

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/7621246>

Cost-of-illness analysis. What room in health economics?

Article in *Health Policy* · July 2006

DOI: 10.1016/j.healthpol.2005.07.016 · Source: PubMed

CITATIONS

526

READS

10,864

1 author:



Rosanna Tarricone

Bocconi University

255 PUBLICATIONS 5,342 CITATIONS

SEE PROFILE

Review

Cost-of-illness analysis What room in health economics?

Rosanna Tarricone*

Bocconi University, Milan, Italy

Abstract

Cost-of-illness (COI) was the first economic evaluation technique used in the health field. The principal aim was to measure the economic burden of illness to society. Its usefulness as a decision-making tool has however been questioned since its inception. The main criticism came from welfare economists who rejected COIs because they were not grounded in welfare economics theory. Other attacks related to the use of the human capital approach (HCA) to evaluate morbidity and mortality costs since it was said that the HCA had nothing to do with the value people attach to their lives. Finally, objections were made that COI could not be of any help to decision makers and that other forms of economic evaluation (e.g. cost-effectiveness, cost-benefit analysis) would be much more useful to those taking decisions and ranking priorities. Conversely, it is here suggested that COI can be a good economic tool to inform decision makers if it is considered from another perspective. COI is a descriptive study that can provide information to support the political process as well as the management functions at different levels of the healthcare organisations. To do that, the design of the study must be innovative, capable of measuring the true cost to society; to estimate the main cost components and their incidence over total costs; to envisage the different subjects who bear the costs; to identify the actual clinical management of illness; and to explain cost variability. In order to reach these goals, COI need to be designed as observational bottom-up studies.

© 2005 Elsevier Ireland Ltd. All rights reserved.

Keywords: Cost-of-illness; Human capital approach; Decision-making

Contents

1. Introduction	52
2. Methodological issues to carry out COIs study	52
3. Types of COI studies	52
3.1. Prevalence- versus incidence-based COIs	53
3.2. Top-down versus bottom-up approaches	53
3.3. Prospective versus retrospective COI studies	54

* Tel.: +39 0258362582; fax: +39 0258362598.

E-mail address: rosanna.tarricone@unibocconi.it.

4.	Cost evaluation in COIs	55
5.	Productivity costs	57
5.1.	Informal care	58
5.2.	Theoretical issues behind COIs	59
6.	Do COIs still have a role in health economics?	60
7.	Conclusion	61
	References	62

1. Introduction

Cost-of-illness (COI) analysis represents the earliest form of economic evaluation in the health care sector. The principal aim was to evaluate the economic burden illness impose on society as a whole in terms of the consumption of health care resources and production losses. The implicit assumption was that the economic costs of illness represented the economic benefits of a health care intervention had it eradicated the illness.

COI studies have been widely debated and its usefulness as a decision-making tool has been questioned by many health economists. Nevertheless, COIs are among the commonest economic studies in healthcare in Italy [1] and abroad and are commonly used by organisations, such as World Bank, WHO [2], and the US National Institute of Health [3].

The principal aims of this paper are to shed light on this inconsistency in economic evaluation analysis and try to conceptualise COI within health economics in order to assess whether COI studies are worthy doing. The paper also aims at presenting a systematic picture of COI studies by analysing the methods used to carry out COIs, illustrating the different types of COIs, and discussing the most important issues that have characterised the debate around COIs in the last decades (e.g. Human Capital Approach versus Willingness-to-Pay). It finally tries to position COI in health economics and healthcare management as a useful economic tool for decision-making.

2. Methodological issues to carry out COIs study

Although some COI analysis made their appearance well before the mid-1960s, it is only in that period however that health economists, as Rice [4] first spelled out in great detail the methodology for costing illness.

