Benefit-Risk Integration and Representation

IMI-PROTECT Symposium
Benefit-Risk Integration and Representation Workshop
18th February 2014

PROTECT Work Package 5 and 6
Workshop objectives

The objectives of this workshop are to introduce:

1. the main concepts of benefit-risk assessment
2. a range of methodologies used in connection with benefit-risk assessment
3. key methods via worked examples
4. a range of visual representations to accompany benefit-risk assessment
5. a selection of visual displays via worked examples
6. the role of patient and public involvement in benefit-risk assessment
# Outline

<table>
<thead>
<tr>
<th>Title</th>
<th>Presenters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction to benefit-risk assessment of medicinal products</td>
<td>Deborah Ashby (Chair)</td>
</tr>
<tr>
<td>Structured benefit-risk assessment</td>
<td>Richard Nixon</td>
</tr>
<tr>
<td>Taxonomy of benefit-risk assessment methodologies; and BR metrics</td>
<td>Shahrul Mt-Isa</td>
</tr>
<tr>
<td>Benefit-risk assessment: Concepts and methods</td>
<td>Larry Phillips</td>
</tr>
<tr>
<td>Open source software for benefit-risk modelling and clinical data analysis: ADDIS and Effects Table</td>
<td>Douwe Postmus</td>
</tr>
<tr>
<td>Visualising benefits and risks: concepts and ideas</td>
<td>Christine Hallgreen</td>
</tr>
<tr>
<td>PROTECT resources for further learning</td>
<td>Gerry Downey</td>
</tr>
<tr>
<td><strong>TEA/COFFEE BREAK</strong></td>
<td></td>
</tr>
<tr>
<td>Patients and public involvement in benefit-risk assessment and decision-making: methods and applications</td>
<td>Ed Waddingham</td>
</tr>
<tr>
<td>Eliciting Patient Preferences: Applying decision theory to health research</td>
<td>Andrea Beyer</td>
</tr>
<tr>
<td>Closing remarks and take-home messages</td>
<td>Deborah Ashby</td>
</tr>
</tbody>
</table>
Deborah Ashby, PhD OBE FMedSci

INTRODUCTION TO BENEFIT-RISK ASSESSMENT OF MEDICINAL PRODUCTS
Disclaimer

“The processes described and conclusions drawn from the work presented herein relate solely to the testing of methodologies and representations for the evaluation of benefit and risk of medicines.

This report neither replaces nor is intended to replace or comment on any regulatory decisions made by national regulatory agencies, nor the European Medicines Agency.”
Evidence Based Medicine

“EBM is the conscientious explicit, and judicious use of current best evidence in making decisions about the care of individual patients” ... taking into account...
“individual patients predicaments, rights and preferences using best evidence from clinically relevant research.”

Sackett et al, 1996
Decision makers – hierarchical?

- **Patients**
  - Make decisions for themselves

- **Healthcare providers**
  - Make decisions based on prescribing lists

- **Health technology assessors**
  - Makes decisions on cost-effectiveness

- **Regulators**
  - Makes decisions on quality, safety, efficacy and benefit-risk balance to individuals and public health

- **Pharmaceutical companies**
  - Makes decisions on what to develop for which licenses to apply
Challenges in medical decision-making

- Should we formalise decision-making at all?
- Which quantitative approach(es) to use?
- Whose value preferences take priority – regulators, pharma, physicians or patients?
- How do we find these preferences – simple elicitation, decision conferencing, discrete choice experiments....?
- Do we need stakeholders’ preference a priori, or should we provide tools to allow individual decision-makers to explore their own preferences and the consequent decisions?
- How do we communicate benefits and risks?
A simple example of EBM decision-making

- Decision-maker
- Possible actions
- Uncertain consequences
- Sources of evidence
- Utility assessments

Key reference
Ashby D & Smith AFM, Stats in Medicine, 2000
Trastuzumab
Benefit-Risk captured with a single parameter

• Pivotal study:
  • randomised, open-label comparing trastuzumab and placebo in women with non-metastatic, operable primary invasive breast cancer over-expressing HER2 who had completed … therapy… for primary breast cancer.
# Trastuzumab

**Benefit-Risk captured with a single parameter**

- **Benefit:** Disease-free survival (Placebo vs. Trastuzumab)

<table>
<thead>
<tr>
<th></th>
<th>Placebo</th>
<th>Trastuzumab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion with either disease</td>
<td>22.0%</td>
<td>13.9%</td>
</tr>
<tr>
<td>progression or death (due to any</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cause)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion of death (due to any</td>
<td>2.4%</td>
<td>1.8%</td>
</tr>
<tr>
<td>cause)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Risk:** Cardiotoxicity (Placebo vs. Trastuzumab)

<table>
<thead>
<tr>
<th></th>
<th>Placebo</th>
<th>Trastuzumab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant asymptomatic (NYHA class I) or</td>
<td>0.53%</td>
<td>3.04%</td>
</tr>
<tr>
<td>mildly symptomatic (NYHA class II) cardiac</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dysfunction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptomatic congestive heart failure of NYHA</td>
<td>0.06%</td>
<td>0.6%</td>
</tr>
<tr>
<td>class III or IV or cardiac death</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Example: Trastuzumab for early breast cancer

<table>
<thead>
<tr>
<th>Decision-maker</th>
<th>The woman</th>
</tr>
</thead>
</table>
| **Possible decisions** | • Take trastuzumab  
|                      | • Not take trastuzumab                          |
| **Uncertain consequences** | • Breast cancer recurrence  
|                      | • Death  
|                      | • Cardiotoxicity                               |
| **Sources of evidence** | A pivotal trial                              |
| **Utility assessment** | Increased disease-free survival and cardiotoxicity |

# Simple benefit-risk metrics

<table>
<thead>
<tr>
<th><strong>Number needed to treat (NNT)</strong></th>
<th><strong>Number needed to harm (NNH)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of people to be treated to observe a benefit (or to prevent an adverse event)</td>
<td>Number of people to be treated to observe an adverse event (or to prevent a benefit)</td>
</tr>
<tr>
<td>$\text{NNT} = \frac{1}{\Delta p}$</td>
<td>$\text{NNH} = \frac{1}{\Delta q}$</td>
</tr>
<tr>
<td>$= \frac{1}{\text{Pr}(B</td>
<td>T) - \text{Pr}(B</td>
</tr>
<tr>
<td>where</td>
<td>where</td>
</tr>
<tr>
<td>$\text{Pr}(B</td>
<td>T) = \text{probability of observing a benefit among treated individuals}$; and</td>
</tr>
<tr>
<td>$\text{Pr}(B</td>
<td>T') = \text{probability of observing a benefit among untreated individuals}$</td>
</tr>
</tbody>
</table>
NNT and NNH approach for trastuzumab

\[
NNT = \frac{1}{0.861 - 0.780} = 12.3
\]
= for every 13 patients treated, one would benefit from progression-free survival

\[
NNH = \frac{1}{0.0304 - 0.0053} = 39.8
\]
= for every 40 patients treated, one would experience cardiotoxicity

NNT<NNH is desirable
Trastuzumab
Benefit-Risk captured with a single parameter

- MHRA Assessment Report: “If disease-free survival and primary cardiac events were combined into a single endpoint it would be dominated by the disease-free survival data with the hazard ratio favouring trastuzumab.”
- Benefit: Risk captured with a single parameter assuming equal weight for progression, cardiac event and death from any cause.
- Does further quantification add anything in this type of scenario?
- Could estimate weighting that would need to be given to make the benefit: risk unfavourable, or incidence of cardiac events to make benefit: risk unfavourable given equal weight.
**Recommendation Roadmap**

**Planning**
- critical issues
- think & discuss purpose and context
- documentation
- foundations for future analyses and updates

**Evidence gathering and data preparation**
- relevant evidence
- data collection
- data aggregation
- missing/incomplete data

**Analysis**
- Evaluate data
- Quantify benefits and risks
- Weigh or integrate

**Exploration**
- robustness
- sensitivity
- assumptions and uncertainties
- other consequences
- impact or added value to the RMPs

**Conclusion and dissemination**
- communicate results/consensus
- any influence on future actions
- transparent audit trail
- ensures "big picture" is not lost

Hughes D, et al. Recommendations for the methodology and visualisation techniques to be used in the assessment of benefit and risk of medicines. IMI-PROTECT Website 2014.
FROM QUALITATIVE TO QUANTITATIVE BENEFIT-RISK DECISION-MAKING:
STRUCTURED BENEFIT-RISK ASSESSMENT

Richard Nixon, PhD
Decide on a Multiple Sclerosis treatment

*Three outcomes are important to you*

- For two treatments given over a two-year period the proportion of patients experiencing each of three outcomes is:

<table>
<thead>
<tr>
<th></th>
<th>Treatment A</th>
<th>Treatment B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disability progression</td>
<td>40%</td>
<td>30%</td>
</tr>
<tr>
<td>Flu-like reaction</td>
<td>5%</td>
<td>3%</td>
</tr>
<tr>
<td>PML*</td>
<td>0%</td>
<td>0.5%</td>
</tr>
</tbody>
</table>

- Which treatment would you choose?
  - How often does each outcome occur?
  - How important is each outcome if it occurs?

- In real life the decision is more complex
  - Which outcomes do you choose to make the decision?
  - Which treatments do you choose between?
  - How do you assess how important each outcome is to you?

* PML: Progressive multifocal leukoencephalopathy
Natalizumab – A short history

- Natalizumab was approved in 2004 by the FDA for the treatment of relapsing remitting multiple sclerosis (RRMS).
- In 2005 the drug was suspended because of an associated incidence of progressive multifocal leukoencephalopathy (PML), a rare neurological disorder.
- In 2006 it was re-introduced due to patient demand, but with strict risk minimization measures.
- In 2009, due to occurrence of further PML in monotherapy post marketing, CHMP reassessed the PML risk of Tysabri and confirmed the current approval.
The BRAT* Framework for benefit-risk

*Benefit Risk Action Team

**Built on methods to support decision making**

- **A framework, not a recipe**
  - A tool to support decision makers, not an algorithm to replace them.
  - Helps to develop a common understanding of that is of central importance.
  - Process to structure and analyze information.
  - Visualization tools to communicate benefit-risk.

- **Built on well-established Decision Analysis principles**
  - Promotes traceability, transparency and consistency.

- **Communication tool for decision making**
  - Consolidated view of key benefit and risk outcome measures.
### 1) Define a decision context

*Sets the frame of the structured benefit-risk assessment*

<table>
<thead>
<tr>
<th>Objective</th>
<th>Should natalizumab be kept on the market given that episodes of PML are observed?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indication</td>
<td>Relapsing remitting multiple sclerosis</td>
</tr>
<tr>
<td>Population</td>
<td>Adults with relapsing remitting multiple sclerosis</td>
</tr>
<tr>
<td>Drug</td>
<td>Natalizumab, 300mcg, iv, qm.</td>
</tr>
<tr>
<td>Comparative Treatment Alternative(s)</td>
<td>Placebo, Interferon beta-1a, 30mcg, im, qw Glatiramer acetate, 20mg, sc, qd</td>
</tr>
<tr>
<td>Assessment time point</td>
<td>Two years. For PML fives year as it takes longer to manifest.</td>
</tr>
<tr>
<td>Stakeholder perspective</td>
<td>EMA</td>
</tr>
</tbody>
</table>
2) Identify key benefits and risks

Organize the key outcomes driving the benefit-risk in a value tree

Benefits
- Relapse
  - Outcome
    - 2-year relapse rate

Treatment
- Oral od, s c od, i.m. qw, iv. qm

Risks
- Infection
  - PML
  - Flu-like reactions
  - Seizures
  - Infusion/injection reactions
- Reproductive Toxicity
- Liver Toxicity
- Neurological Toxicity
- Other

Benefits
- Convenience
  - Outcome
  - % w/event in 2yrs

Risks
- Transaminases elevation
  - Outcome
  - % w/event in 2yrs

Outcome
- Relapse
  - % w/event in 2yrs

Reproductive Toxicity
- Toxicity
  - Outcome
  - % w/event in 2yrs

Liver Toxicity
- Neurological
  - % w/event in 2yrs

Neurological Toxicity
- Other
  - % w/event in 2yrs

Other
- Flu-like reactions
  - % w/event in 2yrs

Reactivation of serious herpes viral infections
- PML
  - % w/event in 2yrs

Congenital abnormalities
- Transaminases elevation
  - Outcome
  - % w/event in 2yrs

