Some issues in applied statistics in clinical restorative dental research

Tobi, H.

Citation for published version (APA):
Chapter 6

An outline of cost-effectiveness analysis in dentistry

H Tobi¹², CM Kreulen¹, H Vondeling², MS Maningky³, WE van Amerongen¹

¹ Department of Cariology Endodontology Pedodontology, Academic Centre for Dentistry Amsterdam, The Netherlands
² Department of Epidemiology and Biostatistics, Vrije University Amsterdam, The Netherlands
³ Department of Oral and Maxillofacial Surgery, Academic Hospital Vrije University Amsterdam, The Netherlands
Abstract

Objective: The increasing concern over costs of public health care oblige critical resource allocation decisions. The results of economic evaluations help decision makers allocate the limited funds. The literature shows that economic evaluations in dentistry face several problems. The aim of this paper is to explain and appraise guidelines for economic evaluation for researchers in dentistry, to promote further quality improvement.

Basic Research Design: Guidelines for economic evaluation published in the British Medical Journal are briefly described and followed by a detailed description of how this is dealt with in an economic evaluation alongside a controlled clinical trial comparing the preventive extraction of mandibular third molars to delayed extraction on indication.

Conclusions: Economic evaluations are likely to gain from adherence to guidelines for economic evaluation. Community dental health care can benefit from the joint effort by dentists, economists and statisticians to yield information of sufficient quality to base decisions on.
Introduction

The increasing concern over costs of public health care made decision makers seek information for allocating the limited funds to the most effective treatments of human diseases. Of every dollar spent in health care, probably 4 to 6 cents are being spent on dental care (Garrison, 1996; CBS, 1997). Therefore, critical allocation decisions are required in dentistry too. As a result, the economic evaluation of interventions in dentistry has grown in importance. So far, the number of studies is limited, and most of them have been published in the last ten years (Elixhauser, Halpern et al., 1998). These studies go under a confusing range of labels and vary greatly in quality (Adams, McCall et al., 1992, Antczak-Bouckoms and Tulloch, 1992; Udvarhelyi, Colditz et al., 1992; Lewis and Morgan, 1994; Mason and Drummond, 1995).

In restorative dentistry, for example, clinical studies used to focus on relative effectiveness of different treatments in terms of behaviour and longevity of the restorations. Analyses of costs were usually not incorporated. Of the published reports on both costs and effectiveness, Hendriks and colleagues (1985) extrapolated the half-time of restorations from the literature and used observed treatment times from their clinic as an estimate for differential costs of amalgam and composite restorations. Since then, several papers were published that postulate treatment time as the main determinant of costs (Plasmans and van't Hof, 1987; Kreulen, van Amerongen et al., 1991; Advokaat, van't Hof et al., 1992; Kreulen, van Amerongen et al., 1992; Plasmans, van't Hof et al., 1992; Meijering, Creugers et al., 1995). These studies investigated costs and although some of these studies claimed to address cost-effectiveness, effectiveness estimates were often lacking.

Mjör (1992) compiled longevity data from the international literature to get effectiveness estimates and used fees for dental treatment in Norway as approximation of costs. Smales and Hawthorne (1996) used records from general dental practices to estimate longevity of five dental restorative materials in various classes of cavity preparations. Costs were based on survey figures of dental fees obtained in the same area and year, as the effectiveness data. The information on durability from several clinical studies and Dutch dental fees were used by Creugers and Käyser (1992) to compare cost-effectiveness of adhesive and conventional bridges.

The studies above represent a moving field in medical sciences. Each study is part of the development of the field of economic evaluation in dentistry. Economic evaluations of health programmes have become increasingly comprehensive, but there are a number of frequently encountered shortcomings (Adams, McCall et al., 1992; Udvarhelyi, Colditz et al., 1992; Lewis and Morgan, 1994; Mason and Drummond, 1995). For instance,
definitions of the population, treatment indications and inclusion criteria are frequently not clear. From the methodology point of view, information on costs and effectiveness should preferably be gathered from the same population. Although textbooks on economic evaluations are available, an outline and worked example for cost-effectiveness research specially designed for use in dental studies would be appropriate. In the present paper the guidelines given in the British Medical Journal (Drummond and Jefferson, 1996) are described and appraised for their use in economic evaluations in dentistry, with reference to restorative dentistry. Nine of the ten guidelines are briefly explained, followed by a detailed account of how that point is dealt with in a controlled clinical trial at the Academic Hospital of the Vrije University on the cost-effectiveness of the preventive extraction of mandibular third molars compared to extraction on indication. Particular attention will be paid to methodological and statistical issues. The tenth guideline concerning the presentation of results also contains valuable information but is here ignored because it is less important in the design and planning stages of a quality economic evaluation.

