the benefits nor the risks are easily or appropriately compared quantitatively (CIOMS Working Group IV, 1998). Hence, the word 'ratio', which is a mathematical term, is incorrectly used in this context (Edwards *et al*., 1996). 'Benefit–risk balance' is a more general and therefore more appropriate term, which is used for example in the new EU pharmaceutical legislation and defined as 'an evaluation of the positive therapeutic effects of the medicinal product in relation to the risks', with risks being defined as 'any risk relating to the quality, safety and efficacy of the medicinal product as regards the health of patients' (Council of the European Union, 2003). The terms 'risk–benefit ratio' or 'risk–benefit assessment' are also sometimes used in the literature, although it can be argued that putting risks before benefits is inappropriate and that benefit–risk is a more logical order.

In this book, the terms 'benefit', 'risk' and 'benefit–risk balance' will be used preferentially because they are most frequently used in the literature and by regulatory authorities, despite the fact that it would be more accurate to use the term 'harm' instead of risk.

1.4   VIEWS AND PERCEPTIONS OF BENEFITS AND RISKS OF MEDICINES

Health care professionals' perspectives

In this section, the views and perceptions of the professionals involved in benefit–risk assessment, which are the pharmaceutical companies, the regulatory authorities, and prescribers will be described.

   As a prerequisite, it is important to note that benefit and risk are fundamentally evaluative terms (Veatch, 1993). According to Veatch, they each contain in their very meaning value judgements that, in principle, cannot be made scientifically. Clinical studies, within the limits of their statistical power, can show that a medicine will have an effect. What studies cannot do is determine whether the effect is a benefit and, if so, how beneficial the effect is. Neither can they determine that the effect is a risk (harm) and, if so, how harmful. For example, the science can determine within certain limits what the change in 1-year survival probability is from the use of a medicine, but it cannot determine how valuable it is to survive. In this respect, it has been argued that providing information in terms of extension of lives instead of saving of lives would give a more realistic and balanced assessment of the efficacy of a medicine, and should moreover include an assessment of the quality and quantity of the life (Tan and Murphy, 1999). In general, the science can determine what the change in mortality rate is or what the cause of a rash is, but it cannot determine how bad it is for these events to occur (Veatch, 1993). In fact, it cannot even determine whether the effect is good or bad, that is, whether it is a benefit or a harm (risk). At the extreme, some people suffering from very serious conditions would count on a rapid painless drug-induced death a benefit rather than a harm. Still according to Veatch, benefit and risk (harm) are categories that necessarily involve subjective value judgements that ultimately only patients themselves can make, and this fact has radical implications for how information about benefit and risk judgements is processed by the regulatory authorities. Obviously judgements must be based on facts about what the effects of a medicine are likely to be. However, it requires superimposing evaluative judgements on these facts, and labelling the effects as good or bad, helpful or harmful. This requires drawing on religious, cultural, social and political judgements about how a medicine fits the basic values of society (Veatch, 1993). Veatch provides the example of two hypothetical medicines, one of which has a 50% success rate in curing its target disease but a 0.01% mortality rate. The other has a 60% success rate, but a 1% mortality rate. It is absolutely impossible to determine whether either of these medicines is safe or effective based on these data. If the medicines treat headaches, virtually everyone would make the value judgement that neither is safe enough nor, perhaps, effective enough neither. If the medicines

are used to treat small cell lung carcinoma, perhaps both would be judged safe and effective enough to give patients the right to access, but what if these medicines treat a serious, occasionally fatal condition such as paranoid schizophrenia? This example shows clearly that weighing benefits and risks of a medicine is a value judgement. According to Rawlins (1987), comparing benefits and risks is the problem of comparing apples and pears, or of titrating for example one life saved versus 1000 patients with impotence. A complicating factor is often that long-term benefits have to be weighted against short term risks, or vice versa. For example, with the oral contraceptive pill a benefit of no fear of getting pregnant (tomorrow) must be compared with a risk of venous thrombosis or myocardial infarction (15 to 25 years in the future) (Herxheimer, 2001). However, most people are less concerned about an adverse event that might occur in the future compared with one that might occur in the present, a process known as 'discounting' (Naimark *et al.*, 1997). In contrast, for a statin for the treatment of hypercholester-olaemia, the long-term benefit of a decreased risk of coronary heart disease must be weighted against the risk of adverse effects in the short-term such as myopathy. The time-dependence of benefit–risk has so far not been addressed appropriately (O'Neill, 2008).

Having established this fundamental concept of value judgement, the next question is how pharmaceutical companies, regulatory authorities and prescribers view benefits and risks of medicines. Pharmaceutical companies view benefits of medicines in terms of their ability to demonstrate sufficient efficacy so that: (1) regulatory authorities will approve their product, (2) third-party payers will reimburse for the medicine, (3) physicians will prescribe the medicine, and (4) people will purchase and use the medicine (Spilker, 1994). Risks are usually viewed as adverse experiences relative to existing treatment. Companies may accept increased risks until they reach the point where regulatory authorities will not approve, or physicians will not prescribe the medicine. Another driving force behind many risk assessments is legal liability (Bowen, 1993). Within a company some groups traditionally focus to a greater degree on potential risks (e.g. lawyers, research and development staff), while others tend to focus primarily on benefits (e.g. marketers, public relations, advertising staff).

Regulatory authorities tend to view benefits and risks for the nation as a whole rather than for individuals (Spilker, 1994; Cromie, 1987). Because of the cost of treatment to society and the required trade-offs that have to be made in this respect, this may lead to a conflict between the interests of individual patients and the population (Swales, 1997). Also, regulatory authorities evaluate the benefits and risks of a medicine against available alternatives (Bass, 1987; SCRIP, 2000), and they usually focus more on risks than benefits because of their mandate to protect society's health (Spilker, 1994). Just as the nature of the disease being treated is important for regulatory authorities in balancing benefits and risks of medicines, the

purpose of the intervention will also impact the benefit–risk assessment (Arlett, 2001). Medicines may be used for prevention, treatment (symptomatic, curative, reduction of morbidity and/or mortality, stabilization, improvement of quality of life), or diagnosis. Regulatory authorities require a very well established and favourable safety profile for medicines used for the prevention of a disease, in particular if that disease is non-serious, and for medicines which treat symptoms (Miller, 1993). In general, regulators tend to view patients as incompetent to judge risks and benefits (Bowen, 1993). Finally, since benefit–risk assessment is essentially a value judgement, inevitably it will be prone to a number of biases. Some people who make individual decisions will be risk-prone while others will be risk-averse on a particular issue. Other people might well reach different decisions on another occasion, even when presented with the same data (Rawlins, 1987). According to Asscher (1986), a selection bias is inevitable because of the difficulty to digest all the evidence presented. For example, FDA Advisory Committee members review only a small fraction of documentation included in a New Drug Application, and most members are not provided with a formal training in the interpretation of toxicology data or other important aspects of a safety evaluation. Hence the mechanisms employed by Advisory Committee members to assess safety and efficacy are largely based on clinical judgement (Lewis, 1993).