Seizures
- Hypersensitivity reactions
  - % w/event in 2yrs

Flu-like reactions
- Transaminases elevation
  - Outcome
  - % w/event in 2yrs

Transaminases elevation
- Hypersensitivity reactions
  - % w/event in 2yrs

Transaminases elevation
- Flu-like reactions
  - % w/event in 2yrs

Transaminases elevation
- Flu-like reactions
  - % w/event in 2yrs

Transaminases elevation
- Flu-like reactions
  - % w/event in 2yrs

Transaminases elevation
- Flu-like reactions
  - % w/event in 2yrs
3) Consolidate source data

*Pool clinical data from internal and external studies*

- **Identify**
  - Search strategy
  - Search query

- **Select**
  - Study eligibility criteria

- **Extract**
  - Extraction guidelines

- **Aggregate**
  - e.g. meta-analysis, placebo-calibration

**Study worksheet**
- one row per study

**Data source table**
- one row per study/treatment/outcome

**Data summary table**
- one row per outcome
### 4) Customize and communicate

**Effects table of key benefits and risks**

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Outcome</th>
<th>Natalizumab prob / 1000pts</th>
<th>Placebo prob / 1000pts</th>
<th>Prob difference (95%CI) / 1000pts</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Convenience Benefits</strong></td>
<td>Convenience</td>
<td>-</td>
<td>-</td>
<td>- (-,-)</td>
</tr>
<tr>
<td><strong>Medical Benefits</strong></td>
<td>Relapse (# patients)</td>
<td>276</td>
<td>537</td>
<td>-261 (-326,-195)</td>
</tr>
<tr>
<td></td>
<td>Disability Progression</td>
<td>113</td>
<td>230</td>
<td>-117 (-124,-110)</td>
</tr>
<tr>
<td><strong>Infection</strong></td>
<td>Reactivation of serious herpes viral infections</td>
<td>0</td>
<td>0</td>
<td>0 (-6,3)</td>
</tr>
<tr>
<td></td>
<td>PML</td>
<td>1.51</td>
<td>0</td>
<td>1.51 (0,3)</td>
</tr>
<tr>
<td><strong>Liver Toxicity</strong></td>
<td>Transaminases elevation</td>
<td>50</td>
<td>40</td>
<td>10 (-19,36)</td>
</tr>
<tr>
<td><strong>Reproductive Toxicity</strong></td>
<td>Congenital abnormalities</td>
<td>0</td>
<td>0</td>
<td>0 (-6,3)</td>
</tr>
<tr>
<td><strong>Neurological Disorders</strong></td>
<td>Seizures</td>
<td>5</td>
<td>5</td>
<td>0 (-2,12)</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td>Infusion/Injection reactions</td>
<td>236</td>
<td>0</td>
<td>236 (202,269)</td>
</tr>
<tr>
<td></td>
<td>Hypersensitivity reactions</td>
<td>0</td>
<td>0</td>
<td>0 (-6,3)</td>
</tr>
<tr>
<td></td>
<td>Flu-like reactions</td>
<td>399</td>
<td>399</td>
<td>0 (-90,86)</td>
</tr>
</tbody>
</table>

**Summarize in one place all the benefits and risks data that are driving the decision**
4) Customize and communicate

**Forest plot**

Relapse = Number of patients with at least one relapse
5) Assess outcome importance

**MCDA and the Women's heptathlon**

<table>
<thead>
<tr>
<th>Event</th>
<th>Jessica Ennis</th>
<th>Value</th>
<th>Lilli Schwarzkopf</th>
<th>Value</th>
<th>Tatyana Chernova</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Javelin throw (m)</td>
<td>47.49</td>
<td>812</td>
<td>51.73</td>
<td>894</td>
<td>46.29</td>
<td>789</td>
</tr>
<tr>
<td>High Jump (cm)</td>
<td>186</td>
<td>1055</td>
<td>183</td>
<td>1016</td>
<td>180</td>
<td>979</td>
</tr>
<tr>
<td>200 metres (s)</td>
<td>22.83</td>
<td>1096</td>
<td>24.77</td>
<td>909</td>
<td>23.67</td>
<td>1013</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2963</strong></td>
<td><strong>2819</strong></td>
<td><strong>2781</strong></td>
<td><strong>2781</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5) Assess outcome importance

**MCDA and multiple sclerosis drugs**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weight</th>
<th>Measure</th>
<th>Value</th>
<th>Benefit-risk</th>
<th>Measure</th>
<th>Value</th>
<th>Benefit-risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relapse</td>
<td>8%</td>
<td>1.46</td>
<td>0.27</td>
<td>0.022</td>
<td>0.47</td>
<td>0.766</td>
<td>0.061</td>
</tr>
<tr>
<td>PML</td>
<td>54%</td>
<td>0</td>
<td>1</td>
<td>0.54</td>
<td>0.0015</td>
<td>0.998</td>
<td>0.54</td>
</tr>
<tr>
<td>Infusion reactions</td>
<td>3%</td>
<td>0</td>
<td>1</td>
<td>0.03</td>
<td>0.24</td>
<td>0.764</td>
<td>0.02</td>
</tr>
<tr>
<td>injection reactions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>0.59</strong></td>
<td></td>
<td></td>
<td><strong>0.62</strong></td>
</tr>
</tbody>
</table>
Drill down to the values and the weights

*Incremental benefit-risk of natalizumab – placebo*

This shows which outcomes are contributing most to the total benefit-risk.

Even though the weight given to PML is large, the incidence is small, leading to a small contribution to the benefit-risk.
Waterfall plot

*Incremental benefit-risk of natalizumab – placebo*

- The length of each bar gives the contribution to the overall BR.
- End of the last bar gives the overall benefit-risk.
  - Denominated in the BR of one EDSS progression
- Green = positive BR.
- Red = negative BR.
- The contribution to the overall BR of PML is very small.
Sensitivity analysis on the weights

Incremental benefit-risk of natalizumab – placebo

- The weights are shown under each bar.
  - The base case weight is shown in the middle, with a +/- 20% range given at the ends.
- The weights are changed one at a time.
- The most important weight is the one given to relapses.
Current vs. future benefit-risk communication
From a narrative to a structured framework

“Traditional” benefit-risk communication

- Narrative describing benefits and risks.
- Lacking explicit identification of key benefit and key risk outcomes.
- Limited systematic comparison of active drug vs. comparators for all key benefits and key risks.
- No structured, quantitative summary of all key benefit and key risk outcomes.

Structured benefit-risk leads to communication that is transparent and defensible

- Which key benefits and key risks were considered and why.
- Which comparators were chosen.
- The magnitude of benefit and risk effects.
- Presentation in a graphical/tabular summary together with concise text.
- Written in such a way as to meet the Health Authority reviewer needs and expectations.
Acknowledgements and further details: Special issue of the Biometrical Journal

Biometrical Journal 00 (2015) 0, 1–20 DOI: 10.1002/bimj.201300248

A case study using the PrOACT-URL and BRAT frameworks for structured benefit risk assessment

Richard Nixon*,**,1, Christoph Dierig2, Shahrul Mt-Isa3, Isabelle Stöckert2, Thaison Tong4, Silvia Kuhls2, Gemma Hodgson5, John Pears6, Ed Waddingham3, Kimberley Hockley3, and Andrew Thomson7

1 Novartis Pharma AG, Statistical Methodology and Consulting, CH-4002 Basel, Switzerland
2 Bayer Pharma AG, Global Development, D-42096 Wuppertal, Germany
3 School of Public Health, Imperial College London, St. Mary’s Campus, Norfolk Place, London W2 1PG, UK
4 Health Economics and Decision Science (HEDS), School of Health and Related Research (ScHARR), University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK
5 Qi Statistics, Penhales House, Ruscombe, Berkshire RG10 9JN, UK
6 Astra Zeneca, R&D Global Regulatory Affairs, Mereside, Alderley Park, Macclesfield, Cheshire SK10 4TG, UK
7 Vigilance & Risk Management of Medicines Division, Medicine and Healthcare Products Regulatory Agency, 151 Buckingham Palace Road, London SW1W 9SZ
Shahrul Mt-Isa, PhD

TAXONOMY OF BENEFIT-RISK ASSESSMENT METHODOLOGIES, AND BENEFIT-RISK METRICS
Which benefit-risk methodology?
Methodologies available

**Benefit-risk assessment framework**

- **Descriptive framework**
  - PROACT-URL
  - ASF
  - BRAT
  - FDA BRF
  - CMR-CASS
  - COBRA
  - SABRE
  - UMBRA
  - OMERACT 3x3

- **Quantitative framework**
  - BLRA
  - NCB
  - Decision tree
  - MDP
  - MCDA
  - SMAA
  - SBRAM
  - CUI
  - DI

**Metric indices for B-R assessment**

- **Threshold indices**
  - NNT
  - NNH
  - AE-NNT
  - RV-NNH
  - Impact numbers
    - MCE
    - RV-MCE
    - MAR
    - NEAR

- **Health indices**
  - QALY
  - DALY
  - HALE
  - Q-TWiST

- **Trade-off indices**
  - UT-NNT
  - INHB
  - BRR
  - GBR
  - Principle of 3
  - TURBO
  - Beckmann

**Estimation techniques**

- DAGs
- PSM
- CPM
- ITC
- MTC
- CDS

**Utility survey techniques**

- SPM
- CV
- CA
- DCE
- AHP
- Swing-weighting
- MACBETH

---

Metric indices

- To quantitatively describe and communicate benefit-risk assessment results:
  1. Number Needed to Treat / Harm (NNT/H)
  2. Benefit-Risk Ratios (BRR)
  3. Incremental Net Health Benefit (INHB)
  4. Impact numbers
  5. QALY
  6. Q-TWiST
Benefit-risk ratio (BRR)

- Benefit divided by risk
- Benefit is expressed as multiples of risk
- BRR is a simple idea but can be powerful
- In practice, equilibrium in most cases is not 1
  - Region of equivalence must be established *a priori*
  - Trastuzumab example

\[
\frac{\text{Benefit}}{\text{Risk}} = \frac{\text{NNT}}{\text{NNH}} = \frac{12.3}{39.8} = 0.3 (< 1)
\]
Incremental net health benefit (INHB)

- Specifically difference between QALY gained (benefit) and QALY lost (risk)
  - QALY is the quality adjusted life years based on time spent in certain health state e.g. using EQ5D index
  - Q-TWiST proposed health states for cancer therapy
- More generally, not using health index

\[
INB = (\text{incremental benefit}) - (\text{incremental risk})
= (B_1 - B_0) - (R_1 - R_0)
\]
Incremental net health benefit (INHB)

- In the trastuzumab example:
  \[ \text{INB} = (B_1 - B_0) - (R_1 - R_0) \]
  \[ = (0.861 - 0.780) - (0.0304 - 0.0053) \]
  \[ = 0.0559 \]

- So in this case, the incremental net benefit is 0.0559 in favour of trastuzumab
Impact numbers

- Extend NNT concept to public health perspective
  - Uses background data from the intended population
- “Population Impact Measures (PIM)”
  - Population attributable risk (PAR)
  - Exposure impact number (EIN) \( \equiv \) NNT
  - Population impact number of eliminating a risk factor over time \( t \) (PIN-ER-\( t \))
  - Number of events prevented in the population (NEPP)
- Descriptive measure

Impact numbers: trastuzumab example

- Say we want to know, how many event free survivals (EFS) over one year in 1000 women with breast cancer. 50% of whom already receiving trastuzumab, and we would like to increase the uptake to 75% in the population.
  - attributed to receiving trastuzumab
  - will be prevented by receiving trastuzumab under the new regime
  - Assume baseline EFS rate is 0.780 (rate in control group in e.g.)
### Impact numbers: trastuzumab example

<table>
<thead>
<tr>
<th>PIM</th>
<th>Calculation</th>
<th>Interpretation</th>
</tr>
</thead>
</table>
| PAR   | \[
\frac{0.5 \times 0.104}{1 + (0.5 \times 0.104)} = 0.049
\]                          | 5% EFS are due to trastuzumab in the general population                           |
| PIN-ER-t | n \times r_u \times \text{PAR} = 1000 \times 0.780 \times 0.049 = 38.6       | 39 women of the 1000                                                            |
| EIN   | \[
\frac{1}{0.861 - 0.780} = 12.3
\]                                     | 13 women had to take trastuzumab to see one EFS                                  |
| NEPP  | n \times P_e \times r_u \times (RR - 1) = 1000 \times (0.75 - 0.5) \times 0.780 \times 0.104 = 20.3 | 20 extra EFS when increase intake from 50% to 75%                              |
Remarks

- Recommendations for further testing are toolkit to aid methodology selection
  - Complexity and purpose
- Benefit-risk assessment methodologies are NOT tools that can make choices
- Using metric indices alone does not guarantee structured, transparent and/or robust assessment
- Sufficient for simple decision problems, or as quick initial descriptions
- There is a trade-off between being too simplistic and just being too incomprehensible
FROM QUALITATIVE TO QUANTITATIVE
BENEFIT-RISK DECISION-MAKING:
CONCEPTS AND METHODS

Lawrence Phillips, PhD
By the end of this presentation, you will...

