

# Early stage Health Technology Assessment (HTA) of biomedical devices. The MATCH experience.

L. Pecchia and M. P. Craven

University of Nottingham, Faculty of Engineering, Division of Electrical Systems and Optics, Nottingham, United Kingdom

**Abstract**— Early stage Health Technology Assessment (HTA) of biomedical devices requires different methods than those usually employed for pharmaceuticals. This paper reviews widely-used methods for HTA, discusses their limits for early stage evaluation of biomedical devices and presents two methods for early stage HTA being developed in the Multidisciplinary Assessment of Technology Centre for Healthcare (MATCH) project: Analytic Hierarchy Process (AHP) to elicit user needs; and early stage economic evaluations using Markov Models.

**Keywords**— Health Technology Assessment, Early stage HTA, evaluation methods, biomedical devices.

## I. INTRODUCTION

Health Technology Assessment (HTA) is a multidisciplinary and multidimensional process useful for evaluating alternative and competing medical technologies. In addition to drugs, health technologies, following the definition of WHO, also includes biomedical devices, healthcare processes and healthcare service provision. Nonetheless, HTA methods can be difficult to apply to biomedical devices particularly in early stage assessment. Established in 2003, the Multidisciplinary Assessment of Technology Centre for Healthcare (MATCH) is a long-term research collaboration between four UK universities (Birmingham, Brunel, Nottingham & Ulster) and a cohort of industrial partners, also supported by stakeholders from the NHS and other public sector organizations ([www.match.ac.uk](http://www.match.ac.uk)). The MATCH project is focused on development of methods and tools for early stage HTA. This paper, after reviewing standard HTA methods, presents some methods being developed in MATCH for early stage HTA of biomedical devices. This paper is structured into three sections: a brief review of standard methods of HTA; limits of these methods for biomedical devices and for early stage HTA; two methods, are being investigated through MATCH projects: early stage economical evaluations via Markov models; Analytic Hierarchy process AHP to elicit user needs.

## II. HTA DE FACTO STANDARD

The majority of published HTA reports have a regular structure. This defines a *de facto* standard of methods and tools for HTA. The basic structure of many HTA reports

can be summarized as: definition of the medical goal and decision problem; assessment of consequences using clinical evidence; resource assessment using cost analysis; analysis of incremental cost versus consequences.

Different scenarios may affect the decision definition problem. Medical goals are often driven by contingencies, and change according to the scale of the problem. Therefore it is not easy to identify a gold standard method. Although many authors suggest assessing the consequences of a health technology in several dimensions (economical, technical, ethical etc.) the majority of studies focused mainly on the clinical and the economical dimensions. While this is in theory a limitation, these two dimensions are sufficient for the majority of HTA studies, and these will be the only two considered in this paper.

### A. Consequence assessment: the clinical evidence

The gold standard for the assessment of clinical evidence is the so called pyramid of evidence. According to the existing knowledge on the specific medical problem, clinical evidence is collected as primary data, via clinical trials, or synthesized as secondary data, via systematic review of literature. Primary data collection is usually performed via four different kinds of direct clinical researches of growing evidence: laboratory studies, case series/reports, case control studies, cohort studies, Randomized Controlled Trials (RCT). Systematic literature reviews synthesize and generalize outcomes of direct clinical research. At the top of the evidence pyramid are the systematic reviews of RCTs with meta-analyses. These studies, after individuating relevant literature using well defined research strategies and scientific databases, pool together RCT outcomes, overcoming contextual limitations, which otherwise reduce their evidence (e.g. limited number of patients, single centre patients' enrolment, limited time of enrolment). The data pooling is a weighted average of homogeneous RCT outcomes, performed using statistical methods, basing on two main ideas: RCT outcomes are considered more reliable (weighted relatively more) according to: the number of patients involved, the precision of its findings [1]; heterogeneity among RCTs, may be partially corrected using statistical techniques: fixed effects [1] or random ones [2]. Clinical consequences may be classified in three main groups: effectiveness, utility, benefits. Effectiveness is considered as the efficacy in real world and is measured as the

Table 1 Description of the principal cost-consequence analyses.