Prescribers view the benefits of a medicine in relative terms. They assess or guess whether other doses, other medicines or a non-medicine treatment will enhance or reduce the degree or type of benefit obtained (Spilker, 1994). Risks may offset one another in that a medicine may increase risks of adverse experiences but lower risks attributable to the disease being treated (Spilker, 1994). However, sometimes the risks of a medicine are given greater emphasis by prescribers relative to the risks of the disease being treated because the prescriber feels causally responsible for the former (Edwards *et al*., 1996). Overall, they discuss the benefits of treatment more than the risks, as shown in a study (Elwyn *et al*., 2003). In this respect, it should also be noted that perceptions of benefits and risks by physicians are strongly influenced by the way the data are presented in scientific journals and advertisements (Skolbekken, 1998). Also, it is felt that whilst great progress has been made in obtaining reliable evidence on the beneficial effects of a treatment, reliable evidence on harms is often lacking (Cuervo and Clarke, 2003). Finally, several authors advocate that there should be more education of prescribers on the evaluation of benefits and risks of medicines (Edwards and Hugman, 1997; Risk:benefit analysis of drugs in practice, 1995). This need for more education can be illustrated by a statement with regard to the class of cyclo-oygenase-2 (COX-2) inhibitors, that it is still difficult to give patients an honest, accurate, and understandable account of the balance between their benefits (relief of pain and improved function) and risks (the likelihood of serious adverse effects) (Jones, 2002).

Patients' and the general public perspectives

In its guide to patients, the FDA advises patients that they must decide what risks they can and will accept to get the benefits of medicines they want. For example, in a life-threatening disease, they might choose to accept more risk in the hope of getting the benefits of a cure or living a longer live. On the other hand, if patients are facing a minor illness, they might decide that they want to take very little risk, although the FDA recognizes that the benefit–risk decision is sometimes difficult to make for patients (FDA, 2002). The FDA also advises on some specific ways for patients to lower the risks and obtain the full benefits of medicines: talk to the doctor and other health care professionals, know your medicines, read the label and follow directions, avoid inter-actions, and monitor the medicine's effect. Similarly, Lumpkin (2000) argues that the fundamental goal of risk education must be to help patients understand that it is not a question of risk alone, but of benefit–risk, and how health care professionals know about benefits and risks relates to the individual patient and the decisions he or she has to make regarding their personal health.

According to Spilker, patients view benefits in terms of whether their symptoms improve, by how much and for how long. For medicines that prevent or suppress a disease or problem, benefits are viewed as how well they accomplish those goals. Risks are viewed as the likelihood of having adverse reactions, exacerbations or new episodes of the underlying disease, or other medicine-related problems. Risks that a patient voluntarily accepts (e.g. smoking cigarettes, using oral contraceptives) are often differentiated from those where little or no choice is possible (e.g. cosmic radiation) (Spilker, 1994). In addition, patients often have the misconception that medicines should be completely safe, which may lead them to focus excessively on any risks to which their attention is drawn. Perception of risk may therefore be different from actual risk (Edwards *et al*., 1996). Most people knowingly accept risk if (1) the probability of something harmful happening is small or distant in time, or (2) the possible harmful event is not serious, or (3) the possible benefits are sufficiently strong. For each person this kind of risk assessment is different and each person applies different criteria to different kinds of risks. AIDS patients have different views on what they perceive to be an acceptable risk compared to what regulatory authorities and industry perceive to be acceptable (Bowen, 1993). Perception of risk is affected by many variables from the tiny workings of individual psychology through the sweeping influences of language and culture (The Uppsala Monitoring Centre, 2002b). Overall, patients' perspectives on assessing benefits and risks are not by any means uniform. They often realize that benefits and harms necessarily require very personal, subjective value judgements that can only be based on religious, philosophical, and social values (Veatch, 1993). In a study by Johnson *et al*. (1994), women's views on the benefits and risks of hormone replace-ment therapy were measured using utility analysis. Results showed that women are

reluctant to accept any increase in a risk factor such as a small increase in the risk of breast cancer, even if it reduces their total risk (which includes cardiovascular disease, osteoporosis, uterine cancer, ovarian cancer, breast cancer). A social climate, which is increasingly intolerant of risk, has created a tendency to seek absolute safety and security. This has led to a trend of blaming and punishing those who are seen to be the cause of accident or misadventure. Part of this results from an unwillingness to accept that it is natural for accidents to happen, and that there is always some uncertainty in all human activities, in spite of legislation or good intentions. Part of it results from the pressure on politicians and public servants to show that they are in control of the variables of existence (The Uppsala Monitoring Centre, 2002b).

Many factors influence patients' choices about treatment and its benefits and risks. These include their own attitudes and beliefs, as discussed above, and those expressed in the media, by other health professionals and by family and friends. Previous experience, trust in their physician and the chances of success also affect their decisions (Risk:benefit analysis of drugs in practice, 1995). Finally, perceptions of benefits and risks by patients are strongly influenced by the way the data are presented, which is also the case for physicians (Skolbekken, 1998). However, research has demonstrated that, despite the fact that lay people sometimes lack certain information about hazards, their basic conceptualization of risk is much 'richer' than that of the experts and reflects legitimate concerns that are typically omitted from expert risk assessments (Slovic, 1987).

A survey, whose purpose it was to establish whether patient groups across the EU considered that pharmaceutical companies should provide the public with more information on prescription medicines, revealed that the majority of patients are no longer willing to rely on doctors as the most important source of information about their prescription medicines (PatientView, 2002). Patients prefer a range of information sources about medicines, and the ideal source of information would be, amongst many other features, striking a balance between a treatment's beneficial and adverse effects (Dickinson and Raynor, 2003). Patients often desire more information than is currently provided, and professionals need to support patients in making choices by turning raw data into information that is more helpful to the discussions than the data (Edwards *et al*., 2002). Patients do not want to know all undesirable effects of a medicine, but only those which are relevant to them (Meiners, 2002). Overall, given the difficulties and ramifications of benefit–risk assessment, the challenge of presenting understandable, coherent information to patients is considerable (Edwards and Hugman, 1997). Similarly, regulatory authorities have difficulties in communicating information on benefit–risk assessments to patients because of a lack of information (Beermann, 2002). In this respect, Edwards made a number of recommendations with regard to communication of benefit–risk assessments to patients, and called for increased attention to more transparent, accurate and helpful benefit–risk assessments and expressions.