- see how efficacy and safety data are transformed into benefits and risks
- know the distinctions between qualitative, semi-quantitative and fully quantitative B-R approaches
- appreciate the role of judgement in each approach
- understand how a fully quantitative approach can integrate data and clinical judgement
- recognise how disagreements amongst experts can be synthesised into shared understanding with decision conferencing
- see how frameworks and approaches can help assessors develop insight about a drug’s benefit-risk
Efficacy & Safety ⇒ Benefits & Risks

- **Efficacy & Safety Data**
- **Favourable & Unfavourable Effects**
- **Clinical Relevance of the Effects**
- **Benefits & Risks**

Regulators & medical experts

Judgement required

Physicians & patients
B-R Assessment

- Qualitative
- Partially Quantitative
- Fully Quantitative
No quantitative modelling is used by any regulator anywhere to deal with the massive amount of data—10GB more or less!
Pharma-BRAT framework

Can be applied at any stage of drug development, approval and post-approval.


Missing: Clinical relevance of the metrics and uncertainty of the effects
PrOACT-URL framework

MCDA (Multi-Criteria Decision Analysis)

- An extension of decision theory that covers any decision with multiple objectives.
- A methodology for appraising options on individual, often conflicting criteria, and combining them into one overall appraisal.

Decision Conferencing

- One or more workshops to solve a ‘hot’ problem
- Attended by key players representing diversity of perspectives on the issues
- Facilitated by an impartial specialist in group processes & decision analysis
- Using a requisite (just-good-enough) MCDA model created on-the-spot to provide structure to thinking

Efalizumab (Raptiva) case study

- Drug approved in 2004 for chronic plaque psoriasis
- Emerging safety issues signalled CHMP to give opinion in Jan 2009 on benefit-risk
- Maintain, vary, suspend or withdraw Marketing Autorisation? It was suspended
- PROTECT Task Force developed quantitative model from regulator’s 2009 perspective

Model source for this project: Hiview3, originally developed at the London School of Economics, now available from Catalyze Ltd, www.catalyze.co.uk
Efficacy & Safety ⇒ Benefits & Risks

- Efficacy & Safety Data
- Favourable & Unfavourable Effects
- Clinical Relevance of the Effects
- Benefits & Risks

Regulators & medical experts

Judgement required

Clinicians & patients
Choose favourable & unfavourable effects

- Select only effects that are relevant to the B-R balance.
- Include patients’ views.
- Agree definitions of all effects with key players.
## Summarise information as an Effects Table

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Fixed Upper</th>
<th>Fixed Lower</th>
<th>Units</th>
<th>Raptiva</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Favourable Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PASI75</td>
<td>Percentage of patients achieving 75% reduction in baseline PASI¹ at week 12.</td>
<td>60.0</td>
<td>0.0</td>
<td>%</td>
<td>29.5</td>
<td>2.7</td>
</tr>
<tr>
<td>PGA</td>
<td>Percentage of patients achieving Physician's Global Assessment² clear/almost clear at week 12.</td>
<td>40.0</td>
<td>0.0</td>
<td>%</td>
<td>295</td>
<td>5.1</td>
</tr>
<tr>
<td>OLS</td>
<td>Percentage of patients with Overall Lesion Severity rating of minimal or clear at FT (day 84).</td>
<td>40.0</td>
<td>0.0</td>
<td>%</td>
<td>32.1</td>
<td>2.9</td>
</tr>
<tr>
<td>DLQI</td>
<td>Dermatology Life Quality Index³. Mean percentage of patients showing an improvement.</td>
<td>10.0</td>
<td>0.0</td>
<td>Change score</td>
<td>5.8</td>
<td>2.1</td>
</tr>
<tr>
<td><strong>Unfavourable Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe infections</td>
<td>Proportion of patients experiencing infections serious enough to require hospitalisation.</td>
<td>3.00</td>
<td>0.00</td>
<td>%/100ptyrs</td>
<td>2.83</td>
<td>1.4</td>
</tr>
<tr>
<td>Severe Thrombocytopenia</td>
<td>Number of cases exhibiting severe (grade 3 and above) thrombocytopenia⁴.</td>
<td>10</td>
<td>0</td>
<td>number</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Interstitial Lung Disease</td>
<td>Number of cases of interstitial lung disease.</td>
<td>20</td>
<td>0</td>
<td>number</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>Haemolytic anemia</td>
<td>Number of cases of haemolytic anemia.</td>
<td>25</td>
<td>0</td>
<td>number</td>
<td>24</td>
<td>0</td>
</tr>
<tr>
<td>PML</td>
<td>Number of cases of progressive multifocal leukoencephalopathy.</td>
<td>5</td>
<td>0</td>
<td>number</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Aseptic Meningitis</td>
<td>Number of cases of aseptic meningitis.</td>
<td>30</td>
<td>0</td>
<td>number</td>
<td>29</td>
<td>0</td>
</tr>
</tbody>
</table>

¹PASI is a measure of the average redness, thickness and scaliness of the lesions (each graded on a 0-4 scale), weighted by the body region and the area affected. PASI range is from 0 to 72.

²PGA is a seven point scale with 7 being clear, 6 almost clear, 5 mild, 4 mild to moderate, 3 moderate, 2 moderately severe and 1 severe psoriasis.

³DLQI is a 10-item quality of life index scored by the patient on a four point scale.

⁴As shown in laboratory test results that indicate a decrease in number of platelets in a blood specimen.
Efficacy & Safety ⇒ Benefits & Risks

- Efficacy & Safety Data
- Favourable & Unfavourable Effects
- Clinical Relevance of the Effects
- Benefits & Risks

Judgement required

Regulators & medical experts

Physicians & patients
Scoring clinical relevance of data

Linear conversions of data to preference values

FE: PASI 75

Larger percentages achieving PASI 75 are preferred

UFE: Haemolytic anaemia

Smaller numbers of cases are preferred
Scoring clinical relevance of data: PML

Non-linear conversion to clinical preference values

The 0 – 3 difference in number of PML cases is increased in preference value, representing its clinical relevance.
Weighting clinical relevance of effects

- Swing-weight favourable effects
- Swing-weight unfavourable effects
- Swing-weight most favourable against most unfavourable

"How big is the difference, and how much do you care about it?"
Explore results: benefit-risk differences

Overall, clinical value of Raptiva is greater than the placebo.

Just three favourable effects & one unfavourable effect account for this difference in clinical value.
Consider only PASI75 & PML

```
<table>
<thead>
<tr>
<th>Model Order</th>
<th>Cum Wt</th>
<th>Diff</th>
<th>Wtd Diff</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physicians' ratings</td>
<td>PGA</td>
<td>22.4</td>
<td>61</td>
<td>13.7</td>
</tr>
<tr>
<td>Physicians' ratings</td>
<td>PASI75</td>
<td>28.0</td>
<td>45</td>
<td>12.5</td>
</tr>
<tr>
<td>Patients' ratings</td>
<td>DLQI</td>
<td>20.4</td>
<td>37</td>
<td>7.6</td>
</tr>
<tr>
<td>Physicians' ratings</td>
<td>OLS</td>
<td>7.0</td>
<td>73</td>
<td>5.1</td>
</tr>
<tr>
<td>Observational data</td>
<td>ILDs</td>
<td>1.3</td>
<td>-90</td>
<td>-1.2</td>
</tr>
<tr>
<td>Observational data</td>
<td>Aseptic Meningitis</td>
<td>1.3</td>
<td>-97</td>
<td>-1.3</td>
</tr>
<tr>
<td>SAEs</td>
<td>Serious Infections</td>
<td>2.8</td>
<td>-48</td>
<td>-1.4</td>
</tr>
<tr>
<td>Observational data</td>
<td>Haemolytic anemia</td>
<td>1.6</td>
<td>-96</td>
<td>-1.5</td>
</tr>
<tr>
<td>SAEs</td>
<td>Svre Thrombocytopeni</td>
<td>2.3</td>
<td>-90</td>
<td>-2.0</td>
</tr>
<tr>
<td>Observational data</td>
<td>PML</td>
<td>12.9</td>
<td>-95</td>
<td>-12.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100.0</td>
<td></td>
<td>19.2</td>
</tr>
</tbody>
</table>
```
Sensitivity Analysis on PML

12.9, current weight

19.2, overall difference: drug minus placebo

Placebo preferred to efalizumab

Total weight on PML
Double the weight on PML
Benefits and risks nearly balance

<table>
<thead>
<tr>
<th>Model Order</th>
<th>Cum Wt</th>
<th>Diff</th>
<th>Wtd Diff</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physicians' ratings PGA</td>
<td>17.7</td>
<td>61</td>
<td>10.8</td>
<td>10.8</td>
</tr>
<tr>
<td>Physicians' ratings PASI75</td>
<td>22.1</td>
<td>45</td>
<td>9.9</td>
<td>20.6</td>
</tr>
<tr>
<td>Patients' ratings DLQI</td>
<td>17.7</td>
<td>37</td>
<td>6.5</td>
<td>27.2</td>
</tr>
<tr>
<td>Physicians' ratings OLS</td>
<td>5.5</td>
<td>73</td>
<td>4.0</td>
<td>31.2</td>
</tr>
<tr>
<td>Observational data ILDs</td>
<td>2.2</td>
<td>-90</td>
<td>-2.0</td>
<td>29.2</td>
</tr>
<tr>
<td>SAEs Serious Infections</td>
<td>4.4</td>
<td>-48</td>
<td>-2.1</td>
<td>27.1</td>
</tr>
<tr>
<td>Observational data Aseptic Meningitis</td>
<td>2.2</td>
<td>-97</td>
<td>-2.1</td>
<td>25.0</td>
</tr>
<tr>
<td>Observational data Haemolytic anemia</td>
<td>2.6</td>
<td>-96</td>
<td>-2.5</td>
<td>22.4</td>
</tr>
<tr>
<td>SAEs Svre Thrombocytopeni</td>
<td>3.5</td>
<td>-90</td>
<td>-3.2</td>
<td>19.2</td>
</tr>
<tr>
<td>Observational data PML</td>
<td>22.1</td>
<td>-95</td>
<td>-21.0</td>
<td>-1.7</td>
</tr>
<tr>
<td></td>
<td>100.0</td>
<td></td>
<td></td>
<td>-1.7</td>
</tr>
</tbody>
</table>
Our conclusions

• Benefit-risk balance is favourable for efalizumab
• Conflict with 2009 CHMP decision? Not necessarily
  – Hindsight bias
  – We used only publically-available reports of effects
  – Public health interpretation of data: EPAR reports that 27% of patients achieved PASI75—a ‘modest effect’
• Experts and assessors frequently disagree
• Quantitative modelling within a decision conference provides ‘intellectual technology’ that can enable assessors to achieve shared understanding
Summary

• Judgement is required about safety and efficacy data to assess benefit-risk.

