From data to models : reducing uncertainty in benefit risk assessment : application to chronic iron overload in children

Bellanti, F.

CHAPTER 1

Integration of PKPD relationships into Benefit-Risk Analysis

F Bellanti, RC van Wijk, M Danhof, and O Della Pasqua

Br J Clin Pharmacol – 2015

Summary

Aim: Despite the continuous endeavour to achieve high standards in medical care through effectiveness measures, a quantitative framework for the assessment of the benefit-risk balance (BRB) is lacking prior to drug approval. The aim of this short review is to summarise the approaches currently available for benefit-risk assessment. In addition, we propose the use of pharmacokinetic-pharmacodynamic (PKPD) modelling as the pharmacological basis for evidence synthesis and evaluation of novel therapeutic agents.

Methods: A comprehensive literature search has been performed using MESH terms in Pubmed, in which articles describing benefit-risk assessment and modelling and simulation (M&S) were identified. In parallel, a critical review of multi-criteria decision analysis (MCDA) is presented as a tool for characterising a drug's safety and efficacy profile.

Results: A definition of benefits and risks has been proposed by the European Medicines Agency (EMA), in which qualitative and quantitative elements are included. However, in spite of the value of MCDA as a quantitative method, decisions about BRB continue to rely on subjective expert opinion. By contrast, a model-informed approach offers the opportunity for a more comprehensive evaluation of BRB before extensive evidence is generated in clinical practice.

Conclusions: BRB should be an integral part of risk management and as such considered prior to drug approval. M&S can be incorporated into MCDA to support the evidence synthesis as well evidence generation taking into account the underlying correlations between favourable and unfavourable effects. In addition, it represents a valuable tool for the optimisation of protocol design in effectiveness trials.

1.1 Benefit-Risk Analysis: the current situation

Despite the recognised implications of unmet medical needs and challenges in dealing with new diseases, the current regulatory framework in the European Union has made drug approval a demanding task. This situation is compounded by emerging safety findings, which have led to post-approval withdrawals of more than a dozen products with high therapeutic potential in the past decade (1,2). Such a landscape places regulators, clinical scientists and drug developers with yet another dilemma: how to balance rapid access to new drugs *versus* gathering comprehensive data on efficacy and safety? (3). Currently, regulators make these decisions in an isolated, fragmented, and to a large extent subjective manner.

The decision to approve a new medicinal product is based on the assumption that a systematic review of all available data provides an accurate, unbiased picture of a drug's efficacy and safety. This assumption may, however, not be true for the large majority of drugs; the evidence generated to support regulatory submission does not always account for the overall heterogeneity of the target population, the impact of treatment on disease progression or external confounding factors on treatment response. Moreover, one needs to acknowledge that the information gathered in the context of pivotal clinical trials may not provide evidence that dose selection, dosing regimen, and treatment duration are truly optimal.

Undoubtedly, efficient gathering and use of data are required to answer the clinical questions that arise with new drugs or therapeutic interventions. Among other things one needs to distinguish effectiveness from clinical response. In addition, it is crucial to understand whether there is added value, as compared with other treatments. These are multidimensional questions which require clear understanding of how data will be generated and how benefit and risk will be quantified. Whereas different theoretical considerations and techniques have been used by health technology assessment agencies, a clear framework for benefit risk (BR) assessment is still lacking during drug development and subsequently for regulatory approval. Consequently, decision making at important milestones in R&D and at submission remain empirical, inconsistent and more often than not, non-transparent(1,4–8).

In the past years awareness about the aforementioned issues has increased significantly. Several projects (9–13) have been funded to evaluate some of the available methodologies and better understand the requirements for a more systematic approach to BR analysis. In this context, the work of the Committee for Medicinal Products for Human Use (CHMP) of the European Medicines Agency (EMA) is particularly relevant. Starting in 2006, a working group was installed to examine the issue and provide recommendations about ways to improve BR assessment, including aspects such as transparency, consistency and

communication between stakeholders (9). Among the techniques evaluated by the working group, quality-adjusted life years (QALYs) and number needed to treat (NNT) were found to be the most used concepts in clinical practice, very likely due to their simplicity (9,14). However, these methods are qualitative in nature and as such lack some important features that allow one to make appropriate inferences about quantitative differences, especially when comparing treatment options. There is a clear need for more comprehensive methodologies, which enable better integration of data and facilitate the evaluation of complex clinical scenarios that arise in real life.
Most of these complexities seem to have been addressed by the development of multi-criteria decision analysis (MCDA), an integrative approach that has gained interest from the scientific and clinical community over the last few years. From 2009 to 2011, data can be found for nine products which have been evaluated by MCDA alone, or in combination with simulation, decision trees or Markov modelling (15).

In this review, a brief overview of different techniques for the evaluation of benefit and risk is presented, with especial focus on the contribution of quantitative methodologies to the development and approval of novel medicines. Two main topics are discussed initially. First, the definition of benefit-risk balance and the impact of qualitative and quantitative methodologies on the measurement of benefit and risk during the drug development process (7). In addition, we consider further refinement of the approaches used for assessing BRB by integrating it with pharmacokinetic-pharmacodynamic (PKPD) modelling. It is envisaged that modelling and simulation may account for correlations between therapeutic response and adverse events, providing a biologically plausible basis for BRB. The availability of such an integrated approach may enable better choices regarding treatment selection and dose rationale in special groups or conditions involving small numbers of patients such as rare diseases.

1.2 Methods

Initially, an exploratory literature search was performed to retrieve relevant publications to identify current quantitative approaches for benefit risk assessment (BRA) to improve decision making in drug development. Seven documents (1,4,14,22,23,61,84) were available before the exploratory phase and were used to identify 58 articles, books and reports. Based on this pool of 65 documents 21 quantitative methodologies were identified (see Table S1). This result was used to integrate the available information with a systematic literature search within PubMed, in which the name of the methodology was combined with the term benefit risk assessment/analysis, which was replaced by benefit risk or risk assessment when the query lead to an outcome of 0 publications. This resulted in 253 publications, of which 231 were rejected based on title and abstract information. The resulting 22 publications,

together with the 65 publications from the exploratory search were reviewed. Additionally, 23 publications were added on external advice, for a total number of papers reviewed in the context of current approaches for benefit risk assessment equal to 110. The steps described in this paragraph are summarised in Figure 1.

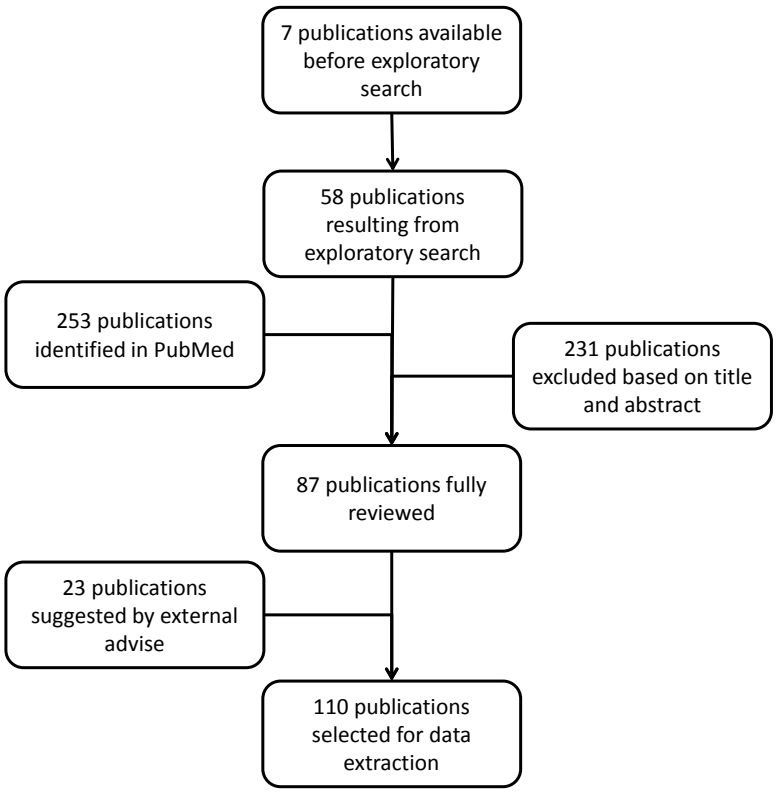

Figure 1. Flow diagram of the literature search.

1.3 Definition of benefit and risk

An important aspect of any BR analysis is the definition of both terms, and more importantly, how to measure or quantify them. Benefit is usually described as a potential effect that moves the condition of the patient from disease towards health, within a given (pre-defined) context (Table 1) (16–19). Risk is the opposite, a potential effect that moves the condition of the patient from health towards disease, also within a pre-defined context. To measure both possibilities, at least two concepts play an important role: the magnitude or severity of the effect, and its incidence or frequency. Benefit or risk is then estimated by the product of these concepts, possibly multiplied by the duration (17) or the reliability of the data (19). The BR assessment, in which the no-treatment option should not be overseen, is simply a ratio of the two components, for which pre-defined acceptance thresholds are stated.

Table 1. Glossary of terms.