All of this is, however, to a certain extent in contrast with how the media cover the benefits and risks of medicines. In a study, coverage by US news media of the benefits and risks of three medicines that are used to prevent major diseases was analysed. The medicines were pravastatin (a cholesterol-lowering medicine for the prevention of cardiovascular disease), alendronate (a bisphosphonate for the treatment and prevention of osteoporosis) and aspirin (which is used for the prevention of cardiovascular disease). A systematic probability sample of 180 newspaper articles (60 for each medicine) and 27 television reports that appeared between 1994 and 1998 were analysed. Of the 207 stories, 83 (40%) did not report benefits quantitatively. Of the 124 that did, 103 (83%) reported relative benefits only, 3 (2%) absolute benefits only, and 18 (15%) both absolute and relative benefits. Of the 207 stories, 98 (47%) mentioned potential harm to patients. It was concluded that news-media stories about medications may include inadequate or incomplete information about the benefits and risks about the medicines (Moynihan *et al.*, 2000). Moreover, it can be assumed that the vast majority of the intended audience will not take the time to read a full article as would be required to accurately understand the complex balance of benefit and risk (Califf, 2004). Other studies had identified an overstatement of adverse effects and risks, with negative clinical trials being highlighted and positive trials being ignored (Lebow, 1999), and that media coverage was driven by moral outrage instead of scientific notions of calculable risk (Brown *et al.*, 1996). Yet another study came to a somewhat different conclusion that in the Netherlands a discrepancy exists between the medical literature and the newspapers, in that the negative consequences of the use of medicines receive proportionally more attention in the professional literature than in the newspapers, whilst this difference does not exist for 'good news' (van Trigt *et al.*, 1995).

1.5   STAGES AND CONCEPTS IN BENEFIT–RISK ASSESSMENT

When is benefit–risk assessment used?

Benefit–risk assessment is used during the various stages of the life cycle of a medicine: during the development of a new medicine, during the approval phase of a new medicine, and for the post-approval re-assessment of marketed medicines. A short description of how benefit–risk assessment is used during these three stages is provided below.

Benefit–risk assessment during the development of a new medicine

The safety and efficacy profile of a new medicine gradually takes shape over the course of the development of a new medicine, but will not even be fully elucidated

by the time that the medicine finishes its development. Also, whilst during the early stages of development more weight is placed on the animal data because these are essentially the only data available, human experience becomes more important as development progresses (Brimblecombe, 1987). The major technique for incorporating benefit–risk considerations during the development stage is through the use of minimally acceptable hurdles (Spilker, 1994). These standards specifically identify the types and magnitudes of activities the compound must have (or must not have) at each stage of the development to further progress. The rationale for the creation of these hurdles, which are essentially benefit and risk hurdles, is that unless a compound meets the minimum standards, there is little or no reason to continue its development. An important factor in this respect is that a company that is developing a new medicine may be unaware of a competitor who has a safer or more effective medicine under development. Introduction of a major competitive product can instantly turn what might have been an important advance in therapy into a product that will not be approved by a regulatory authority or will not be widely prescribed.

Specifically, benefit–risk assessment is initially considered of importance during the phase II dose-ranging studies in selecting the dose or doses for the phase III trials, based on a trade-off of the efficacy and safety observed with each dose. As will be discussed in the next section, some models exist for this purpose. Obviously the phase III data offer the best chance to conduct a meaningful benefit–risk assessment prior to the market entry of the medicine, and they are the basis for the benefit–risk assessment by the regulatory authorities as discussed below. Nevertheless, at least one author suggests that benefit–risk assessment in the development of a new medicine is often done on an *ad hoc* basis, and there are few cases where the user of the phrase was specific about how this assessment should be conducted (Chuang-Stein, 1999).

Benefit–risk assessment during the approval process of a new medicine

Decisions on the benefit–risk balance of a new medicine by regulatory authorities are generally speaking based on the legislation in place, which is comparable throughout the Western world (Bass, 1987). Nevertheless, different regulatory authorities can come to different decisions based on the same data, and this is because of differences in administrative and scientific views and interpretations. In addition, regulatory authorities have to make decisions based on incomplete data, with more safety and efficacy data becoming available during the commercialiation of the medicine. Also, they compare a new medicine with the available alternatives (Bass, 1987), and although not published, the EU Committee for Human Medicinal Products (CHMP) has made it clear that they would not approve a medicine which has a benefit–risk profile which is inferior to existing therapy.

According to Chuang-Stein (1999), benefit–risk assessment remains largely a concept rather than a routine practice in the current regulatory environment. Even though everyone agrees on the need to conduct such an assessment, few can come up with the specifics to actually carry it out. Miller (1992) cited the lack of attention paid to the benefit–risk issues as one pitfall that frequently delays the review and approval process. Nevertheless, pharmaceutical companies rarely receive any guidance from the regulatory authorities regarding the basics and contents of a pertinent benefit–risk assessment package (Miller, 1992; Chuang-Stein, 1999). Walker stated in the mid-1980s that as a result of a more accurate assessment of benefits and risks in the next 10 or 20 years, it may be possible to plot on a matrix system (Figure 1.1) any new chemical entity at the approval stage, or at the post-approval stage (Walker, 1987). Whilst there is no hard evidence that the benefit–risk assessments of new medicines conducted today are superior to the assessments made in the eighties, the basic matrix proposed by Walker is still valid. More recently, it was suggested that the emerging risk management strategies including extensive and earlier epidemiological assessments of benefits and risks of new medicines will eventually create a new standard of evidence for industry and regulators (Andrews and Dombeck, 2004).
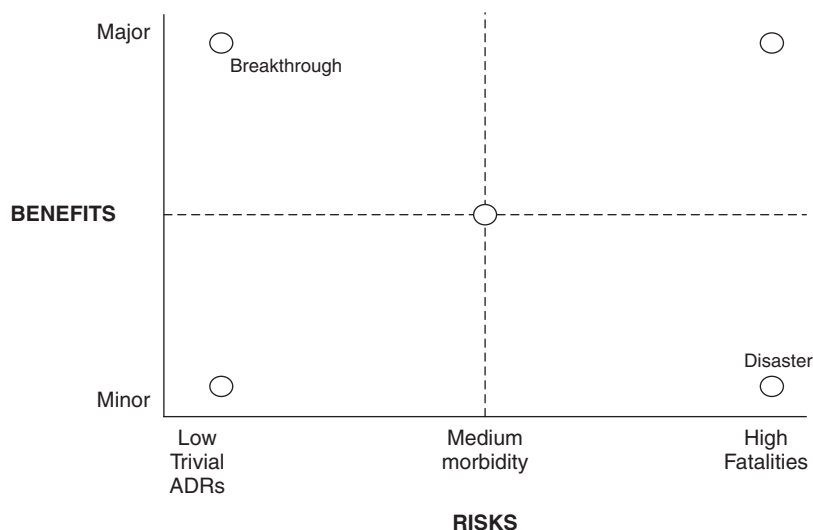

Figure 1.1   Matrix system for plotting the benefit–risk ratio of medicines

Finally, it should be noted that cost-benefit assessment is not part of the approval process of new medicines either in Europe or in the United States. Costs and health economics in a broader sense are not considered during the approval process of new medicines.