  1) Which favourable and unfavourable effects?
  2) How clinically relevant are the data and the effects?

• Application of frameworks such as BRAT or PrOACT-URL are useful ‘best-practice’ approaches to B-R.

• Quantification, partial or full, can enhance understanding, develop insight about the benefit-risk balance and facilitate communication about decisions.
Douwe Postmus, PhD

AGGREGATED DATA DRUG INFORMATION SYSTEM (ADDIS)
AN EVIDENCE-BASED DECISION SUPPORT SYSTEM FOR THE
BENEFIT-RISK ASSESSMENT OF MEDICAL PRODUCTS
The 3 pillars of structured decision making

- Well-defined and transparent process
  - PrOACT-URL (EMA benefit-risk methodology project)
- Guidance on how to conduct the various steps in this process
  - IMI PROTECT benefit-risk group recommendations report
  - IMI PROTECT website and training materials
- Supporting software
  - ADDIS
ADDIS – a brief history

• The development of ADDIS started in 2009 as part of work package 3.2 of the Escher project
• This has resulted in the development of ADDIS 1
• ADDIS 2 is a web-based redevelopment of the previous prototype desktop application
• ADDIS 2 is currently still under heavy development but the software in now becoming useable as an analytical tool
• Both ADDIS 1 and 2 are open source and freely accessible from our website www.drugis.org
ADDIS 2: functional perspective

Quantitative methods

- Define the decision problem
- Synthesize & summarize data
- Assess benefit-risk balance

Workflow

Visualisations

- Value tree
- Effects table
- SMAA descriptive indices

(network) meta-analysis

MCDA/SMAA

Disease progression modelling
ADDIS 2: technical perspective
MCDA WEB INTERFACE
Illustrative case study

- We consider the problem of assessing the benefit-risk balance of regorafenib using the data available at the time of the initial marketing authorization application of this product.
- All data used for this assessment were directly taken from the EPAR of this product (EPAR EMA/CHMP/403683/2013 available from the EMA website).
- The value judgments provided throughout this example are hypothetical and do not reflect the opinion of the CHMP.
Overview of the decision problem

Overview

Problem description
The indication sought for regorafenib was the treatment of adult patients with metastatic colorectal cancer (mCRC) who have been previously treated with, or are not considered candidates for, available therapies. Data in support of this indication were mainly obtained from one pivotal randomized, double-blind, placebo-controlled trial comparing regorafenib plus best supportive care (BSC) versus placebo plus BSC in patients with mCRC who have progressed after standard therapy. A detailed description of the design of this phase III study can be found in the EPAR of this product (EPAR EMA/CHMP/403683/2013 available from the EMA website).

Alternatives
- Placebo
- Regorafenib

Value Tree
- Benefit-risk balance
  - Favourable effects
    - Overall survival
  - Unfavourable effects
    - Hand-foot skin reaction
    - Hypertension
    - Haemorrhage
    - Hyperbilirubinaemia
## Effects table

### Regorafenib - Initial marketing authorization application

#### Effects table

Show alternatives

- Placebo
- Regorafenib

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Description</th>
<th>Units</th>
<th>Placebo</th>
<th>Regorafenib</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Favourable effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall survival</td>
<td>Median overall survival time</td>
<td>Months</td>
<td>4.96</td>
<td>6.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.96 – 5.06</td>
<td>6.44 – 6.44</td>
</tr>
<tr>
<td><strong>Unfavourable effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hand-foot skin reaction</td>
<td>Incidence of grade 3 events</td>
<td>%</td>
<td>0.4</td>
<td>16.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.4 – 6.4</td>
<td>16.6 – 16.6</td>
</tr>
<tr>
<td>Hypertension</td>
<td>Incidence of grade 3 events</td>
<td>%</td>
<td>0.8</td>
<td>7.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.8 – 6.8</td>
<td>7.6 – 7.6</td>
</tr>
<tr>
<td>Haemorrhage</td>
<td>Incidence of grade 3-5 events</td>
<td>%</td>
<td>0.8</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.8 – 6.8</td>
<td>2 – 2</td>
</tr>
<tr>
<td>Hyperbilirubinaemia</td>
<td>Incidence</td>
<td>%</td>
<td>9.5</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9.5 – 9.5</td>
<td>20 – 20</td>
</tr>
</tbody>
</table>
Preference elicitation: scale ranges

Regorafenib - Initial marketing authorization application

Preferences
Default

Scale Ranges
Partial Value Functions
Trade-off Order
Trade-off Ratios

Results

Scale Ranges

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Theoretical Range</th>
<th>Observed Range</th>
<th>Configured Range</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall survival</td>
<td>$0, \infty$</td>
<td>4.96, 6.44</td>
<td>4.6</td>
<td>Months</td>
</tr>
<tr>
<td>Hand-foot skin reaction</td>
<td>0, 100</td>
<td>0.4, 16.6</td>
<td>0.20</td>
<td>%</td>
</tr>
<tr>
<td>Hypertension</td>
<td>0, 100</td>
<td>0.8, 7.6</td>
<td>0.10</td>
<td>%</td>
</tr>
<tr>
<td>Haemorrhage</td>
<td>0, 100</td>
<td>0.8, 2</td>
<td>0.5</td>
<td>%</td>
</tr>
<tr>
<td>Hyperbilirubinaemia</td>
<td>0, 100</td>
<td>5.5, 20</td>
<td>5.20</td>
<td>%</td>
</tr>
</tbody>
</table>

Define Scale Ranges
Preference elicitation: partial value functions

Define Partial Value Function for: Overall survival

What is the value of x such that an improvement in Overall survival from 4 (Months) to x is equivalent to an improvement from x to 8 (Months)?

Adjust the slider:

So that the following statement is true:
The improvement from 4 to 5.4 is equivalent to the improvement from 5.4 to 8.
Preference elicitation: ordinal trade-offs

Regorafenib - Initial marketing authorization application

Ordinal SWING weighting (1/4)

Given the following situation:
- Hand-foot skin reaction = 20
- Haemorrhage = 5
- Hypertension = 10
- Hyperbilirubinaemia = 20
- Overall survival = 4

Which of these improvements is most desired:
- Hand-foot skin reaction → 0
- Haemorrhage → 0
- Hypertension → 0
- Hyperbilirubinaemia → 5
- Overall survival → 8
Preference elicitation: ordinal trade-offs

Ordinal SWING weighting (2/4)

Given the following situation:
Hand-foot skin reaction = 20  Haemorrhage = 5  Hypertension = 10
Hyperbilirubinaemia = 20  Overall survival = 8

Which of these improvements is most desired:
- Hand-foot skin reaction → 0
- Haemorrhage → 0
- Hypertension → 0
- Hyperbilirubinaemia → 5
Preference elicitation: ordinal trade-offs

Regorafenib - Initial marketing authorization application

Ordinal SWING weighting (DONE)

You have given the following trade-offs:

\[ w_1 : \text{Overall survival} (4 \to 8) \]
\[ w_2 : \text{Hand-foot skin reaction} (20 \to 0) \]
\[ w_3 : \text{Hypertension} (10 \to 0) \]
\[ w_4 : \text{Haemorrhage} (5 \to 0) \]
\[ w_5 : \text{Hyperbilirubinaemia} (20 \to 5) \]

\[ w_1 \geq w_4 \]
\[ w_4 \geq w_2 \]
\[ w_2 \geq w_5 \]
\[ w_5 \geq w_3 \]
Results based on ordinal trade-offs

Central Weights

<table>
<thead>
<tr>
<th>Alternative</th>
<th>Confidence</th>
<th>Hand-foot skin reaction</th>
<th>Haemorrhage</th>
<th>Hypertension</th>
<th>Hyperbilirubinaemia</th>
<th>Overall survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo</td>
<td>1</td>
<td>0.17248</td>
<td>0.26829</td>
<td>0.045044</td>
<td>0.10019</td>
<td>0.414</td>
</tr>
<tr>
<td>Regorafenib</td>
<td>1</td>
<td>0.086353</td>
<td>0.20483</td>
<td>0.018686</td>
<td>0.04477</td>
<td>0.64536</td>
</tr>
</tbody>
</table>
Preference elicitation: exact trade-offs

Exact SWING weighting (1/4)

Determining the relative importance of:
Overall survival (4.000 → 8.000)
Haemorrhage (5.000 → 0.000)

Given the following situation:
Overall survival = 4.000, Haemorrhage = 0.000

Adjust the slider:

So that the following alternative is equally desirable:
Overall survival = 5 Haemorrhage = 5.000
Preference elicitation: exact trade-offs

Exact SWING weighting (2/4)

Determining the relative importance of:
Haemorrhage (5.000 → 0.000)
Hand-foot skin reaction (20.000 → 0.000)

Given the following situation:
Haemorrhage = 5.000, Hand-foot skin reaction = 0.000
Adjust the slider:

So that the following alternative is equally desirable:
Haemorrhage = 3 Hand-foot skin reaction = 20.000
Preference elicitation: exact trade-offs

Exact SWING weighting (3/4)

Determining the relative importance of:

Hand-foot skin reaction (20.000 → 0.000)
Hyperbilirubinaemia (20.000 → 5.000)

Given the following situation:
Hand-foot skin reaction = 20.000, Hyperbilirubinaemia = 5.000

Adjust the slider:

So that the following alternative is equally desirable:
Hand-foot skin reaction = 8, Hyperbilirubinaemia = 20.000
Preference elicitation: exact trade-offs

Exact SWING weighting (4/4)

Determining the relative importance of:

Hyperbilirubinaemia (20.000 → 5.000)
Hypertension (10.000 → 0.000)

Given the following situation:
Hyperbilirubinaemia = 20.000, Hypertension = 0.000

Adjust the slider:

So that the following alternative is equally desirable:
Hyperbilirubinaemia = 8 Hypertension = 10.000
Results based on exact trade-offs
Concluding remarks

- Developing quantitative methods that are both theoretically sound and easy to use by decision makers has proven to be far from straightforward.
- Our ultimate aim will be to arrive at methodologies that allow decision makers to simultaneously explore:
  - Imprecision in the preference statements (i.e. shape of the partial value functions, criteria weights)
  - Uncertainty in the effect size estimates
  - Uncertainty in the long-term clinical consequences
- We have started to develop a flexible set of tools to address all these aspects (www.drugis.org)
Christine Hallgreen, PhD

VISUALISING BENEFITS AND RISKS: CONCEPTS AND IDEAS
Many research on visualisations
Lack of use in formal B-R assessment

Graphics and other formats

Verbal Labels Can Triple Perceived Risk in Clinical Trials

The purpose of this study was to assess whether the use of verbal descriptors, such as “common” and “rare,” affects people’s perceptions of the risks involved in clinical trials as well as their likelihood of entering into the trial. Participants were required to imagine that they had a serious skin condition and being asked if they would take part in a clinical trial for a new drug. They were given just the verbal descriptors were significantly less satisfied with the information, perceived risk to be higher (by a factor of three) and benefit to health to be lower, and indicated that they would be significantly less likely to enter the trial. We recommend that patients are informed...

http://intl-dij.sagepub.com/content/40/3/249.refs

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Treatment A</th>
<th>Treatment B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physician's view on HDL cholesterol levels</td>
<td>Mild improvement</td>
<td>No change</td>
</tr>
<tr>
<td>Number of people who experience a 10% weight loss</td>
<td>10 out of 1000</td>
<td>450 out of 1000</td>
</tr>
<tr>
<td>Number of people who experience psychiatric conditions</td>
<td>100 out of 1000</td>
<td>1 out of 1000</td>
</tr>
<tr>
<td>Number of people who experience cardiovascular conditions</td>
<td>1 out of 1000</td>
<td>100 out of 1000</td>
</tr>
<tr>
<td>Number of people who experience gastrointestinal conditions</td>
<td>1 out of 1000</td>
<td>None</td>
</tr>
</tbody>
</table>
Tree diagram – a value tree

- Benefit-risk balance
  - Favourable effects
    - Glycaemic efficacy
    - Microvascular events
  - Unfavourable effects
    - Congestive heart failure (CHF)
    - Cardiovascular (CV) death
    - Non-CV death
    - Myocardial infarction (MI)
    - Stroke
    - MACE
    - Other
      - Weight gain
      - Macular oedema
      - Bone fractures
      - Bladder cancer
## Effects table