Some years later, Hodgson and Meiners [5] provided guidelines for those who intended to undertake COI studies. More recent position papers [2,6,7] on COIs still refer to these authors when the methods of the technique is at issue. These works used to classify the economic costs of disease as direct, indirect and intangible costs. Direct costs refer to healthcare and non healthcare costs. The first have been defined as the medical care expenditures for diagnosis, treatment, continuing care, rehabilitation, and terminal care, while the second relate to the consumption of non healthcare resources, such as transportation to and from health providers, certain household expenditures, costs of relocating and certain property losses, legal and court costs, and informal care, that is the time family members or volunteers spend caring for the patient. The term “indirect” is used in economics to refer to productivity losses related to illness or death. However, the term “indirect” may cause confusion as it has different meanings in different settings. In accounting, indirect costs refer to supporting and overhead activities that need to be shared amongst the user units. For this reason, it has been suggested to substitute the term indirect with productivity costs which are associated with morbidity and mortality [8]. Intangible costs traditionally referred to patients’ psychological pain and discomfort but have never quantified in monetary terms and hence seldom considered in COIs. In this paper costs are classified into direct costs and losses of production/productivity.

3. Types of COI studies

COI studies can be described according to the:

1. epidemiological data used: *prevalence* versus *incidence* approach;
2. methods chosen to estimate the economic costs: *top-down* versus *bottom-up*;

3. the temporal relationship between the initiation of the study and the data collection: *retrospective* versus *prospective* studies.

3.1. Prevalence- versus incidence-based COIs

COI studies can be prevalence or incidence-based [2,9]. Prevalence studies refer to the total number of cases in a determined period of time (usually a year), while incidence studies refer to the new number of cases arising in a predefined period of time.

The prevalence approach involves estimating for any disease or group of diseases the direct costs and production losses attributable to all cases occurring in a given year. The incidence approach involves estimating the lifetime costs of the new cases of a condition or group of conditions which have their onset in a given period.

The underlying rationale of the prevalence approach is that disease costs should be assigned to the years in which they are borne or are directly associated. Under this approach, direct costs and productivity losses resulting from disease are assigned to the years in which they occur. Lost expected future earnings resulting from premature mortality are assigned to the year of death. In contrast, the incidence approach is based on the principle that the stream of costs associated with an illness should be assigned to the year in which the stream begins. All costs, both direct and productivity costs, are present-valued and assigned to the year in which the disease first appears [9,10].

The major differences between the two approaches are that the prevalence-approach results are generally larger than those obtained through an incidence approach [9]. This is generally the case for illnesses that produce long-term sequelae. For conditions that do not produce long-term sequelae, there will be little difference between the two approaches. Owing to the limited duration, the years in which costs are borne will typically be the years of incidence, which also will generally be the years of disease-caused death. The discrepancy between the prevalence and incidence approaches grows with the average duration of the condition. This reflects the fact that some costs which are not discounted in the prevalence approach are discounted in the incidence approach.

Finally, the ratio of prevalence-approach costs to incidence-approach costs for a given disease will gen-

erally be higher (other things being equal): (1) when incidence is declining (the prevalence approach capturing the costs of chronic conditions from the larger incidence cohorts of earlier years); (2) when annual treatment costs and disability losses are declining over time (since the analyst should reflect this trend in calculating the future costs of present incidence cases); (3) when annual treatment costs and disability losses rise over the course of the disease (because the later, greater costs will be discounted in the incidence approach but not in the prevalence approach).

It can be said that prevalence based COI analysis can be particularly useful when the aim of the study is that of [11]:

1. Drawing decision-makers interest for conditions whose burden has been somehow underestimated. Owing to the numerical differences between the prevalence and the incidence approaches, the first serves this purpose better than the incidence approach.
2. Planning cost containment policies. This is because the study provides decision makers with a picture of the global burden and, more importantly, of the major cost components, that is the areas where cost containment policies would have the greatest impact.

Incidence-based COI analyses are particularly useful when the aim is that of:

- considering preventive measures. Incidence-based COIs therefore provide an estimate of the savings that potentially accrue if the preventive measure is implemented.
- analysing the management of the illness from the onset till recovery or death. The incidence approach allows for analysis of disease staging thus showing how costs are distributed while the illness progresses. This could encourage for instance the development of clinical/therapeutic guide-lines aimed at increasing the effectiveness and the efficiency of both the management of the disease as a whole and of each single step of the clinical therapeutic pathway.