Benefit–risk assessment during the post-approval re-assessment of marketed medicines

At the time medicines are initially approved by the regulatory authorities, evidence for the efficacy and safety will have been more or less secured. However, at that stage, efficacy has often only been established in a limited number of patients under a narrow range of conditions, and safety in relatively small number of patients. During the post-marketing period, evidence may accumulate to suggest that the efficacy is considerable in a wider group of patients than was originally indicated or that the medicine has a true impact on clinical outcomes as opposed to surrogate outcomes, and that safety during long-term use in a more heterogeneous population is acceptable. Alternatively, efficacy may appear to be less satisfactory than originally believed and safety becomes pre-judiced (Rawlins, 1987). The UK Committee on Safety of Medicines considered that the following issues might lead to changes in the benefit–risk profile of a marketed medicine (Rawlins, 1987):

1. quality (very unusual)

2. efficacy

    (a) new evidence for lack of efficacy

    (b) uniqueness of therapeutic properties

    (c) uniqueness of efficacy in small patient subgroups

3. safety

    (a) spontaneous reports

        literature

        yellow cards

        other agencies

    (b) cohort studies

    (c) case–control studies

    (d) animal carcinogenicity.

Apart from the new information about the medicine itself, external factors not involving the medicine may change the benefit–risk balance (Spilker, 1994). These

include the introduction of new medicines, combination therapy or other modalities that compete with the original medicine.

Following the benefit–risk assessment of a marketed medicine based on new information, it will sometimes be necessary to take action either to increase the benefit of the medicine by improving its rational use, or reducing the risk usually by restricting its use (Arlett, 2001). With regard to the latter, formal risk management plans are now being established for new medicines, which aim at minimizing the risks while maximizing the beneficial effects of a medicine by ensuring its proper use by the right patients for the right disease conditions (Marwick, 1999; Perfetto *et al.*, 2003). If the overall balance of benefits and risks is judged to be negative, then the medicine may be withdrawn unless risk reduction strategies can be identified which would swing the balance away from risk. In this respect, the following actions may be taken (Arlett, 2001):

- continue to passively monitor the issue: may be appropriate for non-serious adverse events where causality is not established

- actively collect further data: if causality is not established, mechanism unclear, risk factors not identified

- add warnings to the product label: for less serious adverse effects, particularly those that are unavoidable, warning health care professionals and patients may be the only action necessary

- changes to the product label to reduce risk: restrict the indication, contraindicate those at greatest risk, advice monitoring

- monitor effects of action taken: ensure that the action effectively protects public health.

Concepts in benefit–risk assessment

Based on the above, a number of concepts and rules with regard to benefit–risk assessment and benefit–risk balance can be identified, all of which are also discussed in the literature, and some of which (concepts 3 and 4) are related to the fundamental paradigm that benefit–risk assessment and balance are value judgements, as discussed above.

Concept 1: a separate benefit–risk balance for each indication

It is generally recognized that a separate benefit–risk assessment should be conducted for each indication, because medicines are clinically tested for each indication separately (Spilker, 1994; Beckmann, 1999). The efficacy and thus the benefits of a

medicine in one indication are not the same as in another indication, and, similarly, the safety is not necessarily comparable, one of the reasons being that the dosage might be different in different indications.

Concept 2: all available data should be considered in benefit–risk assessment

Several authors emphasize the need that all available and relevant data should be considered in benefit–risk assessment (Bass, 1987; Arlett, 2001). For example, an application to the regulatory authorities for a new medicine must include very comprehensive data and information with regard to the quality of the active and inactive substances and the finished product, the non-clinical studies (animal pharmacology and pharmacokinetics, animal toxicology including single and repeat-dose toxicity, genotoxicity, carcinogenicity, reproductive and developmental toxicity), and the clinical studies (human pharmacokinetic and pharmacodynamic studies, clinical efficacy and safety studies, and post-marketing experience if applicable) (International Conference on Harmonization, 2000). Thus, benefit–risk assessment involves the review and interpretation of a wealth of safety and efficacy information, which must eventually be captured in the benefit–risk balance.

Concept 3: the nature of the disease should be taken into account for benefit–risk balance

The seriousness and prognosis of the disease being treated by the medicine under investigation will have a major impact on the benefit–risk balance (Meyboom and Egberts, 1999; Schiller and Johnson, 2008). For example, if the disease is self-limiting such as influenza in an otherwise healthy young adult, serious adverse effects would be likely to result in a negative balance of benefits and risks. In contrast, with a disease with a high mortality, serious adverse effects may still be outweighed by the benefits afforded by the medicine (Arlett, 2001). In general, the more severe the disease, the more tolerable are the potential therapy risks (Miller, 1993). From a conceptual perspective, the seriousness of the disease can be compensated for in the evaluation of the benefits and/or risks, as suggested above, or alternatively the seriousness of the disease can be taken into account using the paradigm that the less severe the disease, the more the benefits of a medicine should outweigh its risks.

Concept 4: absolute versus relative benefit–risk balance

There seems to be a consensus in the literature that the availability of alternative therapies should be considered in benefit–risk assessment, and that the interpretation

of the benefits and risks always involves a comparison (Bass, 1987; Miller, 1993; Spilker, 1994; Arlett, 2001). For example, the German Drug Law requires consideration of alternative sources of risk, namely: (1) the risk associated with no treatment, and (2) the benefit–risk balance of therapeutic alternatives, if available (Schosser, 2002). The US Federal Regulations do not require that new therapies be superior to or safer than existing therapies (nor do the EU regulations). The experience to date shows, however, that the best case for approval of a new medicine is where the disease is severe and/or life-threatening and for which there are few or no available treatments. This contrasts with the case for a medicine for which alternative therapies exist and for which lingering questions surrounding the safety of medicines in the therapeutic class remain unresolved (Miller, 1993). For example, according to the FDA the benefits of the diabetes compound Rezulin (troglitazone) outweighed its risks at the time of initial approval, but the launch of newer products with better safety profiles had made Rezulin 'outmoded' (SCRIP, 2000). In most cases benefit–benefit and risk–risk comparisons are made between the medicine in question and: (1) a higher or lower dose of the same medicine; (2) another medicine of the same type; (3) a medicine of a different type; (4) a combination of two or more medicines; (5) another treatment modality; or (6) a combination of a medicine plus another treatment modality (Spilker, 1994). As will be discussed in Chapter 2, pivotal clinical trials which aim at establishing the efficacy and safety profile of a new medicine include quite often an active comparator, and hence the comparison of the benefits and risks versus the alternative treatment is inherently part of the evaluation of the results of such trials. Thus, from the above it can be concluded that the benefit–risk balance is a relative and not an absolute concept.

Concept 5: the benefit–risk balance is dynamic and evolves over time

An important concept which is mentioned by many authors in the literature is that the benefit–risk balance is dynamic and changes over time (Rawlins, 1987; Spilker, 1994; Schosser, 2002; Konstam, 2003). The benefit–risk balance evolves as new information becomes available during the development of a new medicine, and the conclusions on the benefit–risk profile by the regulatory authorities at the time of market introduction of a new medicine may change after the medicine is marketed. With regard to the latter situation, the benefit–risk balance may change for two major categories of reasons. Those categories are: (1) new information about the medicine itself, and (2) external factors not involving the medicine (Spilker, 1994). Factors external to a medicine that affect its benefit–risk balance, as explained above in concept 4, include the introduction of new medicines, combination therapy or other modalities that compete with the medicine in question. In addition, the discovery of new problems with existing therapy may improve the benefit–risk balance of a new medicine.