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Fixed Upper</th>
<th>Fixed Lower</th>
<th>Unit</th>
<th>Rosi + adjunct</th>
<th>Adjunct only</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Favourable effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycaemic efficacy</td>
<td>(A surrogate marker of the quality of glucose regulation.) Mean change from baseline in the proportion of Hb in which A1c is greater than 48 mmol/ml.</td>
<td>5.00</td>
<td>-5.00</td>
<td>%</td>
<td>-1.18</td>
<td>0.06</td>
</tr>
<tr>
<td>Micro-vascular events</td>
<td>Incidence of new cases of microvascular events compared to baseline (Retinopathy requiring photocoagulation, vitreous haemorrhage, &amp; fatal or non-fatal renal failure.)</td>
<td>20.00</td>
<td>0.00</td>
<td>%</td>
<td>2.70</td>
<td>3.50</td>
</tr>
<tr>
<td>CHF</td>
<td>Proportion of patients experiencing congestive heart failure during the study period.</td>
<td>4.00</td>
<td>0.00</td>
<td>%</td>
<td>3.69</td>
<td>1.89</td>
</tr>
<tr>
<td>CV death</td>
<td>The proportion of patients who died from any cardiovascular event including stroke.</td>
<td>4.00</td>
<td>0.00</td>
<td>%</td>
<td>2.70</td>
<td>3.19</td>
</tr>
<tr>
<td>Non-CV death</td>
<td>The proportion of patients who died from any non-cardiovascular event including stroke.</td>
<td>4.00</td>
<td>0.00</td>
<td>%</td>
<td>2.97</td>
<td>3.86</td>
</tr>
<tr>
<td>MI</td>
<td>Proportion of patients who experience a non-fatal heart attack.</td>
<td>5.00</td>
<td>0.00</td>
<td>%</td>
<td>3.33</td>
<td>3.01</td>
</tr>
<tr>
<td>Stroke</td>
<td>Proportion of patients who experience a non-fatal ischemia stroke.</td>
<td>5.00</td>
<td>0.00</td>
<td>%</td>
<td>1.94</td>
<td>2.83</td>
</tr>
<tr>
<td>Weight gain</td>
<td>Mean change from baseline in weight gain at 1 yr.</td>
<td>10.00</td>
<td>-5.00</td>
<td>Kg</td>
<td>3.80</td>
<td>0</td>
</tr>
<tr>
<td><strong>Unfavourable Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MACE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI</td>
<td>Proportion of patients who experience a non-fatal heart attack.</td>
<td>5.00</td>
<td>0.00</td>
<td>%</td>
<td>3.33</td>
<td>3.01</td>
</tr>
<tr>
<td>Stroke</td>
<td>Proportion of patients who experience a non-fatal ischemia stroke.</td>
<td>5.00</td>
<td>0.00</td>
<td>%</td>
<td>1.94</td>
<td>2.83</td>
</tr>
<tr>
<td>Weight gain</td>
<td>Mean change from baseline in weight gain at 1 yr.</td>
<td>10.00</td>
<td>-5.00</td>
<td>Kg</td>
<td>3.80</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macular oedema</td>
<td>Proportion of patients who experience macular oedema.</td>
<td>1.00</td>
<td>0.00</td>
<td>%</td>
<td>1.27</td>
<td>0.23</td>
</tr>
<tr>
<td>Bone fractures</td>
<td>Proportion of patients experiencing bone fractures.</td>
<td>3</td>
<td>0</td>
<td>%</td>
<td>8.33</td>
<td>5.3</td>
</tr>
<tr>
<td>Bladder cancer</td>
<td>Proportion of patients contracting bladder cancer.</td>
<td>1.00</td>
<td>0.00</td>
<td>%</td>
<td>0.27</td>
<td>0.22</td>
</tr>
</tbody>
</table>
Patients who will die from any-cause over a course of one year whether they take warfarin or not

Patients who will be saved from dying by any-cause over a course of 1 year by taking warfarin

Patients who will not die from any-cause over a course of one year whether they take warfarin or not
Stacked bar graph

More red, more safe

More green, more benefit
Interactive visual display

Drag sliders to assign weights on criteria

- Non-CV death
- CV death
- Stroke
- Microvascular events
- Bladder cancer
- MI
- Glycaemic efficacy
- Weight gain
- Bone fractures
- CHF
- Macular oedema

Favours rosiglitazone  Favours adjunct alone

Total

Benefit  Safety
Tornado-diagram

- Relapse weight
- Infusion reactions/injection reactions weight
- Disability progression weight
- Convenience weight
- Transaminases elevation weight
- PML weight
- Flu-like reactions weight
- Hypersensitivity Reactions weight
- Seizures weight
- Reactivation of serious herpes viral infections weight
- Congenital abnormalities weight

**Legend:**
- Decreased weight (20%)
- Increased weight (20%)

**Benefit-risk (#relapses):**

- 6.4% decreased weight
- 0.6% increased weight
- 3.6% decreased weight
- 2.4% increased weight
- 4% decreased weight
- 6% increased weight
- 1.2% decreased weight
- 0.8% increased weight
- 13.2% decreased weight
- 8.8% increased weight
- 11% decreased weight
- 65% increased weight
- 43% decreased weight
- 54% increased weight
- 0.8% decreased weight
- 1.2% increased weight
- 1% decreased weight
- 0.8% increased weight
- 1% decreased weight
- 4% increased weight
- 6% decreased weight
- 4% increased weight
- 6% decreased weight
- 4% increased weight
- 6% decreased weight
- 5% increased weight
Remarks on visual representation

- No one visual type fits all
- Different visual types carry information differently
- Different user may prefer different visual representation – cannot always generalise
  - Visual type preference study shows preference towards tables and bar graphs
  - Understanding and/or preferences may still be affected by the actual information being displayed
- Visual representation formats should be tested with the intended audience before actual use
Gerry Downey, MSc CBA
Subhakanta Das, BSc

PROTECT RESOURCES FOR FURTHER LEARNING
Dissemination of results/recommendations arising from PROTECT*

- Publications & Presentations
- PROTECT Web Portal
- The ENCePP network
- Training Programmes (WP7)
- The EMEA Scientific Committees, Working Parties and regulatory activities
- Other possible means

http://protectbenefitrisk.eu/

* PROTECT Full Project Proposal / IMI Call #6 (20th January 2009)
Web design

• Responsive Web Design: “Reponses or addictiveness Quickly and Positively” to the users. It responds to users environment based on screen-size, platform and orientation.

• Offers: Smooth Navigation, Easy reading, Reduces scrolling and zooming, social media integration and excellent user experience – across a good vary of devices (from smartphones to desktops).
• Recommendations tab is organised into five broad stages common to all benefit-risk assessments
• Interactive version of final recommendations report (Hughes et al, Nov 2013)
### Classification of methodologies used in benefit-risk assessment

<table>
<thead>
<tr>
<th>Framework</th>
<th>Metric Indices</th>
<th>Estimation Techniques</th>
<th>Utility Survey Techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>PrOACT-URL</td>
<td>NNT, QALY, INHB</td>
<td>PSM</td>
<td>DCE</td>
</tr>
<tr>
<td>BRAT</td>
<td>NNH, Q-TWIST, BRR</td>
<td>ITC</td>
<td>CV</td>
</tr>
<tr>
<td>ASF</td>
<td>Impact Numbers, HALE, UT-NNT</td>
<td>MTC</td>
<td>CA</td>
</tr>
<tr>
<td>CMR-CASS</td>
<td>AE-NNT, DALY, GBR</td>
<td>CPM</td>
<td>SPM</td>
</tr>
<tr>
<td>COBRA</td>
<td>RV-NNH, Principle Of threes</td>
<td>DAGS</td>
<td></td>
</tr>
<tr>
<td>FDA BRF</td>
<td>NCB, MCE</td>
<td>TURBO</td>
<td>CDS</td>
</tr>
<tr>
<td>SABRE</td>
<td>RV-MCE, BECKMAN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UMBRA</td>
<td>CUI, MAR, NEAR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DI</td>
<td>NEAR</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
DCE (Discrete Choice Experiment)

1. Description
DCE (Discrete Choice Experiment) uses exactly the same principles as Conjoint Analysis (CA) with a more structured guideline to generating the hypothetical scenarios to be used in the elicitation process. [1][2][3][4] DCE can be regarded a framework for eliciting utilities from relevant stakeholders with roots in the random utility theory and a strong foundation in behavioural psychology. In DCE the most important characteristics of a situation are defined and labelled as attributes. Then, each attribute is assigned levels which can be cardinal, ordinal, or categorical. The attributes and levels are then systematically varied to explore all potential configurations of attributes. These are later reduced via fractional factorial designs, where the optimal design would be orthogonal. This results in hypothetical situations, which are then compiled into choice sets that contain two or more hypothetical scenarios. Stakeholders will select the most attractive scenario from the choice set, and it is assumed their selection has the highest utility out of the options provided. From this, it is possible to analyse the value each attribute via logistic regression.

- Taken together, they would be a sufficiently powerful toolbox for most benefit-risk assessments
- Interactive version of systematic review of methodologies (Mt-Isa et al, 2014)
Introduction

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 has been conducted in two stages. The first stage provided a level of evaluation as to the suitability of visuals presented in the application of benefit-risk approaches in PROTECT methodology review. However, external circumstances such as the intended audience, complexity of the benefit-risk problem, time in drug lifecycle, and other factors that are not related to the benefit-risk methodology may influence the type of visual representation to use. The second stage therefore explored and identified suitable visuals to communicate benefits and risks to different stakeholders in different situations. This included the use of visualisations in dynamic and interactive settings.
Seventeen recommendations for the application of visuals at key stages proposed

Interactive version of visual review (Mt-Isa et al, Part 1 & Part 2; 2013)
Each case study applied several methodologies and visual representations

Interactive summary of Case Study Reports:

- Efalizumab (Micaleff et al. Wave 1 & Suppl 1; Phillips et al. Suppl 2; 2013)
- Natalizumab (Nixon et al. Wave 1 and Wave 2, 2013)
- Rimonabant (Juhaeri et al. Wave 1; Mt-Isa et al. Suppl 1; Juhaeri et al. Wave 2, 2011/2012)
- Rosiglitazone (Philips et al. Wave 2, 2013)
- Telithromycin (Quartey et al. Wave 1, 2012)
- Warfarin (Hallgreen et al. Wave 2, 2013)

Introduction

This case study aims to investigate the difficulties that may be encountered when undertaking a benefit-risk assessment for an older medicinal product with well-established use. To assess the difficulties of undertaking a benefit-risk assessment for an older medicine, we applied the BRAT framework to a case study assessing the benefit-risk balance of warfarin for the treatment of atrial fibrillation. The framework ensured that the process was documented, and that the discussions were focused on outcomes relevant to the BR problem. One of the biggest challenges identified related to the large variety of data sources, a result of this was very broadly defined benefit-risk criteria, which can make it difficult to elicit preference values.
What is the Patient and Public Involvement project?

The Patient and Public Involvement project was a working group in PROTECT Benefit-Risk. We were created following a strong interest in patient and public involvement (PPI) from the PROTECT Benefit-Risk case study task forces. Our aim was to develop a toolbox for those who wish to involve patients and the public in medical benefit-risk decision making. Our technical report can be found on the IMI PROTECT website (http://www.imi-protect.eu/benefitsRep.shtml).

Our research focussed on three areas of involving patients and the public: (a) testing formal methods which can be used to value the benefits and risks of medicines, (b) testing out different visual images to see if they are understandable, trustworthy, and useful, and (c) understanding how to communicate the process and results of benefit-risk assessment.

We developed this section of the website to provide information to patients and the public and professionals who are interested and would like to learn more about the benefit-risk assessment of medicines. We would like to thank the following organisations for reviewing the content on this section of the PROTECT benefit-risk website: organisations name should be added here.

