Multiple criteria decision analysis (MCDA) for drug benefit-risk assessment

Gert van Valkenhoef

2014-11-17 © Brown Biostats
Section 1

Introduction
Credits

This is joint work with

- Tommi Tervonen (Erasmus University Rotterdam, NL)
- Douwe Postmus (UMCG)
- Hans Hilleg (UMCG and European Medicines Agency)

Mostly done during my PhD (2009-2012)
Drug benefit-risk decision making

- To choose between alternative treatments
  - prescription
  - marketing authorization
  - reimbursement policy
- Trading off benefit and risks (harms) of treatment
- Example: in anti-clotting agents, trade off
  - Benefit: prevention of thrombosis
  - Risk/Harm: major bleeding
Problems

- Decisions are often unstructured and lack transparency
  - benefit-risk valuation is not explicit
  - even if evidence base is explicit and rigorous
- Lack of adequate methods to take uncertainty into account
- Unwillingness to commit to trade-offs
The Escher Project: scientific evidence and dialogue on the development and regulation of medicines.

The Escher Project is a Top Institute Pharma funded public-private partnership that studies medicine development and the European regulatory system for medicines. By combining expertise from various disciplines and backgrounds, The Escher Project has provided solid scientific evidence and supported a results-oriented dialogue on reform. By contributing to resolving bottlenecks in the system, The Escher Project aims to stimulate innovative medicine development and regulation, and bring safe and effective medicines to patients faster.
Bridging the gap between aggregated clinical data and evidence-based drug regulation using state of the art methods for benefit risk decision making

Implementing in usable software to be deployed not only in the regulatory domain but also in the decision-making domain of e.g. HTA agencies, hospital and community pharmacists, medical specialists, general practitioners and patients
ADDIS: Aggregate Data Drug Information System

Key ingredients:
- Structured database of clinical trials data
- On-the-fly statistics, evidence synthesis
- Benefit-risk decision modelling / decision support

ADDIS 1.x is free/open source software
- Download: https://drugis.org/addis

ADDIS 2.0 under development (IMI GetReal)
- Collaborative web-based platform
- Bridging efficacy/safety to relative effectiveness
Making better use of clinical trials

Health care policy decision makers routinely evaluate the health impact of alternative treatment options. Here, benefit-risk assessment is key, which consists of weighing the favorable effects (benefits) and unfavorable effects (risks) of the alternatives. For example, before a new drug is allowed onto the market, regulators evaluate its benefit-risk balance in comparison to placebo or competing drugs.

Ideally, benefit-risk assessments are based on the best available evidence, typically meaning randomized controlled trials. However, finding the evidence and explicitly linking it for assessment is complicated by several factors. First, the results of clinical trials are mainly made available in text-based documents that cannot be processed automatically. Second, the data from such trials must be combined into a consistent basis for benefit-risk analysis. Third, quantitative decision models are required to directly link decisions to the underlying evidence and to make trade-off decisions explicit.

This thesis addresses these topics through the development of the Aggregate Data Drug Information System (ADDIS), an integrated system for decision support, based on databases of structured clinical trials data. Novel algorithms are presented to automate network meta-analysis to combine clinical trial results and multi-criteria decision models are developed to support benefit-risk assessment.

Complete Dissertation
Overview of the initiatives since 2000

- **FDA**: Federal Drug Administration
- **EMA**: European Medicines Agency
- **CASS**: Taskforce of representatives from Health Canada, Australia’s Therapeutic Goods Administration, Swissmedic and the Singapore Health Science Authority
- **COBRA**: Consortium on Benefit-Risk Assessment
- **PhRMA BRAT**: Pharmaceutical Research and Manufacturers of America Benefit-Risk Action Team
- **IMI PROTECT**: Innovative Medicine Initiative Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium
- **Eu2P**: European programme in Pharmacovigilance and Pharmacoepidemiology
- **ISPOR**: International Society for Pharmacoeconomics and Outcomes Research
- **CMR**: Centre for Medical Research
- **CIRS**: Centre for Innovation in Regulatory Science
- **UMBRA**: Unified Methodologies for Benefit-Risk Assessment
- **ESFP**: European Federation of Statisticians in the Pharmaceutical Industry

ESFP/PSI Benefit-Risk Special Interest Group
Section 2

MCDA
Multiple Criteria Decision Analysis

Structuring and solving decision problems involving multiple criteria
Advantages of MCDA

- Problem structuring
- Making (subjective) value judgments explicit
- Transparently combine value judgments and evidence
- Focus and language for discussion
- Complement and challenge intuition
The process of MCDA