3.2. Top-down versus bottom-up approaches

Another difference between the two approaches is that the incidence approach requires that the analysis be

performed “from the bottom-up”, totalling the lifetime costs of illness. This, in turn, requires that input data be gathered at a level of detail much greater than that employed in the prevalence approach where, in general, the analysis is performed “from the top-down”, allocating portions of a known total expenditure to each of several broad disease category.

Rice fully developed this methodology in 1966. It consists of allocating total national health care expenditures by type of care (i.e. hospital care, physicians’ services, ...) among the 16 disease categories of the International Classification of Diseases (ICD). All expenditures were allocated to the primary diagnosis. Cooper and Rice’s [12] application of the prevalence approach resulted in a comprehensive picture of the economic costs of illness in the United States in 1972. Their analysis made use of pre-existing estimates for that year of total annual direct costs by cost component for the entire ill population and of similar, though less complete, estimates of illness-related productivity losses.

In the bottom-up approach, the estimation of costs can be divided into two steps. The first step is to estimate the quantity of health inputs used and the second step is to estimate the unit costs of the inputs used. The costs are then estimated by multiplying unit costs by the quantities. The data needed and available will vary with the scope of the study. In a comprehensive study, there are usually national surveys which provide reliable data on medical care utilisation. In a more limited study, the investigator may have to collect data which may be a disadvantage even though in doing so he may find more reliable estimates of the different cost components than in nationally collected data.

An often claimed advantage for comprehensive studies is that by allocating total national expenditures among the major diagnostic categories, one can avoid the risk that the sum of treatment costs of individual diseases – estimated through the bottom-up approach – is greater than total health care expenditure in a given country [13]. However, top-down COIs are likely to present misallocation of costs. Firstly, because the use of national health care expenditures may either under or overestimate total direct costs. Second, the exclusion of cost categories that are not included in national health care expenditures (i.e. transports, informal care) also biases the estimates of costs by disease category since different disease categories may absorb different non-health costs. Lastly, another problem with this

method is that all costs are attributed to the primary diagnosis. This is a serious problem if we consider that a relevant part of all hospital discharges involve patients with multiple diagnoses.

3.3. *Prospective versus retrospective COI studies*

COIs can also be prospectively or retrospectively performed depending on the temporal relationship between the initiation of the study and the data collection.

In retrospective COI studies, all the relevant events have already occurred when the study is initiated. This means that the process of data collection must refer to data already recorded. Conversely, in prospective COI studies the relevant events have not already occurred when the study is initiated. This means that the process of data collection needs to be done by following-up the patients over time. Prevalence- and incidence-based COIs can be both performed either prospectively or retrospectively.

The major advantage of retrospective COI is that they are less expensive and time consuming than those performed prospectively because all relevant events have already occurred at the time the study is initiated. Thus, the retrospective design is particularly efficient for the investigation of diseases that have a long duration requiring many years to reach the relevant end points. Retrospective COIs can only be carried out when sufficient data are available. This is not often the case. It may happen that data have been collected for purposes different from those of a COI study. Again, health care resource consumption – for instance – might have been recorded at local level, preventing the use of such data in nationally representative studies. Conversely, in prospective COI analysis, analysts can design data collection systems they require. Data on the illness and the consumption of health care resources are gathered by the analysts on the basis of purposely designed questionnaires submitted to patients and/or providers. This allows for complete data as every action/intervention is registered. Secondly, diaries can be provided to patients for those cost components that are not recorded by health care organisations. In this way non-health costs such as transportation can be more easily measured. Precise estimates of time off work of patients and their relatives can be collected. However, if the disease has a long time

span (hepatitis C has a course of 30–40 years), prospective incidence-based COIs would be enormously costly and time consuming. In this case, retrospective COIs may be more efficient in measuring the burden of illness.