Guidelines

(1) Study Question

A well-defined study question addresses both costs and effects of the treatment under study. It gives a comparison of the alternative treatments and states what perspective is taken for the analysis. The economic importance of the research question should also be outlined. An increasing portion of research grant suppliers demands economics being part of the justification for performing the trial.

Example:
It is well known that clinicians, general practitioners as well as oral surgeons, disagree on whether or not to remove third molars (Kostopoulou, Brickley et al., 1997; Knutsson, Brehmer et al., 1992; Mercier and Precious, 1992). The preventive extraction of third molars involve high direct costs, but the relative effectiveness has not been decidedly established. The literature on morbidity due to the removal of third molars is extensive (de Boer, Raghoebars et al., 1995; for reviews see Daley, 1996; Mercier and Precious, 1992). This morbidity induces additional costs both inside and outside the health care system, for example due to sick leave (Berge, 1997). The above justifies an economic evaluation of the preventive extraction of mandibular third molars compared to extraction on indication from the societal perspective. First, this question is split in two to address treatment effects and costs: What morbidity is associated with the preventive removal of asymptomatic mandibular third molars and what are the effects in the long run of not
removing? What are the direct and indirect costs associated with the preventive removal of mandibular third molars and what are the direct and indirect costs of refraining from preventive removal? Subsequently, costs and effects are related.

(2) Selection of alternatives

In principle, the alternative (control) programme for comparison should be the most cost-effective alternative available. In practice, often the most widely used programme is chosen for comparison. Because the placebo branch of a trial does not reflect practice, placebo studies give no adequate information as to costs of the new treatment. The exception is, of course, the situation in which the alternative is “doing nothing”.

Sufficient detail is needed to enable readers to judge the applicability of the treatments to their own situation and to assess for themselves whether costs or effects have been omitted. This level of detail can be achieved by answering the questions who did what to whom, where, when, how often and with what results (Drummond, O'Brien et al., 1997 p31; White and Antczak-Bouckoms, 1997 p167).

The “to whom” part of this question is extremely important because costs and effects are usually extrapolated from the specific study population to the general population of dental care consumers. In addition, cost-effectiveness may differ between populations (Phelps, 1997). So reporting the inclusion and exclusion criteria for patients, can be regarded part of this guideline.

Example: The only alternative to preventive extraction is to wait and see until pathological reasons for extraction arise.

The inclusion criteria for patients are as follows: between 18 and 30 years of age, no earlier removal of a mandibular third molar, at least one mandibular third molar clinically visible or visible on a radiograph, good general health, no evidence of local pathology, and insured via the National Health Insurance Program (Dutch: Ziekenfonds). In addition, and in compliance with good clinical practice, patients informed consent is needed. The patient is randomly assigned to either the preventive extraction group (Group One), or the no preventive intervention group (Group Two).

For both groups an appointment is made for an intake and orthopantomography.

In Group One, an oral-maxillofacial surgeon removes one mandibular third molar under local anaesthesia and sutured with resorbing material. The patient receives instruction and a leaflet on the prevention of possible consequences and a prescription for an analgeticum (Naprosine). An appointment for an evaluation one week after the operation is made.
Group Two receives no treatment per se and is asked to continue routine visits to the general dentist (once every 6 months). When an indication for removal occurs in a Group Two patient, the general dentist refers the patient to the maxillo-facial surgeons in the University Hospital. Three and six years after randomisation, all patients are invited to the Hospital for an additional check-up, including an orthopantomography, by an oral maxillo-facial surgeon.

