Establishing the value of innovative medicines
While dealing with uncertainty using multicriteria decision analysis (MCDA)

19 April 2011

CAPT Annual Meeting, Ottawa

EVIDEM Collaboration - Board of Directors
Rob Baltussen PhD, Radboud University, Netherlands (formerly WHO)
Renaldo Battista MD, University of Montreal, CHU Ste Justine, Canada
Mireille M Goetghebeur PhD, BioMedCom Consultants inc, CHU Ste Justine, Canada
Paul Kind PhD, University of York, UK
Sharon Kletchko MD, Nelson Marlborough District Health Board, New Zealand
Mark Legault MA, Pfizer Canada
Jacqui Miot PhD, University of Pretoria, South Africa
Donna Rindress PhD, BioMedCom Consultants inc, Canada
Outline

- Overview
- Defining value - MCDA
  - Decision criteria
  - Weighting techniques
  - Scoring scales
  - Mathematical model & qualitative considerations
- Exploring uncertainty - HTA
  - By-criterion evidence synthesis
  - Interactive open web registry
- Users & applications
- Advantages & challenges
- Future developments
Overview - EVIDEM Collaboration

- Not-for profit independent legal entity

- **Object:** promote health and efficient decisionmaking via systematic assessment of evidence and value of healthcare interventions
  - **MCDA based decision making framework tools** freely available under Creative Commons license
  - Collaborative development
    - Open tools **regularly upgraded** based on academic research and feedback from users
    - **Open web** registry

- On-going collaborations
  - Canada, Italy, Netherlands, New Zealand, South Africa, UK, USA
  - Tools used and tested by government agencies and academic centers

- Funding & support: Canadian Institutes of health Research (CIHR), Pfizer Canada (start up), BioMedCom (in kind)

---

Structuring the natural thinking process to define value

→ MCDA
Defining value of interventions
Criteria?

Which criteria define the most valuable healthcare interventions?

- Elicit from evaluators
- Use an existing set of criteria & adapt
  
  - MCDA-based EVIDEM framework
Defining value of interventions
Criteria?

Apply MCDA principles
- criteria should be complete,
- with minimum overlap
- mutually independent
- operationalizable*

www.communities.gov.uk/pub/252/MulticriteriaanalysismanualPDF1380Kb_id1142252.pdf
Defining value of interventions
EVIDEM conceptual approach

» Develop a universal generic tool and a contextualization tool

- MCDA Core Model
- Contextual Tool
**MCDA Core Model**

What **should we do** for sustainable healthcare systems?

15 universally **normative criteria**

- Highest rank/value or priority should be given to interventions
  - For severe disease (D1)
  - For common disease (D2)
  - For disease with many unmet needs (C2)
  - Recommended in consensus guidelines by experts (C1)
  - Conferring major improvement in efficacy/effectiveness over standard of care (I1)
  - Conferring major improvement in safety & tolerability over standard of care (I2)
  - Conferring major improvement of patient perceived health over standard of care (I3)
  - Either conferring major risk reduction (T1) or major alleviation of suffering (T2)
  - That results in savings in treatment expenditures (E1) as well as other medical and non medical expenditures (E3); **cost-effective (E2)**
  - For which there is sufficient data (Q1), that is fully reported (Q2) and valid and relevant (Q3)

---

*Cost-effectiveness is a composite of some elements of other criteria and does not comply with the non-redundancy design requirement of MCDA. It may be included in the framework since many decisionmaking processes currently rely on this composite measure.*
Contextual Tool
What is our context and what can be done?