Term	Definition
ADE	Adverse Drug Effects
Bayesian statistics	Probability-based statistics, concerning parameter values derived from distributions
Benefit	Favourable effect, accounting for uncertainty of that effect [as defined by the EMA]
BILAG-index	British Isles Lupus Assessment Group, a measure for severity of SLE
BR	Benefit Risk
BRAT	Benefit Risk Action Team, operating under PhRMA
CHMP	Committee for Medicinal Products in Human Use, operating under the EMA
Decision tree	Method to aid decision making by visualizing different scenarios as a series of events, and by calculating outcome based on assigned probabilities of the events
DSD	Death or serious disabled, measure of estimated outcome in the swine flu case study
EMA	European Medicines Agency
FDA	Food and Drug Administration [USA]
H1N1	Influenza virus categorized by surface proteins hemagglutinin and neuraminidase (in this case swine flu)
In silico	Experiment in a computer, virtually
In vitro	Experiments in cell cultures
In vivo	Experiment in animals (preclinical)
IPRED	Individual prediction, possible outcome of PKPD modelling prediction variables and parameter values of an individual patient
M&S	Modelling and Simulation, in pharmacology a way of describing data by constructing a validate model and simulate new data, as a virtual experiment
Markov model	Quantitative method of modelling states and transitions between states
MCDA	Multi criteria decision analysis, quantitative method analysing single weighted components of a problem before reassembling it to aid a final decision
NDA	New drug application, to be submitted to the FDA for approval before market access

NNH	Number needed to harm, measure of the number of patients that has to be treated to present a single adverse effect
NNT	Number needed to treat, measure of the number of patients that has to be treated to prevent a single occurrence
PhRMA	Pharmaceutical Research and Manufacturers of America
PKPD	Pharmacokinetics and pharmacodynamics, two disciplines within pharmacology concerning what the body does to the drug and what the drug does to the body, respectively
PrOACT-URL	Qualitative framework by Hammond, Keeney and Raiffa, consisting of Problem, Objective, Alternatives, Consequences, Trade-offs, Uncertainty, Risk tolerance and Linked decisions
QALY	Quality-adjusted life year, measuring the outcome of therapy by the adjustment of a quality life year, in which the patient can fully function (economically)
Risk	Unfavourable effect, accounting for uncertainty of that effect [as defined by the EMA]
RV-NNT	Relative Value adjusted number needed to treat, a type of NNT accounting for patient preference as value function
SLE	Systemic Lupus Erythematosus, an autoimmune disease
SLEDAI	DLE Disease Activity Index, a measure of severity of SLE
TURBO	Transparent Uniform Risk-Benefit Overview

Currently a slightly different definition of benefit and risk has been adopted by the EMA. They are defined respectively, as favourable and unfavourable effects and are at the same time coupled to the uncertainty of both effects (Figure 2) (14). Whereas the reasoning seems intuitive, this situation represents a mathematical challenge, i.e., integrating terms or factors that are measured in incommensurable units and in different time scales. Any reliable product of these factors imposes data manipulation or transformation to ensure that all terms are expressed in the same unit and time scale. However, the illusion of this mathematical precision tends to hide another important conceptual challenge: what is acceptable? (18). This depends on the perception and values of the stakeholders, i.e., the regulator, the clinical experts, and the patients. Procedures have been devised to ensure that perceived benefits and risks are quantified in a systematic manner. This process is known as prior elicitation and involves expert judgment. It is aimed at making subjective opinions more consistent, comprehensive and transparent (16,19,20).

Favourable effects	**Uncertainty of favourable effects**
Unfavourable effects	**Uncertainty of unfavourable effects**

Figure 2. EMA's definition of benefit and risk, where favourable effects are beneficial to the population and unfavourable effects are undesirable for the population. Uncertainty is caused by variation, biased data, limitations of data or methodology etc. Based on [14].

1.4 Current approaches

The assessment of benefit and risk has evolved in a rather empirical manner and still relies on subjective criteria, in that perceived benefits and risks depend on the context in which the treatment is used, i.e., which standard of care is set as reference and whether short and long term consequences of the intervention are considered against the progression of disease and any correlated co-morbidities or complications. Irrespective of the lack of consensus on how to assess and weight any measures associated with benefit and risk, one needs to consider two different dimensions of the problem. First, a qualitative approach is required to allow for explicit contextualisation of the problem. It is crucial to fully understand the main issues before any quantitative analysis starts, i.e., to identify the factors that contribute and/or determine benefit and risk as well as capture the views and differences of opinion from different stakeholders, especially with regard to the perception of risk, in terms of its incidence, severity, chronicity and reversibility. Second, a quantitative approach is needed in which results from the initial (qualitative) evaluation are normalised by means of mathematical and statistical procedures. Such a normalisation implies the availability of sufficient data for those endpoints and measures which have higher weights. It also imposes clear understanding of the trade-offs between benefit and risk, especially of the correlations between outcomes. Whereas these requirements seem obvious, little attention has been paid to the biological or pharmacological basis that determines treatment outcome, i.e. how exposure-response (PKPD) relationships underpin favourable and unfavourable events.

The next paragraphs will provide an overview of the available techniques, including recent examples in which benefit and risk have been evaluated in the context of regulatory

approval and treatment optimisation. Additional details of the methodologies can be found in the supplementary material (Supplementary Figure 1 and 2 and supplementary Table 1 and 2, see Appendix).

Qualitative approaches

A qualitative framework is essential for characterising benefit and risk, as it structures the problem and its context, before any actual assessments are made. It provides clarity about the possible outcomes of the assessment, as well as the input and the process in between, for example by defining which decision criteria are to be used. This framework ensures that no alternative measures or trade-offs are overlooked during the subsequent steps, i.e., during which quantitative methods are applied.

PhRMA BRAT

The Pharmaceutical Research and Manufacturers of America (PhRMA) assigned a Benefit Risk Action Team (BRAT) to create a decision framework. Their framework consists of six steps which are developed and implemented prior to drug approval. Before phase III, focus is given to the definition of a decision frame, identification of relevant outcomes, identification of the data sources, and customisation of the framework for BR analysis. At the time of filing and NDA review, attention is paid to the outcome itself as well as to the quantification and interpretation of key BR metrics (13,14). It should be noted that this framework seems to end with the decision and defence after which a drug is approved or rejected. It does not involve post marketing data, which are known to potentially change BR balance.

EMA PrOACT

The qualitative framework suggested by the EMA is based on Hammond's, Keeney's and Raiffa's PrOACT approach (21), combined with the less known addition of the so-called URL: Problem, Objective, Alternatives, Consequences, Trade-offs, Uncertainty, Risk tolerance and Linked decisions. In this way, the problem is clearly structured and information can be gathered in a consistent way to assist the decision-making process (14). Despite its general nature, the use of PrOACT-URL has proven its success since 1999. In contrast to PhRMA BRAT, the inclusion of uncertainty paves the way for a more statistically sound implementation.

Quantitative approaches

The use of a qualitative framework for assessing benefit and risk may be sufficient when complexity is minimal. This is however not the case in drug development where very complex scenarios arise. To include all data and present a sound overview of all alternatives, consequences and trade-offs, as well as differentiate between objectives otherwise

considered comparable, one or more quantitative techniques are required (1,4,11,14,17,19,22,23). A qualitative framework will still be essential to define the problem and the objectives of the analysis and as such will precede the implementation of a quantitative BR analysis.
In the past decades several methodologies have been developed and used to evaluate the BR balance of a number of drugs. These methodologies present completely different features and their use has been tailored for very specific cases, contributing to an increase in the number of options available when starting an analysis. These specificities have however made them unsuitable for subsequent application in a general BR framework. An overview of these methods (1,4,17,19,22–102, 120-124), including advantages and limitations is provided in Supplementary table 1. By contrast, multi criteria decision analysis (MCDA) in combination with decision trees has been suggested as a plausible quantitative approach that embeds the needed features for a generalised and structured framework for BR evaluation.

MCDA presents several advantages compared to other methodologies: the main one is the simplification of a complex problem by breaking it into smaller pieces and making them comparable by weighting their scores on a single scale; normalizing the different criteria allows comparison on the same ground. In addition, the uncertainty carried by the subjective component, is further reduced by the possibility of performing sensitivity analysis, in which the model provides different outcomes depending on weights variation. There are, however, still limitations. Given the complexity of the scenarios analyzed, it is often expected to observe correlations between the endpoints considered. This is not yet taken into account within the methodology, where each endpoint is analyzed in an independent manner. In the SLE-case, which is discussed in the supplementary material, the immunosuppressive effect of Benlysta and the incidence of infection might very well be correlated in a nonlinear way. This might influence the outcome, leading to biased results. Furthermore, it is a matter of concern how the input data for the decision model is provided. This is not a direct limitation of the methodology, but of how the analysis is implemented. Many quantitative methods are limited by statistics and inclusion of uncertainty, confounding factors, or limited data. The latter concerns both the experimental data, as well as preference values of different stakeholders required for weighting criteria (1). MCDA offers a statistical sound method, where probability and uncertainty are combined with preference. Its limitation lies in the complexity of data required, which is often unavailable, as well as in the subjective judgement that is required and the dependence on risk perception differences. Besides, sequential decisions require data gathering over a longer time period, especially in conditional approval (103).

Despite the aforementioned advantages, MCDA, like any other quantitative method, still relies on subjectivity. This is partly overcome by structuring the analysis in a transparent, consistent manner and by incorporating communication with different stakeholders as a critical step (14,15). In fact, communication with different stakeholders is also accounted for in NNT/NNH. Although applicability of the former to BR assessment in general is very limited because of the lack of preference data, as well as the limited statistical power (57,58), it shows an important issue in communication. Individual patients seem unable to objectively estimate their own chances. In a distribution of 1 out of 20, all 20 patients expect to be the exception, when it comes to a beneficial effect, but not in case of an adverse effect. As a result, the magnitude of risk is misperceived, as the chances of common consequences are underestimated and those of rare consequences are overestimated (8). This problem of risk perception is essential when considering including different stakeholders. Although MCDA does present data in a transparent and consistent way, it is not a technical process, but an effective design of the social processes required for subjective weighting (41).

1.5 Integration of PKPD modelling into BR analysis

Modelling and Simulation (M&S) techniques represent an invaluable resource for drug development. Of relevance for BR analysis is the opportunity that PKPD modelling offers in terms of describing variability in a parametric manner. This allows the characterization and prediction of the time course of treatment response at individual level under physiological and pathological conditions (104,105). The current emphasis on mechanism-based modelling has also the advantages of increased understanding about drug-specific and system-specific properties such as, target site distribution, binding, pharmacokinetic interactions, transduction of signals, pharmacodynamic interactions, homeostatic feedback, tolerance and disease progression (106–108). In addition, model-based simulations can provide insight into conditions that may not have been tested experimentally, unravelling patterns or responses that may represent clinically relevant changes in the BR balance.

From a technical, scientific point of view, M&S ensures for integration of data and knowledge in a continuous, objective and reproducible manner, thereby enhancing the quality of decision making (105). Over the last decade, regulatory perception and role of M&S in drug development has changed. Its relevance in clinical development has been acknowledged and processes are in place to support a more structured use of M&S (106,109).