- A guide for patients and interested members of the public who are new to the benefit-risk assessment of medicines or would like to know more
- Example case-study on “visual communication of the benefits and risks of weight loss interventions”
About Us

Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium (PROTECT), a collaboration amongst private and public sector partners, is a project set up under the Innovative Medicines Initiative (IMI). Its goal is to strengthen the monitoring of the benefit-risk balance of medicines in Europe. This website is developed as part of the PROTECT Benefit-Risk Group who has advanced the understanding of both the integration, communication and visual representations of benefit and risk assessment methodologies.

PROTECT Benefit-Risk aims to provide practical recommendations for benefit-risk decision processes and supporting tools to various stakeholders, particularly the regulators. We advocate for increased transparency and robust decision making by making explicit and effectively communicating the methodologies, assumptions, and outcomes utilised in the assessment of benefit-risk balance in medicine. Our experience makes what we believe is a unique contribution that complements and builds on the efforts of other benefit-risk assessment initiatives.

Our Team

- Professor Deborah Ashby
- Dr Alain Micallef
- Dr Steve Hobbiger
- Dr Diana Hughes
- Dr Joanna Tzouliaki
- Dr Shahrul Mt-Issa
- Gerry Downey
- Dr Ian Hirsch
- Dr Richard Nixon
- Dr Kimberley Hockley
- Ed Waddingham
- Alesia Goginsky
Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approach</td>
<td>The system of methods and principles used in a particular discipline</td>
</tr>
<tr>
<td>Aseptic meningitis</td>
<td>A syndrome characterized by headache, neck stiffness, low grade fever, and CSF lymphocytic pleocytosis in the absence of an acute bacterial pathogen. Viral meningitis is the most frequent cause although mycoplasma and rickettsia infections; diagnostic or therapeutic procedures; neoplastic procedures; septic perimeningeal foci, and other conditions may result in this syndrome. (From Adams et al., Principles of Neurology, 6th ed, p745)</td>
</tr>
<tr>
<td>Aspect ratio</td>
<td>The ratio of the lengths of the two axes on a graph; a square graph has an aspect ratio of 1</td>
</tr>
<tr>
<td>Benefit</td>
<td>The positive results of a given treatment for an individual or a population (i.e., efficacy, convenience, or even quality of life)</td>
</tr>
<tr>
<td>Benefit-risk assessment</td>
<td>An evaluation of medical product either quantitatively or qualitatively taking both benefits and risks of the product into account</td>
</tr>
<tr>
<td>Benefit-risk model</td>
<td>A formal way to analyse benefit and risk consequences and their balances from a set of actions and to aid making choices amongst actions when risk aversion and preferences are specified</td>
</tr>
<tr>
<td>Bias</td>
<td>The systematic tendency of any factors associated with the design, conduct, analysis, and evaluation of the results of a benefit-risk assessment to make the estimate of a treatment effect deviate from its true value</td>
</tr>
</tbody>
</table>

- Links to all published reports from IMI PROTECT Benefit Risk.
- Complete Glossary, Abbreviations and References also provided.
Questions …

An online space has been created so that findings and recommendations can be explored interactively and will continue once PROTECT closes following this symposium.

Thank you from the IMI PROTECT Benefit-Risk Team (http://protectbenefitrisk.eu/)
COFFEE/TEA BREAK
20’
Ed Waddingham

PATIENTS AND PUBLIC INVOLVEMENT IN BENEFIT-RISK ASSESSMENT AND DECISION-MAKING: METHODS AND APPLICATIONS
Decision makers

- **Patients**
  - Make decisions for themselves

- **Healthcare providers**
  - Make decisions based on prescribing lists

- **NICE**
  - Makes decisions on cost-effectiveness

- **EMA/MHRA etc.**
  - Makes decisions on quality, safety, efficacy and benefit-risk balance to individuals and public health

- **Pharmaceutical companies**
  - Makes decisions on what to develop for which licenses to apply
Patient and public involvement

Patient and public:
Clinical trial participants, patients and potential patients, disabled people, parents and guardians, people who use health and/or social care services, carers, members of the public, and the organisations who represent the interests of these consumers.

Involvement:
An active partnership between stakeholders in the research process, rather than the use of people as ‘subjects’ of research. Public involvement in research is often defined as doing research ‘with’ or ‘by’ the public, rather than ‘to’, ‘about’ or ‘for’ them.
Varying degrees of involvement

**Consultation**  
Health professionals elicit the patient and public perspective to inform the decision making process

**Collaboration**  
Health professionals and patients and the public form an active partnership and jointly participate in decision making
At what stage can PPI occur

There is scope for patients and the public to be involved throughout the BR assessment process...

...we focus on explicit methods for eliciting & incorporating patient preferences into the analysis
Preference elicitation

• Well-known methods for preference elicitation:
  - MCDA swing-weighting (multi-criteria decision analysis)
  - MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique)
  - AHP (Analytic Hierarchy Process)
  - DCE (Discrete Choice Experiment)
**Simple weighting**

*Multi Criteria Decision Analysis (MCDA)*

For each outcome category

1. Rank outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infusion/injection reactions</td>
<td>1</td>
</tr>
<tr>
<td>Hypersensitivity reactions</td>
<td>2</td>
</tr>
<tr>
<td>Flu-like reactions</td>
<td>3</td>
</tr>
</tbody>
</table>

2. Relative importance

**Infusion/injection reactions**

**Hypersensitivity reactions**

**Flu-like reactions**

How much more important is it to avoid the top-ranked event compared to the others?
Repeat this process all the way up the value tree

The top ranked outcome in each category is carried up the tree

- Move bottom-up through the tree and compare the top-ranked outcomes from each category
- Finally, the top-ranked benefit is compared to the top-ranked risk
- The individual weights for each outcome can then be calculated
Repeat this process all the way up the value tree

The top ranked outcome in each category is carried up the tree

10% Benefit

Treatmen

100% Infection

100% Relapse

70% x 10% = 7%

100% x 10% = 10%

10% x 10% = 1%

12% x 100% x 100% = 12%

100% x 100% x 100% = 100%

10% x 100% = 10%

20% x 100% = 20%

10% x 100% = 10%

100% x 5% x 100% = 5%

40% x 5% x 100% = 2%

40% x 5% x 100% = 2%
MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique)

Qualitative assessment

- MACBETH is similar to MCDA, except that it provides a different way to get the weights
- **Step 1: Qualitatively** assess relative attractiveness of outcomes on **pairwise** basis
- **Step 2:** Check consistency of answers (eg cannot have A>B>C>A)
- **Step 3:** Compute initial guess at weights with optimization
- **Step 4:** Refine weights while maintaining consistency
## MACBETH

### Qualitative assessment

<table>
<thead>
<tr>
<th></th>
<th>PML</th>
<th>Abortion or congenit</th>
<th>Seizures</th>
<th>Reactivation of sero</th>
</tr>
</thead>
<tbody>
<tr>
<td>PML</td>
<td>no</td>
<td>extreme</td>
<td>extreme</td>
<td></td>
</tr>
<tr>
<td>Abortion or congenit</td>
<td>no</td>
<td>strong</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seizures</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reactivation of sero</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Consistent judgements

- PML: 100.00 (extreme)
- Abortion or congenit: 65.22 (strong)
- Seizures: 47.83 (moderate)
- Reactivation of sero: 34.78 (weak)

[all zero]
AHP (Analytic Hierarchy Process)

Qualitative assessment

- Based on qualitative pairwise comparisons (similar to MACBETH)
- No consistency check, but rather a score
- Qualitative responses are translated to a quantitative scale (integers from 1 to 9)
- Weight is calculated by finding the dominant eigenvector of the corresponding matrix, or by regression
Weighting individual events has its limits

- Examples so far involved trade-offs between individual events (eg 1 relapse vs 1 disability progression vs 1 case PML)
- This implies that events of a given type are all equal in value i.e. linear (partial) value functions

- It can be difficult to trade off events that are very different in importance (eg 1 infusion/injection reaction vs 1 case PML)
Swing weighting (1)

- Set best and worst possible figure for each outcome
- How much more attractive is it to move from worst to best for outcome A vs moving from worst to best for outcome B?
Swing weighting (2)

- Allows non-linear (partial) value functions (these can be elicited in the same way as weights)
- Helps to establish common value scale for events that are different in importance

Errors to watch out for:

- Not communicating swings clearly to participants
- Not accounting for swings correctly during benefit-risk assessment
Discrete Choice Experiments (DCEs)

In a DCE, participants are asked to consider a number of choice scenarios, e.g.:

**Car A**
- Manufacturer: Maserati
- Price: £££££
- Mileage: 0
- Fuel efficiency: Poor

**Car B**
- Manufacturer: Vauxhall
- Price: £££
- Mileage: 10,000
- Fuel efficiency: Good

Which car would you choose?
Natalizumab DCE
6 attributes, 2 levels each

BR Balance

Benefits
  - Relapse
  - Disability progression

Risks
  - PML
  - Serious allergic reactions
  - Mild allergic reactions
  - Flu-like reactions
**Natalizumab DCE: questionnaire (1)**

<table>
<thead>
<tr>
<th>Outcome (measured over 2 years)</th>
<th>Treatment A</th>
<th>Treatment B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of relapses per patient</td>
<td>2 relapses</td>
<td>1 to 2 relapses</td>
</tr>
<tr>
<td>Disability progression</td>
<td>250 patients out of 1000</td>
<td>100 patients out of 1000</td>
</tr>
<tr>
<td>PML</td>
<td>3 patients out of 1000</td>
<td>0 patients out of 1000</td>
</tr>
</tbody>
</table>
### Natalizumab DCE: questionnaire (2)

<table>
<thead>
<tr>
<th>Mild allergic reactions</th>
<th>0 patients out of 1000</th>
<th>500 patients out of 1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serious allergic reactions</td>
<td>0 patients out of 1000</td>
<td>0 patients out of 1000</td>
</tr>
<tr>
<td>Depression</td>
<td>200 patients out of 1000</td>
<td>100 patients out of 1000</td>
</tr>
</tbody>
</table>

**Which would you prefer?**

*(Please tick one)*

- □ Treatment A
- □ Treatment B
DCE design – technical considerations

• Need to specify utility/value model based on multiple attributes - not restricted to linear additive models (unlike other methods such as MCDA)

• The required number of attributes and levels depends on the model that is chosen and the required level of precision

• Balance with reasonable limit on number of questions
Cognitive strain becomes an issue in all but smallest DCEs.

Need to limit number of attributes, alternatives, choice sets.

Plus usual need to ensure task & key background info is understood.

Validation questions can be included.
DCE design – full / fractional factorial

- **Full factorial**: uses all (!) possible combinations of attribute levels
  
  \[ \text{# combinations} = A^L \] if all A attributes have L levels

- **Fractional factorial**: uses a subset of the possible combinations of attribute levels
  - Not all fractional factorial designs are equally efficient
  - Efficient designs exhibit various kinds of symmetry:
    - Level balance
    - Orthogonality
    - Minimal overlap
    - Utility balance

Fractional factorial designs: like working out the dimensions of a box given the locations of some of the corners
## Comparative overview of elicitation methods

<table>
<thead>
<tr>
<th></th>
<th>Swing-weighting</th>
<th>MACBETH</th>
<th>AHP</th>
<th>DCE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Responses</strong></td>
<td>Quantitative</td>
<td>Qualitative</td>
<td>Qualitative or qualitative</td>
<td>Qualitative</td>
</tr>
<tr>
<td><strong>How is consistency measured?</strong></td>
<td>Method ensures consistency</td>
<td>Inconsistencies must be resolved</td>
<td>Computes a consistency score</td>
<td>Reflected in uncertainty of estimates</td>
</tr>
<tr>
<td><strong>Weight calculation</strong></td>
<td>Direct</td>
<td>Linear optimisation (plus tuning)</td>
<td>Principal eigenvector</td>
<td>Regression</td>
</tr>
<tr>
<td><strong>Can be given out as a paper questionnaire?</strong></td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Conclusions

• Eliciting patient preferences in regulatory assessment can add value and lead to more clinically relevant decisions
  – Political legitimacy, transparency, trust, communicability

• A number of formal methods can be used to elicit patient preferences
  – Each methodology has its own features, strengths and weaknesses
  – The PPI work from PROTECT is still ongoing...
ELICITING PATIENT PREFERENCES: APPLYING DECISION THEORY TO HEALTH RESEARCH
Contents

• Why collect patient preferences?
• Decision Analysis
• Visualize Sub-study: eliciting patient preferences
• Study design
• Building Value Function
• Eliciting Weights
• Planned analysis
• Summary
Importance of Patients’ Perception for Treatment Decisions

Regulators’ view:
An increased cure rate in cancer, a potentially life-saving treatment will always outweigh a grade 1 or 2 AE (e.g. (permanent hair loss) - positive regulatory decision

Some patients’ view:
This permanent hair loss is important, severe enough for me to decline the potentially curative and life-saving adjuvant therapy – negative treatment decision

“The mastectomy and loss of breast are NOTHING compared to the loss of my hair.”

