Pharmacoepidemiological Research on Outcomes of Therapeutics by a European ConsorTium

Review of visualisation methods for the representation of benefit-risk assessment of medication: Stage 1 of 2

Shahrul Mt-Isa,[1] Ruth Peters,[1] Lawrence D. Phillips,[2] Kakit Chan,[1] Kimberley S. Hockley,[1] Nan Wang,[1] Deborah Ashby,[1] Ioanna Tzoulaki,[1] On behalf of PROTECT Work Package 5 participants

PROTECT Consortium Work Package 5 members involved in the generation of this report:
Shahrul Mt-Isa,[1] Ruth Peters,[1] Lawrence D. Phillips,[2] Torbjörn Callréus,[3] Kakit Chan,[1] Georgy Genov,[2] Christine E. Hallgreen,[4] Ian Hirsch,[5] Steve Hobbiger,[6] Kimberley S. Hockley,[1] Davide Luciani,[7] Ruth Peters,[1] George Quartey,[8] Sinan B. Sarac,[3] Isabelle Stoeckert,[9] Nan Wang,[1] Alain Micaleff,[10] Deborah Ashby,[1] Ioanna Tzoulaki.[1]

1 School of Public Health, Imperial College London, London, United Kingdom
2 European Medicines Agency, London, United Kingdom
3 Danish Medicines Agency, Copenhagen, Denmark
4 Novo Nordisk A/S, Søborg, Denmark
5 AstraZeneca AB, Macclesfield, United Kingdom
6 GlaxoSmithKline Research and Development LTD, Middlesex, United Kingdom
7 Mario Negri Institute for Pharmacological Research, Milan, Italy
8 Genentech, South San Francisco, USA
9 Bayer Schering Pharma AG, Berlin, Germany
10 Merck KGaA, Geneva, Switzerland

Version 1 Date: 15 Feb 2013 Date of any subsequent amendments below	Person making amendments	Brief description of amendments

Acknowledgements: The research leading to these results was conducted as part of the PROTECT consortium (Pharmacoepidemiological Research on Outcomes of Therapeutics by a European ConsorTium, www.imi-protect.eu) which is a public-private partnership coordinated by the European Medicines Agency.
The PROTECT project has received support from the Innovative Medicines Initiative Joint Undertaking (www.imi.europa.eu) under Grant Agreement n° 115004, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution

Executive summary

Background

Pharmacoepidemiological Research on Outcomes of Therapeutics in a European Consortium (PROTECT) is a project, set up under the Innovative Medicines initiative, with the aim of strengthening the monitoring of the benefit-risk of medicines in Europe. The evaluation of the balance between benefits and risks of drugs is fundamental to all stakeholders involved in the development, registration and use of drugs including patients, health care providers, regulators and pharmaceutical companies. There are many ways in which benefits and risks are presented and communicated. There is an absence of a consensus on which visual representations are most suitable to display benefit-risk profiles.

The visual representation of benefits and risks review is conducted in two stages. This report forms the first of the two-part review which provides a level of evaluation as to the suitability of visuals presented in the application of benefit-risk approaches in PROTECT methodology review. The second stage will explore the use of more innovative benefit-risk visualisation techniques, in particular the interactive and dynamic visuals which are becoming much easier to produce with the current computing technology.

Objective

The objectives of this visual representation and communication appraisal are:

1) To present the visual representations that could be associated with the 13 benefit risk methodologies recommended in the PROTECT benefit-risk methodology review ("A systematic review and classification of methodologies for benefit-risk decision-making in medicines")

2) To provide an initial level of appraisal as to their suitability and based on previously published criteria.

3) To make recommendations of potentially suitable visuals for each benefit-risk assessment approach recommended in PROTECT WP5 methodology report

Methods

We used the generic definition of graphics to classify the visual representations. We have further used Carswell's taxonomy to facilitate our evaluation as to the types of visuals that are likely to be of greater use for the different tasks required in the decision-making. We evaluated the potential of each 'visual' (e.g. line graph, scatter plot etc.) in the context of Wickens's principles of display design which we have redefined with reference to benefit-risk assessment. This resulted in excluding principles within the Wickens's "mental model" domain and the multimodality (audio-visual) principle. Finally we attempted to cross-match the final benefit-risk metrics from the recommended benefit-risk approaches to the most appropriate tasks within their scope that is an attempt to provide a mapping from benefit-risk approaches to suitable visual representations for the required tasks. Recommendations are then made using Cleveland's taxonomy supported by Tufte's data-ink ratio principles, which in effect favours simpler visual representations when there is more than one way to represent certain information visually.

Results

We have classified the visuals into categories according to type. These are standard classifications based on the definitions of the graph types, and have grouped the visuals from the PROTECT methodology review into, the area graph, bar graph, contour plot, distribution plot, flow diagram, dot/forest plot, grids and tables, line graph, network graph, scatter plot, surface plot, tornado diagram, and tree diagram.

We then assessed the suitability of these visual types using Carswell's taxonomy to comment with regard to their usability in terms of performing four different tasks: point reading, local comparison, global comparison, and synthesis judgment. The architectures within the visuals which facilitate the task they are associated to were highlighted. Bar graph, dot/forest plot, line graph and scatter plot were appraised as likely to be the most useful visual representations and are also widely used. However, there are no hard and fast rules as to the most appropriate visual representations of benefits and risks and the choice of visual also depends on the data to be presented in addition to the task and design. Finally, the audience also need to be taken into account and this includes issues such as levels of prior experience, time to evaluate the information, culture, physical, mental and cognitive status.

There are, therefore, many aspects to consider when presenting the results of benefits and risks of medicines including the environment in which they are presented and the audience they are presented to. Formal testing of these additional considerations are beyond the scope of this review, however, in general, given these caveats, we conclude that the current practice of benefit-risk visual representation seems appropriate. Based on our experience we also suggest that the Wickens's principles of display design, as redefined for the context of benefit-risk assessment, may be useful for future work as guidelines to aid the design of better visuals.

As inferred above, the communicability of visual representations are also of great importance. Formal testing of this is beyond the scope of this review. However, we have aimed to evaluate the potential of communicating five elements of risk communication. Risk communication is a vast subject; for further information please refer to the US FDA report ("Communicating risk and benefits: An evidence-based user's guide") published in August 2011 that covers the aspects of communicating risk in greater depths.

Recommendations

To facilitate direct application of visual representations of benefits and risks, the recommendations are categorised by benefit-risk approaches. These are limited to the list of recommendations from PROTECT methodology review.

The PROTECT work stream B recommendations for visual/graphical representations for use in the representation of benefit and risk and to accompany recommended benefit-risk approaches are:

1. PrOACT-URL

   We recommend an effect table for presentation of efficacy and safety data.

2. PhRMA BRAT

   We recommend table, dot/forest plot, and bar graph for presentation of efficacy and safety data. Value tree diagram may be used to represent the model and to develop insight into the decision problem.

3. Multi-Criteria Decision Analysis (MCDA)

We recommend bar graph and the 'difference display' for presentation of benefit-risk results. Additionally, table may be used to display evidence data, value tree diagram may be used to represent the favourable and unfavourable effects considered in judging the benefit-risk balance. Line graph for the sum of utilities versus total weights on a criterion may be used for sensitivity analysis to assess the robustness of an assigned weight.

4. Stochastic Multi-criteria Acceptability Analysis (SMAA)

   Our recommendations are similar to those for MCDA. Additionally, bar graph representing the acceptability indices could be used to represent the uncertainty in the ranking of the alternatives. Connected line scatter plot (effectively line graph) for the central weighting (weights specific to the given results) may also be used to provide decision-makers with an overview of typical criteria weights which contribute to alternative being ranked the way they were in any given SMAA analysis.

5. Benefit-Risk Ratio (BRR)

   We recommend bar graph, dot/forest plot, and line graph for presentation of the magnitudes of the benefit-risk ratios. Additionally, scatter plot and contour plot of the measured effects under changing assumptions may be used for sensitivity analysis. Tornado diagram by three possible states of a treatment being inferior, non-inferior or superior to an alternative for each criterion may be used to encourage absolute judgment.

6. Number Needed to Treat (NNT)

   We recommend dot/forest plot, line graph, and scatter plot for presentation of the number needed to be treated (or harmed) to observe one outcome (benefit or risk). Additionally, contour plot of the NNT under changing rates assumptions may be used for sensitivity analysis. Tornado diagram by three possible states of a treatment being inferior, non-inferior or superior to an alternative for each criterion may be used to encourage absolute judgment.

7. Impact numbers

   Our recommendations for impact numbers are similar to those for NNT above for the number of people affected.

8. Quality-Adjusted Life-Years (QALY)

   We recommend bar graph and dot/forest plot for presentation of QALY values. Additionally, line graph and scatter plot may be used for sensitivity analysis to assess the effect changing assumptions.

9. Quality-adjusted Time Without Symptoms and Toxicity (Q-TWiST)

   Our recommendations are similar to those for QALY above. The visual representations should be done to every health state defined in Q-TWiST.

10. Incremental Net Health Benefit (INHB)

    We recommend line graph and scatter plot for presentation of the incremental net health benefit. Additionally, contour plot may be used for sensitivity analysis to assess the benefit-risk balance for different cut-off points.

11. Probabilistic Simulation Model (PSM)

    There is no visual representation to be recommended with this approach since it does not directly represent benefit-risk profiles. However, network graph may be used to represent the model.

12. Mixed Treatment Comparison (MTC)

    There is no visual representation to be recommended with this approach since it does not directly represent benefit-risk profiles. However, network graph may be used to represent the model.

13. Discrete Choice Experiment (DCE)

    We recommend bar graph for presentation of elicited utilities through appropriate grouping of stakeholders and by criterion. Additionally, line graph and scatter plot may be used for sensitivity analysis to assess the change in assumptions or to assess the robustness of the results.

Conclusion

Our recommendations agree with the visuals that have been originally proposed in benefit-risk approaches such as PhRMA BRAT, MCDA and SMAA. For the visuals without specific visual presentation proposals, our recommendations make the current practice of presenting visuals more explicit. It should be remembered that, the choice of visual types to represent benefit-risk is only the tip of the iceberg; there are other aspects to consider such as tasks, audience, and the physical appearance of the visuals. It is a very difficult but interesting field for research due to diverse scientific issues from statistical to cognitive. We are in agreement with the conclusion of a recent review by the FDA ("Quantitative summary of the benefits and risks of prescription drugs: a literature review") that there is no single visual representation that consistently emerged as being better than others, and visual representations of benefit-risk need to account for the intended audience due to differences in their abilities and other cultural-specific factors.

We hope at this point in time, our contribution to this field may develop insight into visual representations in benefit-risk assessment, may emphasise the need for clear guidelines of visual communications between researchers and stakeholders, and may highlight some research questions to be explored further in future visual representation in benefit-risk assessment research.

Table of contents

Glossary and abbreviations

Terms	Description
Approach	The system of methods and principles used in a particular discipline
Aspect ratio	The ratio of the lengths of the two axes on a graph. A square graph has an aspect ratio of 1.
Cognition	The mental action or process of acquiring knowledge and understanding through thought, experience, and the senses.
Greyscale	The shades in the black and white spectrum with no other colours.
Hue	The dominant colour. Higher hue of a primary colour gives the perception that the object appears with the shades of that colour.
Line pattern	The look of a line which could be solid, dash, dot, etc.
Perception	The way in which something is regarded, understood or interpreted i.e. the translation of sense impressions into meaningful experiences of the outside world.
Preference value	The value or utility associated with a score. Preference values or utilities are judged by assessors to reflect the clinical relevance of effects or outcomes.
Rates	The relative frequency of an event in a given time period
Reference point	An anchor on the visual usually refers to meaningful values on the scale to aid information extraction
Saturation	The purity of primary colours in relation to the wavelengths. Narrower wavelengths are more saturated than wider wavelengths.
Score	A measure of a real world effect or outcome.
Utility	A subjective measurement that describes a person's or group's preferences (satisfaction, risk attitude etc.) for an effect or outcome.
Visual methods/ representation	The principles and procedures to present some numerical features or relations by a graph

Abbreviations	Description
BRAT	Benefit Risk Action Team
BRR	Benefit Risk Ratio
CPM	Confidence Profile Method
CUI	Clinical Utility Index
DAG	Directed Acyclic Graphs
DI	Desirability Index
DM	Decision-Maker
INHB	Incremental Net Health Benefit
ITC	Indirect Treatment Comparison
MAR	Maximum Acceptable Risk
MAUT	Multi Attribute Utility Theory
MCDA	Multi Criteria Decision Analysis
MTC	Mixed Treatment Comparison
NCB	Net Clinical Benefit
NEAR	Net Efficacy Adjusted for Risk
SBRAM	Sarac's Benefit Risk Assessment Method

NNH	Number Needed to Harm
NNT	Number Needed to Treat
PrOACT-URL	Problem, Objectives, Alternatives, Consequences, Trade-offs, Uncertainty, Risk, and Linked decisions framework
PSM	Probabilistic Simulation Method
QALY	Quality Adjusted Life Years
Q-TWiST	Quality-adjusted Time Without Symptoms and Toxicity
SMAA	Stochastic Multi-criteria Acceptability Analysis
TURBO	Transparent Uniform Risk Benefit Overview

Abbreviated name	**Full name**
EMA	European Medicines Agency
FDA	Food and Drugs Administration
IMI	Innovative Medicines Initiative
PROTECT	Pharmacoepidemiological Research on Outcomes of Therapeutics by a European ConsorTium

1 Introduction

1.1 The PROTECT project

Pharmacoepidemiological Research on Outcomes of Therapeutics in a European Consortium (PROTECT) is a project set up under the Innovative Medicines Initiative (IMI). Its goal is to strengthen the monitoring of the benefit-risk of medicines in Europe. This will be achieved by developing a set of innovative tools and methods that will enhance the early detection and assessment of adverse drug reactions from different data sources, and enable the integration and presentation of data on benefits and risks. These methods will be tested in real-life situations in order to provide all stakeholders (patients, prescribers, public health authorities, regulators and pharmaceutical companies) with accurate and useful information supporting risk management and continuous benefit-risk assessment. PROTECT is a collaboration between 31 private and public sector partners and is coordinated by the European Medicines Agency (EMA). This report is the first stage of the second part of the work on integration and representation of data on benefits and risks.

1.2 Visualisation and communication of benefits and risks of medicine

Visualising benefits and risks cannot be separated from their communication. There are currently many initiatives in the field of risk visualisation but these are neither specifically for visualising benefit-risk balance or trade-off, nor specifically linked to the benefit-risk assessment approaches (Cammax Limited, 2011; Gapminder, 2011; IBM, 2011; Spiegelhalter, 2010). In most cases, modern visualisations are moving towards 3-D, dynamic/animated and interactive images. The idea behind these innovative technologies is to add a narrative structure, to the much older and simpler graphics, in order to generate more interest and provide more of the required information. However, dynamic and interactive visuals are not specifically appraised. They are only briefly discussed when their static version is appraised in this review.

The issue in communicating benefits and risks is strongly intertwined with their visualisations. Testing benefit-risk communication is a subject which we cannot attempt to cover formally in any depth within the limited scope of this review, however, we are able to present and comment on simple aspects of visual communication are presented. Thorough discussions on benefit-risk communication in general have been conducted by the US FDA and are published as a user's guide covering huge range of topics (Fischhoff, 2011). Readers who are more interested in the general communication issues should follow the link to the FDA guideline given in the reference.

1.3 Objectives

The objectives of this visual representation and communication appraisal are:

1) To present the visual representations that could be associated with the 13 benefit-risk methodologies recommended in the PROTECT benefit-risk methodology review ("A systematic review and classification of methodologies for benefit-risk decision-making in medicines")

2) To provide an initial level of appraisal as to their suitability and based on previously published criteria

3) To make recommendations of potentially suitable visuals for each benefit-risk assessment approach recommended in PROTECT WP5 methodology report

1.4 Structure of the Report

In Section 2, we introduce the methods which are used in this review. The appraisal criteria for visual representation and communication are defined. The results, in Section 3, define and classify the visuals into physical types and functional tasks. The suitability of the types of visuals to carry certain tasks is discussed. Section 4 discusses this review according to the objectives, and summarises key recommendations of PROTECT with regard to communicating benefit-risk assessment through visual representations for respective benefit-risk approaches.

Whilst we recommend reading this document in its entirety there are several ways to manoeuvre through this document for different purposes. We provide a few suggestions:

i. Section 3.2 can be read to learn or confirm the names for types of visuals as used in this document;

ii. For readers who are clear about the specific tasks (as described by Carswell's taxonomy) they require but are unsure about the most appropriate visual representation of their benefit-risk assessment, Section 3.3 suggests some visuals according to tasks and also points out the architecture that exist on the visuals which facilitates the tasks;

iii. For readers interested in the recommended visuals which are appropriate for their benefit-risk model, Section 4.4 lists them and Section 4.5 proposes some issues to consider when designing the visuals;

iv. For readers who are interested in other suitable visuals, Sections 4.2 to 4.4 provide the conceptual maps of the link between benefit-risk approaches to visual representations;

v. For readers who desire further technical issues associated with different types of visuals, the Appendices are the main point of references which we encourage reading to better understand the justifications of points made. They also contain some guidelines on design of visuals for benefit-risk assessments.

2 Methods

2.1 Introduction

This review is the first stage of the PROTECT visual representation and communication review (or simply "visual review"). In this stage, only visual representations which correspond to the approaches reviewed in the PROTECT benefit-risk methodology review (Mt-Isa, 2011) are appraised. This section lays out the methods for conducting Stage I of the visual review.

2.2 Visual displays available from the methodology review

PROTECT benefit-risk methodology review presents some visual representations (or simply "visuals") of benefits and risks, individually and integrated. The visuals are aimed to communicate the benefits and risks in the most appropriate way. The visuals are classified into generic graphical types. Distinctions are made when the suggested visuals have been enhanced or modified in some way by the proposed methodology.

2.3 Inclusion of visuals identified from reviews

Initially, this review appraises all visual representations from the PROTECT methodology review. Discussion and comment are made on visuals which are methodology-specific or those originating from other methodologies described in the PROTECT methodology review that are not on the recommended list.

2.4 Criteria for visual appraisal

The characteristics of each (generic) visual representation type are described. Any enhancements or variations from the generic types are distinguished and their added values are discussed. The characteristics are shown in Table 1.

Subsequently, each visual representation type is appraised using the 13-principle of display design (Wickens, 2004). These principles are adapted to provide some level of assessment of level of suitability and concept of visual representations as decision support tools in benefit-risk assessment of medicines, and are shown in Table 2. The 13-principle of display design aims to ease cognitive workload of the decision-makers (DM) so that information on a visual can ideally be efficiently communicated to aid decision-making by reducing errors, reducing required training time, increasing efficiency, and increasing user satisfaction. In our evaluation, we have excluded Wickens's principles based on the use of mental models which address how users picture a presented visual in their minds, and the multiple resources principle which address the importance of multimodality presentation of visual with audio because these are out of the scope of this review.

The final stage of appraisal is based on the potential communication aspects of each visual representation type. It is difficult to appraise the communication potential of graphs without formally evaluating this using an audience, but it has been suggested that, to be useful they should be able to convey the five elements shown in Table 3 (Lipkus, 1999).

Table 1 Characteristics of visual representation type

Form[1]	Graphical / non-graphical (illustrations, pictures, symbols, etc.) / Static / dynamic / interactive / 2-D / 3-D
Endpoint	Rank / order / point estimates (absolute, difference, ratio) / region of equivalence
Methodology-specific	Yes / No
Reproduction	Specialised software / specialised commands / generic
Suitable audience	Public / patients / cognitive-impaired patients[2] / general regulators / specialist regulators / pharma

Table 2 Wickens' principles of display design (Wickens, 2004)

	Definition	Description
Perceptual principles		
Legibility (or audibility)	Clarity. It can be seen or heard	Any visual should be visible and legible – e.g. using contrast, colour, angle, illumination, sound, etc. This is necessary but not sufficient.
Absolute judgment	Number of levels of information, "amount of grey area".	Absolute judgment limits should be avoided by presenting DMs with discrete B-R evidence instead of continuous. For instance, a display is less prone to cognitive errors when presented with bars with different colours than when people are presented with gradually changing hues.
Top-down processing	New experience is dependent on recent past experience	Perceived message and thus the interpretation are quickly judged by DMs' recent past experience based on what they expect to perceive. If the new message is presented contrary to expectations, it may not be interpreted correctly.
Redundancy gain	Expressing the information more than once.	A message or information can benefit from more than one representation. In graphs, for example, lines can be colour-coded and also have different patterns. Redundancy gain allows the information to be interpreted correctly when one form of representation is degraded.
Discriminability	Different information should be presented differently.	Similarity causes confusion, thus discriminable elements should be used in a display. In a benefit-risk visualisation, benefits and risks criteria should be discriminated properly especially in the case when there are more than one criteria of benefits and risks. This can be achieved through colour-coding, grouping, etc.

---

[1] The visuals which are included in this review (stage 1) are of static type as they were presented as printed materials. However, if dynamic versions of the graph are available and directly related to the benefit-risk methodologies, they will be referred to and briefly discussed.

[2] Cognitive impairment covers a wide variety of deficits including those that are congenital and acquired, due to injury or disease. It is not possible to take account of this variation within the scope of this report other than to mention this as an area to be aware of

	Definition	Description
Mental model principles a		
Pictorial realism	A display should look like the variable it represents.	An arrangement or representation of the elements in a visual should look like how the variable they represent looks like in the environment. This principle is omitted from the list of evaluation criteria.
Moving part	The movement of elements in a dynamic display	The moving elements of a dynamic display should move in spatial and direction that are compatible with how the DMs think they actually move in physical system. This principle is omitted from the list of evaluation criteria.
Principles based on attention		
Information access cost	The cost in time or effort to "move" selective attention from one display location to another to access information.	The cost should be minimised to reduce the time required and cognitive effort.
Proximity compatibility	The closeness of required related information	Information from two or more sources may be required to complete a task, and should be available nearby. For example, any unfamiliar symbols or patterns are given in a legend within or close to the graph area. In benefit-risk visualisation, it is important that the information for different options is in close spatial proximity to allow them to be compared. If a visual requires mental integration, close spatial proximity is good, but if focussed attention is required, close spatial proximity may be harmful.
Multiple resources	Multimodality in presenting information.	Sometimes it is better to present information as both visually and auditorily. We recognise that auditory/vocal guide from experts can help to improve the understanding and interpretation of all visuals. This principle is omitted from the list of evaluation criteria.
Memory principles a		
Use of existing knowledge of the world	The use of long-term memory from DMs' past experience.	DMs may recall something similar when presented with a visual for benefit and risk. The more agreement there are between DMs' past experience and the newly seen information, the more effective a judgment can be made. However, human memory is much more complex and therefore it is difficult to disentangle and predict which knowledge to be represented would already exist or might be conflicting.
Predictive aiding	Any predictive tasks should be assisted	Predictive tasks, where possible, should be presented as perceptual tasks to reduce information access cost (8). In benefit-risk assessment, predictive aiding has a close analogy to the integration of benefits and risks.
Consistency	Consistency when presenting information in (a series of) displays	It is important to be consistent when representing information because DMs' memory is triggered when seeing something that is expected to be appropriate. This may cause confusion thus increasing processing time. The best approach is to use standard representations (colour, patterns, symbols etc. where possible),

Definition	Description
	and particularly in a same (lengthy) document. For example, many people associate colour red with bad and green with good. In representing benefit-risk, when benefit is represented as red and risk as green, DMs' may get confused and could potentially lead to making incorrect decisions.

[a] Mental model and memory principles can be very culturally specific. Therefore, cultural differences, as well as experience and target audience, should be taken into account when representing benefits and risks visually.