In the next paragraphs we evaluate how the integration of M&S can be advantageous to further improve the existing framework for the evaluation of benefit-risk balance, as suggested by the EMA. To this purpose, we consider three main aspects, namely, the optimisation of evidence that is generated by clinical trials, evaluation of virtual scenarios

and mechanism-based multivariate analysis. The optimisation of the input data available for decision making entails not only the integration of data from different trials, but also the use of optimality concepts for the design of prospective clinical studies. The availability of an integrated model allows for the creation of virtual experiments, which provide a more coherent, biologically plausible basis for performing interpolations and extrapolations. In contrast to current practice, multivariate modelling allows one to establish correlations between therapeutic and adverse events of interest, which are often linked by the very pharmacological nature of the treatment. This overview is complemented by a brief discussion of the issues associated with prior elicitation, which could be better guided by the use of models, rather than empirical distributions. As such, a model-based approach could provide somewhat less subjective weighting and preferences.

Optimizing input data: M&S techniques can be used to optimize the input data available for the BR analysis. PKPD modelling allows the creation of a framework that can be refined and improved throughout the development process, by integrating data from different sources as well as by pooling the information gathered across different phases of development. This iterative process allows one to understand and distinguish drug from system-specific properties. Most importantly, it allows one to identify sources of variation and assess the clinical implications thereof. Among other things, BR analysis could be performed with and without the residual variability or in by inclusion of variability in a stepwise manner. In other words, these procedures increase the value of data whilst decreasing uncertainty (106). On the other hand, M&S can also be used to optimise the design of prospective clinical trials. The quality of the information collected can be considerably improved through optimal design (110–112), enabling the generation of more informative data input for the decision analysis. This is particularly important in special populations where limited evidence is generated, such as in paediatric diseases (106,113,114). The assumptions about the informative value of data obtained from randomised clinical trials are often overlooked. It is assumed that the output or results from a trial are consequence of the drug treatment, rather than the consequence of the interaction between drug properties, disease processes, patient characteristics and experimental protocol.

Evidence from virtual scenarios: A second aspect that could be beneficial for the BR assessment is the use of PKPD modelling for simulation purposes. The availability of a qualified or validated model may provide the opportunity to perform virtual experiments. This allows one to explore scenarios that have not been evaluated during clinical development. Not only efficacy and safety data can be considered, but also the influence of covariates such as disease severity, co-medications, co-morbidities and drug compliance can be evaluated. By inter- or extrapolating, new input data can be generated for a different

population or different dosing regimens. As such these simulated results can be subsequently used as input for BR analysis. As mentioned previously, PKPD modelling may have an even larger impact when considering special populations (114–117).

Correlating multiple endpoints: Thus far we have highlighted the fact that PKPD modelling may reduce the uncertainty in a BR analysis by optimising the information used as input. M&S techniques may overcome another important limitation of BR methodologies, namely the assumption that favourable and unfavourable events are clinically, pharmacologically and statistically independent from each other. This assumption violates our current understanding of the nature and cause of adverse events. Hence, any analysis involving multiple endpoints in a multidimensional system will have to account for the correlations between them. Moreover, we believe that these correlations are often non-linear, requiring some advanced statistical techniques to ensure that interactions between variables and covariate factors are captured accordingly. Multidimensional models can be used to assess quantitatively how endpoints are linked together and how response changes with changes in drug exposure (24).

Advantages from the integration of M&S techniques to BR analysis are not only conceptual. From a technical perspective, PKPD models may contribute to bias reduction during prior elicitation. In addition, it may provide a stronger basis for sensitivity analysis. Although weighting is a subjective procedure, expert opinions can be modelled using prior elicitation. Moreover, if the uncertainty associated with the weights is assessed, it is possible to factor in the impact of each expert's opinion on the overall analysis. Other possibilities exist to weight the experts input, by scaling their precision based on training and experience, or by assigning them to groups of thought that are more or less representative of the common opinion (26,63). PKPD models describing the underlying disease processes as well as the impact of treatment over time through virtual scenarios may facilitate prior elicitation, providing systematic, consistent input for the evaluation of weights and uncertainties.
An example of the impact of M&S concepts on BR analysis is given in Table 2.

Table 2. Impact of M&S on the MCDA approach is visualised and further elucidated by the example of Benlysta. It shows clearly the emphasis on the first part of the methodology; the input data and earlier data evaluation for correlations between parameters and outcomes.

MCDA	**Modelling & Simulation**	**Example: Benlysta**
Step 0: Input data gathering	Step 0.1: Explore and refine the informative contents of data, accounting for variability and uncertainty. Step 0.2: Incorporation of virtual measurements (samples), by evidence generation through simulations.	Step 0.1: Distinguish between-subject variability in relevant parameters from residual error. Step 0.2: Evaluate parameter uncertainty by exploring the implications of different experimental protocol conditions.
Step 1: Defining decision context	Step 1.1: Prioritising elements which affect variability and /or uncertainty.	
Step 2: Identifying options	Step 2.1: Inference by extrapolation, e.g., an additional arm that has not been tested clinically.	Step 2.1: Assess treatment response for alternative dosing regimens than the actual treatment arms in the trial (i.e., 1 and 10 mg doses)
Step 3: Identifying objectives and criteria	Step 3.1: Assess outcomes taking into account the correlation between events.	Step 3.1: The correlation between immunosuppressive effects and incidence of infection can be incorporated into the model, enabling accurate evaluation of the impact of different dose levels on outcome.
Step 4: Scoring	Step 4.1: Estimation of the correlation between events in a parametric manner, thereby avoiding biased scoring of the data.	Step 4.1: Estimation of the parameters describing the nonlinear relationship between immunosuppressive effects and incidence of infection in patients undergoing long term treatment.
Step 5: Weighting factors for differences of opinion	Step 5.1: Prior elicitation of expert opinions can be translated into consistent weighting, including distributions describing differences of opinion (e.g., priors in parameter distributions).	Step 5.1: Simulate outcomes for Benlysta-treated patients taking into account different weighting factors.
Step 6: Combining data	Step 6.1: Simulated scenarios increase the quality of the data and therefore the quality of the overall value, by increasing	Step 6.1: Simulation of different treatment arms to explore the implications of dose selection.

	granularity.	
Step 7: Examining data	Step 6.1: Outcome evaluation is not limited to the data, but to evidence arising from virtual clinical trials, including patients who belong to risk groups (e.g., those who meet exclusion criteria)	
Step 8: Sensitivity analysis	Step 8.1: Irrespective of the decision criteria, model parameters on which the data are based can also be analysed.	Step 8.1: The PKPD model of Benlysta has been evaluated by sensitivity analysis.

1.6 Discussion and conclusion

In this short review, an overview was given of the methodologies currently used for the evaluation of BR balance. Growing consensus suggests that a combined approach involving qualitative and quantitative methods is required to ensure meaningful evaluation and interpretation of benefit and risk data. In fact, this is recommended by the EMA, which suggests the use of PrOACT-URL and MCDA.

Even though a more structured approach is still lacking for BR analysis, MCDA seems to address the need for a multidimensional characterisation of the scenarios that arise in drug development and in the clinical practice. One of its limitations is the way uncertainty is handled; there is the need to further reduce the uncertainty or preferably to capture it accordingly. Attempts have been made to construct stochastic multi-attribute models, also known as stochastic multi-criteria acceptability analysis (SMAA), which incorporates uncertainty regarding the criteria measurements. SMAA provides the possibility to include the sampling variation and to characterize typical trade-offs supporting a drug BR profile without knowing or eliciting the (exact numerical) preferences beforehand (119). An analysis without preference information is valuable when preferences cannot be elicited or when the potential benefits of a drug have to be assessed across a wide range of preferences. This latter situation occurs, for example, when different subgroups of patients are considered. However, stochastic methods do not eliminate discrepancies between perceived risk or benefit and their biological and pharmacological plausibility. Undoubtedly, integration of mechanism-based modelling to multi-criteria decision methods will enhance our ability to characterise benefit-risk balance. It will provide indirect evidence from virtual scenarios in a more effective manner than sensitivity analysis and other statistical techniques have allowed for. Such an integrated approach will also represent an advancement for the field of modelling and simulation, which is often restricted to single endpoints, facilitating the assessment of causality and correlation between favourable and unfavourable events (118). Unfortunately, in literature there are very few examples that present in a clear manner the concepts discussed throughout this manuscript. Among them though, two publications

provide an excellent illustration of these concepts: the work carried out by Bender et al (125) shows how exposure-response relationships quantified through model-based approach for multiple endpoints can be used to explore and assess BR across different dosing regimens in the context of oncology trials. In the same way, the work by Pink et al (126) shows the feasibility of integrating M&S with pharmacoeconomic analysis to inform decision making throughout the whole drug development process and possibly achieve personalised evaluations. Both examples support the fact that PKPD relationships are crucial in the assessment of a drug efficacy and safety and should not be omitted when performing a BR appraisal.

In addition, we propose here the use of PKPD modelling as the pharmacological basis for evidence synthesis and evaluation of novel therapeutic agents. Various methodologies are available for evidence synthesis, and among them network meta-analysis (NMA) has been widely used in BR analyses to combine all available evidence (127-128). These approaches though, rely on very large amount of information and as discussed in this manuscript depend only on the evidence generated. As opposite to a model-based approach, they are not able to provide an understanding or a quantification of the underlying PKPD mechanisms and subsequently cannot be used to anticipate and explore virtual scenarios through Clinical Trial Simulations and/or Not-in-trial Simulations (129). In a post-marketing phase the contribute of NMA is indeed invaluable but in a pre-marketing evaluation where limited data is available PKPD cannot be ignored and to our understanding may be crucial for a comprehensive BR evaluation.

In conclusion, it should be highlighted that models do not make decisions, people do. Ultimately, patients, clinicians, drug developers and regulators need to acknowledge that decisions are better made when data are presented and communicated in a clear, systematic manner. PKPD modelling can complement evidence generation by providing stakeholders the opportunity to explore conditions that have not been experimentally tested at the time of BR analysis. Regardless of the limitations models and simulation scenarios may have, model-based evaluation is likely to outperform gut feeling, which often prevails in clinical decision-making.

References

1. Guo JJ, Pandey S, Doyle J, Bian B, Lis Y, Raisch DW. A review of quantitative risk-benefit methodologies for assessing drug safety and efficacy-report of the ISPOR risk-benefit management working group. Value Health. 2010 Aug;13(5):657–66.

