Understanding cost effectiveness: a detailed review

Andrew F Smith, Gary C Brown

Overview
Given the drive towards the provision of ever better patient care in medicine, and the seemingly contradictory rise of self declared “healthcare crises” in many parts of the world, ophthalmology, like most other medical specialties, is experiencing uncertain, if not turbulent times. Exactly what the result of such changes will be is difficult to predict. It is certain, nevertheless, that increased economic pressures in healthcare financing, coupled with an ever ageing population, euphemistically referred to as the “demographic time bomb”, do little to reassure even the most optimistic of observers. It is against this backdrop that there is an urgent need to provide the general ophthalmic community with a first hand grasp, however simplified, of the key concepts of health economic analysis. Exactly why an understanding of health economics should be important is simple. On the one hand, knowledge of this area informs the wider non-economically literate, ophthalmic community of the key economic techniques used to inform health policy decisions. On the other hand, those who invest in its acquisition will be better able to respond to those pundits who would rather see money spent in other branches of medicine, further impacting the potential level of available funding for eyecare services and research purposes. Most importantly, however, it is hoped that such knowledge and understanding will serve to launch greater and more rigorous investigations into this relatively unexplored, yet important area of ophthalmology.

Although the dawn of health economics is a relatively recent one, its fundamental principles have as their roots the longstanding techniques and concepts of economic science. This said, the aim of the present article is rather more limited and consists of an examination and review of the basic principles of one of the more widely employed forms of health economic evaluation—namely, cost effectiveness analysis. Having accomplished this goal, the secondary aim of this paper is to briefly review those cost effectiveness analyses that have been conducted in ophthalmology during the past decade.

The following section is designed to equip the reader with the fundamental framework necessary to understand the basic concepts of health economic analysis explored in this paper. The first two ideas explored are those of resource scarcity and opportunity cost. Both concepts are important in that they show how an economy’s scarce resources are both distributed and allocated, hopefully in an efficient manner.

Resource scarcity and opportunity costs
Although there may be considerable differences in the accessibility and level of services provided between a wide range of publicly and privately funded healthcare schemes worldwide, given that all resources are finite in nature, all healthcare financing systems operate under some form of budgetary, or resources, constraint. Consequently, the drive to allocate resources in the most efficient and effective manner remains a guiding imperative. In fact, economics might be summed up as concerning itself with the science of making choices between different resource allocation pathways. In this regard, the limited availability of resources, and even our inability to satisfy all our desires and wants, both physical and otherwise, implies that choices must be made constantly.

The above may be more formally stated under the concept of “opportunity cost”, which embraces the notion that by producing more of one good there must be a reduction in the production (or so called “lost opportunity”) of one or more other goods. Theoretically, Garber et al have argued that:

“The real cost to society of a resource consumed or freed up as part of a health service (or as a result of it) is the value of that resource in its next best use to society. Because resources are more scarce than the needs for which they can be used, doing more of a given health service employing more doctors or nurses, utilizing more space and equipment for hospital beds, using more chemical or biological products means forgoing something else of value. In an ideal analysis from the societal perspective, therefore, resources should be valued at an amount equal to their best alternative use—their opportunity cost”.

Generally speaking, when examining the flows of monies within most healthcare systems, be they privately or publicly financed, one is not observing market prices but, rather, prices in the form of charges agreed upon between the purchasers of health care—namely provincial, or state governments or health insurance agencies, and the providers of health care, in this case, physicians, or more particularly ophthalmologists. In fact, as Luce et al suggest such prices or charges might not be such a bad reflection of the true opportunity cost. As they have observed: “... the real cost to society of a given resource is its opportunity cost, the value of the resource in its next best alternative use. For most purposes, market prices provide a reasonable estimate of opportunity cost. For example, the wages of a registered nurse or the charge for an office visit generally provide an adequate measure of the value of the resource consumed”.

In addition, part of the opportunity cost to the patient is the cost of the time during which he or she undergoes the ophthalmic procedure. In this respect, it has been proposed that the “… best approximation of the opportunity cost of time for working age adults is the wage they are, or more correctly could be, making in paid work”. It should be pointed out, however, that this method ignores the inequality in wages between the sexes and various age groups. Further understanding of the steps involved in conducting a cost effectiveness analysis are to be found in an examination of how the specific cost components are identified, measured, and valued.
Derivation of costs in health care

As is to be expected, any cost effectiveness analysis relies on the calculation of both the costs and the effectiveness of the healthcare intervention under investigation. Turning our attention to the first of these tasks there are, in essence, three fundamental stages to deriving the cost components of a healthcare intervention under investigation. Turning on the calculation of both the costs and the components of a cost effectiveness study—namely, (1) identification, (2) measurement, and (3) valuation of cost data. Each involves its own challenges and difficulties.