Table 3 Elements of visual communication (Lipkus, 1999; Lipkus, 2007)

Element	Description
Risk magnitude	How large or how small the magnitude of benefit or risk is. Small probabilities should be communicated with care as the general public has difficulty understanding them. This could substantially affect how people weigh events with small probabilities.
Relative risk[3]	The comparison of the magnitudes of two or more benefits or risks, or the relative magnitude of benefit and risk. Reducing the need to perform complex mental arithmetic can help reduce cognitive workload contributing to better decisions.
Cumulative risk	The estimate of how benefit or risk trends change over time. Benefit or risk magnitude may be very small at any given time, but would add up over time.
Uncertainty	The variability or ranges of the point estimates (magnitude, relative risk, or cumulative risk). As value of the point estimate increases, people's perceived variability decreases i.e. standard deviation is weighed by reciprocal of the mean (Lathrop, 1967). Different sequences also affect perceived variability. Increased of variability may lead to inflation of probabilities, thus affecting decisions.
Interactions among risk factors	The synergy effect on the overall magnitude of benefits or risks. Interaction of multiple risks may contribute to greater risk than the sum of individual risks; and people often underestimate multiplicative risk.

[3] This is loosely used here and not only referred to the relative risk or incidence rates ratio as used in epidemiology

2.5 The functional task of visual representation

The functional task of visual representation is an important element of visual design since the same visual display may not be as effective or useful when presented for different purposes. The effectiveness of a visual is affected by its characteristics, the conditions in which it is presented, the complexity of the data, the task or purpose of presentation, the characteristics of the audience it is presented to, and the criterion for choosing a particular visual (Lipkus, 1999; Meyer, 1997).

It has been suggested that visual tasks can be assessed by taxonomy of four basic tasks: point reading, local comparisons, global comparisons, and synthesis (Carswell, 1992; Lipkus, 1999). An example of point reading is judging a magnitude of a single element of the graph. Local comparisons involve comparing the magnitudes of two elements on the graph. In a global comparison, other quantities on the graphs like the magnitudes and time periods for different elements are put assessed. Synthesis judgments can be made when all data points being presented have been considered, for example when judging whether a disease risk is increasing or decreasing.

2.6 Strategy for preliminary visual recommendation

The recommendations are made based on Cleveland's taxonomy and supported by Tufte's data-ink ratio principles (Cleveland, 1984; Tufte, 2001). In effect, these would favour simpler visual representations over more complex ones if the same information can be conveyed with similar degree of accuracy through the simple visuals.

3 Results

3.1 Introduction

There are two classifications of visual displays presented in this chapter (and review). The first is a classification by visual type (Section 3.2), and the second is a classification by visual task (Section 3.3). The purpose of having two classifications is to help with the structure of the review so that it becomes more digestible. Remarks on the connection between both classifications are made along the way.

In Section 3.2, we introduce the different types of visuals, some of which are more common than others. The same visual types share similar features which provide certain advantages when used in benefit-risk assessment analysis. Because of the features they share, they also have similar drawbacks. The appraisals on these features are discussed in full in the Appendix.

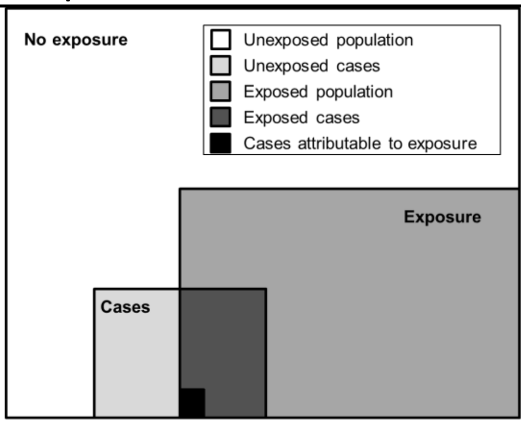
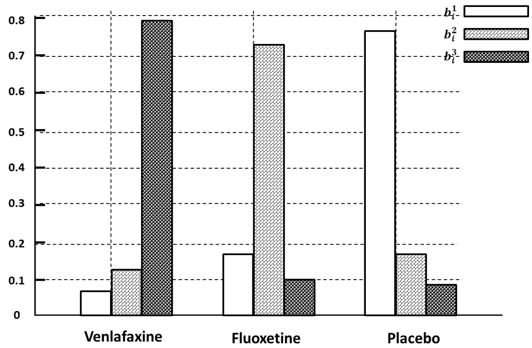
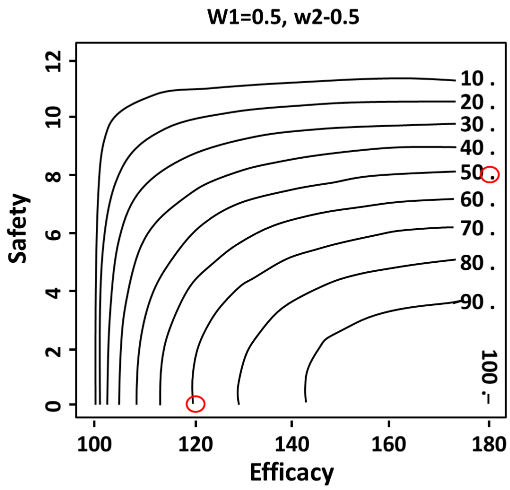
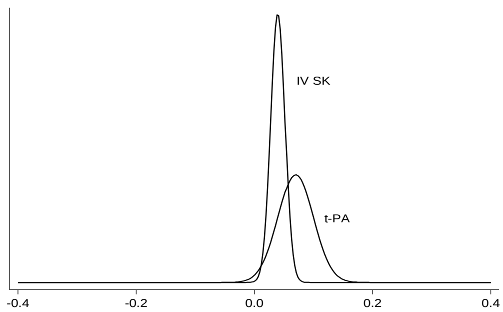
Having introduced the types of visual displays, we then discuss the basic tasks for displaying visuals according to Carswell's taxonomy (Carswell, 1992) in Section 3.3. The types of visuals which are suitable for each task are listed and discussed. Fuller justifications as to why a visual is suitable follows directly from the appraisal of each visual type (see the Appendix).

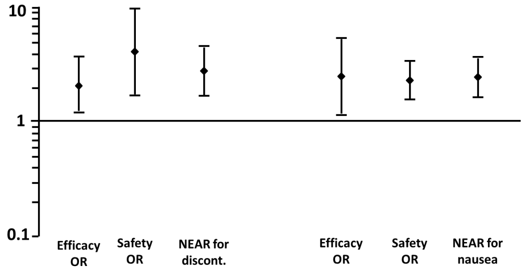
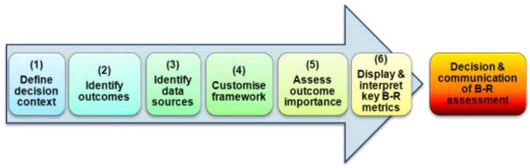
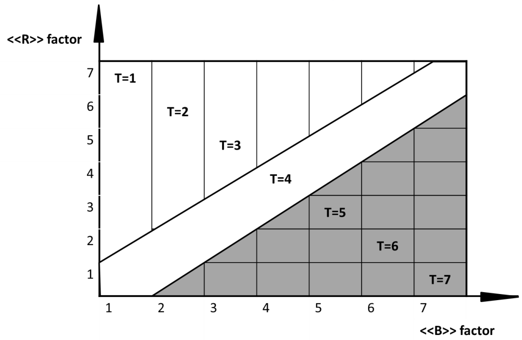
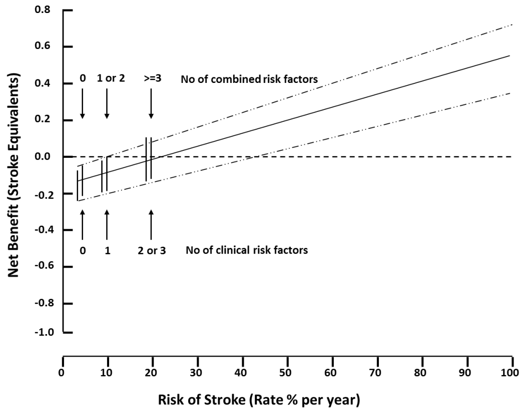
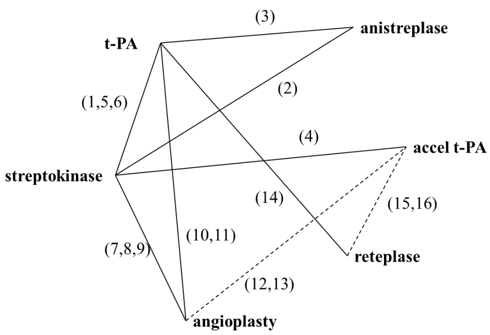
3.2 Types of visual displays

We classified visual displays in the PROTECT methodology review into 13 graphical[4] or visual types in this appraisal. We then appraised the characteristics and capabilities of the visuals according to these graphical types. The definitions and examples of visual types, as well as the benefit-risk approaches they correspond to in the PROTECT methodology review, are given in Table 4.

4 We have not distinguished graphs from diagrams in this review. We use the term "graph" as a collective description of both visual types.

Table 4 Types of visual representations available from PROTECT methodology review (recommended approaches are italicised) [a]

Visual type	Example
Area graph Definition: Information is presented by the size of an enclosed shape against common aligned or unaligned scales. Approaches: *MCDA* (frontier graph), *NNT*, *Impact numbers, Q-TWiST*	
Bar graph Definition: Information is presented by rectangular bars for a number of categories. The position (height of bars) along a common scale is judged supported by the length of the bars. Approaches: *MCDA* (also stacked and colour-coded, and the 'difference display'), *SMAA*, MAR (bar/antenna)	
Contour plot Definition: Information is presented by usually a number of curved lines along common aligned scales. Approaches: CUI/DI	
Distribution plot Definition: Information is presented by the curved line representing the shape of the distributions, and the area under the curves along a common aligned scale. Approaches: NCB (with summary table), CPM (overlapping)	

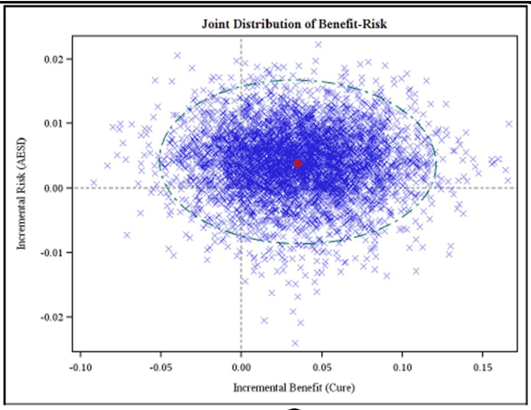
Visual type	Example
Dot/Forest plot Definition: Information is presented as a number of symbols, usually representing the mean effect size along common aligned scale. Each symbol sits on a vertical or horizontal line which usually represents the 95% confidence intervals of the mean effect. Approaches: *BRAT* (with summary table), *NNT* (inc. reversed axis), *Impact numbers*, NEAR	
Flow diagram Definition: Information is presented in a series of ordered tasks. Approaches: *BRAT*	
Grid/Table Definition: Information is presented by the intersection of rows and columns. Written texts are common in tables, but grids make use of common aligned scales. Approaches: PrOACT-URL (as 'effects table'), TURBO, Principle of three, FDA BRF	
Line graph Definition: Information is presented by the position of lines along common aligned scales. Approaches: NCB (threshold, with CI), *MCDA* (also with area), CUI, *QALY*, *INHB*, GBR (with CI)	
Network graph Definition: Information is presented at the ends and on the connecting lines. Approaches: DAGs, CPM, ITC/*MTC*	

Visual type	Example

Scatter plot

Definition: Information is presented as symbols on common aligned scales.

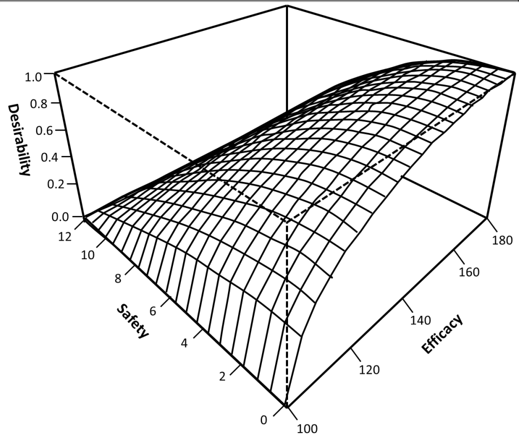
Approaches: *QALY*, *INHB*, *PSM* (with threshold lines)



Surface plot

Definition: Information is presented as wireframe or sheet representing the position of points in the three-dimensional space on common aligned scales.

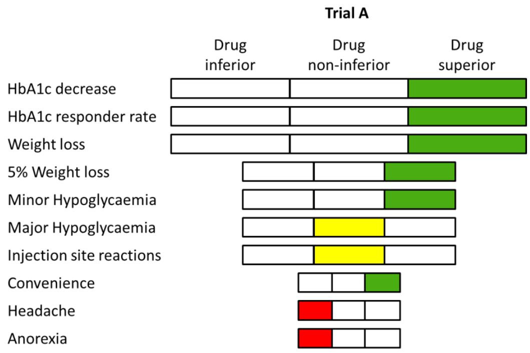
Approaches: CUI, DI



Tornado diagram

Definition: Information is presented as length and position of the rectangular bars on non-aligned scales.

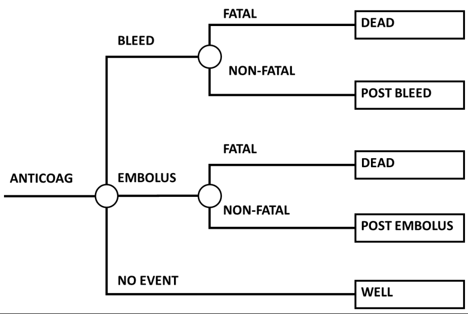
Approaches: SBRAM (also colour-coded)



Tree diagram

Definition: Information is presented at the end of “branches” and on the point where they cross.

Approaches: Decision tree, MDP, *MCDA, BRAT*



[a] The visuals on this table are meant to give a general idea of how each visual representation type may look like and the details are not intended to be legible. See Appendices for full size images and other examples.

3.3 Classification of visuals by task

Carswell's taxonomy specified four basic tasks associated with visual displays (Carswell, 1992). In this section, we define these basic tasks with reference to benefit-risk assessment, and indicate the suitability of visual types for a given task.

3.3.1 Point reading

The simplest task is 'point reading' which involves judging and understanding a particular point on a visual display. In a benefit-risk visual representation, point reading often requires judgment of the magnitude and direction of a benefit or risk criterion independently. Although the task is straightforward and simple, DMs also need to understand the magnitude in the context it is presented including understanding the unit of measurement and how this relates to the DMs. Another aspect of point reading is to understand the "direction"; and it is imperative that DMs properly understand whether a greater magnitude is associated with greater preference or with less preference. In general, the magnitude of a benefit is proportional to the direction of preference, and the magnitude of a risk is inversely proportional to the direction of preference.

The visual representations which promote point reading are listed in Table 5.

Table 5 The architecture of visuals which permits point reading

Visual type	Point reading architecture
Bar graph (Appendix A.2)	The height of the bar read horizontally against the vertical axis for vertical bar graphs or the length of the bar read vertically against the horizontal axis for horizontal bar graphs. This assumes that the widths of the bars bear no additional information.
Dot/Forest plot (Appendix A.5)	The position of the symbol in the middle, and the two ends of each vertical/horizontal line (in forest plot) read horizontally/vertically against the vertical/horizontal axis.
Grid and table (Appendix A.7)	The position of a point on a grid read horizontally/vertically against vertical/horizontal axis; and the written figures in a column/row on a table.
Scatter plot (Appendix A.10)	The position of the symbol read horizontally/vertically against the vertical/horizontal axis.

3.3.2 Local comparison

The next task in the taxonomy is 'local comparison', which may be the most essential task in a benefit-risk assessment for decision-making for the patients and/or carers. Local comparison requires DMs to perform point reading for two alternatives, say treatment options, and to compare them to determine a better alternative from the DMs' point of view at a fixed point in time. Point reading does not need to be accurate to perform a local comparison task since DMs may only compare the relative importance of the criteria. However, if the more accurate point reading is required, cognitive mental processing is increased. A comparison of a benefit (or risk) criterion between two alternatives is an example of local comparison task, as is a comparison of the total benefit-risk balance between two alternatives when the benefit-risk measures are integrated. Comparison of non-integrated benefit and risk criteria requires more cognitive effort than it is required for local comparison task (see Section 3.3.3).

The visual representations which promote local comparison are listed in Table 6.

Table 6 The architecture of visuals which permits local comparison

Visual type	Local comparison architecture
Area graph (Appendix A.1)	The size of an area compared to the size of another area following point reading. However, we acknowledge that area judgment and comparison suffer from perceptual distortion bias (Cleveland, 1984).
Bar graph (Appendix A.2)	The heights of two bars are compared following point reading. This assumes that the widths of the bars bear no additional information.
Dot/Forest plot (Appendix A.5)	The position of the symbol in the middle (e.g. mean), or the two ends (of CI in a forest plot) line from a criteria are compared with those of another criteria following point reading. This assumes that the symbol sizes bear no additional information.
Grid and table (Appendix A.7)	The position of a point on a grid and a figure in a table cell is compared to another point or figure in another cell following point reading.
Line graph (Appendix A.8)	The position of any point on the line is compared with another point (on the same line or on another line) following point reading.
Scatter plot (Appendix A.10)	The position of a symbol is compared to the position of another symbol on the scatter plot following point reading. This assumes that symbol sizes do not bear additional information.

3.3.3 Global comparison

'Global comparison' requires mental arithmetic to be done in order to make the desired comparison. Similar to local comparison, DMs perform point reading on several items on a visual, mentally combine them and then make the comparison. The comparison of non-integrated benefit and risk criteria for different alternatives is an example of global comparison task. Global comparison tasks grow in complexity when there are many criteria involved or when many time points are involved in the decision-making process from the visuals. Most importantly, cognitive efforts are greatly challenged when the criteria to be compared are not presented in the same unit to allow direct trade-off. For many people, simple mental arithmetic may be difficult and mathematical transformation may be beyond what most people are comfortable with in terms of comprehension.

The visual representations which promote global comparison are listed in Table 7.

Table 7 The architecture of visuals which permits global comparison

Visual type	Global comparison architecture
Area graph (Appendix A.1)	Please see Table 6 in Section 3.3.2. Additionally by mentally adding up different areas or by comparing subsets of defined areas.
Bar graph (Appendix A.2)	Please see Table 6 in Section 3.3.2. Additionally by mentally stacking or un-stacking bars to make comparisons.
Dot/Forest plot (Appendix A.5)	Please see Table 6 in Section 3.3.2. Additionally by mentally adding several criteria for comparisons, or by comparing the midpoints (e.g. mean or median) and the lower and upper ends (e.g. confidence intervals or ranges in a forest plot) of two or more criteria. Also to take into account any information bear by the symbol sizes.
Line graph (Appendix A.8)	Please see Table 6 in Section 3.3.2. Additionally by mentally adding lines for comparisons.
Scatter plot (Appendix A.10)	Please see Table 6 in Section 3.3.2. Additionally by comparing more than two points (symbols) or when taking into account any additional information represented in the symbol sizes.
Surface plot (Appendix A.11)	The position of point in the three-dimensional space itself already provides a global comparison for comparing the combined values of any two elements to another. The position of a point can also be compared to another against any of the three axes.
Tornado diagram (Appendix A.12)	The length and position of rectangles on the bars for the discrete benefit-risk balance for one criterion compared to another. Also, several criteria can be combined mentally before making comparison about the combined length and position of the rectangles.

3.3.4 Synthesis judgment

The most demanding task according to Carswell's taxonomy is the 'synthesis judgment' where DMs are required to look beyond the graph itself. Although demanding, it is not necessary to obtain the exact values of the benefit-risk balance, it is sufficient that the visuals allow DMs to think beyond the presented results. In a benefit-risk assessment, this could be extrapolating the information from a presented visual – for example, a DM may want to perceive what the risks of medication are to him/her in long term but only has visual information on short term risks. Synthesis judgment also includes the need for assessing statistical uncertainties involved in a benefit-risk assessment from visuals like scatter plots and line graphs.

The visual representations which promote synthesis judgment are listed in Table 8.

Table 8 The architecture of visuals which permits synthesis judgment

Visual type	Synthesis judgment architecture
Contour plot (Appendix A.3)	The position of the points on the lines on the plot against both axes which can be extended for points outside the plot. The curvature and proximity of the contour lines provided that the line thickness is uniform, otherwise there may be perceptual distortion that may alter judgment.
Distribution plot (Appendix A.4)	The area under the curve beyond certain point on the axis, and the shape and the position of the distributions.
Dot/Forest plot (Appendix A.5)	*Please see Table 7 in Section 3.3.3.* Additionally the length of the error bars represents the amount of uncertainty which may affect judgment. If plotted over time, DMs may judge the effect size outside the presented time range.
Line graph (Appendix A.8)	*Please see Table 7 in Section 3.3.3*. Additionally, the positions of the points and direction of lines outside the graph region.
Scatter plot (Appendix A.10)	*Please see Table 7 in Section 3.3.3.* Additionally the patterns of how the points are positioned allow judgment of uncertainty or correlation to be made.
Tornado diagram (Appendix A.12)	*Please see Table 7 in Section 3.3.3.* Additionally when having to compare more than two alternatives.

3.4 Summary appraisal

3.4.1 Evaluations of visuals

We evaluated each visual type in Section 3.2 using Wickens's principles of visual display. There are inevitably additional surrounding issues and possibilities in designing visuals since visuals are highly dependent on their designers and the tasks. In this section, we describe and summarise our evaluations of the visuals in general, whilst presenting the details of the appraisal specific to visual types in the Appendices.

Our evaluations suggest that flow diagrams, network graphs, and tree diagrams are not likely to be the best methods for presenting the results of benefit-risk assessments. Other visual types may be more appropriate and effective so long as their designs address the three domains in Wickens's principles of display design – "perceptual principles", "principles based on attention" and "memory principles" – to a good extent. Another domain, the "mental models principles" has been excluded (see Section 2.4). We briefly discuss the visual appraisals in each domain in turn below.

"Perceptual principles" is the first domain. The principles in this domain ensure that visual representations can be perceived accurately by the DMs to avoid misunderstanding. Any type of visual display runs a risk of expressing information that results in biased perception. On the positive side, any type of visuals can be customised with appropriate fonts, symbols, colours, contrasts, patterns etc. to communicate the required information for the required task. However, the use of gradually changing hue or greyscale can limit judgment and could be difficult to discriminate from each other; i.e. when used, they should be accompanied by clear boundary lines. In general, perception bias is likely to be less when simpler presentation is used e.g. points or lines in 2-D visuals compared to higher dimension e.g. points or lines in 3-D visuals. Elements of visuals presented by area, volume or angle may also introduce more perception bias. In particular, horizontal lines are easier for human brains to process when compared to vertical lines (Cleveland, 1994). The five principles addressed in perceptual principles domain, which are legibility, absolute judgment, top-down processing, redundancy gain and discriminability can be used as a checklist to help clarify what specifics may minimise any perception bias when DM extracts benefit-risk information.

“Principles based on attention” is the second domain. The principles in this domain ensure that any related information required for the task is clear, intuitive and easy to be found in relation to the visual. This means that for any two values to be compared they should be labelled properly, aligned on the same scales, within close proximity to each other, and accord with perceptual principles. Visuals communicating too much information or data points can be very costly in terms of extracting the correct information, which may affect contour plot, forest plot, scatter plot and surface plot most. Therefore, depending on the tasks and media of presentations, the amount of data points or information to be communicated must be considered carefully.

We omitted the third domain based on “mental model principles” because they are not relevant to the types of visuals reviewed here and are difficult to evaluate.