2. Busch M, Walderhaug M, Custer B, Allain J-P, Reddy R, McDonough B. Risk assessment and cost-effectiveness/utility analysis. Biol J Int Assoc Biol Stand. 2009;37:78–87.

3. Eichler H-G, Pignatti F, Flamion B, Leufkens H, Breckenridge A. Balancing early market access to new drugs with the need for benefit/risk data: a mounting dilemma. Nat Rev Drug Discov. 2008;7:818–26.

4. Holden WL. Benefit-risk analysis : a brief review and proposed quantitative approaches. Drug Saf. 2003 Jan;26(12):853–62.

5. Bjornson D. Interpretation of drug risk and benefit: individual and population perspectives. Ann Pharmacother. 2004;38:694–9.

6. Okie S. Safety in Numbers — Monitoring Risk in Approved Drugs. N Engl J Med. 2005;352:1173–6.

7. Califf R. Benefit assessment of therapeutic products: the Centers for Education and Research on Therapeutics. Pharmacoepidemiol Drug Saf. 2007;16(5-16).

8. Cohen J, Neumann P. What's more dangerous, your aspirin or your car? Thinking rationally about drug risks (and benefits). Heal Aff. 2007;26:636–46.

9. EMA. Benefit-risk methodology project - Project description. 2009.

10. Liberti B, McAuslane N, Walker S. Progress on the Development of a Benefit / Risk Framework for Evaluating Medicines. Regul Focus. 2010;15:32–7.

11. FDA. Structured Approach to Benefit-Risk Assessment in Drug Regulatory Decision-Making - Draft PDUFA V Implementation Plan. 2013.

12. Walker S, McAuslane N, Liberti L, Salek S. Measuring benefit and balancing risk: strategies for the benefit-risk assessment of new medicines in a risk-averse environment. Clin Pharmacol Ther. 2009;85:241–6.

13. Coplan P, Noel R, Levitan B, Ferguson J, Mussen F. Development of a Framework for Enhancing the Transparency , Reproducibility and Communication of the Benefit – Risk Balance of Medicines. Clin Pharmacol Ther. 2011;89:312–5.

14. EMA. Benefit-risk methodology project - work package 2 report: Applicability of current tools and processes for regulatory benefit-risk assessment. 2011 p. 1–33.

15. Phillips L. The European Medicines Agency Benefit-Risk Project (presentation). 2013.

16. Lane D, Hutchinson T. The Notion of “Acceptable Risk”: The Role of Utility in Drug Management. J Chronic Dis. 1987;40:621–5.

17. Edwards IR, Wiholm BE MC. Concepts in Risk-Benefit Assessment. A Simple Merit Analysis of a Medicine? Drug Saf. 1996;15(1):1–7.

18. Ernst E, Resch K. Risk-Benefit Ratio or Risk-Benefit Nonsense? J Clin Epidemiol. 1996;49:1203–4.

19. Beckmann J. Basic aspects of risk-benefit analysis. Semin Thromb Hemost. 1999 Jan;25(1):89–95.

20. Morgan G, Henrion M. Uncertainty - A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press; 1990.

21. Hammond J, Keeney R, Raiffa H. Smart Choices. Harvard Business School Press; 1999.

22. Curtin F, Schulz P. Assessing the benefit:risk ratio of a drug - randomized naturalistic evidence. Dialogues Clin Neurosci. 2011;13:183–90.

23. Puhan MA, Singh S, Weiss CO, Varadhan R, Boyd CM. A framework for organizing and selecting quantitative approaches for benefit-harm assessment. BMC Med Res Methodol. BMC Medical Research Methodology; 2012 Jan;12(1):173.

24. Ahn JE, French JL. Longitudinal aggregate data model-based meta-analysis with NONMEM: approaches to handling within treatment arm correlation. J Pharmacokinet Pharmacodyn. 2010 Apr;37(2):179–201.

25. Airoldi M, Morton A. Adjusting life for quality or disability: stylistic difference or substantial dispute? Health Econ. 2009;18:1237–47.

26. Albert I, Donnet S, Guihenneuc-Jouyaux C, Low-Choy S, Mengersen K, Rousseau J. Combining Expert Opinions in Prior Elicitation. Bayesian Anal. 2012 Sep;7(3):503–32.

27. Ashby D, Smith AF. Evidence-based medicine as Bayesian decision-making. Stat Med. 2000 Dec 15;19(23):3291–305.

28. Barrett J, Skolnik J, Jayaraman B, Patel D, Adamson P. Discrete event simulation applied to pediatric phase I oncology designs. Clin Pharmacol Ther. 2008 Dec;84(6):729–33.

29. Berchialla P, Scarinzi C, Snidero S, Gregori D. Comparing models for quantitative risk assessment: an application to the European Registry of foreign body injuries in children. Stat Methods Med Res. 2013 Feb 19.

30. Brass EP, Lofstedt R, Renn O. Improving the decision-making process for nonprescription drugs: a framework for benefit-risk assessment. Clin Pharmacol Ther. Nature Publishing Group; 2011 Dec;90(6):791–803.

31. Janssen M, Over J, van der Poel C, Cuijpers H, van Hout B. A probabilistic model for analyzing viral risks of plasma-derived medicinal products. Transfusion. 2008;48:153–62.

32. Chauvin P, Josselin J-M, Heresbach D. Incremental net benefit and acceptability of alternative health policies: a case study of mass screening for colorectal cancer. Eur J Health Econ. 2012 Jun;13(3):237–50.

33. Chawla A, Mytelka D, McBride S, Nellesen D, Elkins B, Ball D, Kalsekar A, Towse A, Garrison, LP. Estimating the incremental net health benefit of requirements for cardiovascular risk evaluation for diabetes therapies. Pharmacoepidemiol Drug Saf. 2014;23(3):268–77.

34. Chuang-Stein C. A new proposal for benefit-less-risk analysis in clinical trials. Control Clin Trials. 1994 Feb;15(1):30–43.

35. CIOMS working group I. Benefit-Risk Balance for Marketed Drugs: Evaluating Safety Signals. Report of CIOMS Working Group IV. 1998.

36. Citrome L, Ketter TA, Cucchiaro J, Loebel A. Clinical assessment of lurasidone benefit and risk in the treatment of bipolar I depression using number needed to treat, number needed to harm, and likelihood to be helped or harmed. J Affect Disord. Elsevier; 2014 Feb;155:20–7.

37. Corey-Lisle PK, Peck R, Mukhopadhyay P, Orsini L, Safikhani S, Bell JA, Hortobagyi G, Roche H, Conte P, Revicki DA. Q-TWiST analysis of ixabepilone in combination with capecitabine on quality of life in patients with metastatic breast cancer. Cancer. 2012 Jan 15;118(2):461–8.

38. Corso PS, Hammitt JK, Graham JD, Dicker RC, Goldie SJ. Assessing Preferences for Prevention Versus Treatment Using Willingness to Pay. Med Decis Mak. 2002 Oct 1;22(5 suppl):S92–S101.

39. DeCosse JJ, Cennerazzo WJ. A quality-adjusted time without symptoms or toxicity (Q-TWiST) analysis of adjuvant radiation therapy and chemotherapy for resectable rectal cancer. J Natl Cancer Inst. 1996 Nov 20;88(22):1686.

40. Djulbegovic B, Hozo II, Fields K, Sullivan D. High-Dose Chemotherapy in the Adjuvant Treatment of Breast Cancer: Benefit/Risk Analysis. Cancer Control. 1998;5(5):394–405.

41. Dogson. Multi-criteria analysis: a manual. Communities and Local Government Publications; 2009.

42. Donovan AK, Smith KJ, Ragni M V. Anticoagulation duration in heterozygous factor V Leiden: a decision analysis. Thromb Res. Elsevier Ltd; 2013 Dec;132(6):724–8.

43. Earnshaw SR, Brogan AP, McDade CL. Model-based cost-effectiveness analyses for prostate cancer chemoprevention : a review and summary of challenges. Pharmacoeconomics. 2013 Apr;31(4):289–304.

44. EMA. Benefit-risk methodology project Work package 4 report : Benefit-risk tools and processes. 2012 p. 1–20.

45. Eriksen S, Keller LR. A Multiattribute-utility-function Approach to Weighing the Risks and Benefits of Pharmaceutical Agents. Med Decis Mak. 1993 Jun 1;13(2):118–25.

46. Fahey T, Griffiths S, Peters TJ. Evidence based purchasing: understanding results of clinical trials and systematic reviews. BMJ. 1995 Oct 21;311(7012):1056–60.

47. Felli JC, Noel RA, Cavazzoni PA. A multiattribute model for evaluating the benefit-risk profiles of treatment alternatives. Med Decis Making. 2009;29(1):104–15.

48. Francis GS. Importance of benefit-to-risk assessment for disease-modifying drugs used to treat MS. J Neurol. 2004 Sep;251 Suppl v42–v49.

49. Gelber RD, Gelman RS, Goldhirsch a. A quality-of-life-oriented endpoint for comparing therapies. Biometrics. 1989 Sep;45(3):781–95.

50. Gelber RD, Cole BF, Gelber S, Goldhirsch A. Comparing Treatments Using Quality-Adjusted Survival : The Q-TWiST Method. Am Stat. 2014;49(2):161–9.

51. Geva A, Gray J. A quantitative analysis of optimal treatment capacity for perinatal asphyxia. Med Decis Making. 2012;32(2):266–72.

52. Greving JP, Buskens E, Koffijberg H, Algra A. Cost-effectiveness of aspirin treatment in the primary prevention of cardiovascular disease events in subgroups based on age, gender, and varying cardiovascular risk. Circulation. 2008 Jun 3;117(22):2875–83.

53. Gupta SK. Use of Bayesian statistics in drug development: Advantages and challenges. Int J Appl basic Med Res. 2012 Jan;2(1):3–6.

54. Hauser JR. Conjoint Analysis , Related Modeling , and Applications. Advances in Marketing Research: Progress and Prospects. 2002.

55. Hirsch GB, Edelstein BL, Frosh M, Anselmo T. A simulation model for designing effective interventions in early childhood caries. Prev Chronic Dis. 2012 Jan;9(1):E66.