	cost-minimization	cost-effectiveness	cost-utility	cost-benefits
<b>Costs</b>	• monetary units	• monetary units	• monetary units	• monetary units
<b>Consequences</b>	• equal in both programs	• clinical outcomes	• QALY	• monetary units
<b>Measuring Advantage</b>	• differences in costs ( $\Delta C$ ) • direct measurement • necessary for the other	• ICER • direct measurement • uniform clinical outcomes	• ICUR • indirect measurements • mixed clinical outcomes • multidimensional analysis	• ICBR • indirect measurements • mixed outcomes • multidimensional analysis
<b>Limits</b>	• no consequences	• one-dimensional analysis • data table missing in many national health services	• indirect measurement • data table missing in many national health services	• indirect measurement • monetization of value of life • ethical limits

degree of satisfaction of clinical outcomes (e.g.: reduction of mortality, morbidity, pain, complications etc.). Utility is a multidimensional function of several clinical outcomes, used to estimate the quality of life. The basic idea is that two years lived at 50% are as good as 1 year at 100%. The quality of life is measured using several scales, which specific for pathologies or generally used for different one as the Quality Adjusted Life Year (QALY). In both cases the quality is self-assessed using a visual analog scale (VAS) and/or structured questionnaires like the EQ-5D [3] asking to each patient to judge five dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) using three levels (no problem, some problems, severe problems), defining a 0-1 scale of health of 245 steps,  $3^5+2$  (unconscious or dead). Finally, consequences are measured as benefits if they are expressed in monetary units.

#### B. Resources assessment: cost analysis.

Cost analysis consists in giving a monetary value to each resource used in the process of care. Many authors consider fixed costs and variable costs for technology under comparisons on a plane which has on its x-axis the number of treatments performed in a period of time and on the y-axis the costs for the treatment over that time. This become more complex when resources needed by each patients may vary according to pathology evolution (e.g. worsening may require extra resources). In fact, patients within the same population under assessment could be in different initial stages of pathology (e.g. mild, moderate, and severe) at the start of the study. According to each state, it may vary: the probability to have exacerbation (% have or not exacerbation); the quality of exacerbation (% of mild, moderate, and severe exacerbations); the probability to move from one state to another (e.g.: % of mild becoming moderate) in the next time-step.

Many HTA reports employ Markov models [4] to deal with such scenarios. These models assume that a patient is always in one of a finite number of discrete health states, called Markov states (nodes into the model). All events are represented as transitions from one state to another (edges), with a given probability. A marginal cost is associated with

each event 'k', which represents the total amount of required resources required for that event, including costs for: normal treatment; treatment of exacerbations etc.

#### C. Incremental Cost-consequence analysis

If a new technology is proved to be less effective and more costly than the benchmark, it is rejected with no more analysis required unless there is a high degree of uncertainty in the data. Conversely, a technology is a good candidate for adoption if it more effective and less costly. Often, however, there is an additional cost to achieving more effectiveness. One of the principal reasons is that behind a new technology there are design and prototyping costs together with low volume productions, which often demand higher costs compared with the established product. In these cases the Incremental Cost Effectiveness (or Utility) Ratio (respectively ICER or ICUR) has to be evaluated (table 1). An ICER over a given threshold may suggest that the new technology is more cost-effective, and therefore its adoption is justified since the expected outcomes are considered to be worth the costs incurred.

Also consequences may vary, over a span of time, according to patients' progression of health. Once again Markov models [4] may be used to estimate cost-consequences ratios in patient populations. In this case three steps are usually performed:

1. Each status 'k' of a pathology (including the initial one) is associated to a cost ( $\delta C_k$ ), which represents the total amount of required resources such trading patients in such status, and a marginal consequence, which may be a marginal effectiveness variation ( $\delta E_k$ ) or a marginal utility ( $\delta U_k$ ).
2. The transition from one status to each another is associated with a probability.
3. The expected exacerbations from each state are associated with a probability, for each kind of exacerbation (eg: mild or severe).
4. The total path costs for each final state and the total path effectiveness (or utility) are combined representing the total costs (C) and total consequences (E or U) for groups of patients in each final state.

5. The model is then evaluated dynamically by matrix algebra, as a cohort simulation, or as a Monte Carlo simulation. This generates a distribution of points  $(U,C)_j$ , with  $j=1\dots N$  (number of simulations performed) on a cost-consequence plan.

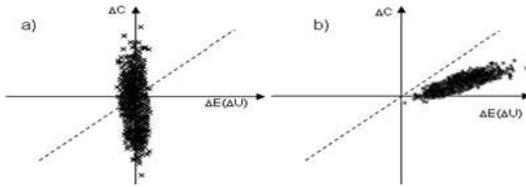


Fig. 1 Didactic examples of cost-consequence analyses: a) equivalence (symmetry with respect to both axes); b) more cost-effective than the benchmark.

