INTERNATIONAL METHODS GUIDELINES FOR ECONOMIC EVALUATION
Where Are We Now?

Michael Drummond
Centre for Health Economics,
University of York
Introduction

• Since the PBAC guidelines in 1992, many jurisdictions have developed sets of methods guidelines for economic evaluation

• The PBAC guidelines are currently under revision

• The report of the second ‘Washington Panel’ is recently published (Sanders et al JAMA 2016; 316(10) :1093-1103)
Outline of Presentation

• Aspects of methods guidelines meriting further discussion
  - perspective for the analysis
  - making comparisons between therapies
  - use of surrogate endpoints
  - value assessment frameworks
• Accompanying decision-making processes
• Conclusions
Perspective for the Analysis

- Most sets of guidelines, including those of NICE and the PBAC, indicate a health care perspective for costs.
- Only the guidelines in Sweden and the Netherlands strongly encourage a ‘societal’ perspective.
- The new ‘Washington Panel’ report (Sanders et al. JAMA 2016) now recognizes the need for the presentation of costs by a range of perspectives.
- Almost no discussion of whether a ‘societal’ perspective also implies changing the basis for the valuation of costs.
- Less agreement on the perspective for benefits (e.g., QALYs gained by family members and carers).
- Some sets of guidelines (e.g., NICE) encourage a broader perspective for public health interventions.
Making Comparisons Between Therapies

• Choice of comparator has been a source of contention since pharmacoeconomic guidelines were first produced
• NICE (‘all relevant alternatives’); PBAC (‘main alternative is the therapy most likely to be replaced’); IQWiG/G-BA seems to favour the alternative with the strongest head-to-head clinical data
• Usually amounts to considering the current standard of care in the jurisdiction concerned
• Much depends on the level of faith one has in indirect and mixed treatment comparisons
Use of Surrogates

- Surrogate endpoints have been defined as ‘a biomarker or intermediate endpoint intended to substitute and predict for a patient-relevant final endpoint’
- They are most frequently used when it would be practically impossible to follow patients long enough in a clinical trial (eg Hba1c in diabetes)
- In cancer they are used in trials in all stages of disease, but are most common in early stages of disease, where patients will move on to other lines of therapy when they progress
- Progression-free survival (PFS) and Overall Response Rate (ORR) are commonly-used surrogates in cancer trials
Surrogate Endpoints: Are they Reliable? (1)

- Surrogate endpoints include: tumour reduction and progression free survival (PFS)
- They are intended to be predictive of the primary clinical outcome, i.e., overall survival (OS)
- The European Network of Health Technology Assessment (EUnetHTA) considers surrogate endpoints to be important and admissible in Relative Effectiveness Assessment (REA), as long as they have been validated
- PFS can also independently be accepted as a relevant outcome due to its impact on patient experience (e.g. lesser symptoms and better QoL).
Surrogate Endpoints: Are they Reliable? (2)

- Systematic review to assess the suitability of progression-free survival (PFS) and time-to-progression (TTP) using three validation frameworks

- Considered evidence in colorectal, lung, breast and ovarian cancer, plus renal cell carcinoma and glioblastoma multiforme

- According to IQWiG’s validation framework, only PFS achieved evidence of surrogacy in metastatic colorectal and ovarian cancer treated with cytotoxic agents

Ciani., O et al. *Int J Tech Asses Health Care.* 2014;30(3)
Analysis of 5 years of FDA Approvals Based on Surrogate Endpoints

- 36 drugs were analysed, 19 of which were approved based on rate of response (RR), 17 based on progression-free or disease free survival (PFS or DFS)
- Based on a median follow-up of 4.4 years, only 5 drugs had demonstrated improvement in overall survival in randomized clinical trials
- 18 had failed to show any improvement and 13 had no results
- Crossover was allowed in 11 of 36, but there no significant difference in eventual overall survival between those with and without crossover
- The authors argue that the FDA should determine a timeline for drugs approved on the basis of a surrogate endpoint to prove their effectiveness

In **Modelling methods**:  
5.7.5 Clinical end points that reflect how a patient feels, functions, or how long a patient survives are regarded as more informative than surrogate end points (such as laboratory tests and imaging findings).  
When the use of 'final' clinical end points is not possible and 'surrogate' data on other outcomes are used to infer the effect of treatment on mortality and HRQoL evidence in support of the surrogate-to-final end point outcome relationship must be provided together with an explanation of how the relationship is quantified for use in modelling.  
The usefulness of the surrogate end point for estimating QALYs will be greatest when there is strong evidence that it predicts health-related quality of life and/or survival.  
In all cases, the uncertainty associated with the relationship between the end point and health-related quality of life or survival should be explored and quantified.
Example of NICE’s Acceptance of a Surrogate Endpoint