56. Hodder SC, Edwards MJ, Brickley MR, Shepherd JP. Multiattribute utility assessment of outcomes of treatment for head and neck cancer. Br J Cancer. 1997 Jan;75(6):898–902.

57. Holden WL, Juhaeri J, Dai W. Benefit-risk analysis: examples using quantitative methods. Pharmacoepidemiol Drug Saf. 2003 Dec;12(8):693–7.

58. Holden WL, Juhaeri J, Dai W. Benefit-risk analysis: a proposal using quantitative methods. Pharmacoepidemiol Drug Saf. 2003;12(7):611–6.

59. Hynes N, Sultan S. A prospective clinical, economic, and quality-of-life analysis comparing endovascular aneurysm repair (EVAR), open repair, and best medical treatment in high-risk patients with abdominal aortic aneurysms suitable for EVAR: the Irish patient trial. J Endovasc Ther. 2007 Dec;14(6):763–76.

60. Ijzerman MJ, van Til JA, Bridges JFP. A comparison of analytic hierarchy process and conjoint analysis methods in assessing treatment alternatives for stroke rehabilitation. Patient. 2012 Jan;5(1):45–56.

61. Inn G, Washington DC, Philips L, Cone M. Benefit-Risk Assessment Model for Medicines : Developing a Structured Approach to Decision Making Report of the Workshop organised by the CMR International Institute for Regulatory Science Stuart Walker. 2006.

62. Khan A a, Perlstein I, Krishna R. The use of clinical utility assessments in early clinical development. AAPS J. 2009 Mar;11(1):33–8.

63. Kinnersley N, Day S. Structured approach to the elicitation of expert beliefs for a Bayesian-designed clinical trial: a case study. Pharm Stat. 2013;12(2):104–13.

64. Kopylev L, Chen C, White P. Towards quantitative uncertainty assessment for cancer risks: central estimates and probability distributions of risk in dose-response modeling. Regul Toxicol Pharmacol. 2007 Dec;49(3):203–7.

65. Korsan B, Dykstra K, Pullman W. A clinical utility index (CUI) openly evaluates a product ' s attributes — and chance. Pharm Exec. 2005;

66. Krishna R. Model-based benefit-risk assessment: can Archimedes help? Clin Pharmacol Ther. 2009 Mar;85(3):239–40.

67. Langley RG, Strober BE, Gu Y, Rozzo SJ, Okun MM. Benefit-risk assessment of tumour necrosis factor antagonists in the treatment of psoriasis. Br J Dermatol. 2010 Jun;162(6):1349–58.

68. Leil TA, Feng Y, Zhang L, Paccaly A, Mohan P, Pfister M. Quantification of apixaban's therapeutic utility in prevention of venous thromboembolism: selection of phase III trial dose. Clin Pharmacol Ther. Nature Publishing Group; 2010 Sep;88(3):375–82.

69. Levitan B, Yuan Z, Turpie AGG, Friedman RJ, Homering M, Berlin JA, Berkowitz SD, Weinstein RB, DiBattiste PM. Benefit-risk assessment of rivaroxaban versus enoxaparin for the prevention of venous thromboembolism after total hip or knee arthroplasty. Vasc Health Risk Manag. 2014 Jan;10:157–67.

70. Loftus E V, Johnson SJ, Wang S-T, Wu E, Mulani PM, Chao J. Risk-benefit analysis of adalimumab versus traditional non-biologic therapies for patients with Crohn's disease. Inflamm Bowel Dis. 2011 Jan;17(1):127–40.

71. Lynd LD, Marra CA, Najafzadeh M, Sadatsafavi M. A quantitative evaluation of the regulatory assessment of the benefits and risks of rofecoxib relative to naproxen : an application of the incremental net-benefit framework y. Pharmacoepidemiol Drug Saf. 2010;19(July):1172–80.

72. Lynd LD, Najafzadeh M, Colley L, Byrne MF, Willan AR, Sculpher MJ, Johnson FR, Hauber AB. Using the incremental net benefit framework for quantitative benefit-risk analysis in regulatory decision-making--a case study of alosetron in irritable bowel syndrome. Value Health. 2010;13(4):411–7.

73. Lynd LD, O'brien BJ. Advances in risk-benefit evaluation using probabilistic simulation methods: an application to the prophylaxis of deep vein thrombosis. J Clin Epidemiol. 2004 Aug;57(8):795–803.

74. Mann J, Ellis S, Waternaux C, Liu X, Oquendo M, Malone K. Classification Trees Distinguish Suicide Attempters in Major Psychiatric Disorders: A Model of Clinical Decision Making. J clin psychiatry. 2013;69(1):23–31.

75. Marino P, Roché H, Moatti J-P. High-dose chemotherapy for patients with high-risk breast cancer: a clinical and economic assessment using a quality-adjusted survival analysis. Am J Clin Oncol. 2008 Apr;31(2):117–24.

76. Meyboom R, Egberts A. Comparing therapeutic benefit and risk. Therapie. 1999;54(1):29–34.

77. Mounier N, Haioun C, Cole BF, Gisselbrecht C, Sebban C, Morel P, Marit G, Bouabdallah R, Ravoet C, Salle G, Reyes F, Lepage E. Quality of life-adjusted survival analysis of high-dose therapy with autologous bone marrow transplantation versus sequential chemotherapy for patients with aggressive lymphoma in first complete remission. Groupe d'Etude les Lymphomes de l'Adulte (GELA). Blood. 2000 Jun 15;95(12):3687–92.

78. Mudd PN, Groenendaal H, Bush M a, Schmith VD. Probabilistic risk analysis: improving early drug development decision making. Clin Pharmacol Ther. Nature Publishing Group; 2010 Dec;88(6):871–5.

79. Mussen F, Salek S, Walker S. Benefit-Risk Appraisal of Medicines. A systematic approach to decision-making. 2009.

80. Mussen F, Salek S, Walker S. A quantitative approach to benefit-risk assessment of medicines – part 1 : The development of a new model using multi-criteria decision analysis y. Pharmacoepidemiol Drug Saf. 2007;(June):2–15.

81. Nixon RM, Hagan AO, Oakley J, Madan J, Stevens JW, Bansback N, Brennan A. The Rheumatoid Arthritis Drug Development Model : a case study in Bayesian clinical trial simulation. Pharm Stat. 2009;(April):371–89.

82. Ohlssen D, Price KL, Xia HA, Hong H, Kerman J, Fu H, Quartey G, Heilmann CR, Ma H, Carlin BP. Guidance on the implementation and reporting of a drug safety Bayesian network meta-analysis. Pharm Stat. 2014;13(1):55–70.

83. Ouellet D, Werth J, Parekh N, Feltner D, McCarthy B, Lalonde RL. The use of a clinical utility index to compare insomnia compounds: a quantitative basis for benefit-risk assessment. Clin Pharmacol Ther. 2009 Mar;85(3):277–82.

84. Ouellet D. Benefit-risk assessment: the use of clinical utility index. Expert Opin Drug Saf. 2010 Mar;9(2):289–300.

85. Pfohl M, Schädlich PK, Dippel F-W, Koltermann KC. Health economic evaluation of insulin glargine vs NPH insulin in intensified conventional therapy for type 1 diabetes in Germany. J Med Econ. 2012 Jan;15 Suppl 2:14–27.

86. Phillips LD, Fasolo B, Zafiropoulous N, Eichler H-G, Ehmann F, Jekerle V, Kramarz P, Nicoll A, Lönngren T. Modelling the risk-benefit impact of H1N1 influenza vaccines. Eur J Public Health. 2013 Aug;23(4):674–8.

87. Poland B, Hodge FL, Khan A, Clemen RT, Wagner JA, Dykstra K, Krishna R. The clinical utility index as a practical multiattribute approach to drug development decisions. Clin Pharmacol Ther. Nature Publishing Group; 2009 Jul;86(1):105–8.

88. Green P, Rao V. Conjoint Measurement for Data Quantifying Judgmental. J Mark Res. 2014;8(3):355–63.

89. Sabin T, Matcham J, Bray S, Copas A, Parmar MKB. A Quantitative Process for Enhancing End of Phase 2 Decisions. Stat Biopharm Res. 2014 Jan;6(1):67–77.

90. Sadatsafavi M, Marra C, Marra F, Moran O, FitzGerald JM, Lynd L. A quantitative benefit-risk analysis of isoniazid for treatment of latent tuberculosis infection using incremental benefit framework. Value Health. 2013;16(1):66–75.

91. Shaffer ML, Watterberg KL. Joint distribution approaches to simultaneously quantifying benefit and risk. BMC Med Res Methodol. 2006 Jan;6:48.

92. Shepherd J, Brodin H, Cave C, Waugh N, Price A, Gabbay J. Pegylated interferon a-2a and -2b in combination with raibavirin in the treatment of chronic hepatitis C: a systematic review and economic evaluation. Health Technol Assess (Rockv). 2004;8(39).

93. Sherrill B, Amonkar MM, Stein S, Walker M, Geyer C, Cameron D. Q-TWiST analysis of lapatinib combined with capecitabine for the treatment of metastatic breast cancer. Br J Cancer. 2008 Sep 2;99(5):711–5.

94. Sherrill B, Wang J, Kotapati S, Chin K. Q-TWiST analysis comparing ipilimumab/dacarbazine vs placebo/dacarbazine for patients with stage III/IV melanoma. Br J Cancer. Nature Publishing Group; 2013 Jul 9;109(1):8–13.

95. Smart A. A multi-dimensional model of clinical utility. Int J Qual Health Care. 2006 Oct;18(5):377–82.

96. Stojadinovic, Alexander Nissan A, Eberhardt J, Chua T, Pelz J, Esquivel J. Development of a Bayesian Belief Network Model for Personalized Prognostic Risk Assessment in Colon Carcinomatosis. Am Surg. 2011;77(2):221–30.

97. Stonebraker JS. How Bayer Makes Decisions to Develop New Drugs. Interfaces (Providence). 2002 Dec;32(6):77–90.

98. Thompson JP, Noyes K, Dorsey ER, Schwid SR, Holloway RG. Quantitative risk-benefit analysis of natalizumab. Neurology. 2008 Jul 29;71(5):357–64.