6 criteria

- Define objectives & priorities – 2 contextual normative criteria
  - Alignment with scope and mission of health care system/plan (Et1)
  - Defining country/institutional priorities for populations & access (Et2)

- 4 Feasibility criteria
  - Exploring opportunity costs (forgone interventions) and affordability (Et3)
  - Verifying system capacity (e.g., infrastructure, skills) and appropriate use of intervention (O1)
  - Assessing political/historical context (e.g. cultural acceptability, precedence) (O2)
    - Impact of intervention on innovation and research?
  - Realizing pressures/barriers from healthcare stakeholders (O3)
Overall EVIDEM framework structure

Clustering criteria

**MCDA core model**

Universally normative criteria

- **Disease impact**
  - Disease severity (D1)
  - Size of population affected by disease (D2)

- **Context of intervention**
  - Clinical guidelines (C1)
  - Comparative intervention limitations (C2)

- **Intervention outcomes**
  - Improvement of efficacy/effectiveness (I1)
  - Improvement of safety and tolerability (I2)
  - Improvement of patient reported outcomes (I3)

- **Type of benefit**
  - Public health interest (e.g., prevention, risk reduction) (T1)
  - Type of medical service (e.g., symptom relief, cure) (T2)

- **Economics**
  - Budget impact on health plan (cost of intervention only) (E1)
  - Impact on other spending (e.g., hospitalization, disability) (E2)
  - Cost-effectiveness of intervention (E3)

- **Quality/uncertainty of evidence**
  - Adherence to requirements of decisionmaking body (Q1)
  - Completeness and consistency of reporting (Q2)
  - Relevance and validity of evidence (Q3)

**Contextual tool**

Context & feasibility criteria

- **(quantitative)**

**(qualitative)**

- **Ethical framework**
  - Utility – Goals of healthcare (Et1)
  - Fairness – Population priority & access (Et2)
  - Efficiency – Opportunity costs & affordability (Et3)

- **Other system-related criteria**
  - System capacity and appropriate use (e.g., infrastructure, skills) (O1)
  - Stakeholder pressures (O2)
  - Political/historical context (e.g. precedence) (O3)

*Based on three principles; since often conflicting, clearly identify trade-offs and legitimize decision by engaging a broad range of stakeholders & explaining decision; legitimizing decision is key to provide accountability for reasonableness (A4R)
Adaptation
Define value in your context

- **Include priorities** defined using the contextual tool as additional criteria of the MCDA Core Model (e.g., rare diseases)

- **Transfer** other contextual criteria in the MCDA core model

- **Expand** criteria into subcriteria*

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Possible sub criteria</th>
</tr>
</thead>
</table>
  | E3 Impact on other spending           | • Impact on primary care expenditures  
  |                                        | • Impact on hospital care expenditures  
  |                                        | • Impact on long-term care expenditures  
  |                                        | • Impact on productivity  
  |                                        | • Financial impact on patients  
  |                                        | • Financial impact on caregivers  

- **Remove** criteria

  *A number subcriteria are available in tools to expand model - apply MCDA principles*
Not all criterion are equal

Disease severity

? =

Improvement of efficacy
Measuring value
Weight elicitation techniques*

- Capture individual perspective on relative importance of criteria independently of healthcare interventions
- No gold standard

⇒ Simple techniques
- EVIDEM
  - EVIDEM
  - Kepner-Tregoe Analysis (KTA)
  - Direct point allocation

⇒ More complex
- Analytical hierarchy process (AHP)
- Best/worst scaling
- Conjoint analysis

⇒ Adapt to user preference/context

---

Measuring value
Scoring scale?

- **Measure performance** of intervention
- Need to define:
  - Type of scale/number of options
  - Scale anchors for each criteria
- Simple approach
  - EVIDEM

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Scoring scale</th>
</tr>
</thead>
</table>
| Example Disease severity | ■ 0 - not severe  
|                        | ■ 1                               |
|                        | ■ 2                               |
|                        | ■ 3 - very severe                 |

- More complex (e.g., more scale options, boolean operators for each option)

- Adapt to user preference/context
Appraising interventions
Mathematics & qualitative considerations

Type of mathematical model
- Simple linear model (combine normalized weights and scores) to calculate perceived value of intervention

Ranking of healthcare interventions
- Quantitative evaluation
- Combined with qualitative impact of context

<table>
<thead>
<tr>
<th>Contextual criteria</th>
<th>Qualitative impact on appraisal/ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example</td>
<td>❑ Negative</td>
</tr>
<tr>
<td>Political/historical context</td>
<td>❑ Neutral</td>
</tr>
<tr>
<td></td>
<td>❑ Positive</td>
</tr>
</tbody>
</table>
Finding the evidence
(scientific and colloquial)
& assess uncertainty