The other reason for changes in the benefit–risk balance of a marketed medicine is new information about the medicine itself, which may lead to either increased or decreased benefits and/or risks. Increased risks occur regularly and are, for example, due to the appearance of new, rare side effects, or to a deterioration of the safety profile in populations with co-morbidities and/or concomitant medications which were not studied in the clinical development programme. An amelioration of the risk profile of a marketed medicine can also occur, an example being the demonstration in the Vioxx Gastrointestinal Outcomes Research (VIGOR) trial of significantly fewer clinically important upper gastrointestinal events with the COX-2 inhibitor rofecoxib than with naproxen (Bombardier *et al*., 2000), although later the increased risk for cardiovascular effects became apparent. Similarly, the benefit of a marketed medicine may be reduced based on the results of new studies. For example, the angiotensin II antagonist losartan was originally approved in a number of countries for the treatment of heart failure based on the results of the Evaluation of Losartan in the Elderly (ELITE) trial, but the subsequent ELITE II trial showed that all-cause mortality (which was the primary endpoint of the trial) was numerically higher in patients treated with losartan than in patients treated with the angiotensin-converting enzyme (ACE) inhibitor captopril, which is standard treatment (Pitt *et al*., 2000). There is also an inverse example in the field of congestive heart failure. Enalapril, an ACE inhibitor, which was originally approved for the treatment of heart failure based on exercise-tolerance studies only, was shown in the Studies of Left Ventricular Dysfunction (SOLVD) trial to reduce mortality and hospitalizations for heart failure (The SOLVD Investigators, 1991).

Schosser (2002) developed a decision matrix (Table 1.1) to determine the need for taking regulatory action in response to a change in the benefit–risk balance of a medicine. Classifying the change of either the risk or the benefit into one of the three categories 'increased', 'unchanged', or 'decreased' resulted in a 3 × 3 decision matrix. According to Schosser, action has to be taken whenever the benefit–risk balance falls below 'the bounds considered justifiable in the light of the knowledge of medical science', such as in the scenarios of increased risk with reduced or unchanged benefit, and reduced benefit with unchanged risk.

Table 1.1 Decision matrix to determine the need for taking regulatory action in response to a change of the benefit–risk balance of a medicine

Benefit	Risk		
	Reduced	Unchanged	Increased
Reduced	Conditional	Yes	Yes
Unchanged	No	No	Yes
Increased	No	No	Conditional

1.6   BENEFIT–RISK ASSESSMENT: THE CURRENT REGULATORY ENVIRONMENT

The withdrawal of the lipid-lowering medicine Baycol (cerivastatin) in 2001 and the withdrawal of the COX-2 inhibitor Vioxx (rofecoxib) in 2004 have drawn the increased attention of the stakeholders such as patients and physicians and the media to the regulatory processes for approving new medicines and for re-assessing marketed medicines. These two withdrawals raised major questions about the assessment of the safety of medicines and thus about their benefit–risk profile. Whilst the withdrawal of cerivastatin because of a non-acceptable risk for rhabdomyolysis was in many aspects different from the voluntary withdrawal of rofecoxib because of an increased risk of myocardial infarction after long-term use, both triggered the regulatory authorities to take a more cautious and conservative approach with regard to the safety and risk of medicines.

1.7   BENEFIT–RISK ASSESSMENT IN OTHER DISCIPLINES

Apart from the specific benefit–risk evaluations, benefits and risks of medicines are also considered, in addition to costs, in cost–benefit analysis and cost-effectiveness analysis. Cost–benefit analysis is considered the pre-eminent method of economic evaluation of medicines because it allows for the direct comparison of different types of outcomes resulting from a variety of actions (Messonnier and Meltzer, 2003). It provides the most comprehensive monetary measures of the positive (beneficial) and negative (costly) consequences of a possible course of action, such as recommending the routine vaccination of restaurant employees against hepatitis A. Because the results of cost–benefit analysis are expressed in monetary terms, both health and non-health programmes can be directly compared. Cost–benefit analysis has a long history in the study of environmental effects and natural resource use in the United States. Cost-effectiveness analysis has become the predominant method for evaluating the costs and effects of alternative health-care strategies (Gift *et al.*, 2003). It is a method used to evaluate the relationship between net investment and health improvement in health-care strategies competing for similar resources. Because cost-effectiveness analysis is limited to the comparison of strategies designed to improve health, no attempt is made to assign a monetary value to disease averted beyond the cost of care for persons with these conditions and costs of lost productivity from morbidity and mortality. Rather, results are presented in the form of the net cost per health outcome, such as 'cost per case prevented' or 'cost per life saved'. The decision-maker is left to make value judgements about the intrinsic value of the health outcomes. Cost-effectiveness

analyses are designed to produce a cost-effectiveness ratio, the most commonly used summary measure for comparing health interventions (Gift *et al*., 2003). The numerator of the cost-effectiveness ratio is the cost of the health intervention (including costs of adverse effects of the intervention and other costs induced by the intervention) minus the costs of the health conditions that are prevented as a result of the intervention. The denominator is the improvement in health produced by the intervention. This improvement in health may be measured in natural units, for example cases of measles prevented, or in terms of the impact on health-related quality of life, for example quality-adjusted life years. Cost-effectiveness analysis is becoming very important as new medicines must increasingly show evidence of cost-effectiveness before obtaining reimbursement. Faced with increasing health-care costs, the reimbursement authorities in many countries are now requiring proof of cost-effectiveness, the so-called fourth hurdle (Taylor *et al*., 2004). It seems that these fourth hurdle policies are contributing to more cost-effective use of medicines.

In the health care sector, benefit–risk assessment is also applied to medical devices. Securing FDA approval of a new device before marketing requires that the manufacturer provides reasonable assurance that the device is safe and effective when used for the purpose for which the approval is sought (Kessler *et al*., 1987). Safety and effectiveness are assessed with specific reference to the uses for which the device is intended, as detailed in the labelling on the device. Safety is evaluated by weighing the probable benefits to health against the probable risks of injury. The benefit–risk ratio must be acceptable, but proof that the product will never cause injury or will always be effective is not required. Similar regulations exist in Europe.

Benefit–risk assessments are also made in a vast number of fields outside the health care sector, although they are usually called risk assessments. They are rather widely applied with regard to energy (nuclear power, non-nuclear electric power), transport (motor vehicles, motorcycles, railways and aviation), food (e.g. the use of preservatives and colouring agents), sport activities (e.g. skiing, mountain climbing), and the environment (e.g. the use of pesticides) (Lee, 1987). In a number of these assessments, risks to human health are predominant. The problems and issues in such risk assessments are quite similar to the issues identified in this chapter for the benefit–risk assessment of medicines, the most important being that it is necessary to consider all benefits and all risks, and not only one aspect of the safety data in isolation (Alexander, 2002; CropLife International, 1997). Other issues that were for example identified for pesticides were that benefit–risk assessments are complex, and that it is inappropriate to undertake a comparative benefit–risk evaluation between two products when both products meet the acceptable standards for approval. Decision analysis techniques such as decision trees, Bayesian statistics, Monte Carlo simulations and sensitivity analyses are frequently used to make risk assessments (Clemen, 1996).