For many decades after their inception, the commonest type of COI study resembled that traditionally spelt out by Rice [4] and Cooper and Rice [12]; that is of retrospectively allocating total national health care expenditures in a top-down fashion by the prevalence data of illness or disease categories. The results of such a study design have however been seen with suspicion because of the large potential biases and high level of approximation. Conversely, prospective incidence-based COI studies with a bottom-up approach are those with the highest informative power even though more demanding in terms of data gathering. More recently, retrospective COI studies are being used more confidently by the investigators since new study designs have been identified that minimise the disadvantages of the traditional framework. First, similarly to the incidence-based design, specific diseases and not entire disease categories are the object of study [6]. This allows more accurate data on the clinical management of the illness to be collected. Second, patients are enrolled in the field – although retrospectively – in order to observe the consumption of health care as well as non-health resources through a bottom-up approach. This allows for the collection of original data and highlights the health care resource consumption that arises from the disease and the disease's comorbidities as well as any other cost component that have arisen because of the illness (e.g. out-of-pocket expenditures, transportation). Third, unit costs are used instead of national health care expenditures to evaluate consumption in monetary terms. This eliminates the misallocation problems envisaged in the allocation of total national expenditures since just the resources strictly consumed because of the disease are evaluated.

However, retrospective COIs can be affected by *recall bias*. Providers as well as patients when questioned about the resources consumed because of the illness might not exactly recall what the consumption actually was, this would misallocate the economic costs of illness. Recall biases can however be at least partially removed by selecting providers with efficient informative systems where patients' records are electronically kept. For those resources that are not routinely recorded

by information systems like for instance transportation, and informal care, diaries can be given to patients and/or caregivers. Again, in order to minimise patients and caregivers' recall bias the investigation can be done either retrospectively for a limited period of time (e.g. did you visit any specialist privately in the last 4 weeks?) or prospectively by assuming that the private consumption of medical services and transportation costs do not change from one period to another [11].

4. Cost evaluation in COIs

The cost-of-illness is estimated by identifying the cost-generating components and by attributing a monetary value to them. The monetary value is the opportunity cost, the value of the forgone opportunity to use in a different way those resources that are used or lost due to illness. Direct costs and production losses are the cost categories that should be valued to assess the total economic costs of illness.

As to direct costs evaluation, the controversy over the inclusion and evaluation of productivity losses that has accompanied the COI analysis since its early inception seems to have overshadowed the important question of the quality of the data on direct costs. As a result, there seems to be a general impression that estimates of direct costs, whether of total direct costs (i.e. total national health care expenditures) or of direct costs of specific illnesses are relatively straightforward and problem-free, quite accurate and reliable.

Having accurate measures of direct costs is not just a matter of accounting nicety. Direct costs may account for a very much relevant part of total costs for many illnesses and their miscalculation or misallocation may bias political and/or managerial decisions.

One of the reasons why cost analysis has seldom been of concern for economic analysts in the health care field might be attributed to the fact that economists are keen to explain the theoretical framework within which economic evaluation would need to be performed but they seldom combine theory with practical applications. As regards to prices and charges in the health care sector, many economists have warned analysts to use them with no adjustments [14] but few of them have shown how to cost services whenever prices do not reflect the opportunity costs. Most studies use mar-

ket prices unadjusted and it has often been remarked that health economists recognise that market imperfections exist in health care, unless they are undertaking an economic evaluation [15]. While for some cost categories like drugs, rehabilitation items (e.g. eyeglasses, hearing aids, speech devices, etc.) and in some cases transportation, there is a market and prices can be taken as indicative of the marginal costs, in many other cases (hospitalisation, outpatient care, tests, primary care, etc.) market values do not reflect the marginal costs as markets do not operate perfectly or do not have a reasonable level of competition [16]. Against this background it is difficult to talk of opportunity cost in the pure sense of welfare economics.

It is here thought that opportunity costs of health care alternatives can be proxied by an analysis of the resources employed in production [11]. This means that a separate analysis would need to be performed in order to assess the economic value (full cost) of these services that do not have a market price, or whose price is not deemed representative of the opportunity costs of resources used. The full cost concept is the one that approximates the long-run marginal cost [17,18] and is therefore put forward as the correct measure for assessing services.