**Problem structuring phase**
- **Goal**: A clear formulation of objectives, reached through consensus between stakeholders. Identification of all aspects relevant to the decision problem
- **Methods**: Structured discussions, focus group meetings. First stage of divergent thinking, followed by convergent phase aimed at structuring the problem
- **Intermediate outcomes**: - Hierarchical structure of problem with decision criteria (decision tree) - Set of decision alternatives

**Scoring phase**
- **Goal**: To identify and incorporate all available data and knowledge on the decision problem
- **Methods**: Searches for existing data in literature and databases. Interviews to elicit expert opinion
- **New information obtained in the scoring phase may require a re-structuring of the decision problem**

**Preference modeling phase**
- **Goal**: To formalize the decision maker's preference structure in order to identify the best alternative or to rank them from best to worse
- **Methods**: The problem is decomposed into a set of smaller subproblems, for which preference information is obtained. Using a mathematical function this is compiled into a preference for the full problem

**Decision gate**
- Does the scoring table indicate a dominating alternative, or is further analysis necessary to support a decision?

**Outcome**
- Decision based on preference model

**Intermediate outcome**
- Table with every alternative scored on each criterium, either cardinal (values) or ordinal (ranking)

**Outcome**
- Decision based on information in scoring table
Focus on (normative) quantitative models
  - Process will not be discussed further
SMAA: taking into account uncertainty
Combining evidence synthesis and SMAA
Optimizing preference elicitation
Section 3

SMAA
Multi-attribute Value Theory (MAVT)

- Simple normative MCDA model
- Alternatives $i \in I = \{1, \ldots, m\}$
- Criteria $j \in J = \{1, \ldots, n\}$
- Partial value functions over deterministic outcomes

$$v_j : X_j \mapsto [0, 1]$$

- Aggregated using additive value function

$$v(x, w) = \sum_{j \in J} w_j v_j(x_j)$$

- Weights non-negative and normalized
MAVT: preferences

- Attractiveness of levels of $x_j$: $v_j(x_j)$
  - Often taken to be linear
  - Otherwise, can be elicited
  - Typically uncontroversial

- Trade-offs between criteria scale swings: $w$
  - More influential
  - More controversial
MAVT limitations / assumptions

- Additive model: criteria are preferentially independent
  - Counter-example: wine and dish choice
  - Remedy 1: non-additive models
  - Remedy 2: reformulate problem
- Does not deal with uncertain outcomes
  - Remedy: MAUT or SMAA
- Requires preferences to be specified exactly
  - Remedy: SMAA
Uncertainty in measurements and preferences

Joint density for the criteria measurements \((m \times n)\)

Density (uniform) over space of feasible weights \(((n - 1)\text{-simplex})\)
Imprecise preference information

\[
\begin{align*}
W &\subseteq W' \\
W_1 &\geq W_2 \\
W_2 &\geq W_3 \\
w_1 &\geq w_2 \\
w_2 &\geq w_3 \\
w_3 &\geq w_1
\end{align*}
\]
Preference elicitation

- Preference elicitation is challenging
- Many types of information are possible:
  - Ordinal information
  - Exact trade-off ratios
  - Imprecise trade-off ratios
  - Pair-wise comparisons
- I will not go into detail in this presentation
Stochastic Multi-criteria Acceptability Analysis (SMAA)

- Stochastic simulation to deal with uncertainty
- Computes rank probabilities (and other metrics)
- Interpretation of probabilities left to DM
  - Risk attitude is not modelled (see MAUT)
SMAA rank acceptabilities

\[
\text{rank}(i, x, w) = 1 + \sum_{k=1}^{m} [v(x^k, w) > v(x^i, w)]
\]

\[
P(\text{rank}_i = r) = \int_{x \in X} \int_{w \in W} f_X(x)f_W(w)[\text{rank}(i, x, w) = r]dw dx
\]

[.] is the Boolean indicator.
SMAA computations

- Done using stochastic simulation
- If all measures and preferences are iid, $10^4$ iterations suffice
- Sampling from $W$: use Hit-and-Run (Tervonen et al. 2012; van Valkenhoef et al. 2014)
Example: anti-thrombolytics

- Which anti-thrombolytic is best?
  - Prevention of thrombosis (DVT, proximal and distal)
  - Increased chance of bleeding

  - Solved using cost-effectiveness techniques
  - Ignored distal DVT
  - Equivalent to SMAA limited to 2 dimensions
  - Here reproduced using SMAA