IDENTIFICATION OF COSTS

With respect to the first stage the identification of costs may be further divided into three categories—namely, (1) health system costs, (2) patient based costs, and (3) external and intangible costs. Health system costs are those associated with the “…organisation and operating costs within the health care sector (for example, health professional’s time, supplies, equipment, power, capital costs)". Health system costs may also be thought of as direct costs, since these are relatively easily measured in comparison with patient based costs (see Table 1).

As stated above, direct costs account for such items as hospitalisations, drugs costs, physician’s fees, laboratory costs, rehabilitation, and long term care costs. Often direct costs come in the form of charges and the true medical costs may be obscured, or difficult to measure, since they do not empirically measure the forgone opportunity cost of using these resources for other purposes. Another feature of calculating direct costs derives from whether they are fixed (for example, land or capital) or variable costs (for example, labour). Such costs as hospital buildings, for example, are assumed to be inflexible in the short run and thus are fixed. Variable costs, by contrast, are more flexible in the short term—that is, they can be increased or reduced with much greater ease, as in the case of hospital staffing levels.

Patient based costs are those derived from “…costs borne by patients and their families, (and include) out of pocket expenses, patient and family input into treatment, time lost from work, and ‘psychic costs’ attributable to pain and suffering”. Non-medical costs such as transport and support for ancillary workers, homecare workers, and other out of pocket expenses may all be included here to gain an overall picture of the costs of a given healthcare intervention from the patient’s perspective (see Table 2). The next step is to measure as precisely as possible the costs of the healthcare interventions or programme.

MEASUREMENT OF COSTS

Measuring costs is an exacting process and relies upon clearly defining that the cost inputs selected for analysis are “…measured in appropriate physical and natural units”. Tallying up all the cost components may yield overlapping areas of similar resource use, such as two variously busy clinics (one a very busy clinic and the other a not so busy clinic). In this case it becomes difficult to disentangle the true proportion of overhead expenses (electricity, heat, rent of hospital space, etc) which is being consumed separately. Under such circumstances, the aim is to make a reasonable estimate of the various amounts involved, including such matters as the number of employees, the size or area of clinic space used, the number and volume of patients seen, etc. As for the measurement of natural, or health, outcomes it is equally important that these are indicated in similar units, whether they be: “…life years gained, or deaths averted; they might relate to morbidity and be measured, for example, in reductions in disability days or improvements on some index of health status measuring physical, social, or emotional functioning; they may be even more specific, depending upon the alternatives under consideration”.

The calculation of specific measures of health outcomes allows comparisons to be made with other healthcare interventions. The final cost category is bound up in the appropriate valuation of costs—namely, attempting to measure as precisely as possible the cost of all healthcare inputs, whether these are incurred in the present or the future.

VALUATION OF COSTS

Valuation of healthcare costs is achieved using local currency and local prices for goods and services, and is normally approximated by healthcare charges and factors set by healthcare authorities or private insurers through negotiations between the providers of health care and government or private agencies. Both current and future healthcare costs are valued in constant monetary terms in order to remove the potentially confounding effects of
inflation. The concept of constant dollars, for example, is related to inflation, while discounting is related to time preference and consumption. Thus, if inflation is running at 5% a dollar today is worth more than it will be in a year’s time. To account for this, economists adjust the price accounts with price indexes. Because we would prefer to value goods today and pay later, the time patterns of costs and benefits are important, as costs and benefits are made equivalent in time by the use of discounting. Put another way, in order to reduce the value of future paths of costs and benefits derived from goods and services we discount. Thus $5000 (£3400) today is worth more than $5000 in 3 years because of our time preference. Under such circumstances, a discount rate is used to convert future costs and benefits into equivalent present values. Typically, 5% to 6% rates per annum are used for costs and similar rates per annum are used for benefits. Often a zero discount rate, or a rate lower than that used, is adopted during subsequent sensitivity analysis. Moreover, a lower discount rate is advocated by some health economists, so as not to penalise governments from initiating preventive programmes and because empirical evidence would suggest its use. Overall, the goal in valuing costs is “...to obtain an estimate of the worth of resources depleted by the (health care) programme”.

In summary, the four main approaches to the valuation of costs include: (1) using market prices, be they actual or proxies from some reference point, (2) computing the time lost by patients as some measure of indirect costs, (3) using disability and rehabilitation payments to estimate lost productivity, or (4) reviews of policymakers’ overall perceptions of costs, whenever this is possible. It must be remembered that proxy costs are never 100% of the actual costs. Moreover, it is difficult to be certain that these costs will always represent the actual opportunity costs themselves. One must, therefore, be aware of these limitations when attempting to use the concept of opportunity costs in any analysis.

