Benefit-Risk Assessment of Medicinal Products

Professor Larry Phillips
London School of Economics & Political Science

Risk: Perception, Measurement & Policy
City University
13 December 2013
There is a risk that this new drug won't help your heart risk and there are risks from taking the drug.
There is a risk that this new drug won't help your heart risk and there are risks from taking the drug.

Your heart might malfunction
There is a risk that this new drug won’t help your heart risk and there are risks from taking the drug.

The drug doesn’t work for everybody
There is a risk that this new drug won't help your heart risk and there are risks from taking the drug.

You could suffer side-effects.
There is a risk that this new drug won’t help your heart risk and there are risks from taking the drug.

The drug doesn’t work for everybody

Your heart might malfunction

You could suffer side-effects
Research to improve benefit-risk assessment

**EMA Benefit-Risk: 2009-11**

- WP1: Describes current B-R regulatory practice in EU
- WP2: Reviews of 18 quantitative approaches to B-R
- WP3: Modelling of five field tests of new drugs being considered by the CHMP
- WP4: Recommends tools and processes
- Report on Risk Perception Study Module

Source: [www.ema.europa.eu](http://www.ema.europa.eu). Click on Special topics, Supporting research, Benefit-risk methodology

**IMI Protect, WP5: 2010-14**

- Reviews 47 approaches & recommends 13 for further investigation
- Reviews visualisation methods; two parts, current & proposed
- Investigates approaches for six previously-approved drugs (in 8 reports, totalling 29MB)
- Publishes Recommendations on website (3.85MB)

How do regulators decide? By...

Discussing

Voting

But no quantitative modelling is used by any regulator anywhere in the world to deal with the massive amount of data—10GB more or less!
Benefit-risk assessment by drug regulators

Regulators need to refine their methods of assessing benefit-risk balances and switch from “implicit” to “explicit” decision making—that is, to an approach involving explicit descriptions not only of all decision criteria and interpretations of data but also valuations, such as the weighting factors for potential treatment outcomes.

Ideally, regulators should also shift from the use of qualitative statements to quantitative descriptions of the size of the net health benefits.

Interviews—6 European Agencies

**What is a benefit?**
1. Everything good
2. Improvement in health state
3. Real-world effectiveness
4. Clinical relevance
5. Improvement in illness
6. Suffering reduced
7. Positive action of drug
8. Meets unmet medical need
9. Positive improvement in health state as perceived by patient
10. Safety improvement
11. Value compared to placebo
12. Change in managing patient
37. Statistically significant effect

**What is a risk?**
1. All that is negative
2. Adverse events
3. Reduction in quality
4. Kinetic interactions
5. Side effects
6. Serious adverse effects
7. Bad effects
8. Danger for the patient
9. Tolerance of a drug compared to serious side effects
10. Harm
11. Severity of side effects
12. Frequency of side effects
51. Potential or theoretical risks
From Efficacy & Safety to Benefits & Risks

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

Judgement required
Defining ‘benefit’ and ‘risk’

<table>
<thead>
<tr>
<th>Favourable Effects</th>
<th>Uncertainty of Favourable Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unfavourable Effects</td>
<td>Uncertainty of Unfavourable Effects</td>
</tr>
</tbody>
</table>

These four cells are now included and elaborated in Section 5 of the EMA’s Guidance Document for preparing the 80-day Assessment Report.  
Pharma-BRAT (Benefit-Risk Action Team) framework

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

Current development as UMBRA (Universal Method for Benefit-Risk Assessment) by CIRS (Centre for Innovation in Regulatory Science).

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


- Problem
- Objectives
- Alternatives
- Consequences
- Trade-offs
- Uncertainty
- Risk attitude
- Linked decisions
The FDA rejects a quantitative approach on the grounds that it usually requires judgements of numerical weights “...that are at best debatable and at worst arbitrary” (p. 4). http://www.fda.gov/downloads/forindustry/userfees/prescriptiondruguserfee/ucm329758.pdf
Is there a Gold Standard for a quantitative approach?