- In 2012 NICE assessed dasatinib, nilotinib and standard dose imatinib as first-line treatment of chronic phase chronic myelogenous leukemia (CML).
- A systematic review and meta-analysis was undertaken to quantify the association between complete cytogenetic response (CCyR) and major molecular response (MMR) at 12 months and overall survival in patients with chronic phase CML.
- In this case the decision-maker (NICE) accepted the observational association between the surrogates and overall survival and hence the modelled assessments of QALYs gained.

Ciani, O et al Value in Health 2013;16 :1081-9
Value Assessment Frameworks

• Still a mixture of QALY lovers (eg NICE, CADTH, TLV) and QALY skeptics (eg IQWiG/G-BA, Medicare/PCORI)

• Some recent converts to QALYs (eg France and Japan)

• Recent value assessment frameworks in the US either use QALYs (ACC-AHA, ICER), a points system (ASCO, NCCN) or a profile (Premera)
Restrictions on the Use of Cost-Effectiveness in the US

• “The Patient-Centered Outcomes Research Institute . . . shall not develop or employ a dollars per quality adjusted life year (or similar measure that discounts the value of a life because of an individual’s disability) as a threshold to establish what type of health care is cost effective or recommended. The Secretary shall not utilize such an adjusted life year (or such a similar measure) as a threshold to determine coverage, reimbursement, or incentive programs under title XVIII.”

The Patient Protection and Affordable Health Care Act, 2010
Recent Value Frameworks in the US

PREMERA BLUE CROSS
ICER INSTITUTE FOR CLINICAL AND ECONOMIC REVIEW
Comparative Effectiveness
Public Advisory Council

AMERICAN COLLEGE OF CARDIOLOGY

Memorial Sloan Kettering Cancer Center

DrugAbacus

AMERICAN SOCIETY OF CLINICAL ONCOLOGY
Making a world of difference in cancer care

NCCN National Comprehensive Cancer Network®
**ACC/AHA Practice Guideline**

**ACC/AHA Statement on Cost/Value Methodology in Clinical Practice Guidelines and Performance Measures**

<table>
<thead>
<tr>
<th>Value</th>
<th>Cost/QALY</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>&lt;$50k</td>
</tr>
<tr>
<td>Intermediate</td>
<td>$50k-$150k</td>
</tr>
<tr>
<td>Low</td>
<td>&gt;$150k</td>
</tr>
</tbody>
</table>
• Incremental Cost per Outcomes Achieved
  – Long-term perspective
  – Cost per quality-adjusted life year (QALY) gained

  • Associated with high care value
    ▪ <$100,000/QALY

  • Associated with intermediate care value
    ▪ $100-150K/QALY

  • Associated with low care value
    ▪ >$150,000/QALY
### ASCO Value Framework

<table>
<thead>
<tr>
<th></th>
<th>Clinical Benefit</th>
<th>Toxicity</th>
<th>Bonus</th>
<th>Net Health Benefit*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced Disease</td>
<td>0 to 80 points</td>
<td>-20 to 20 points</td>
<td>0 to 30 points</td>
<td>Max 130 points</td>
</tr>
<tr>
<td>Adjuvant Treatment</td>
<td>0 to 80 points</td>
<td>-20 to 20 points</td>
<td></td>
<td>Max 100 points</td>
</tr>
</tbody>
</table>

- *relative to an RCT comparator
- costs: drug acquisition, patient cost-sharing
## Premera Value Matrix