⇒ HTA
**HTA objectives**

- Inform healthcare decisionmaking/priority setting
- Systematic reviews & synthesis of evidence
- Multidisciplinary and pragmatic
- Dissemination and reaching out

By-criterion HTA report

- Detailed methodology in agreement with good HTA practices*
  - Comprehensive literature review
  - Analyses & synthesis of evidence (scientific and colloquial)
  - Assessment of the quality of studies (clinical, economic, epidemiologic)
  - Validation by experts

- Data synthesized for each criterion (multidisciplinary)
  - Highly synthesized (quick grasp)
  - Details with evidence tables
  - Full text source documentation (hyperlinks)

- Web based (dissemination)
  - Open source software (Tikiwiki)

*Busse et al. IJTAHC 2002; 18(2): 361-422.
Web registry

http://www.evidem.org/evidem-collaborative.php

Demo: Interactive prototype

https://www.evidem.org/tiki/?page=DEMO-main
**Overview of intervention**

**Last Update:** April 2009  
**Disease:** Turner syndrome (TS)  
**Intervention:** Growth Hormone (GH)  
**Setting:** Canada

**Drug class:** Polypeptide hormone  
**Indication:** treatment of short stature in girls with Turner Syndrome  
**Administration:** subcutaneous injection 3 to 7 days a week  
**Intervention duration:** Needs to be established. Initiate as soon as growth failure demonstrated until satisfactory height reached (6 years of treatment startin at 10 years)  
**Comparator(s):** No treatment  
**Economic burden of illness:** No data available

---

**Interactive by-criterion HTA report – high level synthesis - excerpt**

<table>
<thead>
<tr>
<th>Intrinsic criteria</th>
<th>Highly synthesized information</th>
<th>Score</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Disease criteria</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| D1 Disease severity | **Female-specific genetic disorder** characterized by short stature, cardiovascular defects, absence of puberty, infertility, increased risk of diabetes, defects in visual-spatial organization and nonverbal problem-solving, and decreased life expectancy.  
[see details] | ○ 0 Not severe (minor inconvenience)  
○ 1  
○ 2  
○ 3 Very Severe | □ Low Score due to data - specify |
| D2 Size of population | **Prevalence:** 40/100,000 female adults. | ○ 0 Very rare disease  
○ 1  
○ 2  
○ 3 Common disease | □ Low Score due to data - specify |
| **Context of intervention** |                                                                                               |                                                                      |          |
| C1 Clinical guidelines | **International guidelines** for Canadian Guidelines: Candidate GH treatment as seen. | ○ 0 No recommendation |          |
Interactive by-criterion HTA report
Details on evidence - example

D1 - Disease severity

Turner syndrome is a female-specific genetic disorder (complete or partial loss of one of the X chromosome) characterized by short stature and presenting a wide spectrum of abnormalities, including cardiovascular defects (17-45%), lymphedema, gonadal dysgenesis (90% requiring hormone replacement therapy to induce puberty), infertility, miscarriage, hypothyroidism (15-30%), risk of obesity, ophthalmic defects, hearing problems and ear malformations, gastrointestinal and renal manifestations (Bondy 2007, Sybert et al. 2004). Patients are at increased risk of impaired glucose tolerance and diabetes (Hjerrild 2008, Holl 1994). Overall, cancer risk appeared not to be significantly increased; increased risks were reported in some studies for brain and nervous system tumors, and for colon and rectal cancer (Stochholm 2006, Schoemaker 2008, Hjerrilc 2008). Defects in visuo-spatial organization and nonverbal problem-solving affect most patients with TS; in addition, impaired psychomotor and social functioning have been reported (Bondy 2007, Sybert et al. 2004).