A comprehensive method should:

1. Express all effects, favourable and unfavourable, in comparable units
2. Accept any performance measures: measurable quantities, scoring systems, relative frequencies, health outcomes, etc.
3. Distinguish between performance measures (data) and their clinical relevance (judgements)
4. Capture trade-offs among the effects
5. Be based on sound theory, not ad-hockery
## Quantitative models, 16 drugs 2009-2013

<table>
<thead>
<tr>
<th>Product</th>
<th>Indication</th>
<th>Quantitative Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lilly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug X</td>
<td>Idiopathic short stature</td>
<td>MCDA</td>
</tr>
<tr>
<td>rimonabant</td>
<td></td>
<td>MCDA</td>
</tr>
<tr>
<td>certolizumab pegol</td>
<td></td>
<td>MCDA + simulation</td>
</tr>
<tr>
<td>sunitinib malate</td>
<td></td>
<td>Decision Tree + Markov</td>
</tr>
<tr>
<td>lapatinib</td>
<td></td>
<td>MCDA + simulation</td>
</tr>
<tr>
<td>LSE MSc students</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tafamidis</td>
<td>Transthyretin amyloid polyneuropathy</td>
<td>MCDA</td>
</tr>
<tr>
<td>briakinumab</td>
<td>Chronic plaque psoriasis</td>
<td>MCDA</td>
</tr>
<tr>
<td>vandetanib</td>
<td>Inoperable thyroid cancer</td>
<td>MCDA</td>
</tr>
<tr>
<td>tocilizumab</td>
<td>Systemic juvenile idiopathic arthritis</td>
<td>MCDA</td>
</tr>
<tr>
<td>belimumab</td>
<td>Systemic lupus erythematosus</td>
<td>MCDA</td>
</tr>
<tr>
<td>EMA B-R Project (new drugs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>natalizumab</td>
<td>Multiple sclerosis</td>
<td>MCDA, Forest plot</td>
</tr>
<tr>
<td>rimonabant</td>
<td></td>
<td>MCDA, simulation</td>
</tr>
<tr>
<td>telithromycin</td>
<td></td>
<td>MCDA, simulation</td>
</tr>
<tr>
<td>efalizumab</td>
<td>Psoriasis</td>
<td>MCDA</td>
</tr>
<tr>
<td>belimumab</td>
<td>Systemic lupus erythematosus</td>
<td>MCDA</td>
</tr>
<tr>
<td>IMI PROTECT project</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rosiglitazone</td>
<td>Diabetes</td>
<td>MCDA</td>
</tr>
<tr>
<td>warfarin</td>
<td>Atrial fibrillation</td>
<td>MCDA + simulation</td>
</tr>
</tbody>
</table>

**MCDA is emerging as the gold standard.**
“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, *Decision Analysis*, p. 271
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.

A system *not* based on MCDA

MCDA converts all input metrics of outcomes into the common currency of *value added*. 
A Drug Case Study: Benlysta (belimumab)
Establish decision context

- **Indication:** Treatment of active, autoantibody-positive systemic lupus erythematosus (SLE).
- **Use:** Add-on to standard therapy (hydroxychloroquine and corticosteroids) for adult patients with a high degree of disease activity.
- **Efficacy:** Two randomised, placebo-controlled, clinical studies.
- **Safety:** Three open-label continuation trials.
- **Medical Need:** Newer, more-effective and better-tolerated therapies.
Identify objectives & their criteria

Effects Tree

Primary endpoint

Benefit-Risk Balance

SLEDAI
- % Improved 4
- % Improved 6

PGA
- % no worse
- Mean score
- BILAG A/B
- CS sparing
- Flare rate
- QoL
- Potential SAEs
- Infections
- Sensitivity Reaction

FE
- UFE
Identify alternatives (options)