(3) Form of evaluation

Within the field of economic evaluations the following four forms of analysis need to be distinguished: cost-minimisation, cost-benefit, cost-effectiveness and cost-utility. The form of the analysis should be given and justified. Cost-minimisation studies determine how costs of alternative health programmes differ. The effectiveness of the treatments under study is disregarded, usually because the effectiveness of the treatments is established to be equivalent. In cost-benefit analysis, treatment outcomes are expressed in monetary terms and the question whether a given objective be pursued to a greater or lesser extent is addressed. In cost-effectiveness analysis, treatment outcomes are expressed as a health outcome, for example 'tooth years' or 'days without oral pain'. If a certain objective is set and one wonders what the most efficient way is of achieving that objective then a cost-minimisation or a cost-effectiveness analysis is appropriate. When the budget is fixed and the most efficient way of spending that budget is sought, then cost-effectiveness or cost-utility analysis is the correct choice. Cost-utility analyses take the patients' values and experiences into account by explicit adjustment of the treatment outcomes for the quality as perceived by the patient, for example 'quality adjusted tooth year' (suggested by Antczak-Bouckoms, Tulloch et al., 1989).

Example: Because the societal perspective is taken for this economic evaluation and both costs and effectiveness are unknown, a cost-utility analysis seems most appropriate. Because of likely practical limitations cost-effectiveness is also used as an analytical framework.

(4) Effectiveness data

Broadly speaking, there are two ways to get effectiveness estimates for an economic evaluation: a single effectiveness study, preferably a randomised clinical trial, or an overview of a number of effectiveness studies. If the economic evaluation is based on a clinical trial, details of that study should be given on, for example, selection of the study population, method of allocation of treatment to patient, whether or not analysed by 'intention to treat' and
effect size with confidence intervals. In an overview or meta-analysis, keywords and electronic databases used (whether or not supplemented with a hand-search), details on inclusion criteria, and statistical methods used should be described.

The gold standard for assessing effectiveness of a treatment is a double blinded controlled clinical trial. In operative dentistry, the blinding is -of course- next to impossible to achieve, reason to be extra aware of sources of bias associated with patients' and dentists' preferences during follow-up.

Clinical trials may yield insufficient data for the economic evaluation when the relevant events do not occur frequently. In restorative dentistry when comparing two dental materials for the same class of restoration, survival may be that good and follow-up that short that none or few of the restorations need replacement during the trial. In that case one may predict survival by a model based on the few failures that do occur. Of course, assumptions made in the model should be stated explicitly. Another option is to gather additional effectiveness data on intermediate or surrogate outcomes, such as quality of the restoration margin. Results obtained this way should be regarded as preliminary though (Gotzsche, Liberati et al., 1996).

Example: The inclusion criteria described in section (2) are applied to patients who attend general dentists and are interested in study participation. Of course, patients and surgeons are not blinded. A randomised block design is used to allocate patients to treatments. Each block is of size '10' and the odds of being allocated in the preventive treatment group is 1 to 4. Data will be analysed according to 'intention to treat'. The maximum follow-up of patients is six years. The results will be extrapolated to ten years.

(5) Benefit measurement and valuation

In the context of this guideline, by benefit is meant the possible outcomes and consequences of the treatments under comparison. Of course, the primary outcome measures should be clearly stated. If health outcomes are valued, as is the case in a cost-utility analysis, the methods and population used to obtain the valuation should be described in detail.

When identifying consequences of a treatment one can think of health outcomes such as tooth years gained or number of gingivitis cases prevented. In addition, other values may be created, for example reduction of dental anxiety, or resources may be saved, for example due to less gingivitis cases to be treated (Drummond, O'Brien et al., 1997 p33). To help identify the consequences of the treatments a longitudinal flow-chart, a so-called event pathway, may be useful.

Example: The possible outcomes and consequences of (no) preventive extraction of mandibular third molars are listed in Table 1. COOP/WONCA
charts measure general functional health status (Nelson, Wasson et al., 1996). EuroQol is a questionnaire that measures general health status utility (Kind, 1996). Primary outcome measures are the number of complications after removal and the number of symptom free third molar years for patients in Group Two.