Other analyses (cost-benefit, cost-utility, and cost of illness)
Before passing on to a detailed discussion of how to interpret cost effectiveness data, it is worth noting that there are three other main methods of interpreting cost data—namely, cost-benefit, cost-utility, and cost of illness studies. In the case of cost-benefit analyses, the task is to translate the benefits of a particular healthcare intervention into monetary units, so that both the numerator and denominator are in the same units. Secondly, cost utility studies attempt to quantify the cost of attaining a given level of health gain as measured on a utility scale, such as the quality adjusted life year, or QALY. Other utility scales, include the handicap adjusted life years (HALY) and the disability adjusted life years (DALY) scales, both of which have been used in ophthalmology. Here the outcome unit, in this case the QALY, or whichever scale is used, has been standardised and one measures the cost per QALY or other utility unit. Lastly, there is the cost of illness study which attempts to measure the economic burden due to a given disease, such as the economic cost of glaucoma in terms of lost productivity and medical expenses for its treatment. Finally, no cost effectiveness analysis is complete without a discussion of the uncertainties contained in the data themselves.

Sensitivity analysis
Given the potential for actual or accidental uncertainties contained in the information used to conduct cost effective analyses, the data used to derive the information are often subjected to the rigours of a sensitivity analysis, whereby a range of plausible numerical values is run through the economic model in order to simulate real world imprecision both in the quality of the data and that of the economic model itself. Typically, sensitivity analyses are performed to highlight a range of possible economic outcomes which might arise from the analysis itself. Sensitivity analyses are particularly useful in determining the robustness of the overall cost effectiveness analysis. Finally, it is important that cost effectiveness analyses should be situated within an overall study perspective and time frame.

Study perspective
A pivotal feature to take into account when conducting a cost analysis is the perspective from which the costs are measured, be it a national, regional, or municipal government perspective, that of an employer, an insurance company, a health maintenance organisation (HMO), or the individual’s perspective, such as that of physicians or patients. In the main, the perspective adopted in most forms of economic analyses is the societal or governmental perspective since this allows healthcare resources to be allocated to maximise social welfare. Equally, it is important to indicate the time over which the costs of any healthcare intervention or programme are distributed.

Study time frame
The time frame over which a healthcare programme is to be implemented can affect the costs of any intervention. Consequently, it is necessary to determine the so called “analytical horizon”—that is, the time over which the costs and effects of a given healthcare intervention or programme are derived. Costs, for example, may begin before the healthcare intervention, such as those incurred in the construction of new clinics and medical facilities to see patients, while other costs may be ongoing in the form of salaries for medical staff and equipment. In general, the analytic horizon of a given economic analysis should last long enough to capture that portion of time during which individuals are affected by the healthcare intervention or programme and any benefits which such interventions continue to yield in the form of positive health outcomes for those individuals enrolled in the healthcare intervention or programme. Despite the best attempts at conducting as precise a cost effectiveness analysis as possible, several criticisms with this approach to the calculation of costs exist.

Criticisms of cost determination
Calculation of the loss of potential income is often problematic to the degree that people with lower expected lifetime income levels will have lower economic values for their lives than those with higher expected income levels. Equally, if patients believe that they are at an increased risk of a particularly poor health state, they may be more willing to pay for care than those who do not have the same valuation of their current, or future, health status.

While the issues surrounding the use of indirect costs are complex, it is important to acknowledge the potential impact that productivity losses, as a result of (1) the costs associated with lost or reduced ability to function as a “normal” healthy person both on the job and during one’s leisure time, so called “morbidity costs”, and (2) the costs attributed to lost productivity because of early death, so called “mortality costs”, may have upon the calculation of the overall indirect costs of a disease. Morbidity costs typically arise as a consequence of lost productivity due to time spent recuperating or convalescing. Typically too, in the case of a disabled person, there is
the cost of time spent by family members or others caring for the affected individual, a cost that is rarely captured in formal economic analyses.\textsuperscript{3}

Mortality costs arise from changes in overall life expectancy as a result of the presence or absence of a given healthcare intervention or programme. At present, there is, understandably, some debate as to which productivity costs are capable of being easily measured.\textsuperscript{4} An exception to the above should be pointed out, specifically the fact that measuring lost productivity is confined to the manner in which the question is posed and whether or not certain items are included or excluded in the questionnaire as this can significantly influence the reporting of the overall magnitude of productivity losses.

**Interpreting cost effectiveness analyses**

Under the context of cost effectiveness analyses, the cost effectiveness ratio obtained is a measure of the cost per unit of health effect. In their simplest form, health effects might be regarded as the number of life years saved, or more particularly in an ophthalmological context, the number of sight years saved from vision loss and blindness. Thus, when interpreting the results of a cost effectiveness analysis one has a number of possible options. Firstly, assuming that the health effects obtained by each of the two treatment options being considered are equal, then solely cost considerations need to be assessed between both groups. Under such circumstances, the least costly options are the most efficient in terms of the allocation and distribution of scarce resources. Alternatively, the treatment options may be analysed in terms of a cost effectiveness ratio, whereby such cost effectiveness ratios can be scrutinised for those that offer the lowest cost per greatest unit of health effect. This is so because, within any given budget, more health can be produced by selecting this option, provided that there is no infusion of funds into the medical system.