In young patients, psychosocial issues arise: impaired peer relationship, teasing, social isolation, anxiety, shyness, and poor self-esteem (Bondy 2007, Sybert et al. 2004, Schmidt et al, Busschbach et al. 1998). In audiotaped interviews, Turner Syndrome patients reported infertility as their biggest concern (range: 36% of girls aged 7–13 yrs to 74% of adults aged 20–39 yrs; Sutton et al. 2007). However, many Turner Syndrome patients of all ages reported to be bothered by short stature (30% girls, 44% adolescents and adults, 53% mature adults 40–59 yrs; Sutton et al. 2007); 44% of 25 adult Turner Syndrome patients had a height below the 3rd percentile (Busschbach et al. 1998). Short stature is a major physical handicap in adolescents and adults. Life expectancy is decreased in women with Turner abnormalities (Stochholm et al. 2006; Sybert et al. 2004; number of deaths / expected number of deaths =

Return to DEMO Scoring intervention (NCDA Manual)
Return to DEMO Menu

Content of the Collaborative
## Interactive by-criterion HTA report

### Links to quality assessments - example clinical data

<table>
<thead>
<tr>
<th>Intervention outcomes</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improvement of efficacy/ effectiveness</td>
<td>- 4 placebo controlled RCTs (2-year to 11-year treatments; N=42 to 104, 1 in Canada, 3 in USA): Final height of treated patients = 147 cm to 150 cm; difference with untreated = 7 cm. Observational studies (2-year to 8-year treatments, N=26 to 123, 1 in Germany, 1 in Greece, 1 in Israel, 3 in Italy): Final height of treated patients = 148 cm to 151 cm; difference with controls = 2.1 to 6.8 cm. Example of critical analysis.</td>
</tr>
<tr>
<td>Improvement of safety &amp; tolerability</td>
<td>- Common AEs (from RCTs - frequency at least twice of placebo): Surgeries (50%), ear problems (6% to 47%), joint (13.5%) and respiratory (11%) disorders, sinusitis (18.9%). Serious AEs (from registries, no control data): Intracranial hypertension (0.2%), slipped capital femoral epiphysis (0.2 - 0.3%), scoliosis (0.7%), pancreatitis (0.1%), diabetes mellitus (0.2 to 0.3%), cardiac/aortic events (0.3%), malignancies (0.2%). Warnings: Scoliosis, slipped capital femoral epiphysis, intracranial hypertension, ear disorders, cardiovascular disorders, autoimmune thyroid disease, insulin resistance.</td>
</tr>
<tr>
<td>Improvement of patient reported outcomes</td>
<td>- Inconclusive data: 1 RCT (2-year treatment data, N=28, Canada): higher rating on questionnaire by GH treated patients versus untreated for some domains but not for others. 2 observational studies: no significant differences on SF-36 dimensions in one study (5-year treatment, N=568, France) and significant differences in another (7-year treatment, N=29, Holland); other questionnaires, non significant differences. Conveniences: Subcutaneous injection 3 days a week or daily.</td>
</tr>
<tr>
<td>Type of benefit</td>
<td></td>
</tr>
</tbody>
</table>

T1 Public health interest | No data on risk reduction with GH treatment. |
## Interactive by-criterion HTA report
### Quality of evidence assessment - overall clinical data

<table>
<thead>
<tr>
<th>Relevance and validity – clinical data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Disease:</strong> Turner Syndrome (TS)</td>
</tr>
<tr>
<td><strong>Intervention:</strong> recombinant human growth hormone (GH)</td>
</tr>
<tr>
<td><strong>Setting:</strong> Canada</td>
</tr>
<tr>
<td><strong>Series of key studies</strong></td>
</tr>
<tr>
<td>Stephure et al, 2005: Canada - See full assessment</td>
</tr>
<tr>
<td>Quigley et al 2002: US - See full assessment</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of evidence</th>
<th>Question(s)</th>
<th>Rationale</th>
<th>1 [ ] Low relevance/validity 2 [ ] 3 [x] 4 [ ] High relevance/validity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Efficacy/safety data</strong></td>
<td>How relevant is the research program with regard to efficacy and safety? Are conclusions valid over the range of studies (conclusions across studies consistent or conflicting)? Are individual trials relevant and valid? <em>See assessment of individual studies below</em></td>
<td>Overall, randomized controlled trials consistently demonstrate that GH treatment promotes height gain in girls with Turner Syndrome. Some uncertainty remains on extent to which GH may affect final height. High attrition rates were noted in the Canadian clinical trial (Stephure 2005); 2 multiphase trials were missing a control arm (no GH treatment) and chose GH administration mode (frequency of injections) that did not correspond to current practice (Rosenfeld et al. 1998; Quigley et al. 2002). Safety data monitoring is generally limited, despite the numerous warnings and AEs associated with GH treatment in Turner Syndrome populations. (Humatrope PM. 2007; Saizen PM. 2007; Nutropin PM. 2006).</td>
<td>1 [ ] Low relevance/validity 2 [ ] 3 [x] 4 [ ] High relevance/validity</td>
</tr>
</tbody>
</table>
# Interactive by-criterion HTA report