In the case of a comparison of a new health technology (T1) with a benchmark (T0), this procedure is repeated using the same Markov model, but with probabilities, costs and consequences arising from use of the new technology to calculate differential costs and consequences among the two technologies ( $\Delta=T1-T0$ ). In this case, the cost-consequence plane has as x-axis the differences in consequences ( $\Delta E$  or  $\Delta U$ ) and on the y-axis the differences in costs ( $\Delta C$ ). The final evaluation is then performed using available cost and consequence (usually utility) data. In this case the threshold boundary is the line  $y=\alpha x$ , where  $\alpha$  is the relevant ICUR from historical data, where available [5].

### III. THE MATCH EXPERIENCE

#### A. Limits of Standard methods

The majority of HTA studies focus on evaluation of drugs, rather than medical devices. Therefore, the main methods employed are well suited to such a scope. Nonetheless, there are some differences between drugs and medical devices, which impact severely on HTA (Table 2). Moreover, HTA methods do not really inform biomedical product developers on the probability of return on investment, nor do they inform about the market needs and specific requirements of technologies in development [7]. Finally, standard methods of HTA do not allow prioritization of user need.

#### B. AHP to elicit user needs

AHP is a decision-making method, which aims to solve multifactorial and multidimensional problems. This method is particularly effective in quantifying the user opinions, based on their personal experiences, to design a consistent decision framework. AHP consists of defining a hierarchy of elements and prioritizing them by submitting questionnaires in which each respondent, through pair-wise comparison, judges relative importance of elements. By posing redundant questions it is possible to assess the coherence of

Table 2 main differences among drugs and devices impacting on HTA [6]

Devices	Drug
<b>Principal action</b>	
Other than principally drugs	Pharmaco./Immunologic/Metabolic
Mechanical/Electromagnetic/Materials	Chemical based
<b>Product life cycle</b>	
Short life cycle	Long life cycle
Constantly evolving components/parts	Unchanging compound
<b>Clinical evaluation</b>	
Difficult to blind (no placebo)	Easy to blind
Multiple end users	Usually one end users
Long learning curve	Short learning curve
Strongly dependent by settings/users	Less dependent by settings/users
Complex to standardize for RCT	Easy to standardize for RCT
<b>Use issues</b>	
User-dependent efficacy	Efficacy is less user-dependent
Often require intensive training	Usually do not require training
Complication decrease with use	Complication increase with use
<b>Diversity</b>	
Mainly small companies/few large co.	Mainly large multinationals
Diagnostic or therapeutic	Therapeutic
<b>Costs</b>	
Varying overheads/slow return	High overheads with quicker return
Higher distribution costs	Lower distribution costs
Higher maintenance/installation costs	No maintenance/installation

respondent judgments. Applying this method iteratively it is possible to elicit: relative importance of each need into its category (local weights, LW), relative importance of each category (category weights, CW), relative importance of each need compared to all the other individuated (global weight,  $GW=LW \cdot CW$ ). Further details on AHP can be found in the references [8-11]. The AHP method is effective in eliciting user needs, following the next 3 steps:

1. Identification of needs, which the new biomedical devices aim to satisfy. This step involves 1-2 domain experts (e.g. specialized clinicians) and potential users of the device.
2. Design of a tree of needs with nodes (categories) and leaf (needs). Three main categories are: clinical needs, economical needs, technical needs. This step involves 2 experts (one domain expert and one experienced in AHP).
3. Development and submission of questionnaires to elicit user needs. This involves 1 expert of AHP and "n" expert of domain, recruited via scientific societies.

This method was successfully applied to elicit user needs in previous studies: for a new CT scanner purchasing [8]; to choose a maintenance contract [9] according to hospital location and assets; identify the best model of care for heart failure [10], to identify risk factors for falls in elderly home dwelling [11].

#### C. Early stage evaluations via Markov Chain

During development of a new device or technology innovation it may be difficult to obtain data that fully considers stratification of patient by risk or past interventions, or to consider the full range of outcomes in a clinical pathway. Nonetheless, it may suffice at an early stage to consider the main outcomes and to limit stratification into one or two

groups. One approach is to use rarefied Markov models that use a minimum number of states and transitions according to data availability.

Such an approach was taken by Dong and Buxton with total knee replacement (TKR) whereby knee patients are divided into three groups, those who have had a primary TKR operation with no complications, and those who have minor or serious complications. Following the primary TKR, patients with complications may require a revision or other treatments. In the Markov model this requires probabilities between these states and various additional complication states or death, as determined from the clinical literature or best estimates. A comparison of computer-assisted (image-guided) and standard total knee replacement was performed using this basic model with nine states [12].