“Not a day goes by that I don’t regret doing the NN (therapy). Oh, if we could only turn back the hands of time!”

“I never, never, never would have agreed to take NN if I was informed of this 6.3% risk; even a 3% risk...or any risk...”
How to bring patient preferences/values into BR decisions?

- Patients with the specific disease condition know which outcomes and symptoms matter most to them.
- Patients enrolled in regulatory drug trial are (ideally) the target group for treatment once a drug is licensed, yet we do not explore their values and preferences in a systematic way.
- In terms of listening to the patients’ voice, trial patients are an underutilized source.

G. Rasi, AIFA, 2013
Can Decision Analysis Help?

“The spirit of decision analysis is divide and conquer:

decompose a complex problem into simpler problems, get one’s thinking straight on these simpler problems, paste these analyses together with logical glue, and come out with a program of action for the complex problem”

(Howard Raiffa 1968, p. 271)
Objective:
To evaluate the use of the MACBETH (Measuring Attractiveness through a Categorical Based Evaluation) software for the elicitation of patient preferences using a simple pair-wise comparison between treatment outcomes
- determine value functions for disease attributes
- assess weights between disease attributes (trade-offs)

Design
- Web-based study among patients with diabetes, atrial fibrillation
- Supported by the NICR UK, Dutch hospitals, and Laser who recruited patients and healthcare professionals
Participant recruitment

- Target for MACBETH: 1800
- Study population:
  - Patients
  - Healthcare professionals
  - Regulators supporting CHMP & PRAC
- 3 countries
  - United Kingdom
  - The Netherlands
  - France

We need your feedback!

To understand how information on the benefits and risks of medicines to patients and healthcare professionals could be improved.
Steps to building an elicitation procedure*

- Determine the outcomes of interest
- For each outcome determine levels, ranging from best case to worst case
- Create the value elicitation section of the questionnaire
- Create the weighting elicitation section
- Collect data from patients and convert the qualitative responses of patients to quantitative scores
- * Seek patient input/confirmation for steps 1-4
Examples of Treatment Outcomes and Levels for Atrial Fibrillation

<table>
<thead>
<tr>
<th>Treatment outcome</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ischemic Stroke</td>
<td>No patients developing ischemic stroke</td>
</tr>
<tr>
<td></td>
<td>1% of patients developing ischemic stroke</td>
</tr>
<tr>
<td></td>
<td>2% of patients developing ischemic stroke</td>
</tr>
<tr>
<td></td>
<td>3% of patients developing ischemic stroke</td>
</tr>
<tr>
<td></td>
<td>4% of patients developing ischemic stroke</td>
</tr>
<tr>
<td>Myocardial Infarction</td>
<td>No patients developing myocardial infarction</td>
</tr>
<tr>
<td></td>
<td>1% of patients developing myocardial infarction</td>
</tr>
<tr>
<td></td>
<td>2% of patients developing myocardial infarction</td>
</tr>
<tr>
<td></td>
<td>3% of patients developing myocardial infarction</td>
</tr>
<tr>
<td></td>
<td>4% of patients developing myocardial infarction</td>
</tr>
<tr>
<td>Major bleeding</td>
<td>No patients developing a major bleed</td>
</tr>
<tr>
<td></td>
<td>2% of patients developing a major bleed</td>
</tr>
<tr>
<td></td>
<td>4% of patients developing a major bleed</td>
</tr>
<tr>
<td></td>
<td>6% of patients developing a major bleed</td>
</tr>
<tr>
<td></td>
<td>8% of patients developing a major bleed</td>
</tr>
<tr>
<td>Minor bleeding</td>
<td>15% of patients developing a minor bleed</td>
</tr>
<tr>
<td></td>
<td>20% of patients developing a minor bleed</td>
</tr>
<tr>
<td></td>
<td>25% of patients developing a minor bleed</td>
</tr>
<tr>
<td></td>
<td>30% of patients developing a minor bleed</td>
</tr>
<tr>
<td></td>
<td>35% of patients developing a minor bleed</td>
</tr>
</tbody>
</table>
Building a value scale for “Minor bleeding”

15% of patients with minor bleeding
20% of patients with minor bleeding
25% of patients with minor bleeding
30% of patients with minor bleeding
35% of patients with minor bleeding

What is the difference in value between
15% of patients and 20% of patients
with a minor bleeding?
Building a value scale for "Minor bleeding"

What is the difference in value between 15% of patients and 20% of patients with a minor bleeding?
Building a value scale for “Minor bleeding”

- Strong 15%
- Very Strong 20%
- Strong 25%
- Weak 30%
- Weak 35%
Value functions will fit one of these 10 profiles
If you could increase one treatment effect from its worst value (on the bottom) to its best value (on the top), which one would you increase?

- **Ischemic stroke**: 0% of the patients getting an ischemic stroke, 4% of the patients getting an ischemic stroke.
- **Myocardial infarction**: 0% of the patients getting a myocardial infarction, 4% of the patients getting a myocardial infarction.
- **Pulmonary embolism**: 0% of the patients getting a pulmonary embolism, 4% of the patients getting a pulmonary embolism.
- **Fatal bleeding**: 0% of the patients getting a fatal bleeding, 4% of the patients getting a fatal bleeding.
- **Major bleeding**: 0% of the patients getting a major bleeding, 8% of the patients getting a major bleeding.
- **Minor bleeding**: 15% of the patients getting a minor bleeding, 35% of the patients getting a minor bleeding.
How desirable is this improvement?

- **Ischemic stroke**: 4% of patients getting an ischemic stroke
- **Myocardial infarction**: 4% of patients getting a myocardial infarction
- **Pulmonary embolism**: 4% of patients getting a pulmonary embolism
- **Fatal bleeding**: 8% of patients getting a fatal bleeding
- **Major bleeding**: 35% of patients getting a major bleeding
- **Minor bleeding**: 0% of patients getting a minor bleeding

How desirable is this improvement?

- Extreme
- Very strong
- Strong
- Moderate
- Weak
- Very weak

---

**Ischemic stroke**

- 0% of the patients getting an ischemic stroke

**Myocardial infarction**

- 0% of the patients getting a myocardial infarction

**Pulmonary embolism**

- 0% of the patients getting a pulmonary embolism

**Fatal bleeding**

- 0% of the patients getting a fatal bleeding

**Major bleeding**

- 0% of the patients getting a major bleeding

**Minor bleeding**

- 15% of the patients getting a minor bleeding
If you could increase one treatment effect from it's worst value (on the bottom) to it's best value (on the top), which one would you increase?

- Ischemic stroke:
  - 0% of the patients getting an ischemic stroke
  - 4% of the patients getting an ischemic stroke
  - Ischemic stroke

- Myocardial infarction:
  - 0% of the patients getting a myocardial infarction
  - 4% of the patients getting a myocardial infarction
  - Myocardial infarction

- Pulmonary embolism:
  - 0% of the patients getting a pulmonary embolism
  - 4% of the patients getting a pulmonary embolism
  - Pulmonary embolism

- Fatal bleeding:
  - 0% of the patients getting a fatal bleeding
  - 4% of the patients getting a fatal bleeding
  - Fatal bleeding

- Major bleeding:
  - 0% of the patients getting a major bleeding
  - 8% of the patients getting a major bleeding
  - Major bleeding

- Minor bleeding:
  - 0% of the patients getting a minor bleeding
  - 15% of the patients getting a minor bleeding
  - Minor bleeding
How desirable is this improvement?

- **Ischemic stroke**: 4% of the patients getting an ischemic stroke
- **Myocardial infarction**: 4% of the patients getting a myocardial infarction
- **Pulmonary embolism**: 4% of the patients getting a pulmonary embolism
- **Fatal bleeding**: 4% of the patients getting a fatal bleeding
- **Major bleeding**: 8% of the patients getting a major bleeding
- **Minor bleeding**: 15% of the patients getting a minor bleeding

- **Ischemic stroke**: 0% of the patients getting an ischemic stroke
- **Myocardial infarction**: 0% of the patients getting a myocardial infarction
- **Pulmonary embolism**: 0% of the patients getting a pulmonary embolism
- **Fatal bleeding**: 0% of the patients getting a fatal bleeding
- **Major bleeding**: 0% of the patients getting a major bleeding
- **Minor bleeding**: 0% of the patients getting a minor bleeding

**How desirable is this improvement?**

- Extreme
- Very strong
- Strong
- Moderate
- Weak
- Very weak

**154**
How much more desirable is the improvement on the right when compared to the one on the left?

- **Myocardial infarction**
  - 0% of the patients getting a myocardial infarction
  - 4% of the patients getting a myocardial infarction
  - Strong

- **Pulmonary embolism**
  - 0% of the patients getting a pulmonary embolism
  - 4% of the patients getting a pulmonary embolism
  - Very strong
Qualitative swing weighting

<table>
<thead>
<tr>
<th></th>
<th>Ischemic stroke</th>
<th>Myocardial Infarction</th>
<th>Major bleeding</th>
<th>Minor bleeding</th>
</tr>
</thead>
<tbody>
<tr>
<td>30%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40%</td>
<td></td>
<td></td>
<td></td>
<td>10%</td>
</tr>
</tbody>
</table>
Evaluation of actual clinical data using patient values
Building a decision model

Global Results
Table of global and partial scores for each option in each criteria

<table>
<thead>
<tr>
<th></th>
<th>Number of relapses</th>
<th>Time to disease prog</th>
<th>Disease progression</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Treat A</td>
<td>50</td>
<td>92</td>
<td>86</td>
<td>72</td>
</tr>
<tr>
<td>Treat B</td>
<td>6</td>
<td>89</td>
<td>100</td>
<td>52</td>
</tr>
<tr>
<td>Treat C</td>
<td>-6</td>
<td>11</td>
<td>29</td>
<td>6</td>
</tr>
<tr>
<td>Neutral</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Weights: 46% 38% 15%

Results

Global results
Tabela de pontuações globais e parciais para cada opção em cada fator de avaliação

Analysis

Profile Analysis
Pontuações das opções em todos os fatores. Selecione a opção pretendida para ver o seu perfil. A seleção de duas opções permite ver a comparação entre as duas.