1.8   SPECIFIC METHODS AND MODELS FOR BENEFIT–RISK ASSESSMENT

Having described in the previous sections of this chapter the issues and difficulties with regard to benefit–risk assessment as well as the lack of guidance and standardization in this field, the purpose of this section is to provide a general overview of the possible methods and models for benefit–risk assessment that have been described and/or are being used by one or more stakeholders.

Definition of models and their potential value

Patten and Lee (2002) describe a model as 'something that mimics relevant features of a system being studied'. Models can assist with decision-making in a descriptive, heuristic or prescriptive (i.e. identifying optimal choices) sense. A model need not involve mathematical calculations. For example, a roadmap or a geological map can be regarded, by this definition, as a model. In fact, such non-mathematical models are often useful in decision making to clarify issues. However, explicit and analytic representation of a system usually requires a mathematical model. According to Patten and Lee a mathematical model can be defined as 'an abstract, simplified, mathematical construct related to a part of reality and created for a particular purpose'. Models can be used to simulate outcomes under a specified set of assumptions, and to help identify those factors, out of many possible factors, that may have the biggest impact on the outcome. Another advantage of a model is that it can be used to identify the degree to which changes in one or more variables impact the predicted output, which is called sensitivity analysis.

Back in 1985, Walker (1987) stated that the methodology for determining benefit–risk ratios was yet to be fully defined. He added that in the next 10 or 20 years, a stage may be reached when benefits could be accurately assessed, and that better ways and means of determining risks had to be found. More recently, it is still recognized that clinicians, members of formulary and guideline committees, and regulators must try to be much more specific when considering the benefits of treatments and the kinds of harm they may do (Herxheimer, 2001). If health care professionals can explain to each other how they weigh up the pluses and minuses for a particular intervention, then they will also be able to explain and discuss them more clearly and easily with patients. Medical judgement and consensus would have to be built into a methodology for benefit–risk assessment (Schosser, 2002). Since actual decisions concerning the balance between benefits and risks are heavily influenced by the values of the decision makers and difficult to quantify, better methods are needed for quantifying the benefit–risk profile and expressing the values involved in decision making (Hyslop *et al.*, 2002). Asscher (1986) mentioned four reasons why benefit–risk analysis assumes

greater importance than in the past and why further improvement in the decision-making process is required:

1. Modern treatment more often concerns marginal benefits. The more marginal the gain, the more important to undertake benefit–risk analysis.

2. The modern therapeutic armamentarium contains many bolder forms of intervention, both medical and surgical. Partly as the result of the wise use of these, the population, now exposed to these more dangerous treatments, is older than it used to be. As also mentioned by Hyslop *et al.*, in the last 50 years medical science and especially pharmaceutical science has made tremendous progress, dramatically expanding the physician's therapeutic armamentarium. These advances bring the need to be more adept at assessing the benefit–risk balance, from the perspective of both society (the ill population) and the individual patient (Hyslop *et al.*, 2002).

3. Medical litigation is growing, and every decision in medicine involves a benefit–risk analysis which may have to stand up to critical examination by peers, patients, relatives, media and the courts.

4. Objective analysis of benefits and risks plays a pivotal role in regulatory decisions about medicines.

The use of decision analysis techniques has been advocated in the literature for medical decision-making in general, including benefit–risk assessments. Such decision-analytic models should be simple, because the insights derived from a simple model are similar to those derived from a complex formulation (Detsky *et al.*, 1997). To determine the appropriate level of complexity, it must be considered whether the model captures the key issues necessary to fully describe the benefit–risk trade-off. If a key element of the benefit–risk trade-off is missing, the model will not achieve its goal of helping to understand the trade-off. The potential advantages of a formal model based on decision-analysis techniques are as follows (Waller and Evans, 2003):

- increased transparency

- making reasoning explicit, reflecting complexity

- clearly identifying limitations of the evidence and uncertainty

- the impact of all assumptions can be assessed

- better justified decisions.

In the next sections, not only models for benefit–risk assessment will be reviewed, but also any other specific methods for benefit–risk assessment, the term 'method' being somewhat broader than the term 'model'.

Use of models in the literature

The question addressed in this section is whether in articles in scientific journals that review a specific medicine or a specific clinical trial, methods or models are used to describe the benefit–risk profile of the medicine in question. Whilst it would be outside the scope of this book to provide a complete overview of how benefit–risk assessments are described in the scientific literature, the focus of this review will be on scientific journals which specialize in benefit–risk assessment.

The journal *Drug Safety* is the premier international review journal covering the disciplines of pharmacovigilance, pharmacoepidemiology and benefit–risk assessment, and has a Scientific Citation Index Factor of 3.536 (Adis International, 2009). Since, according to the publisher, this journal includes benefit–risk assessments which provide an in-depth review of adverse effects and efficacy data for a medicine in a specific disease to place the benefit–risk relationship in clear perspective, it is of value to review how such benefit–risk assessments are made. For this purpose, all articles published between 1998 and 2002 and which included the terms 'benefit' and 'risk' in the title were reviewed with regard to the presence of a specific method to conduct the benefit–risk assessment. A search revealed that there were 48 articles that fitted these criteria, and that these articles pertained to benefit–risk assessments of specific medicines (24), specific classes of medicines (9), and medicines for specific diseases (15). The title of most of these articles was called 'A benefit–risk assessment of . . .' or 'A risk–benefit assessment of . . .'. From a review of these articles the following was concluded:

- whilst the efficacy and safety data were extensively reviewed in all articles, this review was not systematically extended and extrapolated into a discussion of the benefits and risks

- no systematic trade-off of the benefits and risks was performed

- overall, no systematic method for benefit–risk assessment was used in these articles, and the content and format of the specific sections of these articles which focused on the benefit–risk profile was not harmonized.

A recent review confirmed the previous findings.

In contrast to the above, Bjornson (2004) conducted a systematic review of the literature and came to the conclusion that the measurement of benefits and risks is

consistently reported with similar terminology. However, he also concluded that the interpretation and subsequent balancing of benefits and risks was more problematic.

Use of models by pharmaceutical industry in drug development

Pharmaceutical companies use benefit–risk weighting in the drug development process. During earlier phases of drug development, research concentrates not only on dose-response modelling and identification of the maximally tolerated dose, but also on risk estimation for dose-dependent toxicity (Troche *et al.*, 2000). The goal is to find a dose that satisfies both safety and efficacy requirements. A simple method to obtain trade-off information is to specify, a priori, the maximum toxicity rate that would be acceptable if the new treatment were to produce 100% response, and similarly to specify a minimum response rate that would be acceptable if the treatment produced no toxicities (Conaway and Petroni, 1996). This method, in which the efficacy–toxicity requirements are specified, against which the medicines are then compared, can be applied to dose-finding studies and for other phase I and phase II trials.