In a further study by one of the authors [MC], a four state model was used to perform a ‘what-if’ analysis for a device that aimed at promoting healing in diabetic foot ulcer. The model included a diabetic no-wound state, a wounded (ulcerated) state, an amputee state, and dead. This is a simplification of a more sophisticated model that would include stratifying patients into low and high risk and to consider multiple ulcerations and amputations with different probabilities. The value proposition of the device was examined by increasing the probability of transition from the wounded to the no-wound state e.g. the healing rate, and decreasing the probability of amputation [13].

#### IV. CONCLUSIONS

According to the WHO definition, health technologies include both drugs and biomedical devices. Nonetheless, HTA methods can be difficult to apply to biomedical devices particularly in early stage assessment. The MATCH experience shows that it is possible to develop specific methods to assess biomedical devices at an early stage. In particular AHP is an effective method with which to elicit user needs and Markov models with a minimal number of states can be used to perform ‘what-if’ analyses at an early stage.

#### ACKNOWLEDGMENT

The authors acknowledge support of this work through the MATCH Programme (EPSRC Grant EP/F063822/1) although the views expressed are entirely their own.

#### REFERENCES

1. Sutton, A.J., *Methods for meta-analysis in medical research*. Wiley series in probability and mathematical statistics 2000, Chichester ; New York: J. Wiley. xvii, 317 p.
2. DerSimonian, R. and N. Laird, *Meta-analysis in clinical trials*. Control Clin Trials, 1986. **7**(3): p. 177-88.
3. Rabin, R. and F. de Charro, *EQ-5D: a measure of health status from the EuroQol Group*. Ann Med, 2001. **33**(5): p. 337-43.
4. Sonnenberg, F.A. and J.R. Beck, *Markov-Models in Medical Decision-Making - a Practical Guide*. Medical Decision Making, 1993. **13**(4): p. 322-338.
5. Chapman, R.H., et al., *A comprehensive league table of cost-utility ratios and a sub-table of "panel-worthy" studies*. Medical Decision Making, 2000. **20**(4): p. 451-67.
6. Craven M. P., *Routes and requirements for realizing pervasive medical devices*, Chapter 9, in Bardram J. E., Mihailidis A., Wan D. (Eds.), *Pervasive Computing in Healthcare*, CRC Press (Taylor & Francis Group), November 2006. ISBN 084933621X
7. Ijzerman, M.J. and L.M. Steuten, *Early assessment of medical technologies to inform product development and market access: a review of methods and applications*. Appl Health Econ Health Policy, 2011. **9**(5): p. 331-47.
8. Pecchia L, Ragozzino A., Iovine C., Bracale M. *Health Technology Assessment: needs analysis via Analytic Hierarchy Process*, ISAHP 2011, 15th-18th June, 2011, Sorrento, Naples, Italy
9. Pecchia, L., et al., *Health Technology Assessment for a Service Contract: a new method for decisional tools*. World Congress on Medical Physics and Biomedical Engineering, Vol 25, Pt 12, 2009. **25**(12): p. 105-108.
10. Pecchia L., Bracale U. and Bracale M., *Health Technology Assessment of Home Monitoring for the Continuity of Care of patient suffering from congestive heart failure*. World Congress on Medical Physics and Biomedical Engineering, Vol 25, Pt 12, 2009. **25**(12): p. 184-187
11. Pecchia L., Bath PA, Pendleton N., Bracale M., *Use of the analytic hierarchic process (AHP) for examining healthcare professionals' assessments of the relative importance of risk factors for falls in community-dwelling older people*. Methods of Information in Medicine, 2011, **50**(5): p. 435-444, DOI: 10.3414/ME10-01-0028.
12. Dong, H. and M. Buxton, *Early assessment of the likely cost-effectiveness of a new technology: A Markov model with probabilistic sensitivity analysis of computer-assisted total knee replacement*. Int J Technol Assess Health Care, 2006. **22**(2): p. 191-202.
13. Craven, M. P. and Morgan, S. M. *Early Stage Economic Evaluation with a Small Medical Device Start-up Company using a Markov Model*. Journal of Medical Devices, Volume 5, Issue 2, 027516, June 2011. doi:10.1115/1.3590366.

Author: Leandro Pecchia, Michael P. Craven  
 Institute: University of Nottingham, Faculty of Engineering, Division of Electrical Systems and Optics  
 Street: University Park, Tower Building  
 City: Nottingham  
 Country: United Kingdom  
 Email: <Leandro.Pecchia, Michael.Craven>@nottingham.ac.uk