- Based on data from Geerts et al. (1996)
### Example: anti-thrombolytics

<table>
<thead>
<tr>
<th>Event</th>
<th>Data</th>
<th>Beta distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( r )</td>
<td>( r/n )</td>
</tr>
<tr>
<td><strong>Heparin (( n = 136 ))</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any DVT</td>
<td>60</td>
<td>0.441</td>
</tr>
<tr>
<td>Proximal DVT</td>
<td>20</td>
<td>0.147</td>
</tr>
<tr>
<td>Distal DVT</td>
<td>40</td>
<td>0.294</td>
</tr>
<tr>
<td>Major bleeds</td>
<td>1</td>
<td>0.007</td>
</tr>
<tr>
<td><strong>Enoxaparin (( n = 129 ))</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any DVT</td>
<td>40</td>
<td>0.310</td>
</tr>
<tr>
<td>Proximal DVT</td>
<td>8</td>
<td>0.062</td>
</tr>
<tr>
<td>Distal DVT</td>
<td>32</td>
<td>0.248</td>
</tr>
<tr>
<td>Major bleeds</td>
<td>5</td>
<td>0.038</td>
</tr>
</tbody>
</table>
Benefit-risk trade-off

Incremental benefit–risk of Enoxaparin over Heparin

Benefit = Proximal DVT, Risk = Major bleeds
Model

\[ v_p(x_p) = 1 - \frac{x_p - 0.0}{0.25 - 0.0} \]

\[ v_b(x_b) = 1 - \frac{x_b - 0.0}{0.1 - 0.0} \]

\[ v(x, w) = w_p v_p(x_p) + w_b v_b(x_b) \]

Note: we’re allowing occasional partial values outside \([0, 1]\).
No preferences

First-rank acceptability of Enoxaparin

Probability Enoxaparin is preferred

Weight for Proximal DVT
Ordinal preferences

First-rank acceptability of Enoxaparin

Probability Enoxaparin is preferred

Weight for Proximal DVT
Ratio bound preferences

First-rank acceptability of Enoxaparin

Probability Enoxaparin is preferred

Weight for Proximal DVT
Extending to more dimensions

- SMAA extends naturally to > 2 criteria/alternatives
- The presented plots do not - other visualizations / metrics
- Extended the case to include distal DVT
  - Alternatives are poorly discriminated on distal DVT
  - Preferences indicated it is of marginal concern
  - Original decision not to include it was reasonable
Other metrics - inverse approach

**Rank acceptability index**  share of weights and measurements making each alternative have ranks $1, \ldots, m$

**Central weight vector**  center of gravity of the favourable weight space

**Confidence factor**  probability for an alternative to be preferred when preferences equal its central weight vector
Section 4

MCDA/SMAA and EBM
Challenges of MCDA in EBM

1. Defining non-overlapping criteria
2. Translating relative measures to absolute scales
   - e.g. odds ratios from meta-analysis
3. Defining suitable scale ranges
4. Eliciting preference information
5. Incorporating uncertainty

We will now focus on (2).
Example: anti-depressants

- Problem: choosing the best anti-depressant
- Alternatives: Placebo, Fluoxetine, Paroxetine, Sertraline, Venlafaxine
- Criteria: treatment response, diarrhea, dizziness, headache, insomnia, nausea
- Evidence base: Hansen et al. (2005) systematic review
Evidence network

- Placebo
- Fluoxetine
- Paroxetine
- Sertraline
- Venlafaxine
- Placebo

Each node represents a treatment, and the edges indicate relationships or comparisons between them.
Network meta-analysis results

Top-left to bottom-right: Fluoxetine, Paroxetine, Sertraline, Venlafaxine
Interpreting the results

- Clearly, the BR profiles differ
- But how should these differences be interpreted?
- Trade-offs on the odds ratio scale are meaningless!
Relative to absolute (1/2)

Treatment effects (log-odds ratios) relative to treatment 1:

\[
\begin{pmatrix}
  d_{1,2} \\
  \vdots \\
  d_{1,m}
\end{pmatrix} \sim \mathcal{N}
\left(
\begin{pmatrix}
  \nu_2 \\
  \vdots \\
  \nu_m
\end{pmatrix}, \Sigma
\right)
\]

Externally derived log-odds estimate for treatment 1:

\[
\theta_1 \sim \mathcal{N}(\mu, \sigma^2)
\]
Relative to absolute (2/2)