## Quality of evidence assessment - excerpt single study

### Relevance and validity - clinical data - study 1

**Disease:** Turner Syndrome (TS)  
**Intervention:** recombinant human growth hormone (GH)  
**Setting:** Canada

<table>
<thead>
<tr>
<th>Type of evidence</th>
<th>Question</th>
<th>Rationale</th>
<th>Score</th>
</tr>
</thead>
</table>
| **Efficacy/safety data** | Is the study question relevant (choice of comparator, time horizon, patient population, and outcome)? Is the design appropriate (setting & design, sample size, patient allocation, analyses, statistics)? See dimensions below. | This is the only study reporting final height (standard measure of GH effect; (Baxter et al. 2007) as the primary outcome. Although differences between GH treatment and no treatment are meaningful, high attrition rates, especially in the control arm (45%) might bias the study conclusions (Baxter et al. 2007). Authors report that supportive intent-to-treat analysis with conservative assumptions on missing data confirmed significance but no details are provided. | 1 () Low relevance/validity  
2 ()  
3 [X]  
4 () High relevance/validity |

### Dimension | Question | Comment |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Target population</td>
<td>Is the target population relevant (age, gender, disease stage, comorbidities, inclusion criteria/exclusion criteria, setting etc)? Does it correspond to the actual population in which the intervention is envisioned to be used?</td>
<td>Target population is relevant and corresponds to the actual population in which the intervention is indicated; prepubertal Turner Syndrome girls (mean age: 10.3 ± 1.8 yrs; range: 7–13 yrs) with evidence of growth failure (height &lt; 10th percentile) (Bondy 2007). Canadian setting but there is no mention of number and location of centers involved.</td>
</tr>
<tr>
<td>2 Intervention &amp; comparators</td>
<td>Is the intervention in agreement with expected use? Does the choice of comparators reflect standard of care?</td>
<td>GH dose and schedule are in agreement with indication to treat short stature in Turner Syndrome girls (Humatrope PM. 2007; Saizen PM. 2007; Nutropin PM. 2006). Comparator is no treatment (standard of care).</td>
</tr>
</tbody>
</table>
| 3 Outcome measures | Are the selected outcomes measures (efficacy, safety and PRO) relevant? Are rationales for outcomes selection valid? Are the instruments/methods/units used to measure outcomes (efficacy, PRO) valid? | The primary outcome is final height (cm), which is the gold standard measure of GH effectiveness (Baxter et al. 2007). Other efficacy analyses are relevant to the assessment of short-term growth, and instruments/units used are valid: height age-specific. Turner Syndrome standard deviation score (SDS; allows comparing to normal population height, height while Turner
## Interactive by-criterion HTA report