<table>
<thead>
<tr>
<th>Category</th>
<th>Factor</th>
<th>Evaluation of Relevant Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical Benefit</strong></td>
<td>Research Question</td>
<td>Strength of Evidence</td>
</tr>
<tr>
<td></td>
<td>Safety</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Efficacy</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Effectiveness</td>
<td>B</td>
</tr>
<tr>
<td><strong>Cost-Effectiveness Analysis</strong></td>
<td>Base Case</td>
<td>$20,000-$30,000/QALY</td>
</tr>
<tr>
<td></td>
<td>High Estimate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low Estimate</td>
<td>&lt;$10,000/QALY</td>
</tr>
<tr>
<td><strong>Societal Values</strong></td>
<td>Ethical Issues</td>
<td>Affordability of the ________’s. While cost effective, the costs of these drugs make widespread coverage impossible within a financially responsible manner. Therefore prioritization of treatment given the very high cost is essential. Prioritization needs to guard against discrimination against patients because others disapprove of the behavior that led to infection (needle sharing, etc.)</td>
</tr>
<tr>
<td></td>
<td>Rare Disease</td>
<td>□ No □ Yes __% of the population has ____________________</td>
</tr>
<tr>
<td></td>
<td>Unmet Need</td>
<td>□ No □ Yes More effective</td>
</tr>
<tr>
<td></td>
<td>Other Societal Considerations</td>
<td>Potential for __________________________. Potential societal impact of __________________________ is substantial.</td>
</tr>
<tr>
<td><strong>Budget Impact Analysis</strong></td>
<td>Regulatory Issues</td>
<td>None noted.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pharmacy Budget Impact</td>
</tr>
<tr>
<td></td>
<td>Base Case</td>
<td>SX</td>
</tr>
<tr>
<td></td>
<td>High Estimate</td>
<td>SY</td>
</tr>
<tr>
<td></td>
<td>Low Estimate</td>
<td>SZ</td>
</tr>
</tbody>
</table>
# Global Scoring Systems in France and Germany

## France

<table>
<thead>
<tr>
<th>ASMR</th>
<th>G-BA/ IQWiG Level of Added Benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td>I – Major innovation (“majeure”)</td>
<td>Major (“erheblich”)</td>
</tr>
<tr>
<td>II – Important improvement (“importante”)</td>
<td>Considerable (“beträchtlich”)</td>
</tr>
<tr>
<td>III – Moderate improvement (“modérée”)</td>
<td></td>
</tr>
<tr>
<td>IV – Minor improvement (“mineure”)</td>
<td>Minor (“gering”)</td>
</tr>
<tr>
<td>V – No improvement (“inexistante”)</td>
<td>No added benefit (“kein Zusatznutzen”)</td>
</tr>
</tbody>
</table>

## Germany

<table>
<thead>
<tr>
<th>ASMR</th>
<th>G-BA/ IQWiG Level of Added Benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td>I – Major innovation (“majeure”)</td>
<td>Major (“erheblich”)</td>
</tr>
<tr>
<td>II – Important improvement (“importante”)</td>
<td>Considerable (“beträchtlich”)</td>
</tr>
<tr>
<td>III – Moderate improvement (“modérée”)</td>
<td></td>
</tr>
<tr>
<td>IV – Minor improvement (“mineure”)</td>
<td>Minor (“gering”)</td>
</tr>
<tr>
<td>V – No improvement (“inexistante”)</td>
<td>Non-quantifiable (“nicht quantifizierbar”)</td>
</tr>
</tbody>
</table>

## Innovative

## Non-innovative
Comparisons of Value Assessments by NICE (UK) and HAS (France) on 49 Cancer Drugs

(Drummond et al, *Pharmacoeconomics*, 2014)

Spearman rank order correlation -0.538, p=0.002
Explanations for the Observed Variations in Approach to Assessing Value

• A recent study in the largest 5 EU countries shows how these differences can be explained by cultural differences (e.g., the weights given to equity, efficiency, need, and personal responsibility) (Torbica et al., 2016)

• These factors influence the methods and use economic evaluation directly, or indirectly, by how they shape the financing and organization of health care
Can Culture and Values Explain Differences in Attitudes Towards QALYs?

• ‘Pro-QALY’ jurisdictions are more likely to:
  - have a NHS, operating with a fixed budget
  - have an institutional tradition that requires more transparency
  - place a high value on horizontal equity

• ‘Anti-QALY’ jurisdictions are more likely to:
  - have a social or private insurance system, where budgetary limits are less well-defined
  - be less worried about transparency
  - place a high value on meeting individuals’ needs and wants
Elements of Value Beyond the QALY

• Possibilities include:
  - convenience and access?
  - severity of disease?
  - rarity of disease?