“Memory principles” is the last domain and aims to take account of users’ previous experience directly or indirectly related to the visuals being presented. For these principles, simpler visuals with fewer data points or information may outdo more complex visual presentations. It may be that audiences will have had greater exposure to simpler visuals and this may make them able to adapt to them more easily. A good benefit-risk visual representation should allow users to see the information clearly by exploiting users’ daily experience: an example would be the use of the colour green to mean a good outcome, and the use of colour red to indicate a worse outcome. However, cultural context may mean that these colours are not appropriate and individual differences may mean different weightings for such cues for example a risk-averse user may place a very high weight for any criterion appearing in red without considering its magnitude or the benefits. Choosing colours may also depend on the audience characteristics; for example the use of red-green combination is not suitable for red-green colour blindness audience. Assistance on choosing suitable colour combinations is freely-available (Brewer, 2006). Any benefit-risk information may require complex cognitive process and may benefit from being aided, for example by presenting a composite measure instead of individual measure. The simplest visual representation to communicate composite measures is likely to be the bar graph. In a series of related visuals particularly, consistency plays a big role to avoid confusion or reduce the time required. From our experience carrying out this review, we suggest that consistency should be emphasised when there is a need for comparing more than one figure via use of consistent graphics scheme (colour, patterns, symbols, etc.), consistent alignment of the axes and scales, consistent sizes and aspect ratios, and so on. Any sets of visuals can be made consistent.

Where simpler visuals can communicate the same information when compared to complex ones, we would propose that the simpler ones should be chosen. Tufte’s data-ink ratio can be used when deciding between two types of visuals, with the exception of the bar graph versus scatter plot since it has been argued previously that a bar graph provides a better perception of magnitude than a scatter plot, allowing better decision to be made.

3.4.2 Evaluations of communicability

We evaluated each visual type in Section 3.2 against Lipkus's elements of visual communication.

Almost all visuals reviewed here aim to communicate risk magnitude with the exception of flow diagrams, grids, network graphs, and tree diagrams. Dot plot in particular has been promoted as a very useful graph type when conveying risk magnitudes (Cleveland, 1994; Heiberger, 2004; Robbins, 2005; Tufte, 2001). Risk magnitudes may be more easily communicated through simple visuals such as Cartesian graphics in an attempt to make information extraction easier. In a static presentation of visuals for example when printed on paper, 2-D visuals are likely to be better at communicating magnitudes than 3-D visuals.

When extracting benefit-risk information from visuals, it is not only the magnitude of benefit or risk that matters but also the magnitude in comparison to another. The comparator could be a different benefit, a different risk, the same benefit or risk at different time points, or the same benefit or risk at the same time point in a different scenario, etc. The term "relative risk" is used loosely here to describe the relative magnitudes of any two items. This is equivalent to "local comparison" and "global comparison" where the risk magnitude may not need to be read accurately. Therefore, any visuals that can clearly convey whether an item is higher or lower than another can communicate relative risk to a certain extent.

Communication of cumulative risk is usually associated with a time element. The most obvious way to convey cumulative risks over time is to plot them against time which can then be read directly from the visual representation, thus reducing the efforts for information extraction. This can be done easily as line graph or scatter plot of cumulative risks. Other visuals can also communicate cumulative risk but with variable degree of effectiveness. In general, simpler clean visuals are likely to be better at conveying cumulative risks rather than complicated or even chaotic visuals. For example, a typical forest plot communicates risk magnitudes easily but may struggle to communicate cumulative risks.

In general, communication of uncertainties requires many data points which could be presented visually without overwhelming the user. Uncertainty may be best conveyed together with summary estimates to put things into context. In this case, forest plot may be an obvious choice and is likely to be simple to read. A distribution plot may be the best depiction of uncertainties of a variable but can be difficult to produce because the distribution of a variable is not always Gaussian normal. Other visuals like the contour plot and surface plot can show great details of uncertainties in a variable but extraction of information may become difficult.

Communication of interactions among risk factors is felt to be very difficult through conventional visuals and is not commonly done. Unless the visuals specifically show the effects of interactions, extracting the related information would likely require greater cognitive effort. Dynamic visuals may be the more suitable type of visuals in this case but are not included in this part of review.

4 Discussion and recommendations

4.1 Discussion

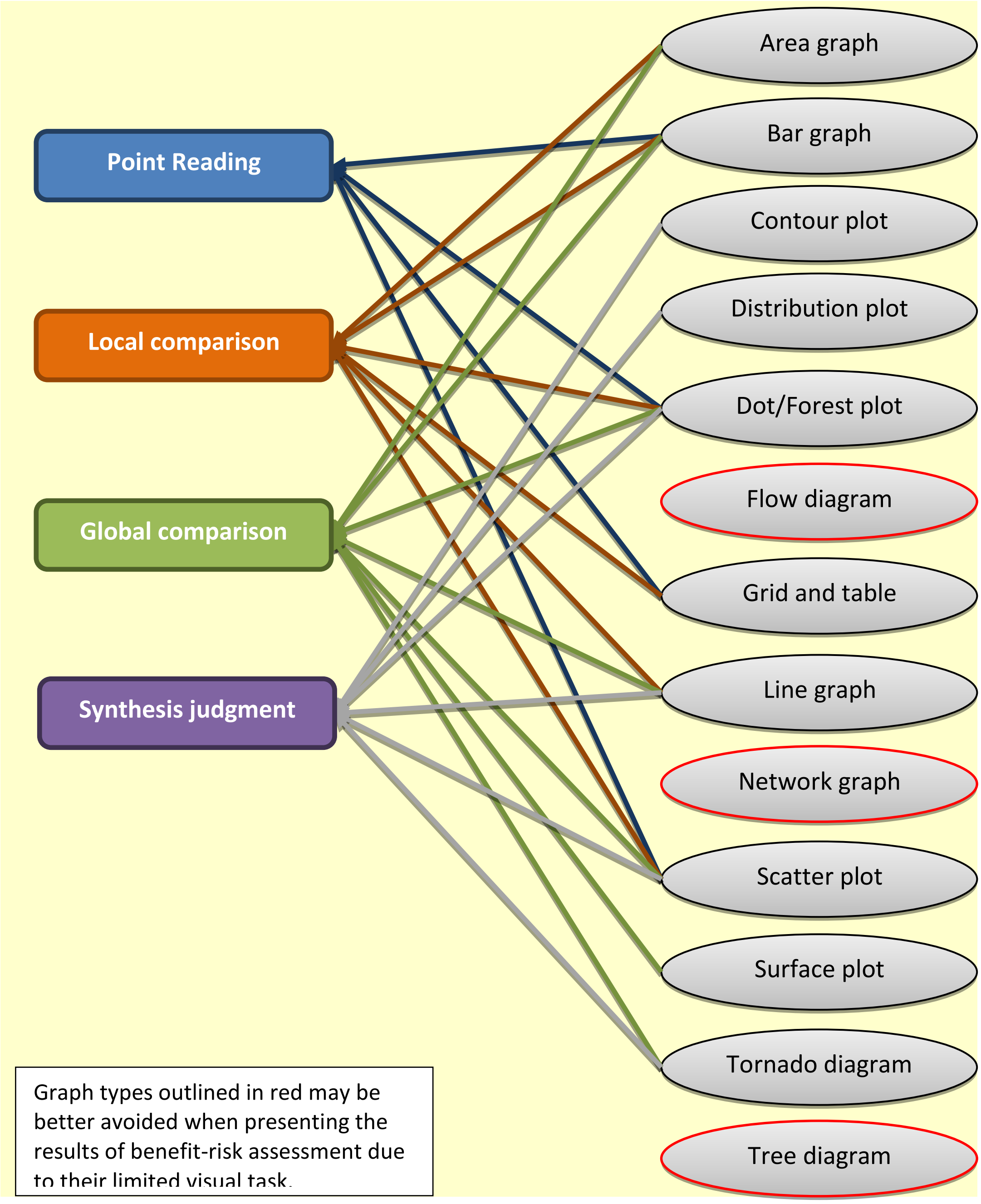
This review appraises the usefulness and usability of visual representations being used in benefit-risk assessment based on the literature. This review only set out to elicit visual representations which are suitable for specific benefit-risk approaches but in due course inevitably touched on visual design issues. Despite this we have, by necessity steered away from specific guidelines for the design of visual representations. The American Statistician published a good cynical commentary in 1984 on guidelines of how to display data badly which should be taken seriously (Wainer, 1984). There are also texts from Tufte on how to display visuals properly (Tufte, 2001).

The recommendations made in this review are typical visuals related to those that have already accompanied the benefit-risk approaches encountered through the review of methodology. At this time, we are not in the position to explore further into the potential of more innovative and modern visual displays which may interest decision-makers in the decision problem and making the decisions to be made more personal.

4.2 Capacity of visual representations and their relationship to tasks

Following the discussion in Sections 3.3.1 – 3.3.4, we summarise how these tasks may be communicated visually in Figure 1 below. Several types of visual can be used to achieve the same tasks but with variable degree of accuracy and complexity in the design and information extraction. Visual types which do not map to tasks are outlined in red; therefore are not suitable to represent benefit-risk assessment results. However, they may be suitable when presenting other aspects of benefit-risk assessment, e.g. to represent the process involved or the relationship of evidence. The Appendices contain detailed evaluations of each visual type.

Figure 1 The relationship between tasks and visual types



4.3 Relationship of benefit-risk approaches to tasks

In the methodology review, we identified three types of resultant metrics which are associated with the benefit-risk approaches:

i. scores (and weighted scores),

ii. rates (and weighted rates),

iii. utilities (and weighted utilities)

None of the approaches which are only scores-based were recommended. Therefore, the visual representations suitable for these approaches are omitted in this review. We speculate that scores may be presented in the same way as rates.

Additionally, the resultant metrics are presented in four simple forms:

i. ratio,

ii. difference,

iii. rank,

iv. sum[5]

Simplifying into these recognisable forms, DMs do not need to know the exact type of metrics from a benefit-risk approach to be able to determine suitable visuals, allowing them to work backwards if needed to.

In order to determine suitable visual representations, benefit-risk assessors have to establish the tasks that are required of the DMs. Clearly, several tasks may be required or need to be considered in one decision problem for an informed decision to be made. In our opinions, not every type of benefit-risk assessment results is effortlessly compatible with every task. Ratios and differences are calculated at different stages of benefit-risk assessment, thus may be compatible with any task. Sums and ranks are positioned at the two extreme ends on the spectrum of being derived from very simple calculations to very complex calculations and process. Either way, they may be only suitable for very elementary tasks such as point reading and local comparison since the complex derivations eliminate the need for more complex tasks. In the situations when the derivations are too simplistic, complex tasks may be too cognitively challenging and are likely to introduce more biases and errors.

Figure 2 shows the compatible pairings as guide to selecting appropriate visuals in the communication of benefits and risks; and Figure 3 shows the relationships between approaches and the results from their analysis, which completes the map.

[5] For simplicity, the face value of one item is also considered as a "summation" i.e. $\sum_{i=1}^{1} x_i$.

Figure 2 Compatible pairings between resultant metric forms and required visual tasks

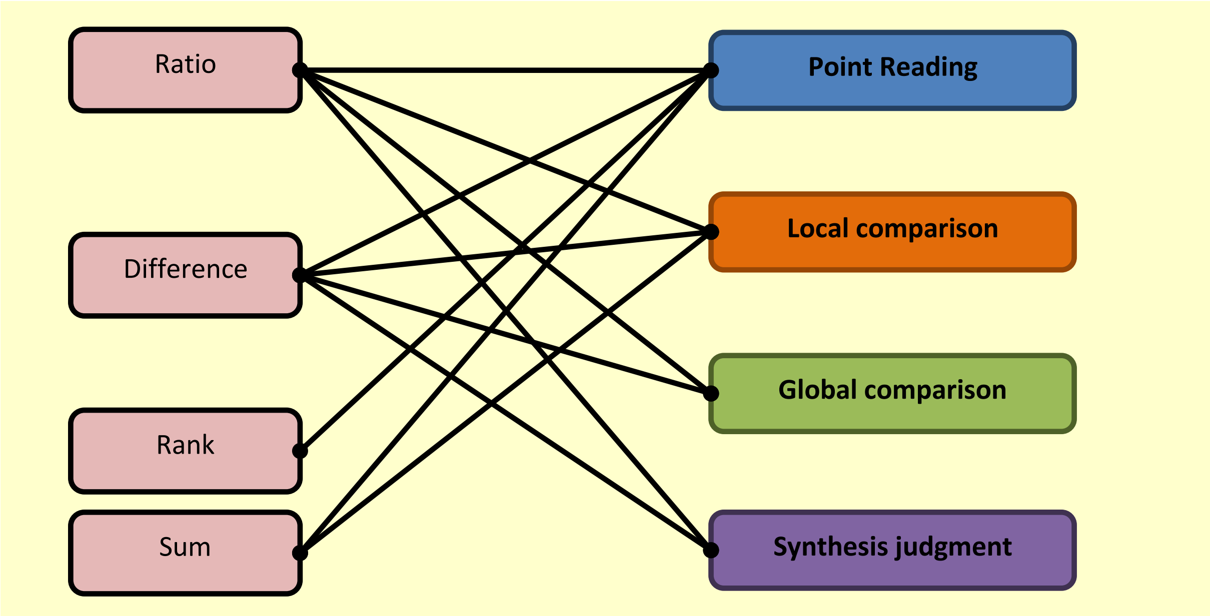
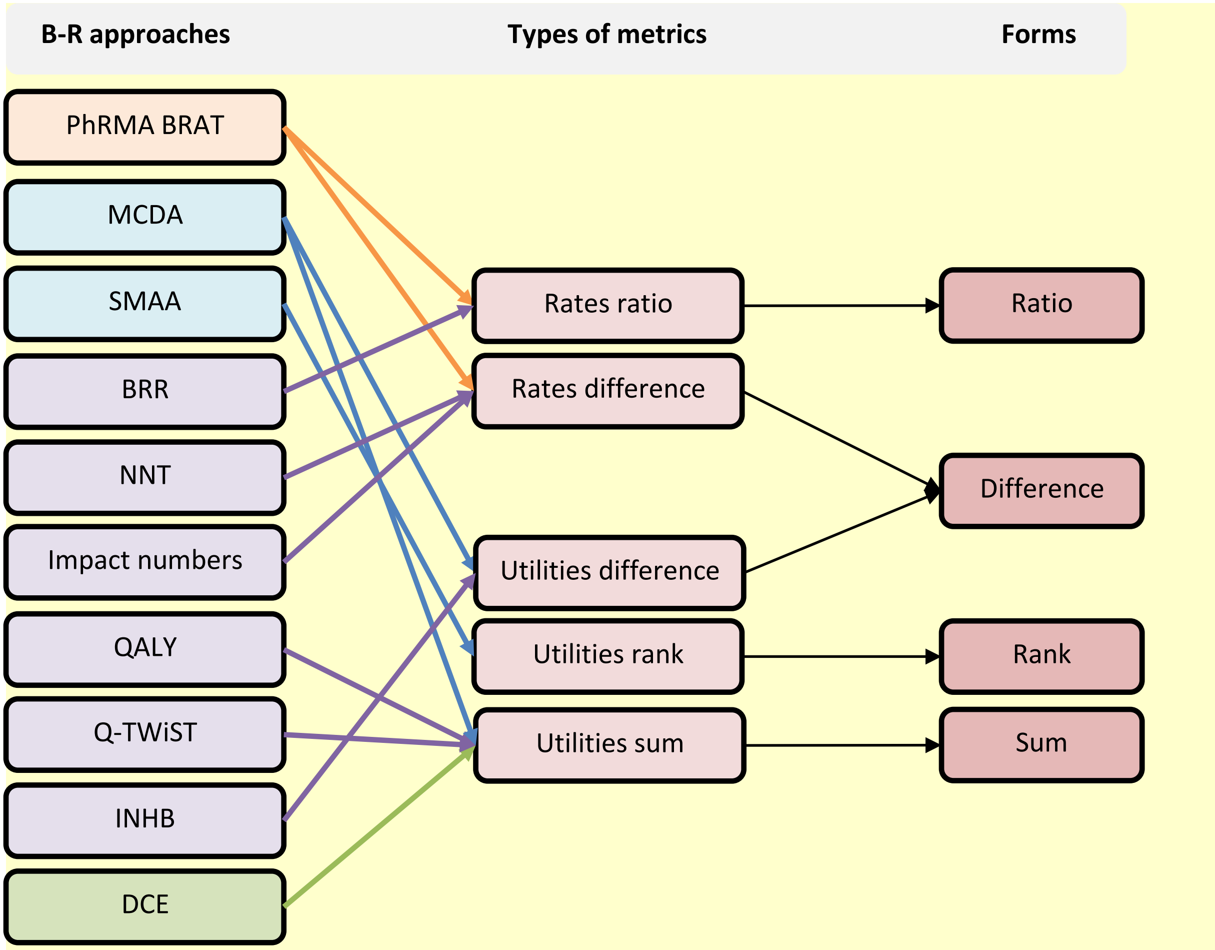


Figure 3 The relationship between approaches and metrics of their results (PSM and MTC are omitted because there are no specific results, whilst PrOACT-URL maps directly to presentation of effects table)



4.4 Benefit-risk approaches and key recommendations

From previous Sections 4.2 and 4.3, we identified many possibilities of presenting the results from a benefit-risk assessment. The possibilities are overwhelming even for very few benefit-risk assessment approaches. To (over-)simplify recommendations at this stage, we recommend at most three visual representation types for each approach to represent the results of benefit-risk assessments. The top three recommendations are based on simplicity; taking into consideration Cleveland's theory of graphical perception for the accuracy of information extraction and Tufte's data-ink ratio (e.g. contour plot is favoured over surface plot) wherever there are ties (Cleveland, 1984; Tufte, 2001).

Table 9 Preliminary recommendations of work stream B for visuals representations of benefit-risk by recommended approaches

Approach	Visual representation of results	Other visual representations of special interest
PrOACT-URL	'Effects' table	n/a
PhRMA BRAT	Table, dot/forest plot, bar graph	Tree diagram to represent model.
MCDA	Bar graph, 'difference display'	Table for evidence data, tree diagram to represent model, line graph for sensitivity analysis.
SMAA	Bar graph, dot/forest plot	Table for evidence data, tree diagram and distribution plot to represent model, line graph and scatter plot for sensitivity analysis.
BRR	Bar graph, dot/forest plot, line graph	Scatter plot or contour plot for sensitivity analysis. Tornado diagram may be suitable to simplify further the results.
NNT	Dot/Forest plot, line graph, scatter plot	Contour plot for sensitivity analysis. Tornado diagram may be suitable to simplify further the results.
Impact Numbers	Dot/Forest plot, line graph, scatter plot	Contour plot for sensitivity analysis. Tornado diagram may be suitable to simplify further the results.
QALY	Bar graph, dot/forest plot	Line graph or scatter plot for sensitivity analysis.
Q-TWiST	Bar graph, dot/forest plot	Line graph or scatter plot for sensitivity analysis.
INHB	Line graph, scatter plot	Contour plot for sensitivity analysis.
PSM	n/a	Network graph to represent model.
MTC	n/a	Network graph to represent model.
DCE	Bar graph	Line graph or scatter plot for sensitivity analysis.

The recommendations made herewith are based on the assumptions that there are resources to produce the recommended visuals. The simple visuals that are favoured in this review are mainly due to the fact that they are likely to take less time for a DM to understand, reduce confusion and that they are presented on paper. Most importantly, all visual representations should follow general graphic design principles, such as labelling of all axes, and specifically adhere to Wickens' principles listed in Table 2.

4.5 Test phase of visual representations

We previously discussed that visual representations do not naturally link to the benefit-risk approaches in a simple manner but in combination with the required tasks. The four basic tasks of Carswell's taxonomy are very general; therefore when designing the visuals, the exact task should be thought about and stated clearly. In addition, the visuals should also be suitable for the means and conditions of their presentations. The last important aspect to consider is the audience for whom the visuals are intended for, which the Wickens's principles of display designs should be carefully addressed. Table 10 presents a brief example of the visual considerations for a hypothetical benefit-risk assessment of drug X versus drug Y; for which a visual is to be designed. In this hypothetical scenario where the benefit-risk assessment could be via BRR, NNT or impact numbers, the possible visuals are bar graph, line graph, scatter plot, and tornado diagram (as proposed in SBRAM) for unassisted presentation, or additionally forest plot and contour plot for assisted presentation.

Table 10 Important aspects to consider when designing visual representations of benefit-risk assessments

Task	To judge which of the two alternatives have smaller threshold for psychiatric adverse events (local comparison)
Means of presentation	Visuals presented on paper
Conditions of presentation	Visuals are presented with assistance from a physician with time-constraint of 15 minutes (physician's available time); or unlimited time without assistance
Audience characteristics	Patients with no cognitive impairment

4.6 Conclusion

As previously mentioned, in order to determine the suitable visuals to benefit-risk assessment approaches, it is inevitable to put aside the visual design principles. A correct visual type may not be as useful if there are design flaws such as illegible symbols. We firmly acknowledge that this review is based on theoretical work in the literature and personal experiences and opinions. Therefore, it is only not possible to be certain which types of visual representations will actually work better, or if the audience with similar characteristics would gain the same understanding from the same visuals. The former, although difficult, may be formally tested in experiments, but the latter is almost impossible to test.

In reality, choosing the visual to represent benefit-risk assessment has substantial subjective elements similar to choosing the correct utility weights in a decision model. Visual designers may simply choose to present certain types of graphs just because they are easy to produce, or appear to be attractive to them. None of the visual representations were found to be superior across all benefit-risk approaches or metrics, and they are likely to be dependent on the intended audience and required tasks. These conclusions concur with findings in a recent literature review commissioned by the FDA (West, 2011). Nonetheless, as in any statistical modelling, the simpler ones would be favoured over the more complex ones if there is no clear advantage for a more complex visual representation.

4.7 Further work

So far this review only appraises the types of visuals which were encountered in the methodology review. There are many other innovative visual representations that are rarely seen in the academic literature for example interactive and dynamic visuals, which are planned for the second stage of PROTECT visual review.

We also acknowledge that scores (and other measured metrics) are also used in benefit-risk analyses. Whilst speculating that their visual representations may be closely similar to those for rates, scores may require different representations to be effective and accurately communicated. Therefore, further work into this aspect may gain some insights into the related issues.

In this review, we did not explicitly discuss the capacity of visual representations when communicating efficacy/safety versus communicating benefit/risk. This is another crucial aspect that needs to be discussed further since efficacy and safety presentation does not directly link to clinical relevance which is required when a decision about benefit-risk balance is to be made. It would be most useful to disentangle the strengths and limitations of one form against another as well as their similarities in the future.

Presenting the results is only one aspect of visual presentation required in any benefit-risk assessment. We also identified in the protocol for visual review four other aspects that might require or could benefit from visualisation but were not explored in this review:

a. How to present the relationship of benefits and risks evidence used in the assessment i.e. input data. This is important to visualise where data are available and of what quality are they. By knowing this, a decision-maker would be able to devise suitable strategy on how to address them.

b. The process of benefit-risk assessment analysis. Some decision-makers might want to know in greater details how the benefit-risk assessment was performed. Suitable visual representations of this process may aid their understanding to enable them to make better decisions. There is also very limited understanding on what need to be presented either graphically or numerically at different stages of benefit-risk assessment which would also benefit from this exercise.

c. The concept and building blocks of the benefit-risk assessment. Some visuals may be required to explain some of the complex mathematical underpinnings of the approaches. We envisage this would involve a series of related visuals which eventually build up to acquire the final visual.

d. The subjective evidence. Subjective evidence may or may not be presented using the same visual representations as objective evidence because of its nature. It is not only the matter of subjective evidence used in a benefit-risk assessment, but also how could visuals be used to collect/elicit the subjective evidence from relevant stakeholders e.g. use of visual analogue scale, effects tables, etc.

The aspects to be presented vary by stakeholders, as well as dependent on their interests. It is difficult to say which stakeholders should be presented with which information, but a survey might be able to give some information on the average preferences. Investigation into graphics or other visual representations which answer the most common benefit-risk questions could help focus future research in this area.

5 References

1 Brewer CA. Colors. *http://www.ColorBrewer.org* (2006). Pennsylvania State University.

2 Cammax Limited. The drugs box - easy-to-use educational tool about drugs and alcohol. *http://www.thedrugsbox.co.uk/* (2011).