99. Tomeczkowski J, Stern S, Müller A, von Heymann C. Potential cost saving of Epoetin alfa in elective hip or knee surgery due to reduction in blood transfusions and their side effects: a discrete-event simulation model. PLoS One. 2013 Jan;8(9):e72949.

100. Verduijn M, Grootendorst DC, Dekker FW, Jager KJ, le Cessie S. The analysis of competing events like cause-specific mortality--beware of the Kaplan-Meier method. Nephrol Dial Transplant. 2011 Jan;26(1):56–61.

101. Weinstein L, Radano TA, Jack T, Kalina P, Eberhardt JS. Application of multivariate probabilistic (Bayesian) networks to substance use disorder risk stratification and cost estimation. Perspect Health Inf Manag. 2009 Jan;6:1-27.

102. Keeney RL, Raiffa H. Decisions with Multiple Objectives. Cambridge: Cambridge University Press; 1993.

103. Walker S, Cone M. The Development of a Model for Benefit-Risk Assessment of Medicines Based on Multi-Criteria Decision Analysis (Summary Report). 2004.

104. Breimer D. PK/PD modelling and beyond: impact on drug development. Pharm Res. 2008;25:2720–2.

105. Breimer D, Danhof M. Relevance of the application of pharmacokinetic-pharmacodynamic modelling concepts in drug development - the wooden shoe paradigm. Clin Pharmacokinet. 1997;32:259–67.

106. Zineh I, Woodcock J. Clinical pharmacology and the catalysis of regulatory science: opportunities for the advancement of drug development and evaluation. Clin Pharmacol Ther. 2013;93:515–25.

107. Jonker D, Visser S, van der Graaf P, Voskuyl R, Danhof M. Towards a mechanism-based analysis of pharmacodynamic drug-drug interactions in vivo. Pharmacol Ther. 2005;106:1–18.

108. Danhof M, de Jongh J, De Lange E, Della Pasqua O, Ploeger B, Voskuyl R. Mechanism-based pharmacokinetic-pharmacodynamic modeling: biophase distribution, receptor theory, and dynamical systems analysis. Annu Rev Pharmacol Toxicol. 2007;47:357–400.

109. Machado S, Miller R, Hu C. A regulatory perspective on pharmacokinetic/pharmacodynamic modelling. Stat Methods Med Res. 1999;8:217–45.

110. Sjörgen E, Nyberg J, Magnusson M, Lennernäs H, Hooker A, Bredber U. Optimal Experimental Design for Assessment of Enzyme Kinetics in a Drug Discovery Screening Environment. Drug Metab Dispos. 2011;39:858–63.

111. Yousef A, Melhem M, Xue B, Arafat T, Reynolds D, Van Wart S. Population pharmacokinetic analysis of clopidogrel in healthy Jordanian subjects with emphasis optimal sampling strategy. Biopharm Drug Dispos. 2013;226:215–26.

112. Lim H, Kim S, Noh Y, Lee B, Jin S, Park H, Kim S, Jang IJ, Kim SE. Exploration of optimal dosing regimens of haloperidol, a D2 Antagonist, via modeling and simulation analysis in a D2 receptor occupancy study. Pharm Res. 2013;30:683–93.

113. Sammons H, Starkey E. Ethical issues of clinical trials in children. Paediatr Child Heal. 2012;22:47–50.

114. Bellanti F, Della Pasqua O. Modelling and simulation as research tools in paediatric drug development. Eur J Clin Pharmacol. 2011;67(Suppl 1):75–86.

115. Cella M, Zhao W, Jacqz-Aigrain E, Burger D, Danhof M, Della Pasqua O. Paediatric drug development: are population models predictive of pharmacokinetics across paediatric populations? Br J Clin Pharmacol. 2011;72:454–64.

116. Manolis E, Osman T, Herold R, Koenig F, Tomasi P, Vamvakas S, Saint Raymond A. Role of modeling and simulation in pediatric investigation plans. Paediatr Anaesth. 2011;21:214–21.

117. Manolis E, Pons G. Proposals for model-based paediatric medicinal development within the current European Union regulatory framework. Br J Clin Pharmacol. 2009;68:493–501.

118. EMA. Benefit-risk methodology project - Report on risk perception study module. 2012.

119. Tervonen T, Van Valkenhoef G, Buskens E, Hillege HL, Postmus D. A stochastic multi-criteria model for evidence-based decision making in drug benefit-risk analysis. Stat Med. 2011; 30(12):1419-28

120. Phillips L. Improving the process of balancing benefits and risks in approving drugs (presentation). EMA Decision Analysis Affinity Group Workshop; 2010.

121. Phillips L. Benefit-Risk modelling of pharmaceuticals: Where are we now? (presentation). EFSPI-FMS-DSBS Benefit Risk Assessment Methodology Workshop 7; 2012.

122. Lam G, Petri M. Assessment of systemic lupus erythematosus. Clin Exp Rheumatol. 2005;23:S120–S132.

123. French S, Xu D. Comparison study of multi attribute decision analytic software. J Multi Criteria Decis Anal. 2005;13:65–80.

124. EMA. EPAR Benlysta: Summary of product characteristics. 2013.

125. Bender BC, Schindler E, Friberg LE. Population pharmacokinetic pharmacodynamic modelling in oncology: a tool for predicting clinical response. Br J Clin Pharmacol. 2013;79(1):56–71.

126. Pink J, Lane S, Hughes DA. Mechanism-Based Approach to the Economic Evaluation of Pharmaceuticals. Pharmacoeconomics. 2012;30(5):413–29.

127. Naci H, van Valkenhoef G, Higgins JPT, Fleurence R, Ades AE. Evidence-Based Prescribing: Combining Network Meta-Analysis With Multicriteria Decision Analysis to Choose Among Multiple Drugs. Circ Cardiovasc Qual Outcomes. 2014;7:787–92.

128. Van Valkenhoef G, Tervonen T, Zhao J, De Brock B, Hillege HL, Postmus D. Multicriteria benefit-risk assessment using network meta-analysis. J Clin Epidemiol. Elsevier Inc; 2012;65(4):394–403.

129. Chain ASY, Dieleman JP, van Noord, Charlotte Hofman A, Stricker BHC, Danhof M, Sturkenboom MCJM, Della Pasqua, O. Not-in-trial simulation I: Bridging cardiovascular risk from clinical trials to real-life conditions. Br J Clin Pharmacol. 2013;76(6):964–72.

Appendix

Decision tree

The theory underlying the use of decision trees is based on visualisation of the decision making process by a branching structure with decisions as roots, and possible outcomes as branches. In this way, decisions, subsequent uncertain events, consequences and multiple criteria are described (14). Complexity arises with addition of extra nodes. In the EMA framework for benefit-risk assessment, this technique has been used as a link between qualitative and quantitative approaches, as it encompasses objectives and possible outcomes, combined with numerical data on frequencies and uncertainties, with which benefit-risk balance can be calculated for each outcome or decision. Figure S1 shows an example of a decision tree. The decision (square node) consists of two alternatives: approving the vaccine against H1N1 swine flu by the end of September, or waiting until October, so more data can be gathered on efficacy and safety. The remaining uncertainties are modelled as events (round nodes), for which consequences the working group determined the probability, mostly based on earlier experience. Disease seriousness for example has a probability of 20% to become severe, based on the historical observation that one in five pandemics becomes catastrophic (Spanish flue). For the delay, this probability increases, as early vaccination can prevent escalation. All other probabilities are determined and the estimated deaths or serious disabled (DSDs) are stated for all 24 outcomes (triangle nodes). The decision tree itself enables back calculation to the actual decision, providing the consequence of each alternative would be. This is achieved by multiplying outcomes with the probabilities as weight. Taking the best case scenario, the working group determined that in the case of early approval, moderate disease seriousness probability is 0.8, probability of efficacy of 75% is 0.3, with probabilities of safety events rate of 1/100,000 and 1/10,000 at 0.9 and 0.1, respectively. The latter outcomes are stated to be 42500 and 87500 DSDs, so calculating back, the average DSD-value of the vaccine in early approval in case of a moderate disease with an efficacy of 75% is 47000 DSDs. Applying these steps for all outcomes and events results in an average DSDs for the two alternative options, in this case 216,500 for September and 291,547 for October, showing that earlier approval is the better option, a decision made after completion of a sensitivity analysis of the chosen probabilities (86). A reconstruction of the complete decision tree can be found in (120). This concept alone is valuable when cases remain relatively simple. Although the tree itself might be too complex in advanced cases, it remains an important building block for more evolved techniques, such as MCDA.

MCDA and the EMA's BR framework

When evaluating very complex scenarios with multiple endpoints and objectives, it becomes crucial to have a clear understanding of the context structure. In very simple words, MCDA allows breaking up the problem and analyzing individual factors, before reassembling each component to provide a thorough overview of the analysis and making a final decision (102). In this structure, it combines the decision tree theory with value functions. In other words, it converts the different inputs for the decision model into preference values, allowing comparing the different endpoints on a common ground. The preference value scale requires probability, utility and the preference of the alternative associated with the highest expected utility. Multiple objectives are evaluated together on different identified criteria and a balance is made after scoring and weighting these criteria, with uncertainty taken into account (14). As highlighted by the EMA BR project, the use of the PrOACT-URL approach in combination with decision tree and MCDA represents a more transparent and consistent assessment of the BR balance (44). The technique consists of eight steps that will be briefly discussed in the following paragraphs. In the next section, these steps are illustrated using Benlysta (belimumab), a drug against Systemic Lupus Erythematosus (SLE), as a paradigm compound (121).

Step 1 to 3: defining decision context; identifying options, eg. study arms; identifying objectives and criteria, eg. maximising benefit and minimising risk, more specified in a decision tree. This part of MCDA overlaps with the PrOACT-URL approach: creating a qualitative framework of objectives and context, as well as with the decision tree, as mentioned earlier (120). Benlysta has been proposed for the treatment of adults with high disease activity, with autoantibody-positive SLE. It should be added to the standard treatment, which consists of hydroxychloroquine and corticosteroids (step 1). The available studies include two randomised, placebo-controlled clinical trials and three open-label continuation safety trials. The dosing regimens used in these trials include either 1 or 10 mg (step 2). There is a medical need for newer, more effective and better tolerated therapies. To specify these criteria, an effect tree is composed, as visualised in Figure S2 (step 3).