• In all cases the key questions are:
  - how much do decision-makers (and society at large) care about these issues?
  - what is the best way of incorporating them in the decision-making process (though analysis or deliberative decision-making)?
DCEs

- Methods well-established, with defined methodological standards (Lancsar and Louviere 2008; Bridges et al 2011)
- Wide range of applications (Clark et al 2014)
- One approach for estimating health state preference values for estimating QALYs
- Can be useful in understanding what attributes of treatments and programmes patients/general public value and their willingness-to-pay
- Use in reimbursement decisions very limited and likely to remain so
MCDA

- Growing in popularity, as a way of introducing a range of factors into the decision
- Better viewed as a way of characterizing the objective function, as opposed to being a resource allocation tool in itself (ie by considering cost, or cost-effectiveness as a component of the MCDA)
- Several applications in LMICs, in determining the contents of the health benefits package, and in high income countries (eg Sussex J et al Value in Health 2013;16(8):1163–9 for ‘orphan’ drugs)
- Decision-makers seem to be divided on whether it’s a useful tool or not (Walker A Value in Health 2016;19:123-4)
Accompanying Decision-Making Processes

- Explicit threshold or not?
- Considering severity of disease
- Performance-based risk-sharing agreements
- ‘Leasing’ technologies for long-term cures
Pros and Cons of Explicit Thresholds

**Advantages**
- Encourages consistency in decision-making
- More transparent
- The threshold can be inferred from decisions in any case

**Disadvantages**
- Hard to estimate the threshold accurately
- Manufacturers may price up to the threshold
- Transparency is valued differently in different cultures
If the therapy:

• is for a small patient population with life expectancy of less than 24 months;
• where the therapy adds three months or more to life expectancy.

Then:

• the QALYs gained should assume full quality of life in the added months;
• in addition the Committee can consider that the QALYs gained should be weighted sufficiently high for the therapy to be approved, given NICE’s current threshold.
• Found, in a review of existing studies, that the amount that individuals were willing to pay for a QALY varied a lot and depends on the context

• Argued that estimates could be useful in informing the debate about the cost-effectiveness (decision-making) threshold
WEIGHTING OF QALYs

• Research commissioned by the Department of Health undertaken by the School of Health and Related Research, University of Sheffield (Brazier et al, 2013).
• Discrete choice experiment using an on-line general population sample (n=3669).
• Presented respondents with patient groups that differed on 4 attributes: life expectancy without treatment, survival gain from treatment, HRQoL before treatment and gain in HRQoL from treatment.
• Strongest preference for survival gains; very small preference for treating those with greater burden of disease, but mainly in relation to improving survival at end-of-life as opposed to improving quality of life.
• The implications of these findings for QALY weighing are the subject of further research.
COMMUNITY SURVEY OF SOCIETAL PREFERENCES


- Choice-based experiment (N=4118 adults in the UK)
- Respondents supported the criteria proposed under the value-based pricing system (for severe diseases, address unmet needs, are innovative provided they offered substantial health benefits, and have wider societal benefits)
- Did not support the end-of-life premium or the prioritisation of children or disadvantaged populations as specified by NICE, nor the special funding status for treatments of rare diseases, nor the Cancer Drugs Fund
Performance-Based Risk-Sharing Agreements

• Vast array of arrangements, going under different names in different countries
• Countries with most experience of performance-based arrangements include Australia, Italy, Sweden and the United Kingdom
• Some tapering off in the number of performance-based arrangements in recent years
• The vast majority of the most recent ‘Patient Access Schemes’ in the UK have involved simple price discounts
Performance-based Schemes by Year

Total Schemes: 209

Source: University of Washington PBRSA Database
Key Success Factors for PBRSAs

• When there is uncertainty about clinical or economic outcomes.
• When outcome targets can be clearly defined and measured.
• When performance-based arrangements are not excessively complicated or costly.
• When the timelines are reasonable.
• When reimbursement and/or pricing decisions clearly follow the outcomes obtained.

Leasing Technologies for Long Term Cures

- Several recent drugs offer long term cures for a big investment upfront (with a relatively short course of therapy)

- Edlin *et al* *Value in Health* 2014;17:438-444 propose a health technology payment strategy replacing the up-front payments with a stream of payments spread over the expected duration of benefit from the technology, subject to the technology delivering the claimed health benefit
Hypothetical Example: Trastuzumab in Early Breast Cancer in the UK

- Drug cost is £21,184 in year 1
- Average relapse-free survival period is 10 years
- Therefore contract for: annual payment is £2,118 compounded each year at 3.5% for patients remaining relapse-free
- Under typical reimbursement model: takes 19.5 years for benefit to outweigh costs
- Under technology leasing reimbursement strategy: takes 6.5 years for benefit to outweigh costs

Edlin et al, 2014
Conclusions

• There have been several developments in economic evaluation methods since 1992

• However, the estimation of incremental cost per QALY gained, using a health care perspective, is still the most prevalent approach and has been adopted by two previously QALY-skeptic countries (France and Japan)

• It remains to be seen whether the recent interest in value assessment frameworks in the US leads to a different approach