3 Carswell, C. M. Choosing specifiers: An evaluation of the basic tasks model of graphical perception. *Human Factors* 34(5), 535-554 (1992)

4 Cleveland, W. S. and Mcgill, R. Graphical Perception - Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association* 79(387), 531-554 (1984)

5 Cleveland, W. S. *The elements of graphing data* (2 edn.). 1-10-1994. Hobart Press,

6 Fischhoff, B. (2011) Communicating risks and benefits: An evidence-based user's guide. *Food and Drug Administration, US Department of Health and Human Services,*

7 Gapminder. Gapminder: Unveiling the beauty of statistics for a fact-based world view. *http://www.gapminder.org/* (2011).

8 Heiberger, R. M. and Holland, B. *Statistical analysis and data display: An intermediate course with examples in S-PLUS, R, and SAS* 2004. Springer, New York, USA

9 IBM. Many Eyes. *http://www-958.ibm.com/software/data/cognos/manyeyes/* (2011).

10 Lathrop, R. G. Perceived variability. *J Exp Psychol* 73(4), 498-502 (1967)

11 Lipkus, I. M. and Hollands, J. G. The visual communication of risk. *Journal of the National Cancer Institute Monographs* 25, 149-163 (1999)

12 Lipkus, I. M. Numeric, verbal, and visual formats of conveying health risk: Suggested best practices and future recommendations. *Medical Decision Making* 27(5), 696-713 (2007)

13 Meyer, J., Shinar, D., and Leiser, D. Multiple factors that determine the performance with tables and graphs. *Human Factors* 39, 268-286 (1997)

14 Mt-Isa, S. *et al.* (2011) A systematic review and classification of methodologies for benefit-risk decision-making in medicines.London (Report PROTECT WP5 Benefit-Risk Integration and Representation.)

15 Robbins, N. B. *Creating more effective graphs* 2005. Wiley, Hoboken, NJ

16 Sarac SB, Rasmussen CH, Rasmussen MA, Hallgreen CE, Soeborg T, Colding-Jorgensen M, et al. Balancing benefits and risks: Data-driven clinical benefit-risk assessment. (2011).

17 Spiegelhalter D. Understanding Uncertainty Website. *http://understandinguncertainty.org/* (2010).

18 Tufte, E. R. *The visual display of quantitative information* (2 edn.). 2001. Graphics Press, Cheshire,CT

19 Wainer, H. How to display data badly. *The American Statistician* 38(2), 137-147 (1984)

20 West, S. L. *et al.* (2011) Quantitative summary of the benefits and risks of prescription drugs: a literature review. *Food and Drug Administration, US Department of Health and Human Services,* Silver Spring, MD (Report 0212305.010.001)

21 Wickens, C. D., Lee, J., Liu, Y. D., and Gordon-Becker, S. *An introduction to human factors engineering.* (2 edn.). 2004. Pearson Prentice-Hall, Upper Saddle River, NJ

Appendices

A.1 Area graph

A.1.1 Description of area graph

Table 11 Characteristics of visual representation type for area graph

Form	2-D static graphics
Endpoint	Sum of probabilities or utilities
Methodology-specific	No
Reproduction	Many software can easily produce area graphs; they include Stata, SAS, R, and Microsoft Excel. Hiview 3 and IDS (MCDA software) produce frontier graph (area with boundary lines) for sensitivity analysis.
Suitable audience	Generic area graphs are suitable for the general audience but frontier graphs may require some technical knowledge to be interpreted correctly.

Figure 4 Area graph showing partitioned survival curve for one treatment in a Q-TWiST analysis

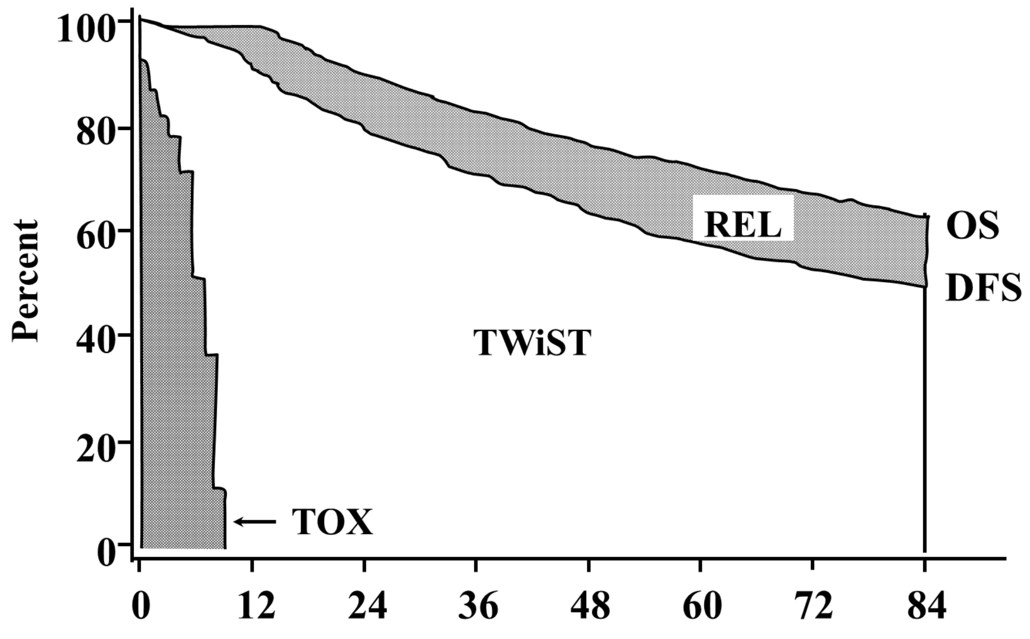


Figure 5 Area graph showing proportions of a particular population from the total population in defining population for impact number analysis

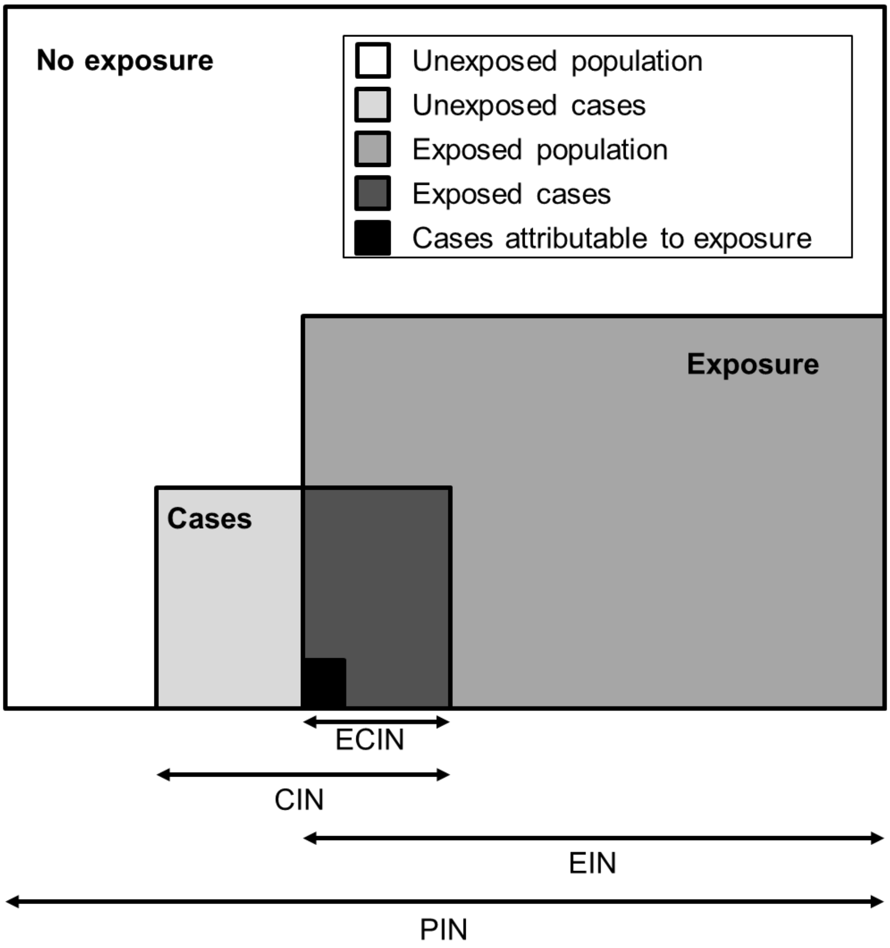
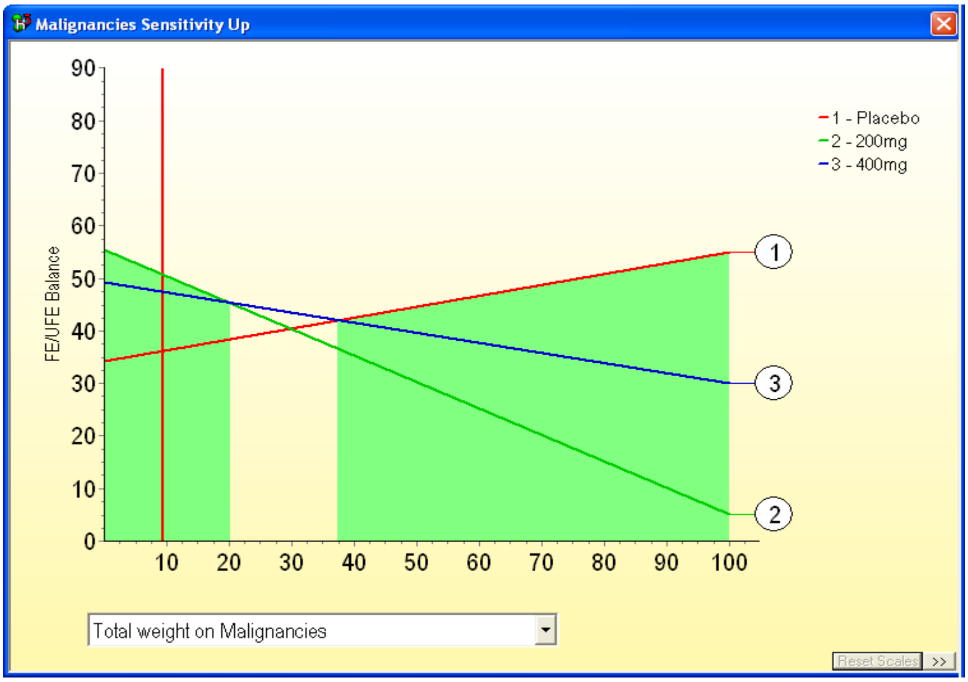


Figure 6 Frontier area graph to assess sensitivity of weighting on malignancies on the optimal treatment choice from Hiview 3 software in an MCDA analysis



A.1.2 Visual evaluation of area graph

Table 12 Wickens' principles of display design for area graph

	Appraisal
Perceptual principles	
Legibility (or audibility)	Most area graphs are legible. It is also important that the boundary of an area on the graph can be clearly seen. For instance, frontier graph produced in Hiview 3 could benefit from having a different colour area since the green line is not very legible on a very similar green hue background (Figure 6).
Absolute judgment	An area is often presented by uniform colour, patterns, or greyscale. They are also enhanced by boundary lines to help with absolute judgment. Fades and gradually changing hue should not be used to avoid absolute judgment limits. The presence of minimum and maximum values of a scale (Figure 6) also allows better judgment.
Top-down processing	Figure 5 demonstrate an aspect of top-down processing where, as DMs analyse the legend, they become accustom to the fact that darker shades refer to more undesirable circumstances. A DM may get confused if there were another shade suddenly contradicts this pattern. However, top-down processing is less material for area graphs in most circumstances.
Redundancy gain	Area graphs bounded by distinct lines can result in redundancy gain. Another form of redundancy gain is also demonstrated in Figure 6 where lines are both colour-coded and numbered.
Discriminability	Areas representing different message or information should be discriminable to avoid confusion. Different colour, greyscale, or patterns should be used to allow discrimination. This is demonstrated well in Figure 5 since human eyes better distinguish contrasts than colour. The area showing toxicity and relapse in Figure 4 may benefit from different contrasts. The vertical red line in Figure 6 should be presented in a different colour to discriminate from placebo line 1.
Principles based on attention	
Information access cost	All area graphs shown here (Figure 4 – Figure 6) minimise information access cost sufficiently by having the required information, e.g. legends and text labels, within the graph area. Figure 4 however has missed title for the x-axis although may be obvious from the context. Figure 6 may benefit from having the drop-down list, which is essentially the x-axis title, to be more centred along the axis and better blended in to be perceived as part of the graph.
Proximity compatibility	Only Figure 6 demonstrates sufficiently close spatial proximity for different options but the primary emphasis of the graph is on the lines and the frontier, not on the areas themselves.
Memory principles	
Use of existing knowledge of the world	Knowledge of the world closely related to area graphs is the size of the area. Larger area is associated with greater magnitude, whether it represents benefits or risks. In Figure 6, the size of the area refers to how probable an option might be an optimal choice – in this case, placebo dominates once weight is greater than ~37.
Predictive aiding	Benefit and risk information on an area graph are best presented as integrated measures because presenting piecewise information can exhaust cognitive process. This is especially true when there are many criteria associated with benefits and risks.
Consistency	Consistency in a single area graph is hard to assess. It may be possible that confusion may arise if an area graph consists of pictorial representations of a

	Appraisal
	variable and they were conflicting with the DMs' previous knowledge of the variable. The most important form of consistency for area graphs is a consistency of the area representation in a series of graphs.

A.1.3 Communicability evaluation of area graph

Table 13 Elements of visual communication for area graph

Element	Appraisal
Risk magnitude	Area graphs convey the magnitude of risks using area size but it is not straightforward since DMs need to calculate the size of the areas mentally. The area has to be on the correct scale and proportion to avoid perception errors in interpretation. For instance, Figure 5 may be confusing or incorrect if the areas are not representative of the real percentages.
Relative risk	Area graphs are better at conveying relative risk when the risk magnitude is of secondary importance i.e. only need to know whether one is less or greater than another. However, DMs may have to perform heavy cognitive tasks in mentally evaluating the relative risk of each area to each other. Use of areas in graphs may also bias visual perception leading to over-estimation or under-estimation.
Cumulative risk	Cumulative risks on area graphs require mental integration of different areas on a graph supposing that the benefits or risks are plotted against time. This task becomes more difficult when the shape of an area is irregular and further complicated when the risks are not linearly additive over time.
Uncertainty	Area graphs cannot convey uncertainty well.
Interactions among risk factors	Area graphs may not be able to communicate effects of interactions effectively.

A.2 Bar graph

A.2.1 Description of bar graph

Table 14 Characteristics of visual representation type for bar graph

Form	2-D static graphics
Endpoint	Sum of probabilities or utilities, or simple probabilities or proportions
Methodology-specific	No
Reproduction	Many software can easily produce area graphs; they include Stata, SAS, R, and Microsoft Excel. Hiview 3 and IDS (MCDA software) produce sum of utilities, and JSMAA produces probabilities of ranking for alternatives.
Suitable audience	Generic bar graphs are suitable for the general audience as they convey the magnitudes in comparison with each other. Bar graphs from Hiview 3 require users with some experience to make full use of their potential.

Figure 7 Area graph showing the expected utility by state in a Q-TWiST analysis

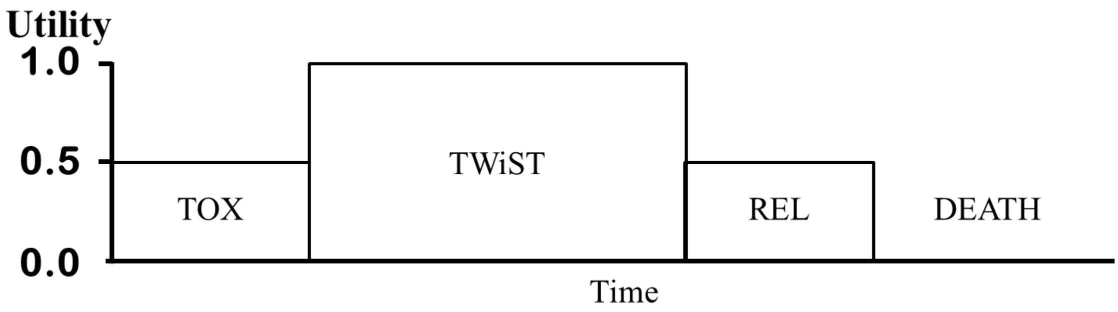


Figure 8 Colour-coded bar graph (the green bars represent benefit and the red bars represent safety) showing the aggregated contributions of benefit and risk criteria after weighting the alternatives in an MCDA analysis in Hiview 3

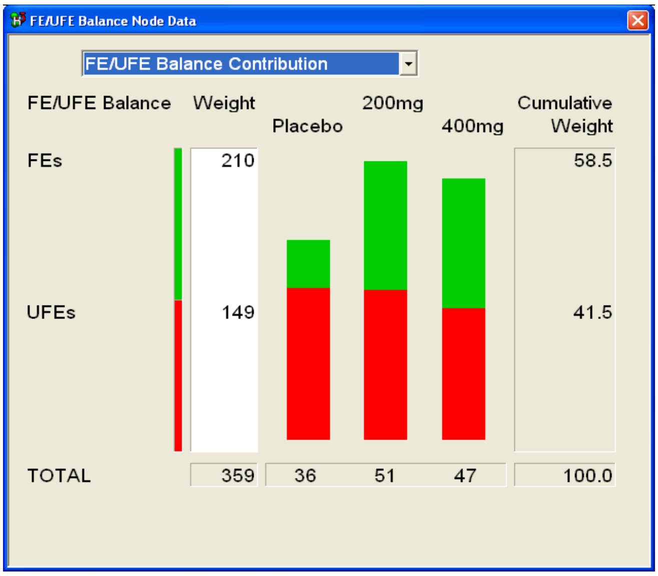
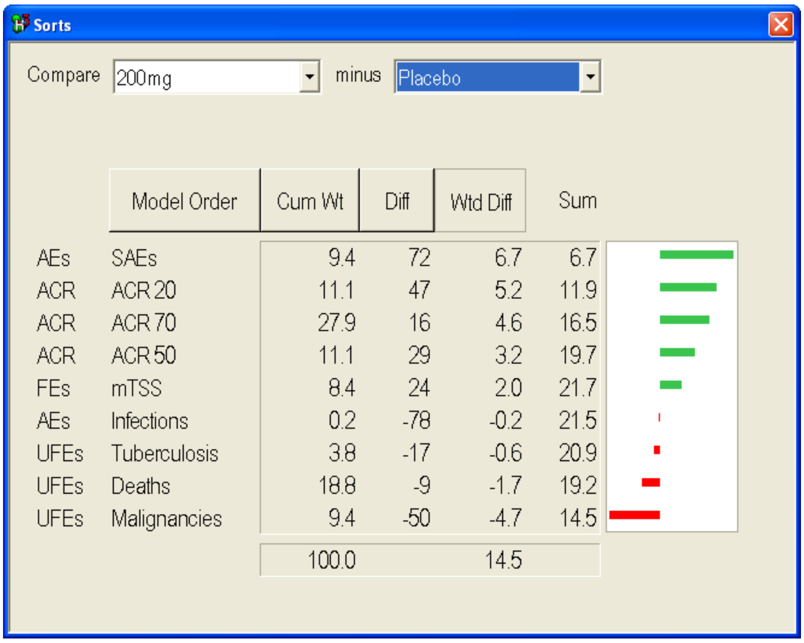


Figure 9 Colour-coded bar graph (a 'difference display') showing weighted differences for benefit and risk criteria of two alternatives in an MCDA analysis in Hiview 3



Sorts

Compare 200mg minus Placebo

	Model Order	Cum Wt	Diff	Wtd Diff	Sum
AEs	SAEs	9.4	72	6.7	6.7
ACR	ACR 20	11.1	47	5.2	11.9
ACR	ACR 70	27.9	16	4.6	16.5
ACR	ACR 50	11.1	29	3.2	19.7
FEs	mTSS	8.4	24	2.0	21.7
AEs	Infections	0.2	-78	-0.2	21.5
UFEs	Tuberculosis	3.8	-17	-0.6	20.9
UFEs	Deaths	18.8	-9	-1.7	19.2
UFEs	Malignancies	9.4	-50	-4.7	14.5
		100.0		14.5	

Hiview 3 interface allows the bar graph displayed to be cumulative weight (representing the relative clinical relevance of a unit of each effect), simple difference (between preference values), or weighted difference (of the data and its clinical relevance). It also allows any two alternatives to be compared directly

Figure 10 Bar graph showing probabilities of an alternative *i* being ranked in r^{th} place (b_i^r) in an SMAA analysis in JSMAA without preference information

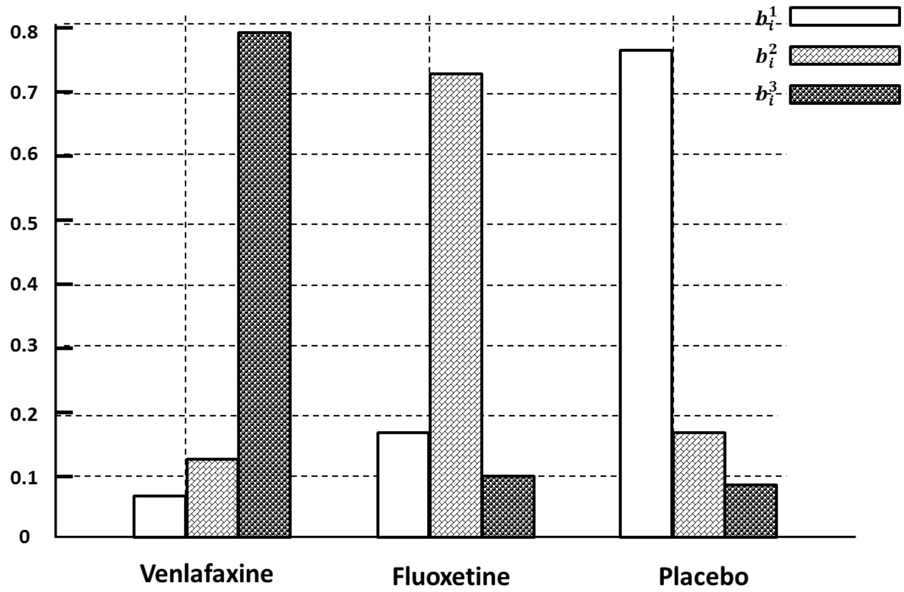
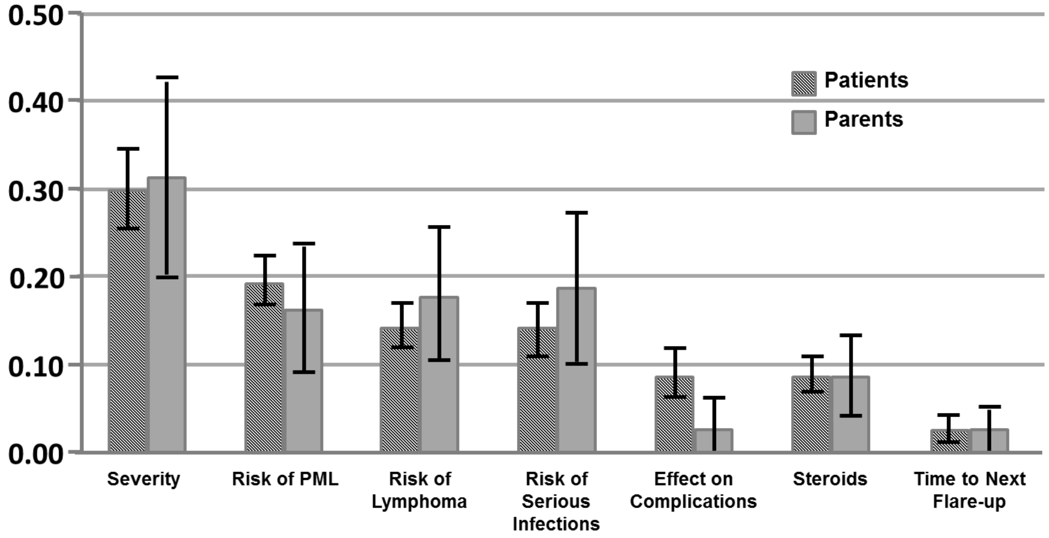