(6) Costing

The costs can be distinguished using two dimensions: (1) costs borne by the health care sector versus costs borne outside the health care sector, and (2) direct costs versus indirect costs. The perspective from the study and the treatment description should provide information to judge whether or not all important costs are identified. Because prices of resources may differ from one setting to the other, resources used should be reported separately from the prices.

Table 1. The outcomes and consequences of (no) preventive extraction

<table>
<thead>
<tr>
<th>Categories</th>
<th>(No) Treatment consequences</th>
<th>Measurement</th>
<th>Valuation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health state change</td>
<td>Peri-operative complications</td>
<td>Standardised questionnaire answered by surgeon</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-operative complications</td>
<td>Interview and tactile sense test one week after surgery</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pain</td>
<td>Visual Analog Scale</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Health Status</td>
<td>COOP/WONCA charts and disease specific questionnaires one and two weeks after surgery</td>
<td>EuroQol</td>
</tr>
<tr>
<td></td>
<td>Need for removal (Group Two)</td>
<td>Third molar years saved</td>
<td>Criteria as applied by general practitioner</td>
</tr>
<tr>
<td>Other value created</td>
<td>Less liability suits</td>
<td>Unnecessary treatments and possibly resulting complications saved</td>
<td></td>
</tr>
<tr>
<td>Resources saved</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example: Since the perspective is societal, costs borne both within and outside the public health sector are incorporated. All costs as identified in Table 2 will be expressed in Dutch Guilders on the level of the final study year. During the operation, a cost form is kept which lists size of the team and time spent on one patient and number of standardised sterile operation sets opened for (partial) use. After the removal of the third molars, a cost diary is given to the patient. In this diary the patient keeps record of additional visits to the general dentist, medication used and other expenses due to the extraction. One week after the operation, the first costs diary and the questionnaire “Labour and Health” (van Roijen, Essink-Bot et al., 1996) is collected. A second set of costs diary and questionnaire is given to the patient, which is to be returned one week later.

Table 2. A list of costs associated with (no) preventive extraction

<table>
<thead>
<tr>
<th>Categories of costs</th>
<th>Example</th>
<th>Measurement Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inside health sector</td>
<td>Personnel costs</td>
<td>Time measurements and cost diary (additional visits to the general dentist)</td>
</tr>
<tr>
<td></td>
<td>Material costs</td>
<td>Cost form (operation equipment)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Costs diary one and two weeks after surgery</td>
</tr>
<tr>
<td>Outside health sector</td>
<td>Patient out-of-pocket expenses</td>
<td>Costs diary one and two weeks after surgery</td>
</tr>
<tr>
<td>Indirect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outside health sector</td>
<td>Productivity loss</td>
<td>Questionnaire ‘Labor and Health’ one and two weeks after surgery</td>
</tr>
</tbody>
</table>

(7) Modelling

Any modelling should be clearly described and details should be given on why a particular model with a particular set of model parameters was used. Modelling may be used for a number of reasons. The two most common are probably, the extrapolation of clinical outcomes beyond the duration of the follow-up in the clinical trial, and the transformation of intermediate outcomes into final endpoints. In section (4), on effectiveness data, the use of few failed restorations to model future survival was an example of the first type of modelling. A model that predicts need for replacement of a restoration based
on the quality of the margin is an example of the second type of modelling. Models can also be used to reflect what might happen in a different clinical setting or population when there is additional evidence from either other trials or a meta-analysis. When an economic analysis calls for a decision analysis, modelling may be required to blend data from different sources (for example: clinical trial, literature, reimbursement protocols).

Example: Indirectly, an incidence and compliance model is used to calculate sample size for Group Two. It is assumed that a patient in Group Two has a chance of 33% to develop an indication for mandibular third molar removal within six years. To obtain a treated group of 100 patients one would need 300 patients. Incorporating that 25% of the patients will not comply (are not traceable for the 6 years the study lasts), the number of patients needed becomes 400.

Modelling techniques will be used to estimate ten years results based on six years data. Specifically, survival modelling techniques will be used to extrapolate the percentage of third molars which show no indication within 6 years after randomisation, to a percentage of third molars which show no indication 10 years after randomisation.