**Evidence from ophthalmology**

Overall, it should be remembered that with the above explanation of the main points of cost effectiveness analysis out of the way, it is useful to consider the main cost effectiveness findings in ophthalmology over the past decade. As shall be seen, little has been done in this area, and what has been conducted has not generally tended to follow rigorous cost effectiveness guidelines. Table 3 summarises the main cost effectiveness studies in ophthalmology conducted over the past decade. As can be seen, vitamin A supplementation,\textsuperscript{5} cataract surgery,\textsuperscript{10–14} and trichiasis surgery for trachoma\textsuperscript{15} are among the most cost effective of all evaluated ophthalmic interventions. In point of fact, the figures for the cost effectiveness of cataract surgery were derived from a detailed examination of the cost of attaining a given outcome—namely, the successful removal of the cataract lens. The next most cost effective ophthalmic interventions are those which involve screening for diabetic retinopathy,\textsuperscript{16} followed closely by screening strategies for glaucoma\textsuperscript{17} and treating threshold retinopathy of prematurity.\textsuperscript{18} Those studies designed to measure the cost effectiveness of screening for diabetic retinopathy have, by and large, used existing epidemiological data on the incidence, prevalence, and overall progression of diabetic retinopathy in the absence of any screening examination and compared the results with the outcome of complying with various diabetic eyecare screening guidelines. It is interesting to note that among the studies presented are to be found the main causes of blindness and vision loss—namely, cataract, trachoma, glaucoma, and diabetic retinopathy. Moreover, it should be pointed out that Laupacis and colleagues have shown that health interventions which are under US$20 000 per QALY are worthy of implementation by society.\textsuperscript{19} Using this guideline of cost effectiveness, it is immediately apparent that all of the eyecare interventions highlighted in Table 3 are highly cost effective. This is especially true if one considers that most of the world’s cataract blindness and trachoma blindness is located in developing countries that must adopt eye healthcare interventions which are highly cost effective. Hopefully, the provision of the results presented here will filter their way into the hands of those attempting to reduce the burden, both social and economic, associated with vision loss and blindness in both developed and developing countries by specifically focusing on these worthwhile areas. Inevitably, as new information becomes available on the cost effectiveness of new ophthalmic interventions, these will be readily welcomed as additions to the fight against blindness.

\begin{tabular}{|l|c|c|c|c|}
\hline
Ophthalmic intervention & Cost effectiveness & Setting & Country & Reference \\
\hline
Vitamin A supplementation & US$1/DALY* & Community based & Not given & Bobadilla et al, 1994\textsuperscript{4} \\
Cataract surgery (conventional extracapsular cataract extraction + posterior chamber intraocular lens) & US$ 1000.85/cataract removal & Specialised cataract unit & Australia & Asimakis et al, 1996\textsuperscript{16} \\
Cataract surgery (phacoemulsification plus posterior chamber intraocular lens) & US$1231.00/cataract removal & Specialised cataract unit & Australia & Asimakis et al, 1996\textsuperscript{16} \\
Cataract surgery (technique not mentioned) & UK£496.90/cataract removal & Specialised cataract unit & England & Cresswell et al, 1996\textsuperscript{11} \\
Cataract surgery (technique not mentioned) & US$5.06/DALY & Specialist cataract unit & Nepal & Marseille, 1996\textsuperscript{12} \\
Cataract surgery (technique not mentioned) & US$23 & Eye Camp & Northern India & Murthy et al, 1994\textsuperscript{13} \\
Trichiasis surgery for trachoma & US$2–15/HALY‡ & Countrywide & Burma & Evans et al, 1996\textsuperscript{16} \\
Diabetic retinopathy screening (type I diabetics) & US$1996/QALY† & Computer modelling & USA & Javitt et al, 1996\textsuperscript{16} \\
Diabetic retinopathy screening (type II non-insulin using) & US$293/QALY & Computer modelling & USA & Javitt et al, 1996\textsuperscript{16} \\
Glaucoma screening (ophthalmoscopy and tonometry, with perimetry when abnormalities detected) & US$3530/QALY & Computer modelling & Canada & Botvin et al, 1996\textsuperscript{16} \\
Glaucoma screening (ophthalmoscopy and tonometry, and perimetry) & Cdn $100 000/year of blindness averted & Computer modelling & England & Tiick et al, 1997\textsuperscript{16} \\
Glaucoma screening (ophthalmoscopy, tonometry, and perimetry) & UK£850/true positive case & Computer modelling & England & \\
\hline
\end{tabular}