Figure 11 Bar and antenna graph showing the relative contribution (proportion) of each criterion by stakeholders in a MAR analysis



A.2.2 Visual evaluation of bar graph

Table 15 Wickens' principles of display design for bar graph

	Appraisal
Perceptual principles	
Legibility (or audibility)	The legibility of bar graphs is generally very good. However, some bar graph variations like stacked (Figure 8) and 3-D bar graphs may limit legibility. The legibility can also be reduced when the orders of magnitude are not the same for all bars resulting in some bars with comparatively small magnitudes to be incomprehensible. Figure 9 demonstrates illegibility for criterion "infections" because of its weighted difference close to zero (the graph can be stretched in Hiview 3 to lengthen the bars for more visibility – also see 2 below).
Absolute judgment	Bar graphs do not often face absolute judgment limit problems since the boundary of each bar can be seen clearly. Very infrequently bar graphs have been used with changing hue or fading greyscale which may limit absolute judgment. Such applications can be seen when a bar graph is made to communicate uncertainty. There may be a slightly different issue with absolute judgment limit in Figure 9 since Hiview 3 allows the graph to be stretched which consequently stretches the length of the bars. The relative effects are consistent but the question remains whether judgments are altered when stakeholders are presented with the same graph with different aspect ratios.
Top-down processing	Current practice of producing bar graph promotes good top-down processing ability. Bars describing the same variables are designed with the same pattern or colour (Figure 8 - Figure 11), and also have the same order when variables when it is presented stratified by another variable – for example risk at the bottom and benefit at the top of the stacked bars for each option (Figure 8), the order of ranks from one to three running from left to right for each option (Figure 10), and relative contributions of patients on the left and of parents on the right for each risk criterion (Figure 11).
Redundancy gain	Two redundancy gain types are demonstrated in bar graphs above. One is the use of boundary line for each bar together with pattern or colour. The second type is not very obvious, which is the use of bars itself. Tufte's data-ink principle suggests that bars are redundant since the only the line at the top of the bar is meaningful, reducing it to scatter plots (Tufte, 2001). However, the existence of the bars may promote redundancy gain.
Discriminability	As long as the patterns or colours used to represent bars are sufficiently different from each other, the discriminability of bar graphs is also sufficient. Unfortunately, as far as we know, the extent to how difference they should be are not yet tested. But when the number of bars on a bar graph becomes larger, the bar graph would have diminished capacity of discriminability.
Principles based on attention	
Information access cost	Information access cost of bar graphs is minimal since many bar graphs tend to separate different groups by gaps (Figure 10 and Figure 11), tend to present bars in the same order for different groups (Figure 8, Figure 10, and Figure 11), and contain legends to bar patterns (or colours) within the graph area (Figure 10 and Figure 11)
Proximity compatibility	Proximity compatibility is demonstrated through having the legend within the graph area (Figure 10 and Figure 11), and through grouping together bars belonging to the same groups (Figure 9 – Figure 11).
Memory principles	

	Appraisal
Use of existing knowledge of the world	The bars are drawn in different directions to depict the direction of effects for each criterion, where most people would associate right with positive (values) and left with negative (values) (Figure 9). Figure 8 and Figure 9 also demonstrate the use of existing knowledge that DMs would also associate green bars with good outcomes (benefits) and red bars with bad outcomes (risks).
Predictive aiding	The use of bar graphs for predictive task is cognitively challenging. By definition, bar graphs present information separately by categories. Cognitive workload is increased in Figure 7 because DM has to mentally work out the expected utility of an option, and further increased having to compare different options. Figure 8 contains an element of predictive aiding by stacking the bars for benefit and risk for each alternative to aid DMs in comparing the total weighted scores contributed by the FE/UFE balance criterion to determine the best alternative for that criterion. The information gained can then be incorporated into the final decision.
Consistency	The consistency within a bar graph has been discussed previously in the consistent ordering and patterns or colours of bars across different categories. In a series of bar graphs, consistency is achieved by maintaining the order and patterns used. The graphs produced in Hiview 3 for MCDA analyses demonstrate great consistency with using consistent red/green colour scheme for the bars, stacking benefits on top of risks (Figure 8), and showing the direction of risks bars to the left and benefits bars to the right (Figure 9).

A.2.3 Communicability evaluation of bar graph

Table 16 Elements of visual communication for bar graph

Element	Appraisal
Risk magnitude	Bar graphs can communicate risk magnitudes with ease by measuring the top of the bars against the y-axis of the graph as long as the thickness of the boundary lines of the bars is acceptable. Thick lines can make reading off the magnitudes difficult.
Relative risk	Relative risks are not as well communicated in bar graphs. DMs are required to mentally compare the height of the bars to estimate the relative risks. This exercise is cognitively challenging especially when there are many mental arithmetic operations to be made for example when many bars are involved and in the situations when there are many two-way comparisons. It may also be perceptively difficult to compare risks for items which are more distant on the graph, when the boundary lines are thick, and when small numbers are involved.
Cumulative risk	Cumulative risk can be shown on bar graphs by adding together bars along the x-axis (Figure 7) or by stacking the individual values as shown in Figure 8.
Uncertainty	It is uncommon to communicate uncertainty using bar graphs, but one way to do this is through that demonstrated in Figure 10, where the probabilities of an option being ranked first, second or third are plotted as a bar graph.
Interactions among risk factors	The effect of interactions can be communicated on bar graphs as stacked bars. Figure 8 is an example of stacked bar graph but may not be the appropriate example for communicating interactions. A stacked bar graph on Hiview 3 showing contributions of criteria towards the final weighted utility would be a better example in this situation.

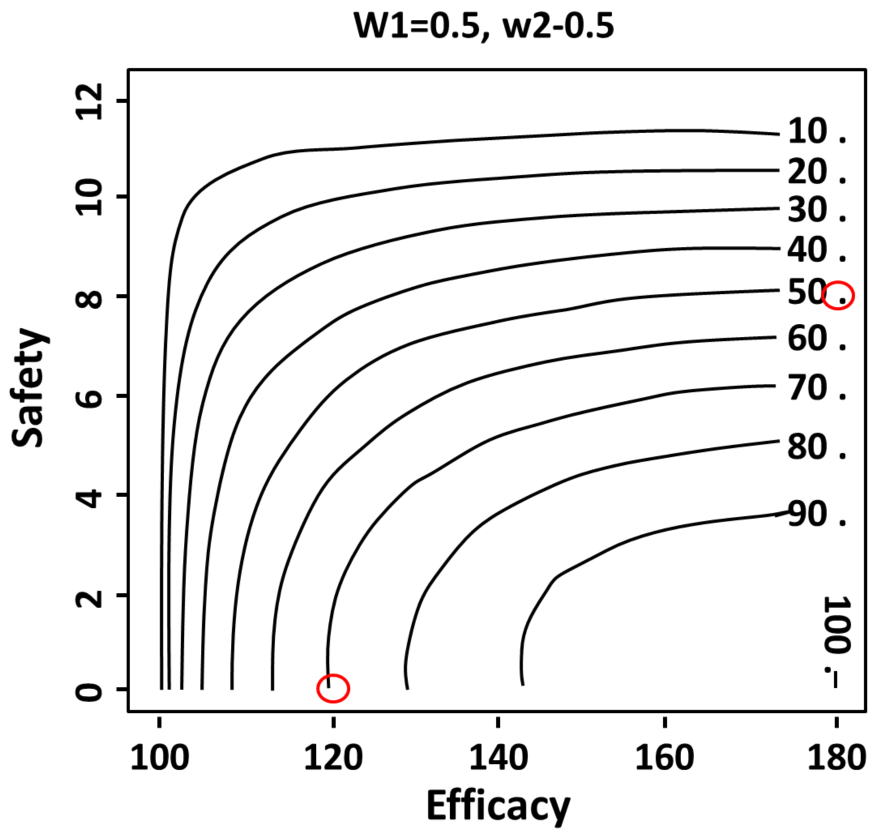
A.3 Contour plot

A.3.1 Description of contour plot

Table 17 Characteristics of visual representation type for contour plot

Form	2-D static graphics
Endpoint	Any point estimate of utilities or scores
Methodology-specific	No
Reproduction	Software like Stata (as part of version 12, or user-written programme for earlier versions), SAS, R, and Microsoft Excel can produce contour plots.
Suitable audience	Contour plots are intended for more specialist audience with knowledge of contour/surface plots as well as the background data underlying them. General public may be able to understand contour plots but only to a limited degree without assistance.

Figure 12 Contour plots of the effects of varying weights in a desirability index analysis



A.3.2 Visual evaluation of contour plot

Table 18 Wickens' principles of display design for contour plot

	Appraisal
Perceptual principles	
Legibility (or audibility)	The legibility of contour plot is variable but it is more dependent on the purpose of its presentation since the legibility of the actual contour lines themselves may not matter. The more important aspect of contour plots is the closeness of the contour lines to create and emphasise "dense" parts on the graph. Contour lines are often labelled at various points to bring attention to the points, but too much labelling can diminish the legibility of contour plots further.
Absolute judgment	Absolute judgment is limited on contour plots in the same way as when fading greyscales are used on area graphs (see Appendix A.1.2). It is quite common when colours or greyscales are used in contour plots, the contour lines are omitted which can decrease the absolute judgment capacity. Points labelling on contour plots are often done to improve on absolute judgment limits but we acknowledge that this may affect legibility.
Top-down processing	Top-down processing is difficult and limited in contour plots because the behaviour of the contour lines may not follow the same patterns as previously encountered lines. For example in Figure 12, it is only possible to learn that the gap between lines widens as "efficacy" increases for a fixed low "safety" value, but this is not true for other higher "safety" values. This is not a weakness of contour plots but a known feature to picture possible values.
Redundancy gain	Redundancy gain in contour plots is often achieved by using both area shading (with colours or greyscales) and the contour lines.
Discriminability	Labelling (Figure 12) and colours are the two common ways to increase discriminability of contour plots.
Principles based on attention	
Information access cost	Contour plots attempt to minimise information access cost by conveying the extra dimension using density of contour lines or the intensity of colours used. Another more straightforward feature is the contour lines themselves, where each line represents different piece of information available within relatively close proximity of each other.
Proximity compatibility	Contour plots have one good aspect of proximity compatibility in terms of comparing the information between contour lines on the same plot. Figure 12 exhibits good practice when more contour plots with different parameters (e.g. weights) are to be compared by aligning them side by side.
Memory principles	
Use of existing knowledge of the world	Some knowledge of line graphs can help when it comes to extracting and interpreting information from contour plots. Otherwise, contour plot may not be a familiar image to the general public with the exception to those who follow meteorological and other seismic activities predictions which often use contour maps (contour plots superimposed on geographical maps).
Predictive aiding	"Grey" information such as the regions between solid lines, greyscales and colour saturation is commonly presented on contour plots. These presentations may bear some information but may not be a good predictive aiding tool since they limit absolute judgment (see point 2 on absolute judgment).
Consistency	On a single contour plot, consistency may not matter very much. But since there

	Appraisal
	are a lot of information being presented on a contour plot, a series of contour plots need to be very consistent to enable DMs to extract information and make the best judgment. This includes ensuring that the line patterns used for the contour lines, the position of labels, the aspect ratios of the plots, the use of any reference points, and other aesthetic features should be consistent throughout.

A.3.3 Communicability evaluation of contour plot

Table 19 Elements of visual communication for contour plot

Element	Appraisal
Risk magnitude	The extent to which risk magnitude can be communicated depends heavily on the presentation of contour plot, and the points on the plot that actually matter in a decision.
Relative risk	It is possible to communicate relative risk on a contour plot by either plotting the values of relative risk, or by mental estimate of the risk magnitudes.
Cumulative risk	It is possible to communicate cumulative risk over time when the time axis is plotted against the risks.
Uncertainty	Uncertainty is shown by the various contour lines on the plot.
Interactions among risk factors	It is difficult to communicate interactions using contour plots.

A.4 Distribution plot

A.4.1 Description of distribution plot

Table 20 Characteristics of visual representation type for distribution plot

Form	2-D static graphic
Endpoint	(Cumulative) Probabilities of scores
Methodology-specific	No
Reproduction	Stata, SAS and R can easily produce distribution plots but limited to the probability distribution functions available. Microsoft Excel can also produce distribution plots provided that users know how to generate random values based on the properties of the distributions. Reproduction of empirical distributions requires simulation.
Suitable audience	

Figure 13 A series of distribution plots showing the probabilities of benefit for different number of risk factors

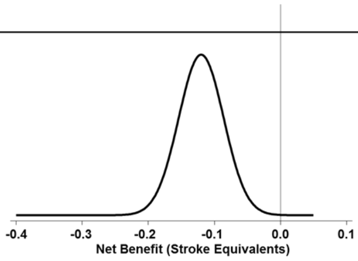
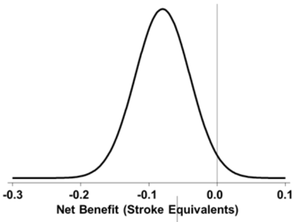
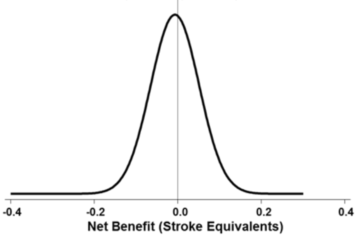
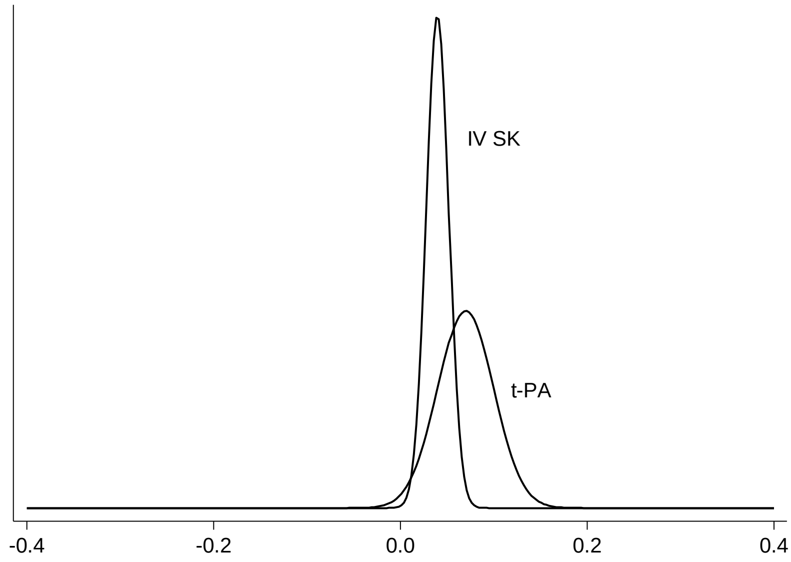
Number of risk factors	*Median (95% CrI)*	*Probability of Benefit >0*	*Simulated Posterior Distribution*
Clinical risk factors			
0	-0.12 (-0.21 to -0.05)	0.04%	
1	-0.08 (-0.18 to -0.001)	1.39%	
2 or 3	-0.007 (-0.12 to 0.11)	44.74%	

Figure 14 A distribution plot comparing two distributions



A.4.2 Visual evaluation of distribution plot

Table 21 Wickens' principles of display design for distribution plot

	Appraisal
Perceptual principles	
Legibility (or audibility)	Recognising the shape of the distributions depends heavily on the knowledge of the DMs. This is important in the legibility of distribution plots because overlapping distributions are not often distinguished by different line colours or patterns (Figure 14). Therefore DMs with no experience of the shape of distributions may not be able to disentangle the path of a line or distribution. Not being able to do so may limit the amount of information to be gained and decisions to be made.
Absolute judgment	There are two important pieces of information from distribution plots: the points on the curved line, and the area under the curved line. The former is much easier to disentangle when read across the horizontal axis but the latter is much more difficult. The absolute judgment limit is thus more limited when determining an area under the distribution line. It can be enhanced by, commonly done, vertical reference line to indicate the size of the area to the left or right of the line to increase absolute judgment.
Top-down processing	For many distribution plots, bell-shaped curves are the most common. Unfamiliar DMs can easily adapt to the fact that these distributions are symmetrical (and maybe extends to $\pm\infty$). With such distributions being presented, top-down processing becomes easier even when the distributions are truncated. Figure 14 shows two overlapping distributions with different means and variance. Top-down processing allows DMs to make comparison of the mean and the variability in one distribution to another by comparing the shift and height of the peaks, and also the width of the distributions.
Redundancy gain	Distribution plots do not require much redundancy gain as the information conveyed is succinct. A redundancy gain that is possible is by labelling the vertical axis to show that the area under each curve equals one (not demonstrated in figures above). Although vertical axis labelling is usually

	Appraisal
	redundant, it also allows DMs to estimate the area under the curve from the graph.
Discriminability	Colour coded and pattern coded lines can be used to increase discriminability of distributions of different variables.
Principles based on attention	
Information access cost	Figure 13 shows how information access cost can be minimised for distribution plots by presenting the plots alongside the information from the distribution e.g. the magnitude of and the statistical uncertainty around the effect, the area under the curve describing the probability of beneficial effect, and the endpoint they relate to. Comparison of different endpoints (or clinical risk factors in Figure 13) is also made simpler by presenting the in intuitive order and close to each other, minimising the information access cost required to make the comparison.
Proximity compatibility	Two typical ways of presenting distribution plots are shown in Figure 13 and Figure 14, where each distribution curve is presented as separate plots (Figure 13) and on the same plot (Figure 14). In both presentations, the degree of proximity compatibility varies according to whether mental integration is required or whether focussed attention is required. Cognitive effort to perform mental integration to compare distributions when they are presented separately is increased, but the separate presentation allows for better judgment of the magnitudes; and vice versa. In Figure 13, cognitive effort is minimised by aligning the centre (peak) of the distributions. Another way to improve cognitive effort is by aligning the vertical lines corresponding to zero effect to provide better perception of the position of the curves against each other. However, it is not possible to test which approach is more optimal at this stage.
Memory principles	
Use of existing knowledge of the world	Distribution plot is a more mathematical than natural concept; therefore there may not be non-mathematical existing knowledge that could really help with reading the plots.
Predictive aiding	Predictive aiding using distribution plots can be achieved by presenting the required information that matters most in the decision alongside the plots. This is demonstrated in Figure 13.
Consistency	Consistency does not apply for a single distribution plot, but standard colour coding may be useful e.g. red for risk and green for benefit. For a series of distribution plots, consistency can be achieved in the same way or by using line patterns to represent the same event or a group of similar events. Using the same scales and positioning the scale consistently in a series of distribution plots can also help DMs when extracting information from the plots.

A.4.3 Communicability evaluation of distribution plot

Table 22 Elements of visual communication for distribution plot

Element	Appraisal
Risk magnitude	The risk magnitude on a distribution plot is the value on the x-axis.
Relative risk	Relative risk of two distributions can be mentally calculated when both distributions are presented. It is also possible that the distribution of the relative risk itself being plotted to facilitate the information extraction.
Cumulative risk	It is difficult to represent cumulative risk over time on a distribution plot.
Uncertainty	The strength of distribution plots is in communicating the uncertainty in an estimate.
Interactions among risk factors	Unless the distributions of the interactions themselves are plotted, it is unlikely that interactions among risk factors can be communicated effectively using a distribution plot.

A.5 Dot and Forest plot

A.5.1 Description of forest plot

Table 23 Characteristics of visual representation type for forest plot

Form	2D static graphic
Endpoint	Point estimates and their confidence intervals or ranges. This could be any appropriate metric indices but most commonly used with mean and 95% confidence intervals.
Methodology-specific	No
Reproduction	Figure 15 can be reproduced using a BRAT Framework Tool, currently a beta release in the form of Microsoft Excel spreadsheet package. The package however is not yet publicly available. There is no specific software available to automatically create dot/forest plot as used in the benefit-risk context shown in Figure 16 and Figure 17. There are some specialised meta-analysis software packages which produce forest plot as part of meta-analysis e.g. Stata *metan* command, and the Comprehensive Meta-Analysis software (http://www.meta-analysis.com/). However, the plots are easily reproduced in many statistical packages like Stata, SAS, and R.
Suitable audience	Any audience with understanding of uncertainty can easily understand dot/forest plots.

Figure 15 A forest plot of the risk difference in a PhRMA BRAT analysis

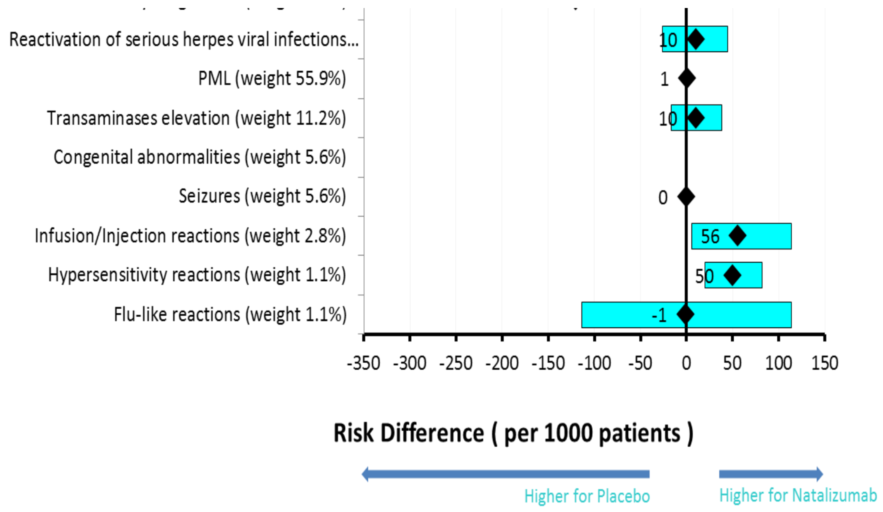


Figure 16 Forest plot showing comparison of odds ratios for efficacy and safety and the combined NEAR odds ratios for ADRs leading to continuation and for nausea separately

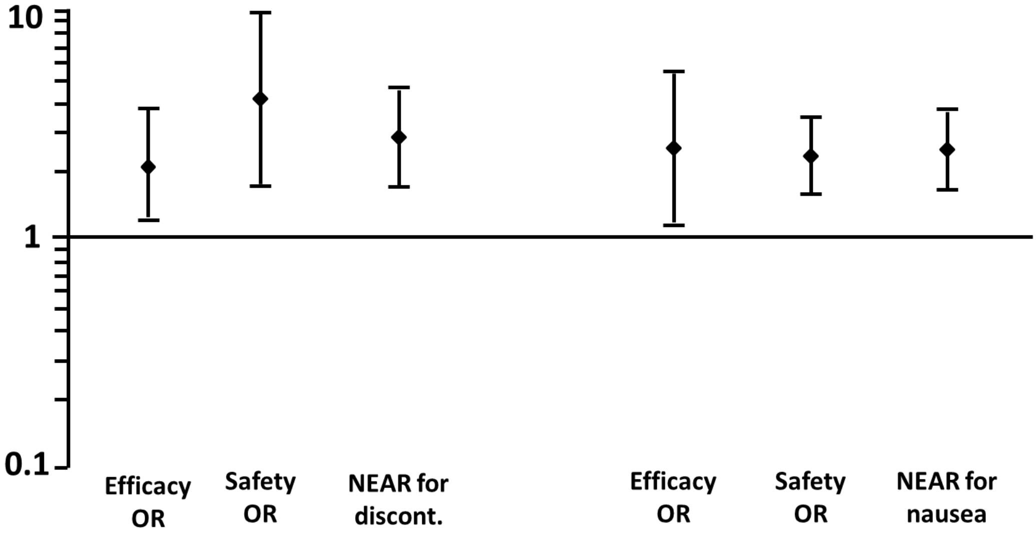
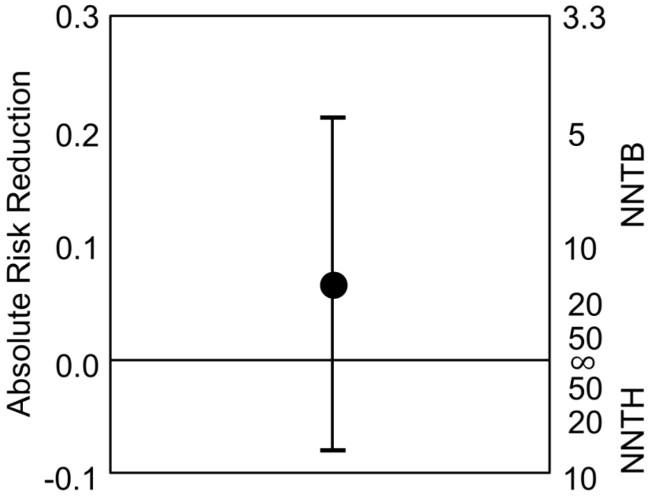


Figure 17 Single forest plot (mean and CI) showing mean absolute risk reduction and the 95% confidence intervals alongside the NNT and its 95% confidence intervals covering both benefit and harm



A.5.2 Visual evaluation of forest plot

Table 24 Wickens' principles of display design for forest plot

	Appraisal
Perceptual principles	
Legibility (or audibility)	The legibility of forest plot is generally good but can bad when the confidence interval lines are too short or too long. This issue usually surfaces when plotting forest plot with mixture of very narrow and very wide confidence intervals. The common practice of truncating the lower and upper ends of long intervals (Figure 15) should be thought of carefully and should only be done when they do not play any role in or affect decision-making.