Two methods can be used to cost services: the micro-costing and the gross-costing approach [8]. With the first method, the cost of a service is assessed by summing up each single cost component (input) that has contributed to the provision of the service. If – for instance – the service to be evaluated is “hospital admission”, resources (e.g. personnel, medications, tests, meals, ...) used to produce the service need to be identified, measured and evaluated to be summed up. This means that the micro-costing is a “bottom-up” approach, i.e. the calculation of production (full) costs consists of transferring inputs to outputs. Conversely, with the gross-costing approach, the cost of a service (e.g. hospital admission) is assessed in a “top-down” fashion, that is by dividing total costs of the service unit (e.g. hospital ward) by the total number of services (e.g. admissions) produced in a period of time. Both the two methods aim at assessing unit costs of the services, however the level of precision attained by them is quite different. The result of the micro-costing approach is the “actual” cost of the service, while the “average” cost is the product of the gross-costing approach [11]. The micro-costing approach is very accurate and can

somehow be considered the “gold standard” for cost assessment. Nevertheless, it is costly and time consuming and its extensive use must be outweighed against the benefits derived by such a detailed analysis. As a general rule, it can be stated that micro-costing is preferred because it spells out the production and cost functions related to the service under study and allows others to see how well the analysis matches their situation, where patterns of care may differ. Micro-costing is suggested when amongst the aims of the research analysis is that of highlighting the cost differences related to the service under consideration. In the Italian study of stroke for instance [19], the investigators also aimed at analysing the cost differences related to the first hospitalisation by type of provider (i.e. general medicine ward, neurology ward, stroke unit). For this reason, the micro-costing approach was used to evaluate patients’ hospitalisation in each of the three models of care, since the use of the national charge like DRG would not have fulfilled this goal. Special attention must be however paid to the representativeness and comparability of such costs [20–22]. As opposite to tariffs and DRGs, it is very much likely that they differ across healthcare providers. Some relevant work has been done aimed at ascertain the transparency and comparability of costing data coming from either gross-costing (top-down) and micro-costing (bottom) approaches in multi-centre studies [23]. Findings reveal that overall the micro-costing approach provides greater consistency and transparency than the gross-costing since it details the cost components [23]. Nevertheless, since the micro-costing approach is time consuming it is important to evaluate when it is worthy adopting. There are technologies, procedures and treatments whose main components are less sensitive to how budgets are formed and spread in healthcare centres. This means that micro- and gross-costing approaches would tend to similar results, thus making not a case for the micro-costing approach. That was the major finding of the cost-effectiveness analysis of the dialysis therapy for end-stage renal disease where consumables represented the largest component of resources [23]. Conversely, for technologies/treatments/procedures with a significant component of staff input and overheads, and significant sharing of staff or facilities between patient groups, a micro-costing approach is the best approach since by detailing cost components it enhances comparability.

5. Productivity costs

Productivity costs measure production lost because of morbidity and mortality. The theoretical grounds on which productivity cost estimation can be traced to is the human capital approach (HCA). Expected future earnings are used on the assumption that they reflect the individual's potential contribution to the economy, or more precisely, that a worker's wage equals the value of his marginal product. Although its basic principles have been used at least since the 17th century [24], the substantial development of the HCA occurred in the early 1960s, at the time when interest among economists was turning to human resources as a neglected and under-supported component of the US economy [25]. The main work using the HCA in the health care field to evaluate the potential for growth was Mushkin's article "health as an investment" [26].

The approach suggested by Mushkin is that of using earnings as a measure of labour product added. The rationale is that wages and salaries are paid in direct return for productive services and correspond to the individual's contribution to production.

Since the earliest COI analysis, morbidity costs were estimated by applying average earnings by age and sex to time off work [4]. Mortality costs measure production lost because of premature death. Calculation of mortality costs considers earnings over a lifetime rather than a single year since, if an individual had not died, he would have continued to be productive for a number of years according to his life expectancy.

Adjustments were also made for those who were out from the labour market. The category that received more attention was that of housewives services [4,27]. Housewives' services were estimated on the basis of their corresponding value in the market, that is by using the replacement value approach. Another approach was that of the opportunity cost. It assumes the economic value of unpaid work to be at least as much as the wage rate that the same person would command in the market place. Basically, if a woman chooses housework over employment, the housework must be equal to or greater than the value of the employment [28,29]. However, it was argued that this approach was inconsistent with that used to value the employed population where what one does is valued rather than what one could be doing [9].