(8) Adjustments for timing of costs and benefits

The time horizon over which costs and benefits are considered should be long enough to yield the desired information. Although often the time horizon should be a life-time, this has to be shortened to satisfy practical purposes. To enable the comparison of costs made over a long time discounting is needed. In discounting, the costs at the long-term time horizon are expressed in today's money, by

\[
\text{Present value} = \frac{\text{Amount to spend in } n \text{ years}}{(1 + d)^n}
\]

where \(d\) is the annual discount rate (Horngren and Foster, 1991). The discount rate should be given. If costs and effects are not discounted this should be explained. The assumption underlying discounting is that one prefers positive health effects now over the same positive health effect later and one rather delays expenditure.

Example: Because this study has a follow-up of six years the discounting of costs is important. An annual discount rate of 4% will be used, following advice of the Ministry of Finance. This means that one hundred guilders in 1996 is 126.53 guilders in the year 2002 (126.53=100\times1.04^6). Likewise, tooth-years gained will be discounted using a similar discount rate.
(9) Allowance for uncertainty

Costs and effects may be regarded as either deterministic or stochastic. When one considers costs as deterministic one assumes that costs can take on only one (true) value and any difference in costs seen between patients is merely the result of a random measurement error. Effects are stochastic: the differences in effects are really there and not only the result of measurement error. Usually, when the economic evaluation is done alongside a controlled clinical trial both costs and effects are regarded stochastic variables. Of course, two confidence intervals, one for the difference in mean costs and one for the difference in mean effectiveness can be estimated but in an economic evaluation one aims to combine costs and effectiveness in the cost-effectiveness ratio:

\[
\text{Cost-effectiveness ratio} = \frac{\text{Costs(A) - Costs(B)}}{\text{Effect(A) - Effect(B)}}
\]

which expresses the mean additional cost per unit gain in effect. There are two ways of handling uncertainty that complement each other: sensitivity analysis and confidence interval approach.

In a sensitivity analysis, one variable in the cost-effectiveness ratio is varied by a meaningful and plausible amount to investigate the sensitivity of the ratio (Manning, Fryback et al., 1996; Drummond, O'Brien et al., 1997). For instance, the impact of the discount rate on the cost-effectiveness ratio is investigated using a range of values for the discount rate. Note that in sensitivity analysis uncertainty is not explicitly related to sampling variation and the resulting imprecision of the estimators for mean costs or effects. Although the results of sensitivity analyses will underestimate the uncertainty of the cost-effectiveness ratio, they are very useful for getting a grasp on the cost-effectiveness structure.

The methodologically best situation, in which both effectiveness and costs are estimated from a sample in a prospective study, is also statistically the most complex one: a confidence interval for the incremental cost-effectiveness ratio, a ratio of two stochastic variables is difficult to obtain. Methods used to construct a confidence interval for the cost-effectiveness ratio are Taylor series expansion, Fieller's method and bootstrapping methods (O'Brien, Drummond et al., 1994; Chaudary and Stearns, 1996). As yet, it is not clear which method is to be preferred in what situation.

Example: Two cost-effectiveness ratios will be calculated. The first ratio considers the incremental average costs and the incremental average effects for patients in both groups according to the intention-to-treat. Depending on the results, this ratio could express costs saved by refraining from preventive removal. The second cost-effectiveness ratio considers the incremental
average costs and the incremental average effects for the treated patients in both Group One and Group Two. This ratio considers a possible lower chance of complications with preventive removal compared to indicated removal and relates these events to the costs of treatment. A sensitivity analysis will be performed for both cost-effectiveness ratios to identify the major determinants of the cost-effectiveness ratios. The discount rate, the chance of complications after surgery, the chance of need for removal based on indication will be among the variables considered.

Discussion

The emphasis of the guidelines proposed by Drummond and Jefferson (1996) is on improving the clarity of economic evaluations. Of course, no division of such a complicated process as designing and planning an economic evaluation, is exhaustive and without overlap. To ensure due attention for the ethical and methodological side of economic evaluations a supplement is advisable.