	Appraisal
Absolute judgment	The absolute judgment may be limited when varying symbol sizes are used on the same forest plot. This is commonly encountered on a meta-analysis forest plot where the size of the middle symbols (e.g. mean) is varied to reflect the weight contribution to the overall analysis. This can be more detrimental when the symbols are large and covering the confidence interval lines. The use of capping at the end of the lines may also limit absolute judgment since it is difficult to tell clearly the values they represent, and this gets worsened when thicker lines are used as caps. Many forest plots are used to represent ratios (odds ratios, relative risks) therefore the vertical axis scale is often given in multiplicative scale (Figure 16). Therefore, unfamiliar DMs may not be able to judge the changes appropriately. Another aspect of absolute judgment limitation is the inverted plot proposed (Figure 17) to represent NNT and NNH. In Figure 17, it is difficult to judge a value that is closer to the infinity line, as well as the conceptual representation of NNT or NNH values when it crosses the line.
Top-down processing	In the simplest case of single “forest” plot (Figure 17), this criterion is not applicable. In others, top-down processing can come in several ways on a forest plot depending on the scales of the axes, the metrics being presented, or the order of the criteria being presented. We expect DMs to quickly adapt that in the situation when multiplicative scale (e.g. log scale) is used on the axis, the interpretation of a unit change also changes to being multiplicative. In a forest plot, only the same metrics of the same units must be used. This increases its top-down processing capability, and DMs would rapidly understand that it is acceptable to visually compare the three main point estimates on the plot: the mean of OR (or RR, median etc.), the lower and upper ends of 95% confidence intervals (or 25th and 75th quantiles) for example. Although Figure 15 arranges the plot in a rather “cyclical” fashion, ordering them by the effect sizes could promote better top-down processing ability as a DM move down or up the criteria list.
Redundancy gain	This comes in through the symbols used for the point estimate at the two ends (e.g. 95% CI). The short line caps are not necessary but allow redundancy gain when drawn. However, forest plot designers should carefully think about how the lines appear when presented to DMs as they may also limit absolute judgment.
Discriminability	There is potential for confusion depending on the way the results of the benefit-risk analyses are presented depending on the clarity of the direction and the meaning of the results.
Principles based on attention	
Information access cost	There is likely to be some ‘cost’ in terms of understanding the positioning of the point estimate and range/95% confidence intervals, particularly if several lines are present. Clear legends and labelling may reduce this.
Proximity compatibility	Related lines/estimates may be grouped together facilitating appropriate processing. However, the converse is also true.
Memory principles	
Use of existing knowledge of the world	Users with expertise may benefit most. A forest plot is likely to be a relatively novel concept to a non-expert audience and may be difficult to interpret.
Predictive aiding	Predictive aiding can be achieved by presenting forest plot of the benefit-risk balance instead of separate plots for benefits and risks.
Consistency	Potential for high levels of consistency if scales/direction of results are presented in the same direction and with clear labelling.

A.5.3 Communicability evaluation of dot/forest plot

Table 25 Elements of visual communication for forest plot

Element	Appraisal
Risk magnitude	Yes this could be via the point estimates and ranges/95% confidence intervals if these are displayed
Relative risk	Yes this could be via the point estimates and ranges/95% confidence intervals if these are displayed
Cumulative risk	The inclusion of cumulative information is possible via a summary point estimate, however, the inclusion of time would be difficult.
Uncertainty	This can be represented to a certain degree via the ranges/95% confidence intervals
Interactions among risk factors	These are harder to present simply in a forest plot and individual outcomes may be likely to be easier to understand in this setting.

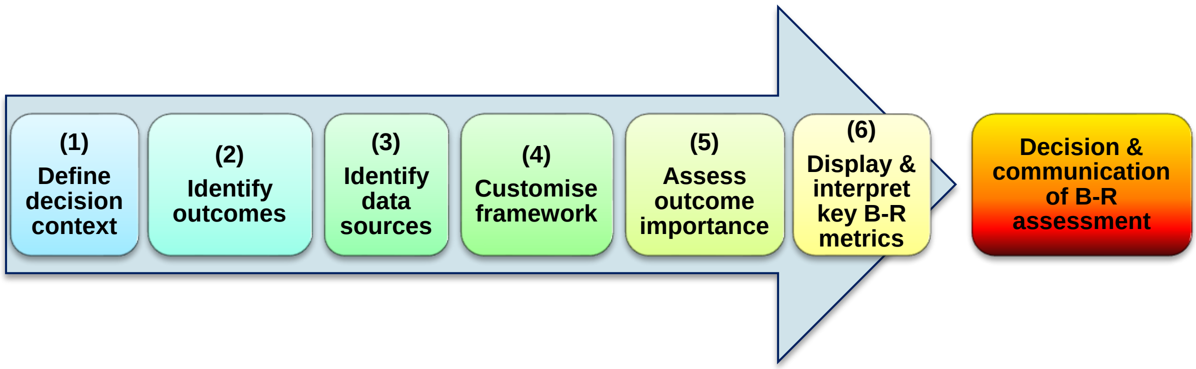
A.6 Flow diagram

A.6.1 Description of flow diagram

Table 26 Characteristics of visual representation type for flow diagram

Form	2D static graphic
Endpoint	Not applicable
Methodology-specific	No
Reproduction	There are many software packages which can produce flow diagrams. Most of them require manual drawing of the diagram but Microsoft Word and Microsoft Visio have easy to use templates for producing flow diagrams. Other software are also listed on http://alternatives.rzero.com/flow.html
Suitable audience	Flow diagrams are intuitive and therefore may be suitable for the general public. However, to fully understand the contents of a flow diagram may require more specific knowledge of the subject it is depicting.

Figure 18 Flow diagram showing the key steps in the PhRMA BRAT framework



A.6.2 Visual evaluation of flow diagram

Table 27 Wickens' principles of display design for flow diagram

	Appraisal
Perceptual principles	
Legibility (or audibility)	It is possible to ensure that flow diagrams are legible but care must be taken with regard to complex information, numbers of boxes or categories, connectors and directionality i.e. the legibility can be diminished when too many elements are presented in a single flow diagram. This is also true when arrow connectors are used and crossing each other too frequently.
Absolute judgment	There should be no absolute judgment limit in a flow diagram when each element is presented clearly and fully self-contained within its context. The absolute judgment limit is introduced in a flow diagram when there are unclear or undefined elements.
Top-down processing	Flow diagram is a visual representation of stepwise elements by indications of the direction of the flow, commonly using arrows or pointed tips. Top-down processing is easy as DMs will quickly learn that elements at the beginning of an arrow precede those at the other end, even when numbering is not used (Figure 18).
Redundancy gain	Flow diagrams can reinforce or convey additional information using size, colour and direction of connecting arrows.
Discriminability	There is the potential for a lack of discriminability if the direction of the 'flow' is unclear or mixed and/or if the number of boxes is too large for the information that the diagram is aimed to convey
Principles based on attention	
Information access cost	Should be fairly minimal if flow of information presented logically
Proximity compatibility	A flow diagram can be presented so that spatial proximity of components is appropriate.
Memory principles	
Use of existing knowledge of the world	The idea of a logical flow of information has the potential to be a familiar concept for a substantial proportion of readers
Predictive aiding	As flow diagram shows process, there is no predictive aiding to be made.
Consistency	Possible to present information consistently

A.6.3 Communicability evaluation of flow diagram

Table 28 Elements of visual communication for flow diagram

Element	Appraisal
Risk magnitude	Difficult without text related information
Relative risk	Difficult without text information or the use of additional information such as changing box size.
Cumulative risk	May be possible if timelines can be taken into account
Uncertainty	Difficult without text related information
Interactions among risk factors	Difficult to represent without complex links between boxes of information and text

A.7 Grids and tables

A.7.1 Description of grid and table

Grids and tables are the most flexible of the visuals presented here since they can convey many forms of information within them.

Table 29 Characteristics of visual representation type for grids and tables

Form	2D static graphic
Endpoint	There is no specific endpoint to be used with grids and tables
Methodology-specific	In general, these are not methodology-specific but grid as shown in Figure 20 is specific to TURBO approach.
Reproduction	Tables can be reproduced easily in many packages including Microsoft Office, Stata, SAS, R, etc. Although grids are conceptually easy to construct, it may be a little tricky to reproduce them since some grids are required to be on certain scale. There is no software available to reproduce grids, but the reproduction is straightforward using statistical packages like Stata, SAS and R using the graph functions.
Suitable audience	General public

Figure 19 A table template showing the required elements of the FDA Benefit-Risk Assessment guideline

Decision Factor	Evidence and Uncertainties	Conclusions and Reasons
Analysis of Condition	Summary of evidence:	Conclusions (Implications for decision):
Unmet Medical Need	Summary of evidence:	Conclusions (Implications for decision):
Benefit	Summary of evidence:	Conclusions (Implications for decision):
Risk	Summary of evidence:	Conclusions (Implications for decision):
Risk Management	Summary of evidence:	Conclusions (Implications for decision):
Benefit-Risk Summary and Assessment		

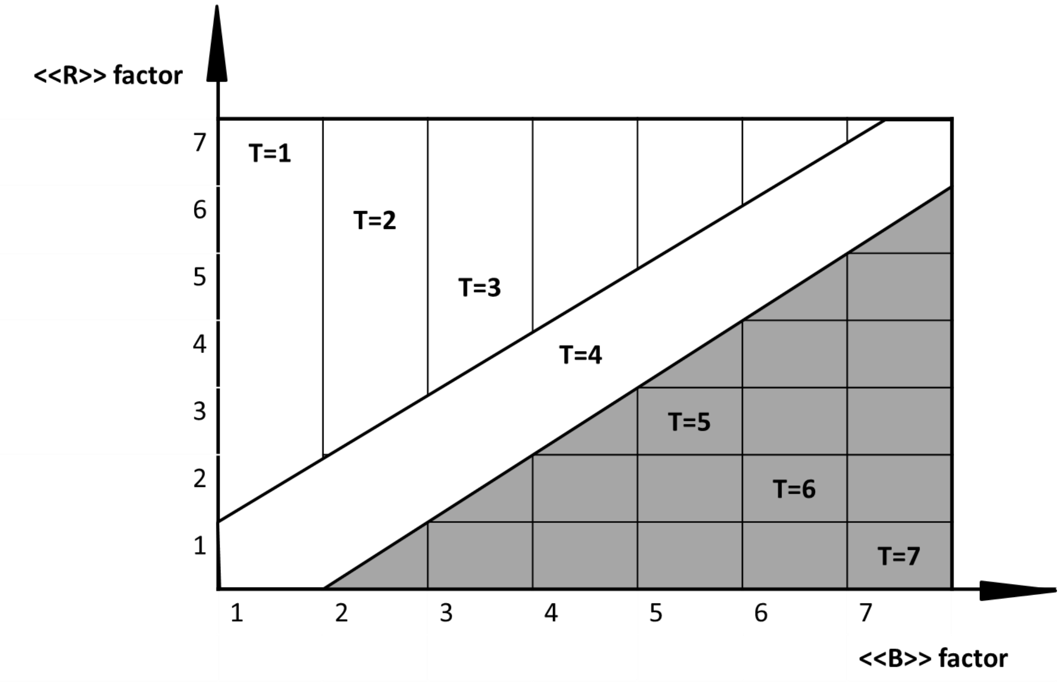


Figure 20 A grid for TURBO scores

Figure 21 A table showing the scores tabulation in principle of three

	Disease	Effectiveness	Dominant ADR
Seriousness	3	3	3
Duration	3	3	2
Incidence	1	0	2
Total	7	6	7

Figure 22 A table showing the benefit and risk criteria and their corresponding values in PhRMA BRAT analysis

		Outcome	Natalizumab Risk / 1000 pts	Placebo Risk / 1000 pts	Risk Difference (95% CI)/ 1000 pts	
Benefits	Convenience Benefits	Convenience (weight 0.6%)	-	-	-	(-, -)
	Medical Benefits	Relapse (weight 3.9%)	280	540	-260	(-326, -195)
		Disability Progression (weight 5.6%)	110	230	-120	(-, -)
Risks	Infection	Reactivation of serious herpes viral infections (weight 6.7%)	80	70	10	(-26, 45)
		PML (weight 55.9%)	1	0	1	(-, -)
	Liver Toxicity	Transaminases elevation (weight 11.2%)	50	40	10	(-16, 38)
	Reproductive Toxicity	Congenital abnormalities (weight 5.6%)	-	-	-	(-, -)
	Neurological Disorders	Seizures (weight 5.6%)	0	0	0	(-, -)
	Other	Infusion/Injection reactions (weight 2.8%)	236	180	56	(6, 114)
		Hypersensitivity reactions (weight 1.1%)	90	40	50	(20, 82)
		Flu-like reactions (weight 1.1%)	399	400	-1	(-114, 114)

Higher for Natalizumab
Higher for Placebo

A.7.2 Visual evaluation of grid and table

Table 30 Wickens' principles of display design for grids and tables

	Appraisal
Perceptual principles	
Legibility (or audibility)	Grids and tables are generally clear and concise. The visual aspect of grids and tables are the presentation of texts at the intersection of the designated rows and columns. However, since the contents are variable and less stringently governed by the design feature than other forms of visuals, the degree of legibility also varies greatly.
Absolute judgment	The values in grids and tables allow good absolute judgment within the designated rows and columns. However, absolute judgment immediately becomes very limited when there is a need to perceive a value outside the contents of the grid or table.
Top-down processing	Grids and tables do not exhibit good top-down processing capability since the contents of grids and tables in general do not always represent the same information. However, grid like Figure 21 has similar characteristics to a scatter plot (see Appendix A.10), therefore allowing better top-down processing. Also, should the criteria appeared in the rows or columns of a table be arranged in a more intuitive order e.g. effect sizes, the table would possess better value for top-down processing.
Redundancy gain	Redundancy gain in grids and tables can be achieved by using colour-coded cells as demonstrated in Figure 22, and to a lesser extent in Figure 20.
Discriminability	Colour-coding of the cells in a table or grid also provides better discriminability among different groups of criteria (Figure 22), which can help DMs to focus their attention to the required parts.
Principles based on attention	
Information access cost	This very much depends on the complexity and design of the table. Judicious placing and labelling of rows and columns is likely to help reduce this.
Proximity compatibility	This very much depends on the complexity and design of the table. Judicious placing and labelling of rows and columns is likely to help reduce this. Figure 20 demonstrates grouping of appropriate outcomes allowing spaces within a figure to facilitate processing.
Memory principles	
Use of existing knowledge of the world	Tables/grids are likely to be familiar ways of presenting information and a wide audience is likely to be able to extract information from this format. If this is to be useful in terms of communication of benefit and risk the choice and labelling of the columns/rows/grid may gain from following more standard designs.
Predictive aiding	There is some potential for predictive aiding, for example in linking information across rows, but it is limited.
Consistency	Use of colour, shading, fonts, symbols, column and row size can be used to facilitate consistency.

A.7.3 Communicability evaluation of grid and table

Table 31 Elements of visual communication for grids and tables

Element	Appraisal
Risk magnitude	Potential for this to be presented by text or by proportion eg coloured cells/cells containing symbols e.g. to represent individuals.
Relative risk	This can be represented by text or coloured/shaded cells
Cumulative risk	It is possible to incorporate time and cumulative information in tables or grids
Uncertainty	This can be represented by text of colour/shading to illustrate levels of uncertainty.
Interactions among risk factors	Tables and grids have the potential for grouping of information to represent interactions but this would be difficult to demonstrate without multiple grids and that would bring additional complexity and issues of proximity and information access cost.

A.8 Line graph

A.8.1 Description of line graph

Table 32 Characteristics of visual representation type for line graph

Form	2D static graphic
Endpoint	Point estimates over a range of values of another variable. Figure 24 and Figure 27 also exploit line graphs to depict the uncertainty of the point estimates through displaying the lower and upper bounds of 95% confidence intervals over a range of values of the other variable.
Methodology-specific	No
Reproduction	Many software can reproduce line graphs with ease so long as they have the capability to reproduce Cartesian (X-Y) graphs. These include Stata, SAS, R, Microsoft Excel, etc.
Suitable audience	

Figure 23 A line graph showing the threshold when benefit-risk balance changed

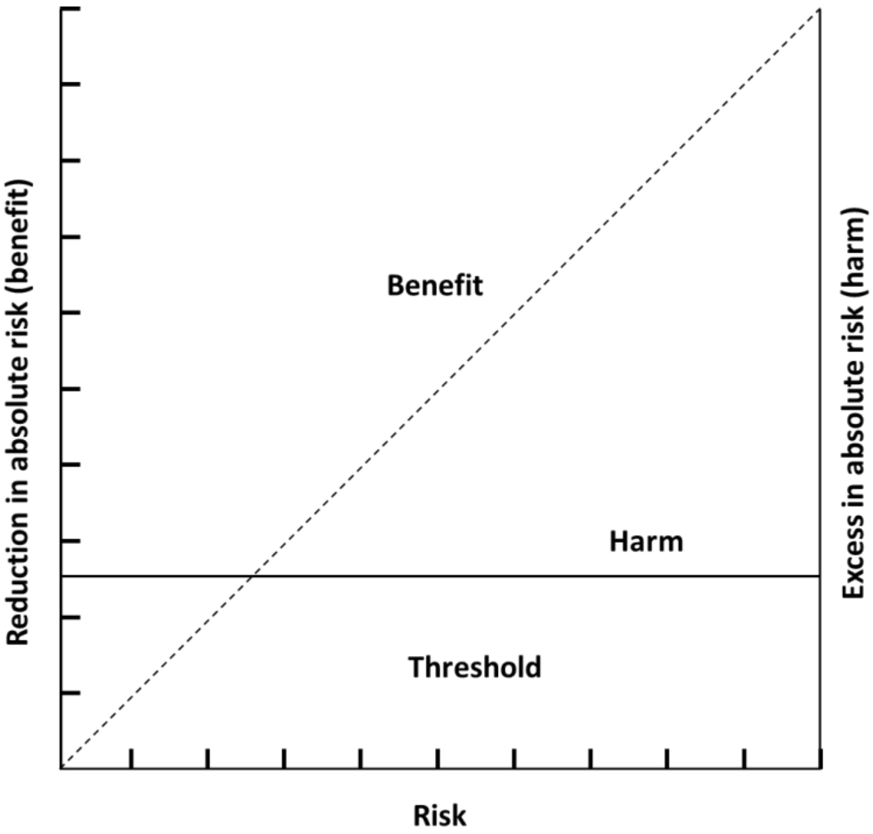


Figure 24 A line graph accompanied by 95% confidence intervals lines for net clinical benefit

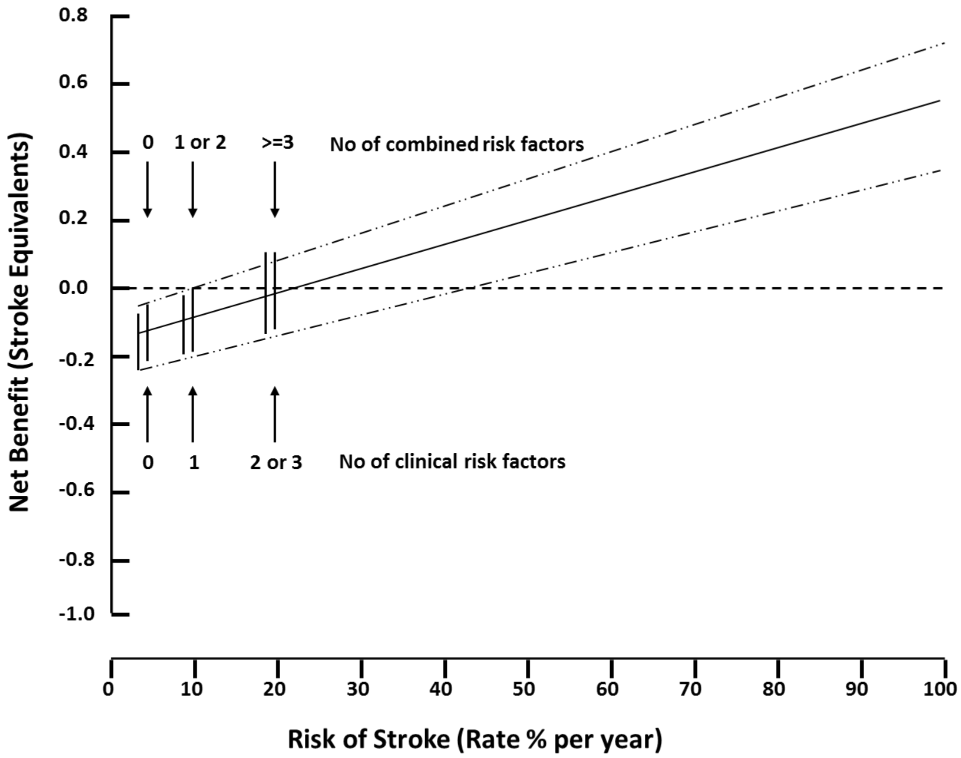


Figure 25 A line graph showing piecewise linear value preference over a range of flares per person-year

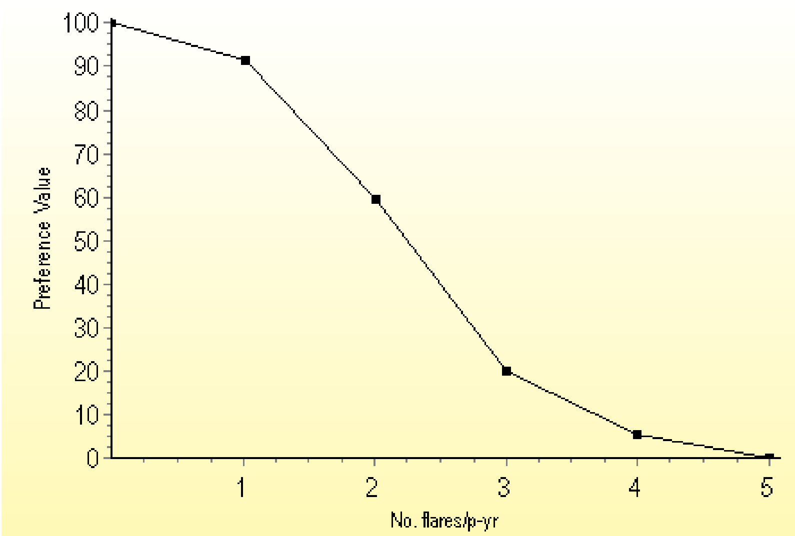


Figure 26 A line graph showing the changing in efficacy and toxicity over a range of exposure values