There is a lack of available techniques during the drug development process to conduct relevant benefit–risk assessments to determine the acceptability of a new medicine across the dose range used in the therapeutic setting (Miller, 1992). Chuang-Stein (1994) developed a risk-adjusted benefit measure in which the intensity of side effects is quantified through a set of scores that are summarized into the risk component experienced by the individual. This approach of subtracting a multiple constant of the weighted risk component from the expected crude benefit is further described in Chapter 3.

Measurement of patient well-being and preference has become a focus of research in the area of clinical decision making (Troche *et al.*, 2000). Much of this work is motivated by the observation that increased attention to patients' quality of life results in greater levels of adherence to medication, satisfaction, and conformance to the principle of informed consent. The Committee for Medicinal Products for Human Use (CHMP) issued a reflection paper on the use of health-related quality-of-life (HRQL) measures in the evaluation of new medicines, which hinted at the increasing importance of such HRQL measures in randomized clinical trials and in the approval of new medicines (Committee for Medicinal Products for Human Use, 2004). One approach, q-TWiST (Quality-adjusted Time Without Symptoms of disease and Toxicity of treatment), is a method that compares therapies in terms of achieved survival and quality-of-life outcomes (Gelber *et al.*, 1996). Q-TWiST assesses both the likelihood of a given health state and the quality weighting to be assigned to that state. The result is a preference measure that captures both quality of life and time-linked data along a single index (Martin *et al.*, 1996).

Another approach is based on the multi-attribute utility theory (MAUT), a method that is founded on the theory of choice under uncertainty (Clemen, 1996). MAUT provides a quantifiable summary score and uses all available information on every defined, attainable health state (Troche *et al*., 2000). MAUT permits direct comparisons across health interventions and even across therapeutic classes. For example, Eriksen and Keller (1993) developed a MAUT function to combine preferences for efficacy and toxicity of a medicine, and applied it to the evaluation of anti-glaucoma treatment. Their model is further mentioned in Chapter 3.

Another method, the probability trade-off technique (PTT), also emerges from the theory of choice. PTT involves a search for thresholds of indifference (Llewellyn-Thomas *et al*., 1996). Descriptive and probabilistic information about potential benefits and side effects of two therapeutic alternatives are presented to patients and a point is located where the patient is indifferent between the two options. The method can determine how much benefit patients would require to accept a new therapy, which carries a particular risk of toxicity over an existing therapeutic intervention (Troche *et al*., 2000). The resultant scale provides a relative measure of respondents' strength of preference for one option over another (Naylor and Llewellyn-Thomas, 1994). Troche *et al*. developed such a patient-specific drug safety-efficacy index to choose a specific therapeutic intervention and dosage, when known risks and benefits are reconciled against patient-specific preferences among an array of therapeutic alternatives. Their method is briefly discussed in Chapter 3.

All the above models measure the benefit and risk components using individual patient data, making benefit–risk trade-offs for individual patients. A few other models have been developed as well as an aid with regard to other aspects of drug development (i.e. approvals of clinical trials by Ethics Committees, selection of clinical trial endpoints and drug discovery targets, and interpretation of efficacy and safety results of clinical trials), and are discussed below.

Lilford and Jackson (1995) introduced the concept of 'equipoise', the condition which applies when there is no preference between the treatment options to be compared, within the context of conducting ethical randomization in clinical trials and the approval of such trials by Ethics Committees. Using probability estimates, 'effective equipoise' occurs when the most likely results of a proposed trial are thought to be an improvement sufficient, but only just sufficient, to compensate for the disadvantage of the treatment with the greatest 'costs'. These 'costs' are side effects when viewed from the point of view of the individual patient, or the combination of those side effects and monetary costs when viewed from the perspective of society.

Bjornsson (1997) developed a systematic classification for the categorization of therapeutic effects of individual medicines based on their relationships to the underlying disease processes being treated or prevented, rather than on the pharmacological or biochemical effects of the individual medicines. This matrix method is based on six categories of therapeutic specificities and three levels of therapeutic effectiveness.

Category 0 includes medicines or medicine uses that prevent the development of a disease when none exists; Category I, those that affect aetiological factors of a disease; Category II, those that affect specific disease processes; Category III, those that affect specific disease manifestations; Category IV, those that affect non-specific disease manifestations or symptoms; and Category V includes medicines used to diagnose or facilitate treatment of a disease. Level 1 of therapeutic effectiveness applies to medicines that involve excellent, major or maximum effectiveness; Level 2 applies to medicines that involve good, substantial or medium effectiveness; and Level 3 applies to medicines that involve poor, marginal or minimum effectiveness. By providing a framework for comparisons across different therapeutic agents, be it based on benefits but ignoring risks, this matrix evaluation method may have implications in choosing clinical trial endpoints and selecting drug discovery targets.

Shakespeare *et al*. (2001) presented a new concept for the calculation and interpretation of the efficacy and safety results of a clinical trial. As discussed further in Chapter 3, they proposed to use confidence intervals and plotting of clinical significance curves and risk-benefit contours.

Use of models by regulatory authorities

At the outset, it is important to note again that benefit–risk analysis undertaken by regulators is not based on a precise mathematical equation and is invariably judgemental, making the term 'benefit–risk ratio' a misnomer (Waller, 2001). Already in 1973, it was acknowledged that regulatory authorities were poorly equipped to undertake the sophisticated benefit–risk analyses, and that better methodology was required (Wardell, 1973). As will be shown below, this is still the case today, and no standardized method or model is available for regulatory authorities.

Only three models for benefit–risk assessment of medicines for use by regulatory authorities are described in the literature: (1) the 'Principle of Threes' grading system developed by Edwards *et al*. (1996), (2) the TURBO model developed by Amery (CIOMS Working Group IV, 1998), and (3) the Evidence-based Benefit and Risk Model developed by Beckmann (1999). The first two methods are also described in the Report of the CIOMS Working Group IV entitled 'Benefit/risk Balance for Marketed Drugs: Evaluating Safety Signals'. All three models have been developed either within the scope of the evaluation of the benefit–risk balance of marketed medicines because of a safety concern, and/or by people who are primarily involved in pharmacovigilance activities. The primary focus of these models is therefore more on safety than efficacy, and more on the re-evaluation of marketed medicines than the evaluation of new medicines.

Edwards set out to establish a simple and fast methodology to evaluate the benefit–risk balance of a medicine. He proposed a 'Principle of Threes' structure to evaluate the

benefit–risk balance of medicines. The model is based upon the concepts of seriousness, duration and incidence, as related to disease indication, disease amelioration by a medicine, and the adverse effects ascribed to the medicine. The TURBO model is a quantitative and graphical approach to benefit–risk analysis, developed by Amery. This method tries to quantify the benefits and risks of a medicine in a given indication, and both factors are then displayed on the TURBO diagram. Beckmann described a model in which the evidence-weighted benefit is balanced against the sum of all evidence weighted risks. All three models are reviewed in detail in Chapter 3. It is however important to note that none of these models have been well-accepted nor are they used in practice by regulatory authorities.