Sensitivity Analysis
Análise da sensibilidade dos resultados a variações nos pesos dos fatores
Summary

- Method can be used to collect patient preferences in a remote setting
- Can be easily extended to patients within clinical trials (advanced PRO)
- Complies with decision theoretic principles
- Further research is needed to assess aggregation of the data
Deborah Ashby, PhD

CLOSING REMARKS, TAKE-HOME MESSAGES AND ACKNOWLEDGEMENT
Benefit-risk assessment methodologies

Non-quantitative

Benefit-risk assessment framework

Descriptive framework

Quantitative framework

PROACT-URL
ASF
BRAT
FDA BRF
CMR-CASS
COBRA
SABRE
UMBRA
OMERACT 3x3

Metric indices for B-R assessment

NNT
NNH
AE-NNT
RV-NNH
Impact numbers
MCE
RV-MCE
MAR
NEAR

Health indices

QALY
DALY
HALE
Q-TWiST

Threshold indices

UT-NNT
INHB
BRR
GBR
Principle of 3
TURBO
Beckmann

Trade-off indices

DAGs
PSM
CPM
ITC
MTC
CDS

Estimation techniques

Utility survey techniques

SPM
CV
CA
DCE

-----

AHP
Swing-weighting
MACBETH

Visual representations of benefit and risk

<table>
<thead>
<tr>
<th>Short name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name/hubic</td>
<td>Tornado plot</td>
</tr>
<tr>
<td>Created in</td>
<td>R.</td>
</tr>
<tr>
<td>Message</td>
<td>The tornado plot shows how the changes in the natalizumab outcome measure affect the incremental benefit-risk score. It displays the relative importance of criteria via one-way sensitivity analysis of changing a fixed amount of the measured outcomes.</td>
</tr>
<tr>
<td>Intended audience</td>
<td>Statisticians and regulators. Not for physicians and patients.</td>
</tr>
<tr>
<td>Knowledge required</td>
<td>Some knowledge on the use of sensitivity analysis and uncertainty. Some understanding of the incremental benefit-risk concept. Some knowledge on how to extract information from tornado diagrams.</td>
</tr>
<tr>
<td>Unintentional massage</td>
<td>The legend of &quot;high/low&quot; is not intuitive and could be misleading.</td>
</tr>
<tr>
<td>Message not communicated</td>
<td>It is unclear which of the options are benefits and which are risks. The colour-coding is not intuitive and difficult to interpret.</td>
</tr>
<tr>
<td>Proposed Improvement</td>
<td>Horizontal axis should be made wider to accommodate benefit-risk values. To re-label legend items to more intuitive terms. The tornado plot could also be accompanied by text annotations to aid interpretation.</td>
</tr>
</tbody>
</table>
Patient and public involvement

Consultation
Health professionals elicit the patient and public perspective to inform the decision making process

Collaboration
Health professionals and patients and the public form an active partnership and jointly participate in decision making
## Benefits and risks of formalising benefit-risk assessment

### Benefits
- Puts benefits and risks on same page
- Gives a framework to include patients’ views
- Transparency facilitates discussion
- It’s fun!

### Risks
- Trade-off between being too simplistic or just incomprehensible
- Can be seen as a ‘black box’
- Pharma want to know what regulators want
ACKNOWLEDGEMENT
Support

- 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](http://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 Medicine Initiative Joint Undertaking ([www.imi.europa.eu](http://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.
IMI-PROTECT Benefit-Risk Group

Deborah Ashby, Alain Micaleff, Steve Hobbiger, Ioanna Tzoulaki, Diana Hughes, Shahrul Mt-Isa.


Subhakanta Das, Jane Okwesa, Emily Thompson.
ADDIS

ACADEMIA:

*University Medical Center Groningen:* Hans Hillege, Andrea Beyer, Gert van Valkenhoef, Joël Kuiper, Daan Reid, Connor Stroomberg

*Erasmus University Rotterdam:* Tommi Tervonen

*University of Groningen:* Bert de Brock

EMA BENEFIT-RISK METHODOLOGY PROJECT:

Francesco Pignatti, Andreas Kouroumalis

FUNDING SOURCES:

The MCDA web interface was initially funded by TI Pharma project Escher and integrated in ADDIS 2 with funding from IMI GetReal. Further development and the creation of training materials is supported by IMI PROTECT.
WP6 Eliciting Patient Preferences

• Hans-Georg Eichler
• Larry Phillips
• Carlos Bana e Costa
• Hans Hillege
• NICR UK
• Bana Consulting
REFERENCES AND BACKUP SLIDES
Structured assessment

Paper in press for a special issue of the Biometrical Journal:

A case study using the PrOACT-URL and BRAT frameworks for structured benefit risk assessment

Richard Nixon*,1, Christoph Dierig2, Shahrul Mt-Isa3, Isabelle Stöckert2, Thaison Tong4, Silvia Kuhls2, Gemma Hodgson5, John Pears6, Ed Waddingham3, Kimberley Hockley3, and Andrew Thomson7

IMI PROTECT Benefit- Risk integration and representation Reports

ADDIS


Benefit-risk is central to key decisions

<table>
<thead>
<tr>
<th>Decision Perspective</th>
<th>Decision</th>
<th>How to value decision?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physician</td>
<td>Which drug for this patient?</td>
<td>Best benefit-risk profile</td>
</tr>
<tr>
<td>Regulator</td>
<td>Which patient for this drug?</td>
<td>More benefit than risk</td>
</tr>
<tr>
<td>Payer</td>
<td>Which drug in which patient population?</td>
<td>Comparative cost-effectiveness</td>
</tr>
</tbody>
</table>

Eichler 2011 - Bridging the efficacy–effectiveness gap: a regulator's perspective on addressing variability of drug response
4) Customize and communicate

Re-visit key benefits and risks

- The benefit-risk process can be iterative.

- The key benefits and risks may need to be “tuned”.
  - Changes outcomes in value tree if data are not available.
  - Outcome measures may be refined in response to how data are measured.

- Guard against bias.
  - Changing the value tree in response to observed data could bias the benefit-risk balance.
6) Benefit-risk communication

Visualization of benefit-risk. Functional and perceptual tasks

- **Carswell (1992) taxonomy of functional tasks**
  - Point reading (reading one value on a graph)
  - Local comparison (reading and comparing two values on a graph)
  - Global comparison (reading and comparing more than values simultaneously on a graph)
  - Synthesis judgment (extrapolating information beyond what is explicitly shown on a graph)

- **Cleveland and McGill’s (1984) perceptual tasks**
  - Position on common aligned scale (e.g. bar charts)
  - Position on common non-aligned scales (e.g. scatter plots)
  - Length (e.g. stacked bar charts)
  - Angle (e.g. pie charts)
  - Area (e.g. circles, blobs)
  - Volume (e.g. cubes)
  - Colour (e.g. coloured circles)

- **Tufte (2001)** - Ink should be reserved for data
# Required natalizumab effect on outcomes to reach a neutral benefit-risk vs. placebo

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weight</th>
<th>Current Tysabri Effect</th>
<th>Required Tysabri Effect</th>
<th>Required Change (Absolute)</th>
<th>New BR</th>
</tr>
</thead>
<tbody>
<tr>
<td>PML</td>
<td>54%</td>
<td>0.15%</td>
<td>6.36%</td>
<td>6%</td>
<td>0.00</td>
</tr>
<tr>
<td>Transaminases elevation</td>
<td>11%</td>
<td>5%</td>
<td>36%</td>
<td>31%</td>
<td>0.00</td>
</tr>
<tr>
<td>Relapse</td>
<td>8%</td>
<td>0.47</td>
<td>1.31</td>
<td>0.84</td>
<td>0.00</td>
</tr>
<tr>
<td>Reactivation of serious herpes viral infections</td>
<td>6%</td>
<td>0%</td>
<td>56%</td>
<td>56%</td>
<td>0.00</td>
</tr>
<tr>
<td>Seizures</td>
<td>5%</td>
<td>1%</td>
<td>68%</td>
<td>67%</td>
<td>0.00</td>
</tr>
<tr>
<td>Congenital abnormalities</td>
<td>5%</td>
<td>0%</td>
<td>67%</td>
<td>67%</td>
<td>0.00</td>
</tr>
<tr>
<td>Disability progression</td>
<td>5%</td>
<td>11%</td>
<td>78%</td>
<td>67%</td>
<td>0.00</td>
</tr>
<tr>
<td>Infusion reactions/injection reactions</td>
<td>3%</td>
<td>24%</td>
<td>100%</td>
<td>76%</td>
<td>0.21</td>
</tr>
<tr>
<td>Flu-like reactions</td>
<td>1%</td>
<td>40%</td>
<td>100%</td>
<td>60%</td>
<td>0.55</td>
</tr>
<tr>
<td>Hypersensitivity Reactions</td>
<td>1%</td>
<td>0%</td>
<td>100%</td>
<td>100%</td>
<td>0.47</td>
</tr>
<tr>
<td>Convenience</td>
<td>1% iv qm hosp</td>
<td>sc od</td>
<td>NA</td>
<td>0.53</td>
<td></td>
</tr>
</tbody>
</table>
Step 5: Assess outcome importance

Linear Additive models

- Linear Additive Models with Swing Weights
  - Value functions: Within outcome importance
  - Swing weights: Between outcome importance

Outcome: 2-year relapse rate

Measure = 0.47

Value(measure) = 0.77

Elicited Weight = 8%

BR Contribution = 0.062

Value = 0.77

2-year relapse rate

0%

1
Step 5: Assess outcome importance

*Three common methods for weight elicitation that use linear additive models*

- Multi-criteria Decision Analysis (MCDA)
- MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique)
- AHP (Analytic Hierarchy Process)
Step 5: Assess outcome importance

**MCDA**

For each outcome category

1. Rank outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infusion/injection reactions</td>
<td>1</td>
</tr>
<tr>
<td>Hypersensitivity reactions</td>
<td>2</td>
</tr>
<tr>
<td>Flu-like reactions</td>
<td>3</td>
</tr>
</tbody>
</table>

2. Relative importance

How much more **important** is it to avoid the top-ranked event compared to the others?

- Infusion/injection reactions
- Hypersensitivity reactions
- Flu-like reactions
Repeat this process all the way up the value tree

The top ranked outcome in each category is carried up the tree

- Move bottom-up through the tree and compare the top-ranked outcomes from each category
- Finally, the top-ranked benefit is compared to the top-ranked risk
- The individual weights for each outcome can then be calculated
Compute the overall weights

Treatment
- Benefits
  - Relapse
  - Disability Progression
  - Convenience
- Infection
  - Reactivation of serious herpes viral infections
  - PML
- Reproductive Toxicity
  - Congenital abnormalities
- Liver Toxicity
  - Transaminases elevation
- Neurological
  - Seizures
  - Infusion/injection reactions
- Other
  - Hypersensitivity reactions
  - Flu-like reactions

Weights

PML is 10x worse than disease progression

Note that as the weight for a relapse is for a value function with the measure scale with a range from 0 to 2, then actual weight of a single relapse is half that shown here.
Example question to assess between outcome importance

- Imagine a clinical trial of 1000 patients with 1 patient developing PML in the treatment arm.
- How many patients would need to have an EDSS progression prevented for you to be indifferent about the benefit and harm caused by the treatment?
MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique)

**Qualitative assessment**

- MACBETH is similar to MCDA, except that it provides a different way to get the weights

- **Step 1: Qualitatively** assess how much more attractive it is to move from worst to best for outcome i vs. moving from worst to best for outcome j and keeping everything else at the worst measure (Do this for each pair of criteria)

- **Step 2**: Check consistency of answers

- **Step 3**: Compute initial guess at weights with optimization

- **Step 4**: Refine weights while maintaining consistency
MACBETH

Qualitative assessment

<table>
<thead>
<tr>
<th></th>
<th>PML</th>
<th>Abortion or congenit</th>
<th>Seizures</th>
<th>Reactivation of sero</th>
</tr>
</thead>
<tbody>
<tr>
<td>PML</td>
<td>no</td>
<td>extreme</td>
<td>extreme</td>
<td>e</td>
</tr>
<tr>
<td>Abortion or congenit</td>
<td>no</td>
<td></td>
<td>strong</td>
<td>v</td>
</tr>
<tr>
<td>Seizures</td>
<td>no</td>
<td></td>
<td></td>
<td>m</td>
</tr>
<tr>
<td>Reactivation of sero</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[all zero]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Consistent judgements

Decision context:
- Identify outcomes
- Source data
- Customize
- Outcome importance
- B-R metrics
AHP (Analytic Hierarchy Process)

**Qualitative assessment**

- Weights are elicited by making pairwise comparisons between criteria
- “How much more important is outcome i vs. outcome j?”
- Must provide number from 1 to 9 on relative scale
- Weight is calculated by finding the dominant eigenvector of the corresponding matrix
- Value functions are computed in a similar manner (do not necessarily come from linear function)
- No consistency check, but rather a score (<0.2 is okay)
Two way sensitivity analysis on weights

Incremental Benefit-Risk of Tysabri – Placebo

- Vary the PML weight (x-axis) and the relapse weight (each line).
- Green line in the middle is the elicited weight. Change by +/- 30%.
- Again the BR is robust to these changes.
Probabilistic sensitivity analysis of the measures

*Incremental Benefit-Risk of Tysabri – Placebo*

- 80% CI are included in the waterfall plot.
- The uncertainty in the overall BR is robust to uncertainty in the outcome measures.
- The components of the uncertainty can be seen.