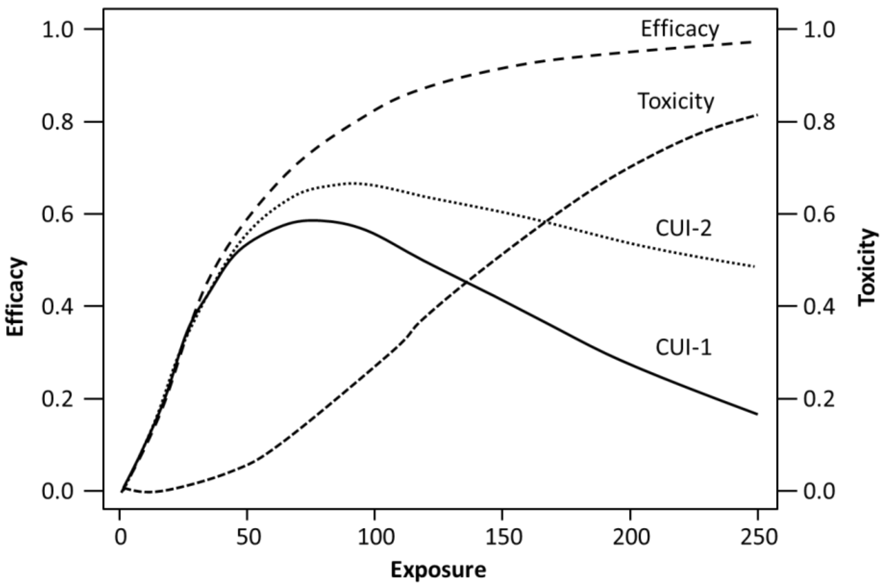
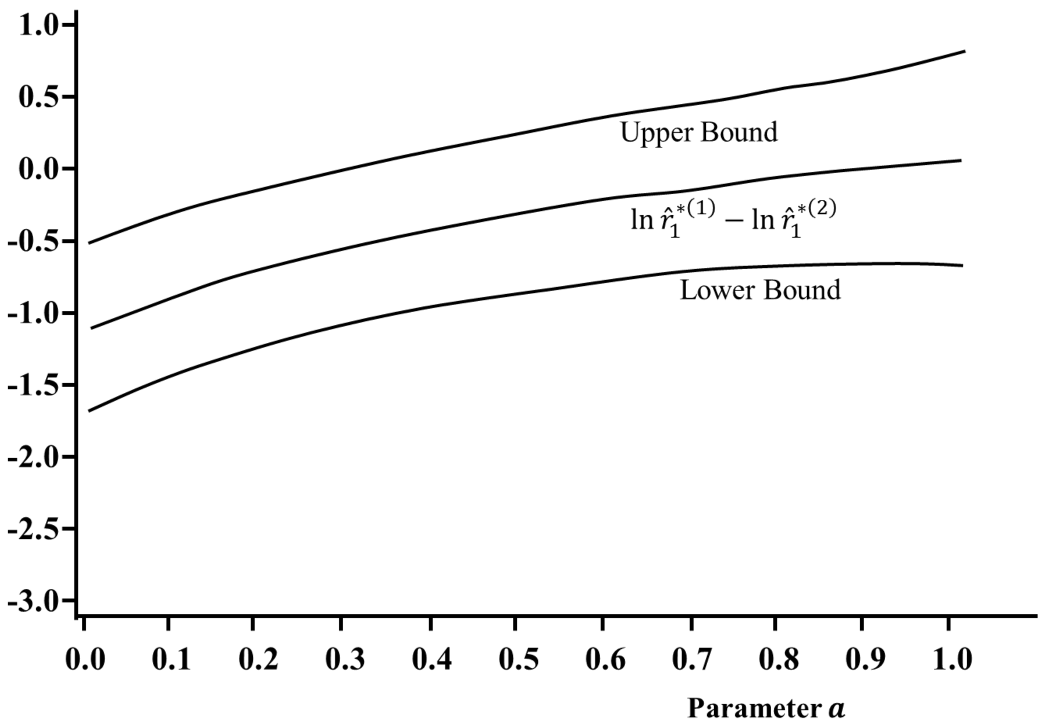


Figure 27 A line graph accompanied by upper and lower bound over a range of model parameter a



A.8.2 Visual evaluation of line graph

Table 33 Wickens' principles of display design for line graph

	Appraisal
Perceptual principles	
Legibility (or audibility)	The legibility of line graphs is good as long as the choice of line patterns or colours is acceptably discriminable.
Absolute judgment	Most of the times the lines drawn are at an angle (not horizontal or vertical) or curved. This reduces the sharpness of the lines as well as limiting the absolute judgment. However, this is not an issue easily resolved due to the resolution on a computer screen or the printer. Points can be marked for example with symbols as reference points (Cleveland, 1984) to help with absolute judgment limits as well as increasing accuracy to extract information (Figure 25).
Top-down processing	Audiences are likely to have relatively high levels of exposure to line graphs. Any message that is presented in this format would therefore benefit from taking account of the cultural norms are likely to exist. For example, it may be that a rising line is widely understood to represent an increasingly positive value.
Redundancy gain	Colour coding and shading are likely to aid accessibility of information.
Discriminability	Likely to be high, assuming clear lines/axes/labels and legends.
Principles based on attention	
Information access cost	This has the potential to be low but will depend on the number of lines present and whether they change direction and cross each other.
Proximity compatibility	Appropriate placing of lines on applicable scales is likely to facilitate proximity compatibility.
Memory principles	
Use of existing knowledge of the world	Audiences are likely to have relatively high levels of exposure to line graphs. Any message that is presented in this format would therefore benefit from taking account of the cultural norms are likely to exist. For example, it may be that a rising line is widely understood to represent an increasingly positive value. Given the wide varieties of likely exposure to line graphs care must also be taken if any predictions are to be made as to likely interpretation (over and above the most general) if based on prior knowledge.
Predictive aiding	This is possible but may be limited, for example, via trajectories of line. In some cases benefit and risk criteria are presented as separate lines to convey their magnitudes at a given exposure value across a range of exposure values (Figure 26). This facilitates DMs to determine an acceptable benefit-risk balance.
Consistency	Consistent representation of lines, areas legends and axes are possible.

A.8.3 Communicability evaluation of line graph

Table 34 Elements of visual communication for line graph

Element	Appraisal
Risk magnitude	Possible to display risk magnitude via line graphs.
Relative risk	Possible to represent relative risk via line graphs
Cumulative risk	Possible to represent cumulative risk via line graphs and to include the element of time.

Uncertainty	Possible to include some form of confidence interval representation either via lines, shaded area or specific values at data points. However, the use of shaded area may limit absolute judgment.
Interactions among risk factors	It is not simple to represent interactions but it may be possible to add this information in a benefit/risk setting in a simple way via the use of multiple lines.

A.9 Network graph

A.9.1 Description of network graph

Table 35 Characteristics of visual representation type for network graph

Form	2D static graphic
Endpoint	Not applicable
Methodology-specific	No but benefit-risk approaches like the Directed Acyclic Graphs (DAGs), Indirect and Mixed Treatment Comparison (ITC and MTC), and Confidence Profile Method (CPM) are heavily dependent on network graph to formulate the benefit-risk problems.
Reproduction	DAGs software like GeNIe (http://genie.sis.pitt.edu) can reproduce (and fit DAGs model) network graphs as part of the analysis. WinBUGS can also be used to build network graphs through “doodle” facility. However, in most cases, network graphs can be constructed in word editor like Microsoft Word – Figure 30 was produced in Microsoft Word 2010.
Suitable audience	More scientific audience would benefit more from network graph.

Figure 28 A network graph showing how nodes are connected

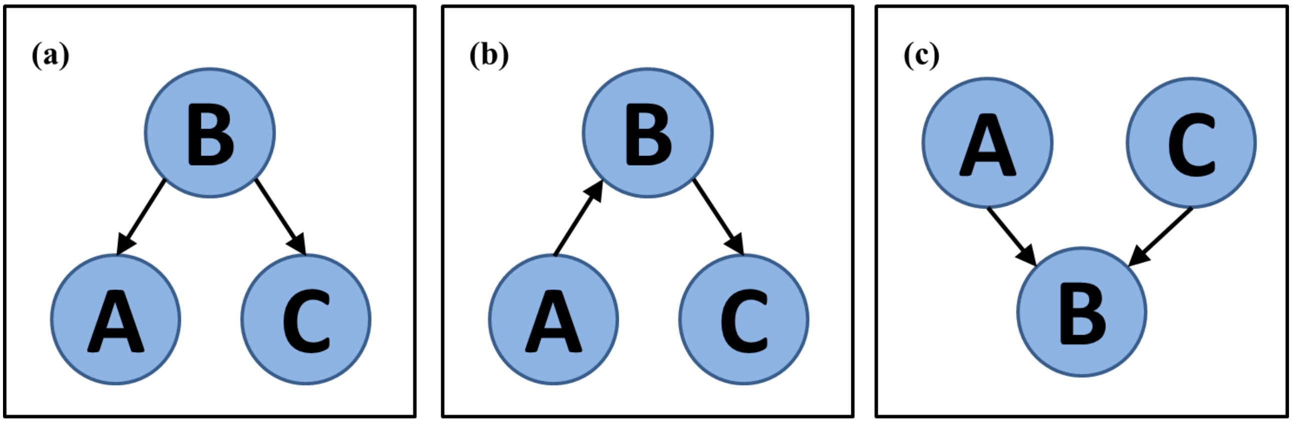


Figure 29 A network graph showing comparisons for thrombolytics and angioplasty in acute myocardial infarction in a mixed treatment comparison analysis

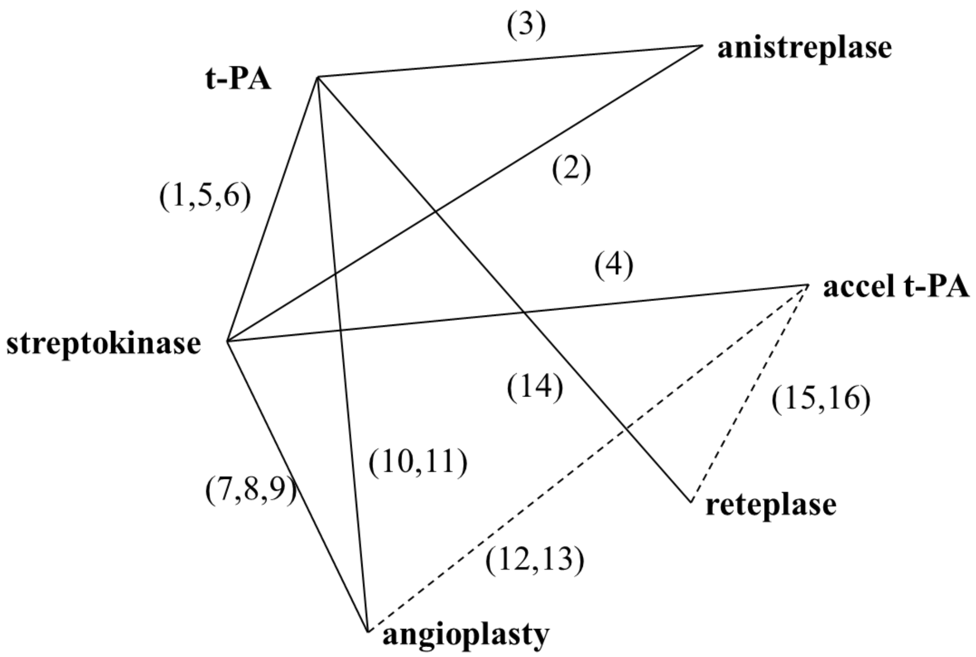
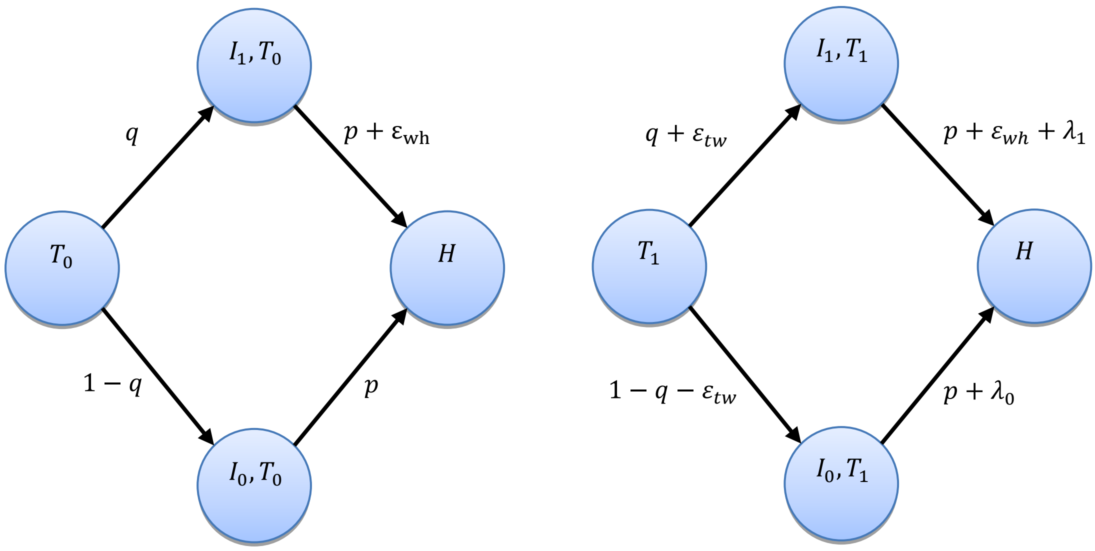


Figure 30 A network graph showing the derivation of formula used in confidence profile method



A.9.2 Visual evaluation of network graph

Table 36 Wickens' principles of display design for network graph

	Appraisal
Perceptual principles	
Legibility (or audibility)	Many network graphs have poor legibility due to the fact that there are many connecting lines involved and intersecting each other. Sometimes, this is unavoidable because of how the nodes on the network are connected. But in many cases this can be avoided by carefully arranging the nodes to minimise intersecting lines therefore increases its legibility. In the case when there are many nodes involved, the nodes and labels are often made smaller in order to accommodate them on a page. This also results in poor legibility.
Absolute judgment	Since network graphs are used in benefit-risk assessment to show the relationship between evidence. In most cases the nodes clearly represent the evidence data for a model. In terms of absolute judgment, there should be no confusion as to which elements constitute the evidence data (the nodes), and which do not (the connecting lines and anything else outside the nodes).
Top-down processing	The top-down processing occurs in network graphs when DMs connect different nodes. If there is an obvious pattern is how they are connected, then a sudden change in this pattern can cause confusion and can increase processing time.
Redundancy gain	Redundancy gain can be achieved through multiple use of coding to highlight similar or different elements on the graph (see Discriminability below).
Discriminability	This can be achieved through colour coding nodes/lines, nodes outline patterns, line patterns, or rearrangement and grouping of nodes.
Principles based on attention	
Information access cost	Figure 29 and Figure 30 minimise information access cost by labelling the arrows to represent how the two connected nodes are related. However, the cost for accessing this information is affected by the complexity of the relationship, and the use of familiar/unfamiliar notations to the DMs. The cost is further increased if the explanation of the extra information cannot be found easily within the document.
Proximity compatibility	The nodes on network graph can be arranged in such a way that nodes representing similar type of information are grouped together, for example by drawing nodes for benefit criteria closer together, and similarly for risk criteria. The grouping of nodes can also be used to show the data source from which the data came from. Any additional information added on network graph e.g. arrow labels on Figure 29 and Figure 30 should be available nearby.
Memory principles	
Use of existing knowledge of the world	Network graphs are commonly used in transport network map such as trains and ferries. General public who have the exposure to these transport network maps may find network graphs in benefit-risk assessment more useful.
Predictive aiding	It is possible to enhance the appearance of network graph using colour coding for example to show which nodes influence another more, or using varying nodes (or arrows) sizes; otherwise there are limited possibilities for predictive aiding.
Consistency	Consensus on the use of symbols etc. on network graphs should be established to allow greater consistency, such as using different symbols to represent different types of element e.g. circles for probabilistic nodes, squares for deterministic, etc. (see Appendix A.13).

A.9.3 Communicability evaluation of network graph

Since network graph is not suitable for communicating results of benefit-risk assessment, the communicability is limited to showing the relationship of the nodes on the diagram.

Table 37 Elements of visual communication for network graph

Element	Appraisal
Risk magnitude	n/a
Relative risk	n/a
Cumulative risk	n/a
Uncertainty	n/a
Interactions among risk factors	n/a – it is likely that there may be interactions between neighbouring nodes.

A.10 Scatter plot

A.10.1 Description of scatter plot

Table 38 Characteristics of visual representation type for scatter plot

Form	2D static graphic
Endpoint	Any point estimates including mean, median, and percentile.
Methodology-specific	No
Reproduction	Many software can reproduce scatter plots with ease so long as they have the capability to reproduce Cartesian (X-Y) graphs. These include Stata, SAS, R, Microsoft Excel, etc.
Suitable audience	

Figure 31 A scatter plot showing the relationship between incremental harm and incremental benefit in a probabilistic simulation

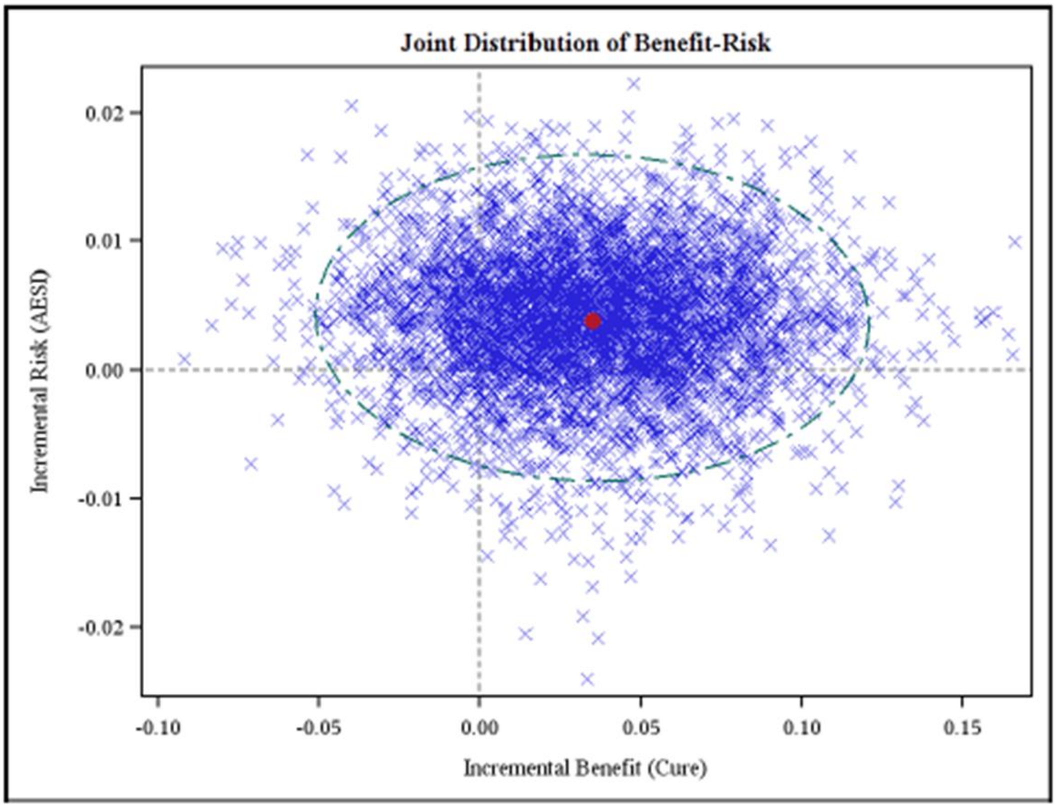
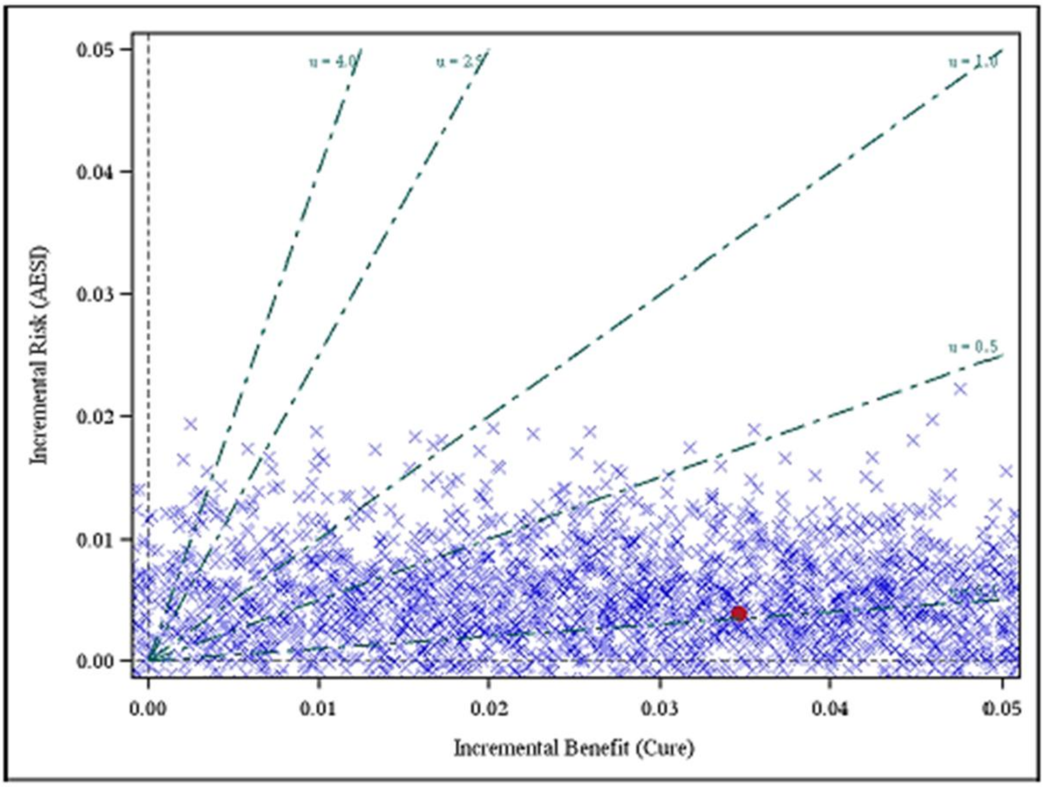


Figure 32 A scatter plot showing incremental risk against incremental benefit and various thresholds in a probabilistic simulation analysis of the incremental net health benefit



A.10.2 Visual evaluation of scatter plot

Table 39 Wickens' principles of display design for scatter plot

	Appraisal
Perceptual principles	
Legibility (or audibility)	Although scatter plot is one of the common visual displays, its legibility is poor particularly when there are too many data points being plotted. However, when summary results, for example means, are plotted instead of individual data points, the legibility is as good as that of bar graph (minus the bars).
Absolute judgment	Absolute judgment of scatter plots depends on the symbols used. It is recognised that the centre of the symbol represents the value of the information to be conveyed. However, the larger the symbols are or the more irregular shapes they are, the more absolute judgment is limited.
Top-down processing	Some top down processing may occur due to familiarity with line graphs and potentially via ability to interpret direction of scatter.
Redundancy gain	Potential for use of colour or shading/symbols to identify different 'populations' within the scatter but otherwise redundancy gain more limited.
Discriminability	Clear use of axes, lines, legends and symbols within the scatter can be used to facilitate discriminability.
Principles based on attention	
Information access cost	Potential for higher cost with larger numbers of points, particularly if more than one population is represented in one plot.