In their work on ‘Principles of health economics’, White and Antczak-Bouckoms (1997) distinguish eight stages for the performance of an economic evaluation, one of which is “Address ethical issues”. After all, when a health programme that benefits a group at increased risk, is somewhat less cost-effective than a programme that benefits a group less at risk, equity and equality concerns demand contemplation in addition to economic observations. White and Antczak-Bouckoms stage “Address ethical issues” after “Allowance for uncertainty” and before “Presentation and interpretation of results”. We suggest contemplation of ethical issues in the first two points, “Study question” and “Selection of alternatives”, because these contain extremely relevant information for medical ethical committees of hospitals and grant suppliers.

In the guidelines as published in the British Medical Journal, the actual data-analysis fits neither “Modelling” nor “Allowance for uncertainty”. We suggest to replace “Modelling” and “Allowance for uncertainty” by “Statistical analysis”. This section would incorporate an estimation of the desired study size and the associated power, a description of the actually planned data-analysis, as well as the sensitivity analysis.

The effectiveness part of economic evaluation will often be investigated by means of a clinical trial. Two concepts are pivotal for quality of clinical research: reliability and validity. Costs and effects are measured reliable, when the individual assessments are accurate and reproducible. This implies that inter- and intra examiner variation has to be examined and that ways sought to limit observer variation and increase reliability should be described.
Two major types of validity are internal and external validity. Internal validity is concerned with whether one measures what one intends to and is threatened by a wide range of sources of bias. The danger to internal validity caused by omitting an important consequence or cost, can be limited by using a longitudinal flow-chart. Often effectiveness will contain several treatment outcomes, for example when comparing programmes for prevention of caries: the prevalence of caries and fluorosis. A difficulty arises when the same treatment is not best on all effect measures. This may be tackled by prioritising one of the effect measures, or adding up the effects with or without some weighting procedure. When listing consequences as in Table 1, one could state the relative importance of each of the treatment outcomes.

External validity concerns generalisability: the question whether the economic evaluation has meaning in general, outside the particular study setting. In that respect section (2) “Selection of the alternatives” contains most relevant details. There is concern over the external validity of studies in an academic dental setting for the general practice. Data obtained in general practices or in general hospitals may be more generalisable than data obtained in an academic setting. However, in general practice randomisation is hardly feasible and internal validity is at risk. Here a paradox has come to surface: to guarantee internal validity a controlled clinical trial is the best choice and to ensure the generalisability a more pragmatic protocol is desirable. It is important to realise that economic evaluation relies on the assessment of the effectiveness. So when the relative effectiveness of the treatments under study is not yet known, the assessment of effectiveness and, hence, the internal validity should prevail.

The main two sources of concern for reliability in the third molar example are the number of patients and the indication criteria used for removal. The underpinning of the sample size is insufficient to feel at ease. The expected unambiguous use of a list of criteria formulated by the American Association of Oral and Maxillofacial Surgeons (AAOMS, 1994) is too optimistic and might endanger both reliability and validity. A translation into Dutch accompanied by some worked ‘paper’ patients could have ensured that the general dentists are indeed familiar with this list of criteria.

The extraction of the third molars takes place in an academic hospital. To ensure validity, additional information from general hospitals on their equipment, management of operation theatres and time taken to perform similar extractions has to be gathered and used in a sensitivity analysis. With respect to the data-analysis, any statistical model or statistical test should include the number of third molars at risk of removal, for example by means of stratification. Hence, the patient should be used as computational unit and not the mandibular third molar.
The worked example on preventive removal of mandibular third molars shows how complicated the economic evaluation of dental treatments can be. Guidelines can be extremely helpful but are not an airtight guarantee for splendid quality. Only interdisciplinary teams in which dental researchers, statisticians and economists collaborate can let go at the challenging problems involved in the design and realisation of an economic evaluation in dentistry.

Acknowledgements: We would like to thank dr. PD Bezemer, Chair of the Department of Epidemiology and Biostatistics, and dr. LE Smeele of the Department of Oral and Maxillofacial Surgery for their helpful comments and suggestions. The example of the mandibular third molar study is an initiative of the Department of Oral and Maxillofacial Surgery and the Department of Epidemiology and Biostatistics, supported by grant 28-2788 of the Dutch Praeventiefonds.

References


An outline of cost-effectiveness research