Meyboom and Egberts (1999) suggested that in an ideal situation, when extensive or even complete qualitative and quantitative data are available, it should be possible to calculate the positive and negative effects of a medicine, using quality-of-life assessment methodology. The cumulative risk of a medicine can be calculated as the Drug-Attributed Loss of Quality-Adjusted Life Years (DALQALY). Similarly, the benefit can be presented as Drug-Attributed Gain of Quality-Adjusted Life Years (DAGQALY). If benefit and risk are expressed in the same measure (standard unit), the net benefit of the medicine can be calculated by subtracting the risk from the total benefit. Otherwise an equation can be made (of risk units per unit of benefit). However, the collection of the body of evidence needed for such assessments requires large-scale, intensive and continuous data-gathering made into a routine in medical and pharmaceutical practice, which is not a reality today. Nevertheless, quality-of-life measures such as SF-36 are increasingly used for measuring health outcomes in clinical research (Garratt *et al*., 2002).

It should be noted that regulatory authorities seem to increasingly advocate the use of 'numbers needed to treat' (NNT) in product labelling, and hence in regulatory decisions. NNT is defined as the reciprocal of the absolute risk reduction due to a given therapy and conveys the effort that must be expended on average to accomplish a single, tangible, and positive outcome in a patient (Mancini and Schulzer, 1999). A quantitative method for benefit–risk assessment based on an adjusted NNT approach has been reviewed (Holden, 2003). The concept and applicability of NNT will be discussed in detail in Chapter 3.

Another quantitative approach called minimum clinical efficacy (MCE) was first introduced by Djulbegovic *et al*. (1998) as a means of choosing high-dose or low-dose chemotherapy for the adjuvant treatment of breast cancer, and was fine-tuned for general use and applied by Holden *et al*. (2003a; 2003b). MCE of a new treatment is the minimal clinical efficacy needed for it to be worth considering as an alternative treatment after taking into account: (1) the efficacy of the standard treatment; (2) the adverse event profiles associated with the standard treatment and the new treatment; and (3) the risk of disease associated with no treatment. This method will be reviewed in Chapter 3.

In Japan, the use of the YUYOSEÏ instrument in clinical trials and hence in marketing authorization applications is mandatory. The YUYOSEÏ instrument is an assessment whereby each patient is considered as a whole with their safety and efficacy experience in a trial combined to obtain an overall outcome measure (Chuang-Stein, 1999). It is a global assessment given by the treating physician to reflect, in the physician's opinion, an individual's overall response to treatment. By its very definition, YUYOSEÏ has objective and subjective components, since the sense of well-being of a patient can be quite subjective. However, research has demonstrated that this utility evaluation system is unreliable because of poor intra-rater as well as inter-rater reproducibility (Naito, 1994).

In the absence of general standard models for benefit–risk assessment by regulatory authorities (except in Japan), it should also be noted that there are no such models for specific categories of medicines either. There is for example no standardized method of incorporating toxicity and tolerability information into decisions regarding the regulatory approval of new cancer treatments (Rothenberg *et al*., 2003). Hence there are no easy answers to questions such as whether a new agent that improves median survival time or time-to-tumour progression by a statistically significant 6-week margin be approved, even if it is associated with substantial rates of treatment-induced morbidity or mortality. Arlett suggested that for thrombolytic agents for the treatment of acute myocardial infarction, an estimate of the number of serious cases of haemorrhage per infarct prevented could be calculated when evaluating the balance of benefits and risks (Arlett, 2001). Methods for assessing the benefit–risk profile of oral triptans, vaccines and selective oestrogen receptor modulators are reviewed in Chapter 3.

Use of models by prescribers

It is recognized that prescribers need validated and interpreted data on the benefit–risk assessment of medicines, which should clearly define the benefit–risk relationship and put it in perspective with other therapeutic means. In addition, such data are required for ensuring that patients receive accurate and transparent information on the benefits and risks of medicines (Lehner *et al*., 2001). The intended use of benefit–risk assessments in the setting of clinical trials, be it individual clinical trials or all available clinical trials for a given medicine, is usually as a population-based assessment (Costantino, 2001). In contrast, in the clinical practice setting the intended use of benefit–risk assessment is as an individual patient-based assessment (Spilker 1994; Edwards and Hugman, 1997).

Costantino developed a tool for communicating to individual patients the results of the benefit–risk assessment of selective oestrogen receptor modulators (SERMs) like tamoxifen. He was of the opinion that without such a tool the communication

of a benefit–risk assessment to an individual woman can be a difficult task because of the complicated interplay of numerous potential beneficial and detrimental effects of SERMs. The properties of an ideal tool were identified as a tool that: (1) avoids to the extent possible the use of probabilities or relative risks as these are entities that can be confusing to those not familiar with such statistical measures; (2) includes information on all potentially affected endpoints with consideration of effects grouped by the relative severity of the different endpoints; and (3) provides a presentation that summarizes the results of the benefit–risk assessment that is limited to one page. The approach to providing information in the tool is based on describing the effects of therapy in terms of the number of cases expected to occur over 5 years of follow-up among 10 000 women of the same race, age group, and predicted breast cancer risk as the individual being assessed. The tool provides a comparison of the number of cases expected if none of the 10 000 was treated with the number of cases that may be prevented or caused if all 10 000 women were treated. Although the methodology was specifically developed and applied to the use of tamoxifen, it is a generalized procedure that can be readily modified and applied to other forms of therapy including other SERMs.

Loke *et al*. (2003) developed a benefit–risk analysis of aspirin therapy which allowed quantifying the reduction in cardiovascular events and the increase in gastrointestinal haemorrhage in an individual patient, according to the individual risk profile. The absolute treatment effects were calculated from estimates of baseline risks for cardiovascular events and gastrointestinal haemorrhage, and relative risks of these events with treatment. Baseline cardiovascular risks were derived from existing risk scores, and baseline risk of gastrointestinal haemorrhage from an observational cohort study. Changes in relative risk were obtained from clinical trial data. The use of a graphical technique incorporating these baseline risks should enable doctors and patients to assess easily the benefit–risk ratio of aspirin in a particular individual.

Conclusions on the existing models

Based on this review of the literature, it is apparent that there are relatively few general methods and models for benefit–risk assessment, in particular for use by regulatory authorities. None of these methods seem to be widely accepted and applied. In general, there appears to be a lack of a methodological framework and guidance with regard to the relevant benefit and risk criteria to be considered for benefit–risk assessments, and with regard to trading-off the benefits and risks of medicines. Furthermore, there appears to be room for a better general, broadly applicable method for benefit–risk assessments of medicines.

1.9   DISCUSSIONS WITH STAKEHOLDERS ON THE CONCEPTS AND MODELS FOR BENEFIT–RISK EVALUATION

The above discussed framework, concepts and models for benefit–risk evaluations were extensively discussed with stakeholders, including representatives from the pharmaceutical industry and from health authorities, in two workshops organized by CMR International Institute for Regulatory Science. The conclusions of both workshops, which further elaborate on and illustrate the content of this chapter, are included in Appendix 1.