	Appraisal
Proximity compatibility	It is possible to maximise proximity and information as to overlap between groups of interest via scatter plot but this may become confusing if focus on one group/outcome is required.
Memory principles	
Use of existing knowledge of the world	Experience with line graphs may transfer to interpretation of scatter plots.
Predictive aiding	Potential for this given the multiple sources of information contained within the scatter plots.
Consistency	Possible to standardise colour, shading, size, symbols for use in the plots.

A.10.3 Communicability evaluation of scatter plot

Table 40 Elements of visual communication for scatter plot

Element	Appraisal
Risk magnitude	Difficult to represent without some summary text or line, unless the points that are plotted represent magnitude of risk.
Relative risk	Possible to represent different outcomes and populations with different relative risk but likely to need summary text/line to represent relative risk for whole populations/groups.
Cumulative risk	Possible to incorporate time and cumulative risk
Uncertainty	A line and 95% confidence intervals or similar representation of range around the line can be added to a scatter plot. Overall it could be said that the spread of the scatter itself conveys a partial understanding of uncertainty.
Interactions among risk factors	Possible to include representation of these at a simple level, once multiple risk factors/interactions are included the scatter may become complex to interpret.

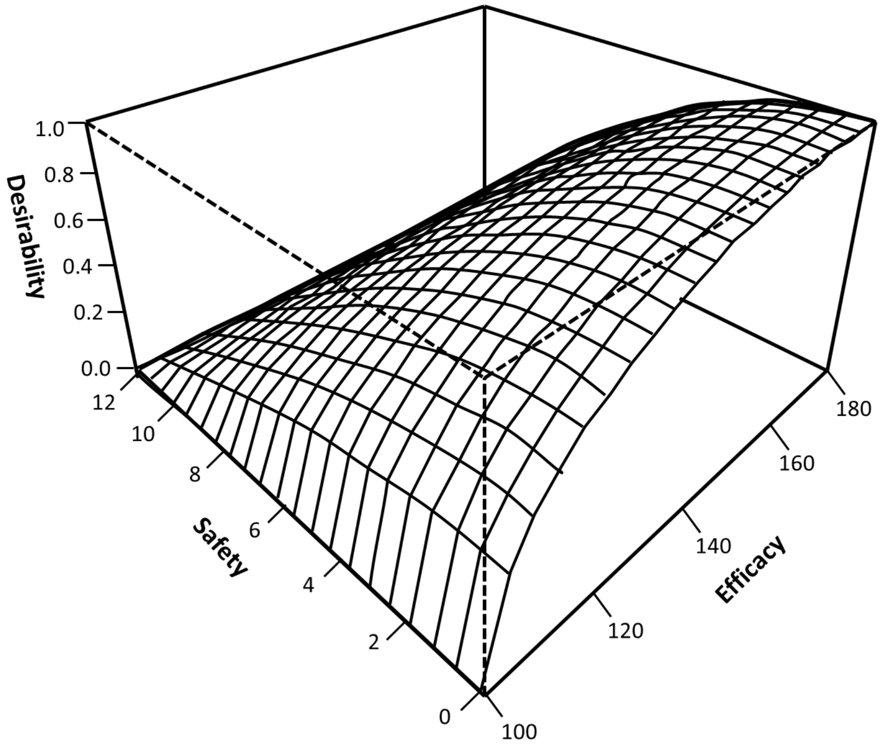
A.11 Surface plot

A.11.1 Description of surface plot

Table 41 Characteristics of visual representation type for surface plot

Form	3-D static graphic
Endpoint	Any point estimates which are continuous
Methodology-specific	No but Figure 33 is specific to desirability index.
Reproduction	Statistical packages like R and SAS have built-in commands to produce surface plots. "Annotate" function in SAS may need to be used to enhance the appearance of the surface plot produced. There is a user-written command in Stata to produce wireframe surface plot (http://ideas.repec.org/c/boc/bocode/s448501.html) but it lacks the aesthetic appearance. A version of surface plot also known as contour plot can be more easily produced in R, SAS, and Stata version 12. Microsoft Excel can also produce excellent surface plots.
Suitable audience	As with any 3-D graphs, surface plots are not easily understood, and very hard to be perceived correctly. Therefore, surface plots may only be suitable to specialist groups of statisticians, regulators, and other decision-makers.

Figure 33 A surface plot showing the desirability index for a combinations of safety and efficacy values



A.11.2 Visual evaluation of surface plot

Table 42 Wickens' principles of display design for surface plot

	Appraisal
Perceptual principles	
Legibility (or audibility)	Since surface plots communicate three-dimensional information on a two-dimensional surface, the legibility can be very poor and highly dependent on the rotation and angle of the plot when it is presented. The legibility of surface plot can be increased when it is presented interactively, or when a series of surface plots are presented representing different angles and rotations of them to the DMs on a two-dimensional medium. However, colourful (colour-coded) surface plots may have increased aesthetic values, could attract more attention, and perhaps may be more engaging.
Absolute judgment	Absolute judgment is limited. DMs may not be able to determine a point on the graph due to "hidden" surface cause by inappropriate rotation or angle. In Figure 33, it is only possible to tell that desirability is low when efficacy and safety are high but not the actual desirability index. Interpolation is difficult in Figure 33 but it may not matter for the purpose it was presented for. The transparent cube in Figure 33 surrounding the plot acts as reference points to help with absolute judgment. In other examples not presented here, the third dimension is presented by fading greyscales or changing colour hue, which clearly limit the absolute judgment since it is difficult to determine the beginning and end values and what they represent.
Top-down processing	This criterion can be difficult to fulfil for the same reasons as contour plot (Appendix A.3) with the added complexity of the third dimension axis. In a series of surface plots, it would help DMs if the angle and rotation are the same across all plots to minimise top-down processing errors. For example, when another surface plot of, say, another drug's benefit-risk profile with different rotation such that the "safety" and "efficacy" axes are swapped6 is compared to the one in Figure 33, DMs may end up making incorrect comparison.
Redundancy gain	Redundancy gain can be achieved as shown in Figure 33 by placing reference points as red circles to focus DMs to these points, although the wireframe structure on the surface plot can already convey them. Like contour plots (Appendix A.3), colour saturation and greyscale are also often used to augment parts of the plot with higher or lower values.
Discriminability	Discriminability is rather poor in surface plot due to potentially poor legibility and limited absolute judgment.
Principles based on attention	
Information access cost	The information access cost is increased in surface plot due to the complexity in extracting the required information. The complexity varies with different designs of contour plots in the sense of whether it is 2-D or 3-D, the use of colour saturation, hue, or greyscale, the rotation, and the angle.

[6] Most computer packages automatically determine the "best" angle and rotation for surface plots based on data which may result in inconsistent presentation of surface plots. The direction of the axis-labelling and how they appear may also be affected. Therefore surface plot designers need to ensure that any two surface plots are directly comparable.

	Appraisal
Proximity compatibility	The third dimension plotted on surface plot is a feature which brings in the extra information required into the same plot, making it easy to perform mental integration of the presented information.
Memory principles	
Use of existing knowledge of the world	It is likely that more expert DMs would gain more from the use of surface plot in benefit-risk assessment. This is similar to contour plot (see Appendix A.3).
Predictive aiding	As in Figure 33, it is easy enough to help DMs with predictive tasks by plotting the benefit-risk integration results on the additional axis against benefit and risk axes.
Consistency	The consistency in surface plot can arise from standardised use of colours, shading, patterns, etc.

A.11.3 Communicability evaluation of surface plot

Table 43 Elements of visual communication for surface plot

Element	Appraisal
Risk magnitude	Risk magnitude is displayed on a surface plot across three dimensions. If the third dimension is an intuitive combination of the other two, then the risk magnitude is easier to extract, as in Figure 33. If the third dimension is another completely different variable, then it is likely that there may be some difficulties in obtaining the correct information, and with increased "cost". However, it is probably best when the benefit and risk criteria are combined within their domain to facilitate this task. Otherwise, presenting the criteria separately on the same surface plot may cause poor legibility and make information extraction incomprehensible.
Relative risk	It is possible to communicate relative risk on a surface plot, but with increased "cost".
Cumulative risk	It is possible to communicate cumulative risk on a surface plot as in line graph (Appendix A.8) but with increased "cost".
Uncertainty	It is difficult to show uncertainty on a surface plot, and likely to be confusing if done.
Interactions among risk factors	It is difficult to communicate interaction using a surface plot.

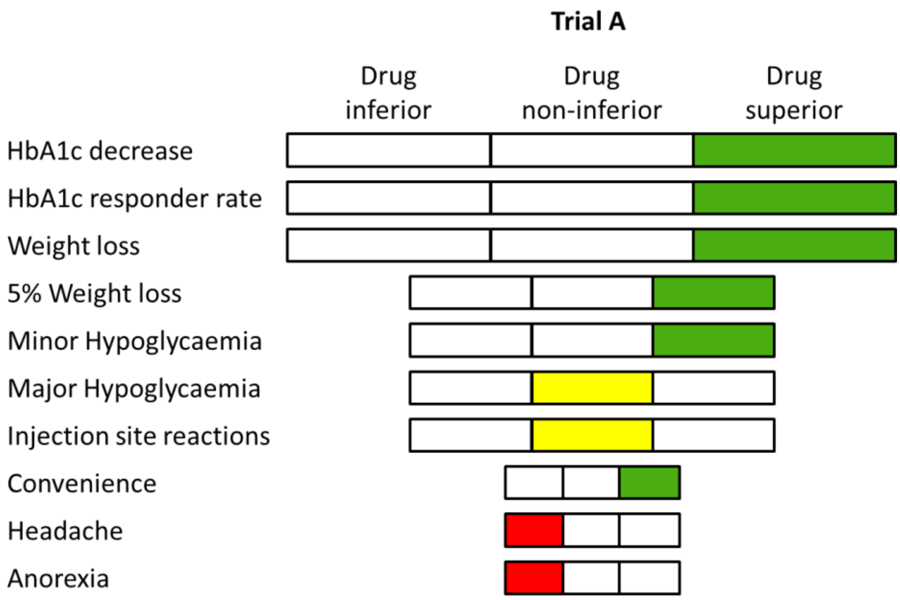
A.12 Tornado diagram

A.12.1 Description of tornado diagram

Table 44 Characteristics of visual representation type for tornado diagram

Form	2-D static graphic
Endpoint	Any point estimate
Methodology-specific	Not in general but Figure 34 is specific to the Novo Nordisk Sarac's Benefit-Risk Assessment Method (SBRAM).
Reproduction	There is no software or specific graphic commands to produce tornado diagrams but it can be achieved by manipulating bar graph commands in statistical software like Stata, SAS, and R. Microsoft Excel can also produce tornado diagrams with some formatting to the bar chart facility, or manually by filling colours or shades to the cells.
Suitable audience	Due to being an uncommon way of presenting information visually, tornado diagrams may only be suitable to specialist groups of statisticians, regulators, and other decision-makers

Figure 34 A tornado diagram showing the discrete outcomes for criteria in a SBRAM analysis



A.12.2 Visual evaluation of tornado diagram

Table 45 Wickens' principles of display design for tornado diagram

	Appraisal
Perceptual principles	
Legibility (or audibility)	The important aspect of legibility in a tornado diagram is the aligned centre and the length of the bars. Other aspects of its legibility are the same as bar graph (see Appendix A.2).
Absolute judgment	The tornado diagram proposed through the SBRAM approach (Sarac, 2011) as

	Appraisal
	shown in Figure 34 increases absolute judgment compared to other typical tornado diagram. This is simply because the outcome for each criterion can take only three discrete values – being inferior, non-inferior, and superior. In general, the lengths of the bars correspond to variability ("volatility") of the item it presents; but it is unclear in Figure 34 whether the lengths correspond to the actual weights they bear.
Top-down processing	The criteria on the tornado diagram are arranged in such a way that the more weighted criteria come at the top and the least weighted criteria come at the bottom. DMs can easily adjust to understand that criteria of the same length carry the same weight, and that going down tornado diagram, the criteria are less important. It is also worth pointing out that the benefit-risk balance is more sensitive to the changes to the criteria at the top than changes to the criteria at the bottom.
Redundancy gain	Redundancy gain is achieved in Figure 34 by colour-coding the three categories of superiority in red, yellow and green together with presenting them from left to right, and also in three blocks. In other tornado diagrams, it might not be the case.
Discriminability	Figure 34 exhibits good discriminability in terms of judging the superiority since the bars are discrete and consistently colour-coded. However, the discriminability between benefits and risks criteria is weak since they are not clearly labelled. In this case, it is difficult to tell which direction they are expected to go.
Principles based on attention	
Information access cost	The information access cost is minimised in the sense that the required information for making decision is presented within close proximity.
Proximity compatibility	Tornado diagram has good proximity compatibility for the importance of criteria since the criteria appear in a values-sorted order. However, this gives a perception that criteria of the same length belong to the same group which can be confusing. As a result, the benefits and risks criteria are mixed up, making them more difficult to evaluate.
Memory principles	
Use of existing knowledge of the world	The SBRAM is the first benefit-risk approach that we encounter proposing a tornado diagram to represent its results. It is likely that DMs with more expertise would benefit more from the diagram.
Predictive aiding	Although tornado diagram is good in aiding prediction for individual criterion, it does not aid very well in synthesising a judgment on the overall benefit-risk balance. Having to mentally combine the "imaginary" values on the proposed tornado diagram (Figure 34) can be both very perceptively and cognitively challenging.
Consistency	Consistent colour codes for the three possible outcomes and also consistent placement for the three outcomes increase the diagram's consistency. In any tornado diagram, a consistent feature is the order of most sensitive to the least sensitive criteria from top to bottom.

A.12.3 Communicability evaluation of tornado diagram

Table 46 Elements of visual communication for tornado diagram

Element	Appraisal
Risk magnitude	In a typical tornado diagram, the risk magnitude is communicated as the aligned centre of the bars on the diagram. However in the proposed simplified tornado diagram in Figure 34, risk magnitude cannot be communicated.
Relative risk	Relative risk on a typical tornado diagram can be mentally calculated if the magnitudes are labelled on the diagram and when the comparator is also depicted, otherwise it is impossible to do so. In Figure 34, the relative risk of a criterion has been estimated and represented as being inferior, not inferior or superior to comparator. However, the actual magnitude of relative risk is not presented.
Cumulative risk	It is difficult to see how cumulative risk can be communicated in a tornado diagram.
Uncertainty	Uncertainty is represented on a tornado diagram by the length of the bar. However, uncertainty is not represented in Figure 34.
Interactions among risk factors	Interactions of risk factors can be communicated on a tornado diagram by adding a bar corresponding to the desired interaction to be displayed. In Figure 34, it is more difficult to communicate interaction since the comparative results have been categorised by criteria.

A.13 Tree diagram

A.13.1 Description of tree diagram

Table 47 Characteristics of visual representation type for tree diagram

Form	2-D static graphic
Endpoint	Not applicable
Methodology-specific	No but benefit-risk approaches like Decision Tree, MCDA, and Markov Decision Process rely heavily on tree diagrams (including value tree) in their formulation.
Reproduction	Decision Tree software such as DPL (http://www.syncopation.com/) and MCDA software such as Hiview 3 (http://www.catalyze.co.uk/) can facilitate the production of tree diagrams. Microsoft Word can also be used to manually produce tree diagrams as shown in Figure 35 and Figure 36.
Suitable audience	

Figure 35 A tree diagram showing the utility of consequences for possible scenarios

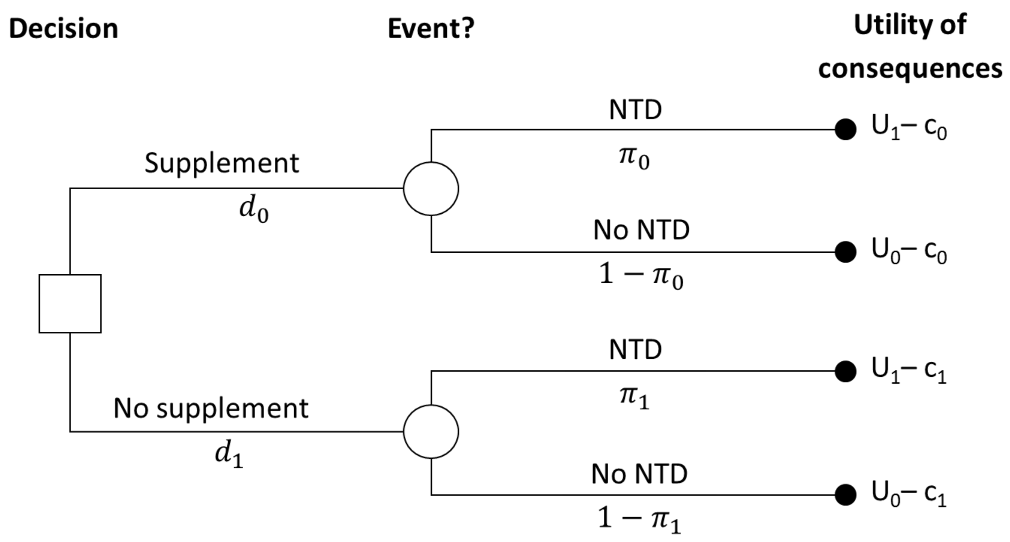


Figure 36 A tree diagram showing the utilities of possible outcomes due to uncertain clinical events

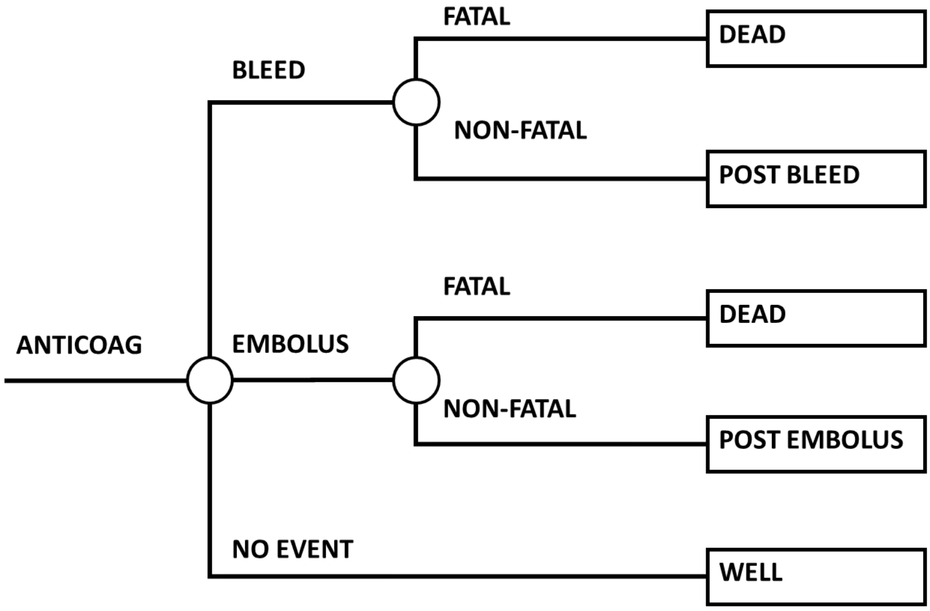
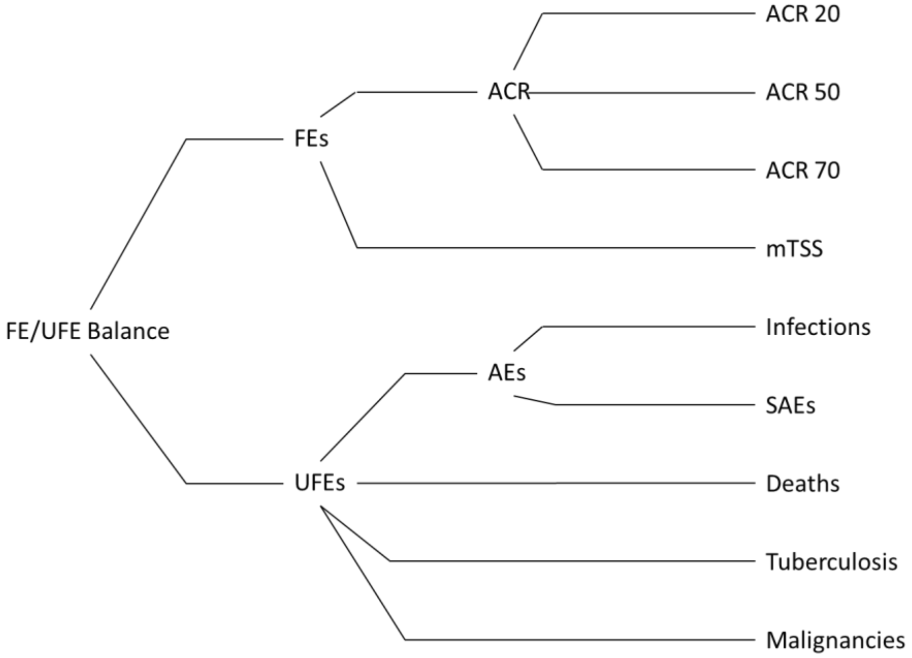


Figure 37 A tree diagram showing the criteria contributions to the FE/UFE balance in a MCDA analysis



A.13.2 Visual evaluation of tree diagram

Table 48 Wickens' principles of display design for tree diagram

	Appraisal
Perceptual principles	
Legibility (or audibility)	Tree diagram is similar to network graph but with better legibility since the nodes on tree diagram are connected in one direction – left to right or top to bottom (or vice versa). However, this convenience of presenting a tree diagram can hinder its legibility when there are too many nodes and hierarchies involved as the diagram will then grow smaller and smaller to fit within the same area. This issue can be rectified by presenting multiple tree diagrams for different parts of the main tree but doing so can increase cognitive costs.
Absolute judgment	Each possible state or those deemed important are represented in the tree diagram as nodes with the branches indicating how they are connected. It is also easy to see which nodes contribute to the expected utility when folding back. Therefore the absolute judgment limit is minimal here in that sense. However, it is difficult to determine a value on a node because in general the displayed probability associated with a node is a conditional probability. Therefore they are not independent of the earlier criteria (nodes).
Top-down processing	It is a common convention to represent stochastic events by circles and decision by squares (e.g. Figure 35). If the convention is adopted in all tree diagrams, it would have better top-down processing capability.
Redundancy gain	The use of symbol at the nodes introduces redundancy gain since they are not necessary but are there for clarity. The examples of introducing the redundancy gain in this manner are shown in Figure 35 and Figure 36. Figure 37 shows that it conveys as much without them.
Discriminability	Discriminability is achieved through having the branches. Criteria that belong to the same group, e.g. the benefits, branch out from the same parent nodes.
Principles based on attention	
Information access cost	The cost to extract benefit and risk balance from a tree diagram is increased since they do not convey this information. To minimise information access cost, probabilities associated with the possible events can be displayed next to the respective nodes.
Proximity compatibility	Because of the "branching" of the nodes in a tree diagram, elements of similar concepts or levels, e.g. the benefits or the risks, consequently fall into close proximity of each other.
Memory principles	
Use of existing knowledge of the world	Users and DMs who are used to organisational charts, at workplace for example, are likely to be the ones who would gain more understanding of a tree diagram.
Predictive aiding	The potential for predictive aiding when presenting tree diagrams is limited.
Consistency	One consensus when presenting a tree diagram is to use squares for decision nodes, i.e. when decision has to be made, and circles for probabilistic nodes i.e. when a random event occurs. However, this is not always adopted in all presentations; therefore the degree of consistency can vary.

A.13.3 Communicability evaluation of tree diagram

Since tree diagram is not suitable for communicating results of benefit-risk assessment, the communicability is limited to showing the relationship of the nodes on the tree and the possible events.

Table 49 Elements of visual communication for tree diagram

Element	Appraisal
Risk magnitude	n/a
Relative risk	n/a
Cumulative risk	n/a
Uncertainty	The uncertainty is conveyed through the probabilities of an event happening which can be labelled on the connectors, as in Figure 35.
Interactions among risk factors	n/a

-- End of document --