### Capture uncertainty

#### Interventions limitations

<table>
<thead>
<tr>
<th>Intervention outcomes</th>
<th>Evidence limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>11 Improvement of efficacy/effectiveness</strong></td>
<td>- 4 placebo controlled RCTs (2-year to 11-year treatments; N=42 to 104; 1 in Canada, 3 in USA): Final height of treated patients ~ 147 cm to 150 cm; difference with untreated ~ 7 cm - Observational studies (2-year to 8-year treatments, N=26 to 123; 1 in Germany, 1 in Greece, 1 in Israel, 3 in Italy): Final height of treated patients ~ 148 cm to 151 cm; difference with controls ~ 2.1 to 6.8 cm. <strong>Example of critical analysis</strong></td>
</tr>
<tr>
<td><strong>12 Improvement of safety &amp; tolerability</strong></td>
<td>- Common year problems (18.9%) - Serious AEs: slipped capital femoral epiphysis (0.2 - 0.3%), scoliosis (0.7%), pancreatitis (0.1%), diabetes mellitus (0.2 - 0.3%), cardiac/aortic events (0.3%), malignancies (0.2%) - Warnings: Slipped capital femoral epiphysis, Intracranial hypertension, ear disorders, cardiovascular disorders, autoimmune thyroid disease, insulin resistance.</td>
</tr>
<tr>
<td><strong>13 Improvement of patient reported outcomes</strong></td>
<td>- Inconclusive data: 1 RCT (2-year treatment data, N=28, Canada): higher rating on questionnaire by GH treated patients versus untreated for some domains but not for others. 2 observational studies: no significant differences on SF-36 dimensions in one study (3-year treatment, N=568, France) and significant differences in another (7-year treatment, N=29, Holland); other questionnaires, non significant differences. - <strong>Example of critical analysis</strong></td>
</tr>
<tr>
<td><strong>Type of benefit</strong></td>
<td>- Public health interest: No data on risk reduction with GH treatment.</td>
</tr>
</tbody>
</table>

#### Specify evidence limitations

- 0 Lower efficacy/effectiveness than comparators presented
- 1
- 2
- 3 Major improvement/efficacy/effectiveness

- 0 Worse patient reported outcomes than comparators
- 1
- 2
- 3 Major improvement

- 0 No risk reduction

EVIDEM
Value of innovation and uncertainty

→ Strike a balance using a framework
Advantages

- Define & measure value
- Identify criteria at play in healthcare decisionmaking
- Allow simultaneous consideration of a wide range of criteria
- Stimulate reflection on perspectives, values and priorities
- Systematize judgment

- Transparent multidisciplinary evidence in a by-criterion HTA report
- Interactive

Challenges

- Criteria selection
- Perception of complexity
- Integration in existing processes
- MCDA estimate may be used as a formula
- Perceived difficulty of breakdown of evidence by criteria
# Users & applications

<table>
<thead>
<tr>
<th>Users</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>❖ Decisionmakers</td>
<td></td>
</tr>
<tr>
<td>Policy (macro/meso)</td>
<td>➢ Priority setting</td>
</tr>
<tr>
<td></td>
<td>➢ Reimbursement (Advisory committees)</td>
</tr>
<tr>
<td>Physicians &amp; healthcare professionals</td>
<td>➢ Clinical practice guidelines (CPGs)</td>
</tr>
<tr>
<td></td>
<td>➢ Seamless access to evidence</td>
</tr>
<tr>
<td>Patients</td>
<td>➢ Access to digested &amp; validated information</td>
</tr>
<tr>
<td>❖ HTA developers</td>
<td>➢ HTA report at criteria level</td>
</tr>
<tr>
<td></td>
<td>➢ Web-based multilevel evidence</td>
</tr>
<tr>
<td>❖ Research</td>
<td>➢ Identify research questions/data needs</td>
</tr>
<tr>
<td></td>
<td>➢ Research planning</td>
</tr>
<tr>
<td></td>
<td>➢ Explore the decisionmaking process</td>
</tr>
<tr>
<td>❖ Developers of new healthcare interventions</td>
<td>➢ Gap analysis</td>
</tr>
<tr>
<td></td>
<td>➢ Positioning</td>
</tr>
<tr>
<td>❖ All</td>
<td>➢ Communication (evidence and values)</td>
</tr>
<tr>
<td></td>
<td>➢ Knowledge translation</td>
</tr>
</tbody>
</table>
Future developments

- **Collaborative studies/applications**
  - Field testing & implementation
  - Methodological development

- **Web registry**
  - Interactive open access web resources
  
  ➔ **Optimize resources, decisions, priority-setting and health**
Acknowledgments: Active members for their contribution to the EVI DEM Collaboration

Thank you

www.evidem.org