Step 4 and 5: scoring; weighting. Scoring and weighting are the most important steps as their aim is to normalise the raw input data for the decision model by translating them into preference values. Scoring means scaling each criterion (input data characterised by different units and time scales) by assigning a new range, which is usually set between 0 and 100. Within this range, different outcomes are directly or indirectly scored, where the ratio of difference is the most important. Scoring for Benlysta, as visualised in Table S2, is performed following defined clinical scales, like SLE Disease Activity Index (SLEDAI). First, the two extremes are evaluated, best and worst with corresponding units, after which the three options are considered within this range.

The weighting step normalise all measures into one preference scale, judging which criteria is more important and allowing comparing the different options into one common level (103). This procedure allows translating the scoring into preference values, which carries the subjective component. Weighting can be done linear, direct or inverse, or non-linear. Finally, swing weights are assigned, based on trade-offs among favourable or unfavourable effects, or between the most important favourable and unfavourable effects. In other words, if objective A is twice as important as objective B, the score doubles on that scale. These swing weights depend on the subjective choice considering the relative difference in original scale and the importance of the corresponding objective to the whole. Considering as an example buying a car, limiting costs is an objective of importance. If, however the difference between alternatives in this criterion is only small, the impact of that objective becomes limited (79). It is also important to take into account the possibility of single events that are multiple times considered. In this specific case, the SLE assessment scores SLEDAI and BILAG-index have similar criteria, like psychosis or vasculitis. If this is not corrected by the assignments of weights, these events have double impact on in this case the unfavourable effects (122).

Step 6 and 7: combining data to overall value; examining results. The overall score is simply the sum of the product of the score and weight per criterion, as stated in equation 1, where Si is the overall score per option i on criterion j, with sij as preference score of the option and wi as the weight of the criterion (41).

$$S_i = w_1 s_{i1} + w_2 s_{i2} + \ldots + w_n s_{in} = n_{j=1} \sum w_j s_{ij} \qquad \text{eq.1}$$

This aggregation is performed by software; several are currently available for this methodology (e.g., HiView, V.I.S.A., Web-Hipre, Expert Choice, Logical Decisions) (123). Cumulative weights are calculated based on the normalized weight; overall weighted scores per options are visualized graphically (Figure S3).

Step 8: perform sensitivity analysis. The sensitivity analysis is important to identify possible judgments of serious impact, thus reducing uncertainty. Displaying the variation of weights on each criterion allows identifying possible crossovers at which a change in the relationship between weight and criterion might be observed for the different options.

Table S1. Overview of quantitative methodologies to assess benefit-risk balance, as given by the CHMP [14].

Method	Advantages	Limitations	References
Bayesian beliefs networks	Network of nodes representing risks, benefits, observations and assessments, connected by conditional arrows, which input probabilities result in probability distribution for all nodes. Inclusion of both objective data and subjective expert opinion. Visualisation of effect of factors on each other.	Requires structural similarity across cases, which in BR might only be appropriate for similar indications. Probability input as a subjective element remains unsupported. Uncertainty of indirect effects introduces bias in their impact on the outcome.	(14,29,82,96,101)
Bayesian statistics	Prior and posterior probabilities based on available evidence. Tgether describe the likelihood of an effect and its uncertainty, combined with utility function in the Bayesian approach. Methodology improves as more data are gathered, as it involves iterative learning.	Significance levels state something about data, not hypotheses, so cannot directly be included into a formal BR assessment. The model itself doesn't include multiple criteria. Mathematical models can get complex.	(14,27,53,64,81,89)
Clinical Utility Index	Multi-attribute utility analysis with weighted trade-offs. Utility function introduces clinical meaning to the assessment. CUI is flexible over different indications and endpoints. Transparent method with possibility of sensitivity analysis.	In case of limited applicable data, complex modelling with high variability and uncertainty is required. Subjective discussion on clinically relevant factors remains unsupported. More useful for a no-go than for a go-decision	(45,56,62,65,68,83,84,87,95)
Conjoint analysis	Covers preferences of different stakeholders, utility weight is based	Labour intensive if all stakeholders are included. Weight might not be	(14,54,60,88)

	on preferred trade-offs. Realistic method helpful in weighting.	independent from methodological decisions. Does not account for uncertainty.	
Contingent valuation	Benefits are translated to financial values by enquiring the prize patients are willing to pay for it.	Not focused on BR assessment	(14)
Decision tree	Overview of all possible outcomes with their probabilities, calculated using the branches and nodes leading to said outcome. The decision tree is a useful framework.	Too simple for complex cases. Uncertainties are only limited covered, as probabilities are often empirically determined.	(14,30,74,80,86,97,124)
Discrete event simulation	Detailed simulation based on differential equations and continuous variables. Ability to handle multiple assumed characteristics and simultaneously assess impact of multiple effects on health economics.	Complexity, complicate adaptability, lack of transparency and validation. Risk of underestimation in case of prediction limited to short term effects. No clear assessment of unfavourable effects.	(14,28,31,51,66,71,85,99)
Evidence-based BR model	Model visualised as a set of scales, including the benefit 'box' with efficacy, including responder rate and evidence and the risk 'boxes', for each ADE, with seriousness, frequency and evidence. The method correlate to EMA's definition, as the first two criteria of either box are (un)favourable effects and the third includes uncertainty of effects.	Simplified multi-criteria model with limited (three) criteria each. There is no application supporting the translation of effects into one unit. Preference weights are not accounted for.	(14,19)
Incremental net	Incremental net health benefit is the	Although this is a version of a multi-	(1,14,23,32,33,72,90)

health benefit	difference between unfavourable effects and favourable effects derived from the treatment options, where all effects are normalised into one unit. The method is transparent and theoretically sound, including uncertainty and extrapolation in time.	criteria model, such as MCDA, translation to a single unit requires another methodology (e.g., value-adjusted life years, or QALY). QALY can only be transferred to health benefit when costs are not considered, in other words the willingness to pay is infinite. Weighting of effects is also dependent on another methodology, like conjoint analysis. This methodology on itself is incomplete. Subject to bias by confounders.	
Kaplan-Meier estimation	Function of survival over time, impact measured in ratio of differences, useful in Markov models.	Limited representation of (un)favourable effects, for example in non-fatal indications. It does not account for uncertainty and cumulative probabilities can be misleading due to lack of correlation structure (e.g., competing events).	(14,49,100)
Markov model	Describes time-dependent dynamic processes, using transitions between health states and their probability distributions.	Probability data might be sparse before approval. Complex health states might be oversimplified.	(14,42,43,52,98)
Minimum Clinical Efficacy	The method allows incorporating risks and benefits into one single metric. In addition, relative utilities can be considered during the analysis.	The statistical properties are not yet fully understood and the methodology does not allow characterising the uncertainty around the benefit-risk measurements.	(1,40,57,58)
MCDA	Multi-criteria method breaking up	Might be too comprehensive for	(1,12,14,22–

	the problem, followed by scoring and weighted assessment of benefits and risks as most representative presentation of data. Sensitivity analysis prevents unwanted impact. Incorporates uncertainty.	simple analysis. Does not account for possible correlations between endpoints. Preference value determination is accounted for in the weighting step.	24,26,34,44,47,63,79,80,102,103)
NNT	Easy understandable measure used in the clinic, stating the number of patients required to treat one occurrence of the disease (or to have one more ADE in NNH). Patient preferences can be included using Relative Value Adjusted NNT (RV-NNT).	Limited statistical power and because of lack of preference data, misinterpretation by different risk perceptions, as well as by using the same scale without proper weighting effects. Ratio of NNT/NNH assumes independence and similar timescale.	(1,4,14,22,23,36,46,48,57,58,67,69)
Principle of threes	Simplified method in which only three criteria per risk/benefit are scaled with three possible outcomes (e.g., low, medium, high), benefit and risk are summed up.	Very limited in number of criteria. No weighting of the criteria.	(14,17,35,76)
Probabilistic simulation	Complementary to point estimate statistics, as it states the impact of risk and benefit as a probability distributions based on simulated random draws from study data. More precise, accounts for uncertainty in trade-offs. Can account for correlation, if suitable data is available.	Limited if using non-validated or non-representative probability distributions for simulation. Benefits or risks are not weighted, shown by the fatal adverse event in the adalimumab-study, which did not seem to affect the simulation analysis.	(1,14,23,70,72,73,78,91)
QALY	Multiple dimensions are scored and	Limited in uncertainty and unique	(14,25,92)

	weighted for preference, outcome measured in life years on population level.	(disease/patient) data representation, more focussed on health- and pharmacoeconomics. Threshold is debatable.	
Q-TWiST	The method is used to convert time into QALYs; time lost due to an ADR is subtracted to time gained from receiving the treatment. Q-TWiST allows comparing benefits and risks into a single metrics. Furthermore, allowing the inclusion of patients' preferences is considered a valid tool for individual BR assessment.	Although valid for individual assessments, it gives more difficulties to evaluate BR on a population level. Does not allow measuring uncertainty around QALYs. The data needed for the analysis might be difficult to acquire. In addition QALYs might have a major influence on the BR outcome.	(1,22,37,39,49,50,59,75,77,93,94)
Stated preferences	Collection of methods using preference values to determine utility functions of different stakeholders. Measures e.g., the extent patients are willing to experience unfavourable effects to achieve favourable effects.	Empirical method that does not account for uncertainty or weighting. Overlaps with conjoint analysis. Gathering of individual patient data is time consuming.	(1,14,23,38)
System dynamics	Account for non-linearity using feedback and time-delays, both short and long term. Possibility of input data from different sources.	No recorded use in drug development. Focus on pharmacoeconomics. No consideration of (weighted) unfavourable effects, such as ADEs.	(14,55)
TURBO	Simplified method in which only two criteria per risk/benefit are scaled up with five possible outcomes. Pairs of outcomes are weighted and assessed. Frequency, probability,	Very limited in terms of the number of criteria. There is no way of knowing prior to assessment which criteria to choose. Choice might be arbitrary. No theoretical basis.	(14,23)

	severity and extent are included into the choices of criteria.		