Combine into estimate of absolute log-odds for any treatment:

\[
\begin{pmatrix}
\theta_2 \\
\vdots \\
\theta_m
\end{pmatrix}
| \theta_1 \sim \mathcal{N}
\begin{pmatrix}
\theta_1 + \nu_2 \\
\vdots \\
\theta_1 + \nu_m
\end{pmatrix}, \Sigma
\]

Which is transformed to the absolute risk:

\[
p_i = \logit^{-1}(\theta_i)
\]
Example: continued

- Estimated baseline effects for all criteria
- Sampled from absolute effect distributions
- Set all ranges to [0, 1]
  - This simplified preference elicitation
  - But decreased value of incomplete preference information
  - Criteria ranges were already very wide
Results: preference free

Fluoxetine  Paroxetine  Placebo  Sertraline  Venlafaxine

- Rank.1
- Rank.2
- Rank.3
- Rank.4
- Rank.5
Results: central weights

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAM-D</td>
<td>0.00</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>0.05</td>
</tr>
<tr>
<td>Dizziness</td>
<td>0.10</td>
</tr>
<tr>
<td>Headache</td>
<td>0.15</td>
</tr>
<tr>
<td>Insomnia</td>
<td>0.20</td>
</tr>
<tr>
<td>Nausea</td>
<td>0.25</td>
</tr>
<tr>
<td>Fluoxetine</td>
<td>0.30</td>
</tr>
<tr>
<td>Paroxetine</td>
<td>0.35</td>
</tr>
<tr>
<td>Placebo</td>
<td>0.00</td>
</tr>
<tr>
<td>Sertraline</td>
<td>0.12</td>
</tr>
<tr>
<td>Venlafaxine</td>
<td>0.57</td>
</tr>
</tbody>
</table>

Fluoxetine 0.12
Paroxetine 0.57
Placebo 0.99
Sertraline 0.55
Venlafaxine 0.63
Results: severe depression (ordinal preferences)
Section 5

Optimizing preference elicitation
Optimizing preference elicitation

- There is an effort / precision trade-off
- “Traditional” elicitation methods can be very hard
- Pair-wise comparisons are easier, but never-ending
- How much preference information is enough?
- Uncertainty due to measurements or preferences?
Entropy as a framework

- Entropy can quantify uncertainty over decision metrics
  - And how much it can be reduced
- This allows to compare quality of questions
- And to determine when no more questions should be asked
Decision entropy

\[ H(Y : w \in W') = - \sum_{y \in Y} p(y|w \in W') \log p(y|w \in W') \] .

Where \( Y \) is the space of all possible decision outcomes (e.g. all \( m! \) rankings or the \( m \) possible best alternatives).
Pair-wise comparisons

- A pair-wise comparison question $Q$ compares two (fictional, deterministic) alternatives
- The answer bisects the current weight space $W'$:
  
  $$A(Q) = \{W_A, W_B\} ; \quad W_A \cup W_B = W'$$

- The ‘quality’ of a question is entropy conditional on the answer:
  
  $$H(Y|A(Q)) = \sum_{W'' \in A(Q)} p(W'')H(Y : w \in W'')$$

- The optimal question minimizes $H(Y|A(Q))$
Mutual information quantifies how much uncertainty is due to the weight vector:

$$I(Y; w : w \in W') = H(Y : w \in W') - H(Y|w : w \in W')$$

Which can be normalized to an uncertainty coefficient:

$$R = I(Y; w : w \in W')/H(Y : w \in W')$$

- $R = 1$: more information can eliminate all uncertainty
- $R = 0$: uncertainty can not be reduced further
Example results
Limitations

- Primarily useful for pair-wise comparison questions
  - Could also be used to stop structured elicitation early
- Computationally expensive
  - But many optimizations still possible
Section 6

Conclusion
### MCDA for benefit-risk assessment

- Structure the problem
- Make value judgments explicit
- Combine value judgments and evidence
- Support, not dictate decisions
SMAA for benefit-risk assessment

- Account for uncertainty in measurements
- Allow imprecise / partial preferences
- Propagate uncertainties to results
But don’t overdo it

Don’t use MCDA when

- the decision is clear-cut
- one criterion dictates decision
- decision makers aren’t convinced of usefulness
Discussion points

- How do discrete choice experiments compare to preference elicitation with pair-wise comparisons?
- What is the role of the QALY?
  - Is the QALY "a form of MCDA"?
  - Should all decisions discount effects over time?
  - Can the QALY be a criterion in a broader MCDA?
Thank you! Questions?

https://drugis.org/publications