1. Benlysta 1mg
2. Benlysta 10mg
3. Placebo
## Summarise data as an Effects Table

<table>
<thead>
<tr>
<th>Effects</th>
<th>Name</th>
<th>Description</th>
<th>Best</th>
<th>Worst</th>
<th>Units</th>
<th>Placebo</th>
<th>10 mg</th>
<th>1 mg</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Favourable Effects</strong></td>
<td>SLEDAI % Improved ≥ 4</td>
<td>Percentage of patients with at least 4 points reduction in SLEDAI</td>
<td>100</td>
<td>0</td>
<td>%</td>
<td>41</td>
<td>53</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>SLEDAI % Improved &gt; 6</td>
<td>Percentage of patients with more than 6 points reduction in SLEDAI</td>
<td>100</td>
<td>0</td>
<td>%</td>
<td>23</td>
<td>37</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>PGA % no worse</td>
<td>Percentage of patients with no worsening in Physician’s Global Assessment (worsening = an increase of less than 0.3 points)</td>
<td>100</td>
<td>0</td>
<td>%</td>
<td>66</td>
<td>75</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>PGA Mean score</td>
<td>Overall mean change of PGA score from baseline for the study population</td>
<td>1.0</td>
<td>0</td>
<td>Difference</td>
<td>0.44</td>
<td>0.48</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>BILAG A/B</td>
<td>Percentage of patients with no new BILAG A/2B</td>
<td>100</td>
<td>0</td>
<td>%</td>
<td>69.0</td>
<td>75.2</td>
<td>70.1</td>
</tr>
<tr>
<td><strong>Secondary Endpoints</strong></td>
<td>CS Sparing</td>
<td>Percentage of patients that reduced the dose of corticosteroids by more than 25% and to less than 7.5 mg/day</td>
<td>100</td>
<td>0</td>
<td>%</td>
<td>12.3</td>
<td>17.5</td>
<td>20.0</td>
</tr>
<tr>
<td></td>
<td>Flare rate</td>
<td>Number of new BILAG A cases per patient year</td>
<td>0</td>
<td>5</td>
<td>Number</td>
<td>3.51</td>
<td>2.88</td>
<td>2.90</td>
</tr>
<tr>
<td></td>
<td>QoL</td>
<td>Mean change in the total score of SF 36 (Short Form)</td>
<td>0</td>
<td>100</td>
<td>Difference</td>
<td>3.5</td>
<td>3.4</td>
<td>3.7</td>
</tr>
<tr>
<td><strong>Unfavourable Effects</strong></td>
<td>Potential SAEs</td>
<td>Potential for developing tumour, adverse interactions with vaccines and AE on pregnancies</td>
<td>100</td>
<td>0</td>
<td>Judgement</td>
<td>100</td>
<td>0</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Infections</td>
<td>Proportion of patients with serious infections that are life-threatening</td>
<td>0</td>
<td>10.0</td>
<td>%</td>
<td>5.2</td>
<td>5.2</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td>Sensitivity Reaction</td>
<td>Proportion of patients with hypersensitivity reactions at any time in the study</td>
<td>0</td>
<td>2.0</td>
<td>%</td>
<td>0.10</td>
<td>0.40</td>
<td>1.30</td>
</tr>
</tbody>
</table>

*Statistical help needed to pool data from multiple studies!*
How do you put it all together?

SLEDAI % Improved ≥ 4
SLEDAI % Improved > 6
PGA % no worse
PGA Mean score
BILAG A/B
CS Sparing
Flare rate
QoL
Potential SAEs
Infections
Sensitivity Reaction

**MCDA modelling + Social process = Smart Decisions**

Phillips’ Law: Never rely on a single expert!
Social process: *Decision Conferencing*

- One or more workshops
- Attended by key players representing diversity of perspectives
- 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

Source: [http://www.lawrencephillips.net/Decision_conferencing.html](http://www.lawrencephillips.net/Decision_conferencing.html)
Describe the consequences

Linear direct conversion to preference values
Describe the consequences

Linear inverse conversion to preference values
Describe the consequences

**Non-linear conversion to clinical preference values**

The 10mg - Xmg difference in number of flares is increased in preference value, representing its clinical relevance.

A hypothetical Xmg dose with a 1.25 flare rate
Trade-offs: assess swing-weights

1. Trade-offs among the favourable effects

2. Trade-offs among the unfavourable effects

3. Trade-off between the most important favourable effect and the most important unfavourable effect

“How big is the difference, and how much do you care about it?”
A major error in weighting

“How important is this effect?”

“How big is the worst-to-best difference, and how clinically relevant is it?”

Aggregation of data & judgements

- Meehl’s 1954 book dropped a bombshell in clinical psychology.
- His survey of 20 studies showed that simple, linear, additive models consistently out-performed clinical predictions of behaviour.
- By 1996, of 136 comparative studies, just 8 favoured clinical prediction.
- He identified integration of multiple pieces of data as the problem, not the judgements about the pieces.

Aggregation: Let the computer do it!

- Calculate weighted preference values (WPVs) for the alternatives on each effect:
  \[ WPV = \text{criterion weight} \times \text{alternative preference value} \]
- Apply the formula to ensure comparability of WPVs.
- Sum them, from right to left on the value tree.
- Normalise the sums at each node to provide cumulative weights at each level of the effects tree.
- Display overall sums of WPVs for the alternatives.