Table S2. Scoring of Benlysta according to the different criteria, as visualised in the decision tree. FE and UFE are favourable and unfavourable effects, respectively. SRI is SLE Response Index, SLEDAI is SLE Disease Activity Index, PGA is Physician's Global Assessment, BILAG is British Isles Lupus Assessment Group, where A indicates severe disease and B less active disease. Secondary favourable endpoints are CS, corticosteroids, Flare rate meaning number of new BILAG A cases and QoL measured as mean change in the total score of Short Form 36. SAE are serious adverse events, such as tumour development, opportunistic infections or progressive multifocal leukoencephalopathy (PML). *> 25% and to less than 7.5 mg/day, ** per patient year. Based on (15).

Effects		**Name**	**Description**	**Best**	**Worst**	**Units**	**Placebo**	**10 mg**	**1 mg**
Favourable	SRI	SLEDAI	Improved ≥ 4	100	0	%	41	53	48
		PGA	No worsening	100	0	%	66	75	76
		PGA	Mean change	1	0	Difference	0,44	0,48	0,45
		BILAG	No new A/2B	100	0	%	69	75,2	70,1
	Secondary endpoints	CS sparing	Dose reduction*	100	0	%	12,3	17,5	20
		Flare rate	New BILAG A cases**	0	5	Frequency	3,51	2,88	2,9
		QoL	Mean change SF36	0	100	Difference	3,5	3,4	3,7
Unfavourable		SAE	Potential	100	0		100	0	90
		Infections	Life-threatening infections	0	10	%	5,2	5,2	6,8
		Sensitivity reaction	Hypersensitivity reactions	0	2	%	0,1	0,4	0,3

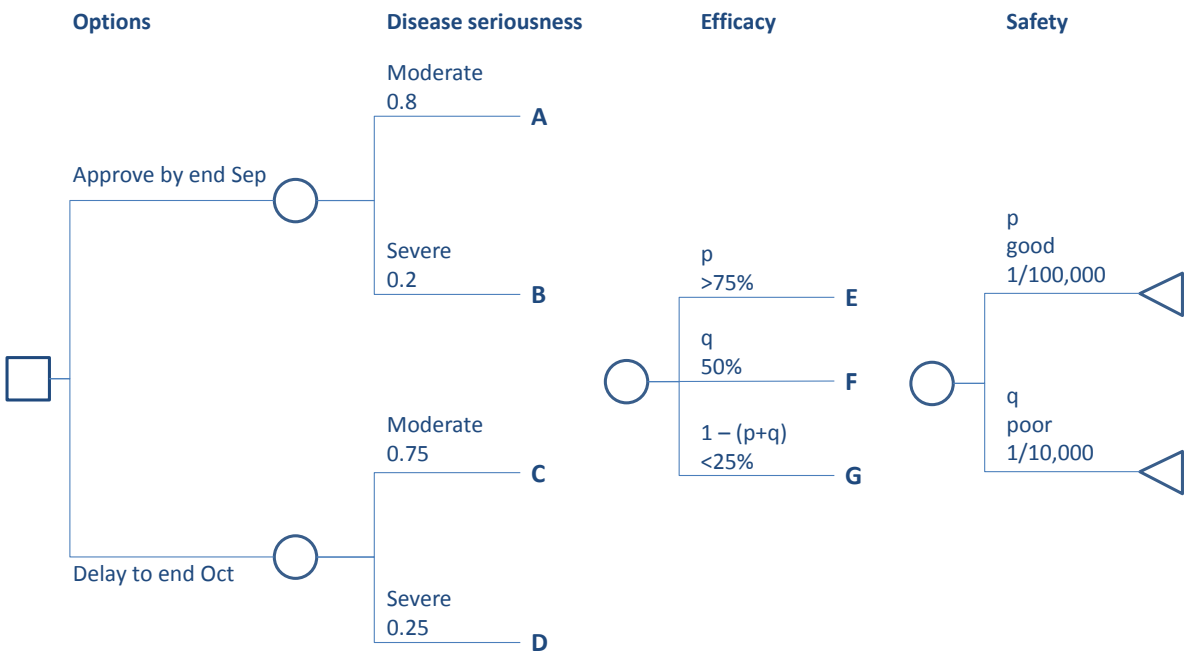

Figure S1. Example of decision tree concerning approval of swine flu vaccine in 2009, where decision of approval planning is followed by the consequences for disease seriousness. Efficacy branches attach to A through D, whereas safety branches to E through G resulting in 24 scenarios with calculable event outcomes. Based on (116).

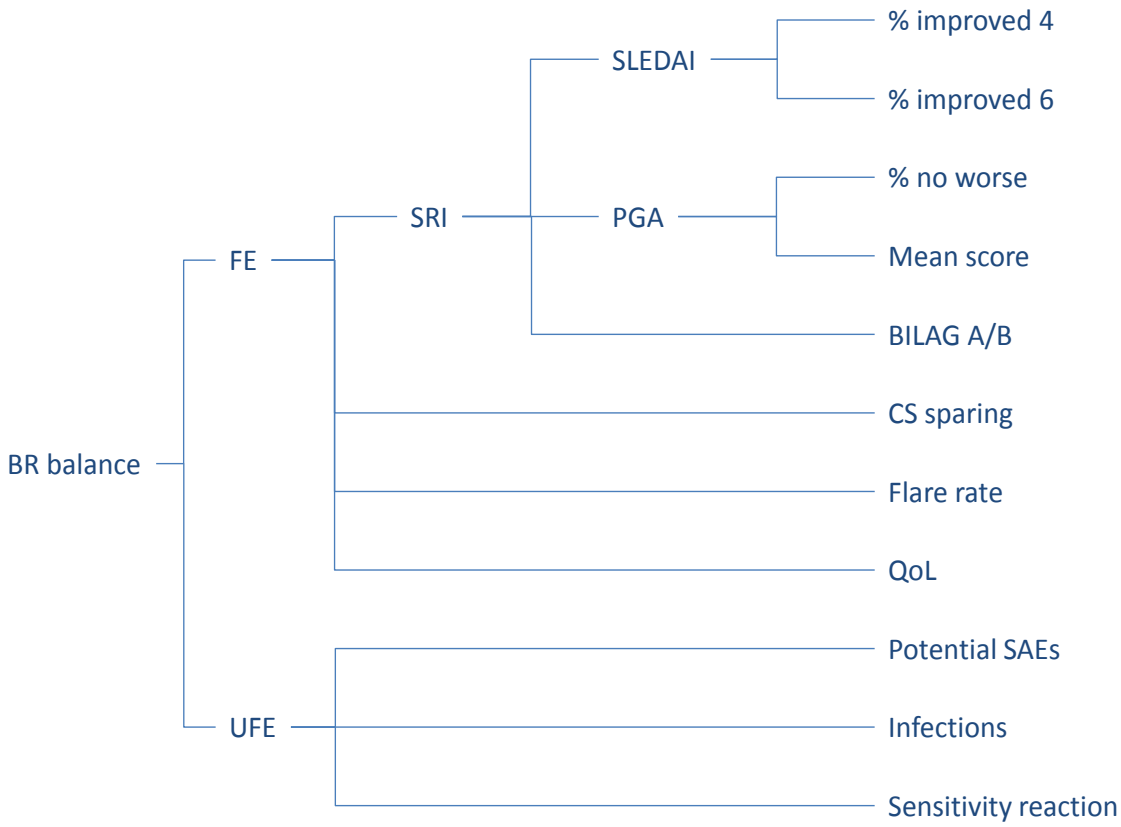

Figure S2. Outcomes tree based on identified criteria used for the Benlysta example. FE and UFE are favourable and unfavourable effects, respectively. SRI is SLE Response Index, SLEDAI is SLE Disease Activity Index, PGA is Physician's Global Assessment, BILAG is British Isles Lupus Assessment Group, where A indicates severe disease and B less active disease. Secondary favourable endpoints are CS, corticosteroids, Flare rate meaning number of new BILAG A cases and QoL measured as mean change in the total score of Short Form 36. SAE are serious adverse events, such as tumour development, opportunistic infections or progressive multifocal leukoencephalopathy (PML). Based on (15).

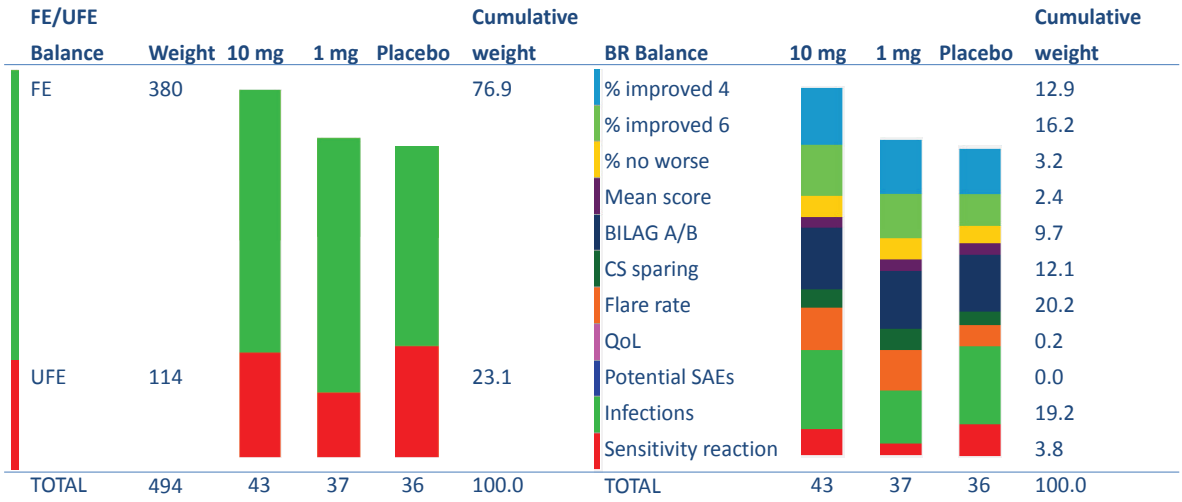

Figure S3. Different data presentations to evaluate benefit risk balance, in which the cumulative weight is calculated and the overall weighted scores are visualised. Left, impact of favourable effects (FE) and unfavourable effects (UFE) are shown in green and red, respectively. On the right, all different criteria are shown with their impact, which results in a more informative presentation of the data. For example, sensitivity reaction to 1mg has decreased impact as compared to placebo or 10mg. Based on (15).