(It’s not rocket science!)

Software here was Hiview3: LSE-developed, available from www.catalyze.co.uk.
Examine results—stacked bar graph of WPVs

(Assuming zero weight on the criterion Potential SAEs)
Show results—difference display

<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>FE Flare rate</td>
<td>20.2</td>
<td>12</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>SLEDAI % Improved 6</td>
<td>16.2</td>
<td>14</td>
<td>2.3</td>
<td>4.8</td>
</tr>
<tr>
<td>SLEDAI % Improved 4</td>
<td>12.9</td>
<td>12</td>
<td>1.6</td>
<td>6.3</td>
</tr>
<tr>
<td>FE CS sparing</td>
<td>12.1</td>
<td>5</td>
<td>0.6</td>
<td>6.9</td>
</tr>
<tr>
<td>SRI BILAG A/B</td>
<td>9.7</td>
<td>6</td>
<td>0.6</td>
<td>7.5</td>
</tr>
<tr>
<td>PGA % no worse</td>
<td>3.2</td>
<td>9</td>
<td>0.3</td>
<td>7.8</td>
</tr>
<tr>
<td>UFE Potential SAEs</td>
<td>0.0</td>
<td>-100</td>
<td>0.0</td>
<td>7.8</td>
</tr>
<tr>
<td>UFE Infections</td>
<td>19.2</td>
<td>0</td>
<td>0.0</td>
<td>7.8</td>
</tr>
<tr>
<td>FE QoL</td>
<td>0.2</td>
<td>-0</td>
<td>-0.000</td>
<td>7.8</td>
</tr>
<tr>
<td>PGA Mean score</td>
<td>2.4</td>
<td>-4</td>
<td>-0.1</td>
<td>7.7</td>
</tr>
<tr>
<td>UFE Sensitivity Reaction</td>
<td>3.8</td>
<td>-15</td>
<td>-0.6</td>
<td>7.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>100.0</td>
</tr>
<tr>
<td>7.2</td>
</tr>
</tbody>
</table>

Advantages of Placebo

Advantages of 10mg
Uncertainty: Sensitivity analysis

Vary the weight on an effect or node (here, UFE) over its entire range from 0 to 100.

Crossover indicates a change in the most preferred option.

The cumulative weight on UFES would have to increase from about 23 (vertical red line) to nearly 80 before the placebo would be preferred to the 10mg dose.
Risk Attitudes of European assessors

Web-based questionnaire

80 European assessors

3 mock clinical dossiers: cardiology, oncology, CNS

8 benefit-risk rating scales

Effect of experience on perceived level of risk

More experience, more averse to risk

Female assessors less averse to risk than male assessors

The Outcomes

- 9 March 2011: FDA approves the drug.
- 19 May 2011: EMA’s CHMP issues a positive opinion.
- 20 September 2011: NICE “provisionally unable to recommend”.
- 26 April 2012: NICE issue draft guidance “belimumab could not be considered a good use of NHS resources compared with current clinical practice”.
- 18 July 2012: Appeal Panel hears appeals from GSK, Lupus UK and Primary Care Rheumatology Society and upholds two from GSK. NICE remits appraisal to the Appraisal Committee.
- 23 July 2013: The appeal fails to change the position. A second public consultation is initiated (closes 13 August). Drug cost per administration to 65-76kg patient is estimated at £769.50.
To sum up ...

- **MCDA does not give the ‘right’ answer.** Nothing can because there is no right answer.
- **MCDA does provide a useful tool for thinking, a structure for rational debate and a serious guide to decision making.**
- It is a model that ‘illuminates’; it provides clarity.
- **MCDA enables rapid exploration of different perspectives on the issues and of uncertainty in the data.**
- **MCDA can be expanded with related model types**
- However, **MCDA requires careful design of social processes: engaging the right people in the right way at the right time.**
When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the state of science, whatever the matter may be.

Lord Kelvin, 1883
A guide to further reading


MCDA in Chapter 3, prioritisation and resource allocation in Chapter 14.


Download free from [http://eprints.lse.ac.uk/12761](http://eprints.lse.ac.uk/12761)

*.MCDA is described in Chapter 6.*


*Shows how to articulate values and make wise decisions.*