People who were too sick to work or keeping house were considered as following the same labour force

experience as the general population. The underlying hypothesis was that a theoretical influx of these persons into the labour force would not depress the employment rate and the level of earnings [9]. This point was not shared by Smith and Wright [30] in their review of COI studies on mental illness. Firstly, they did not assume full employment as Rice [4] and Mushkin [26] did, and secondly, they questioned the fact that losses of production of patients with chronic mental illness should be considered in the calculus of productivity costs. They doubted that people suffering from chronic mental illness would have worked had they been no longer ill, given high rates of unemployment.

It can be argued however that both the two positions are rather extreme. If the assumption of full employment is not realistic even in years of economic growth, the hypothesis of no employment for chronic mental ill patients (had they not been ill any longer) is similarly rather pessimistic. A fair solution could be that of applying the national unemployment rate at those patients who are too sick to work – although potentially in the active labour force – and assuming that the losses of production refer to the remaining portion of patients only [11]. Production losses can then be estimated as follows [1]:

$$PLNE = NEP \times a \times (1 - u) \times AAW \quad (1)$$

where PLNE is the production losses by not-employed, NEP the total number of not-employed patients (theoretical labour force), a the percentage to be applied to the theoretical labour force to sort out the active labour force, u the national unemployment rate, and AAW is the (general population) national average annual wage.

Other health economists [13,31–35] have questioned the reliability of the HCA in estimating losses of production to society because of illness. In particular, they all pointed to the fact that the HCA will overestimate indirect costs for patient groups in an economy with less than full employment. Koopmanschap and van Ineveld [36] explicitly say that the way in which the HCA is usually applied leads to an estimate of the value of *potential* lost production as a consequence of disease, whereas the *actual* loss for society may be much smaller. By and large, production losses may stem from short and long term absences from work.

For short term absences, a person's work may be covered by others or made up by the sick person on his return to work. Secondly, employers usually have excess capacity in the labour force to cover absenteeism. Thus, short term absences may be a cost to the individual but not to society. For long term absences (such as in the case of premature death), an individual's work can be covered by someone drawn from the ranks of the unemployed. Therefore, while absence from work may cost the individual, or that person's employer, it may not cost society very much [13]. What is needed is to estimate the period necessary to replace a sick worker, or to reorganise the production process. This period has been called the friction period [36]. The length of this time span depends on the availability of qualified personnel within firms and on the labour market and on the level of unemployment. During this period, the production falls or remains equal to extra costs. Consequently, the productivity costs are proportional to the length of the adaptation period.

Hutubessy et al. [37] assessed the productivity costs of back pain in the Netherlands through the HCA and the friction cost method (FCM). The results of this study show that the productivity costs estimated by the HCA were more than three times as high as the productivity costs estimated by the FCM. However, it should be noticed that: (i) the FCM requires a huge amount of information (i.e. elasticity for annual labour time versus labour productivity) that are unlikely to be available at country level, and that (ii) the results may change over time within the same country since the friction period depends on the macroeconomic context. In the study of back pain for instance, the authors clarify that when the study was performed (1990) the unemployment rate was 8.2% in the Netherlands, resulting in an average friction period of 2.8 months. More recently, unemployment has decreased to 4.5% implying that the current friction period may be longer and the actual productivity costs may be higher than those estimated in 1990 [37].

5.1. *Informal care*

Informal care is care provided by lay people. Specifically, informal caregivers can be defined as family, friends, acquaintances, or neighbours of a patient who provide care for which they do not have to be financially compensated [38]. Informal care is a component

of direct costs but its assessment follows the methods used for productivity costs.

The valuation of informal care is rather difficult as caregivers do not lose just part of their time, but they also experience intangible effects like fatigue, leisure time forgone and fewer social contacts. The first problem is assessing the exact quantity of time spent on informal caregiving. During the time providing informal care, many normal activities often can continue as usual [38]; thus joint production occurs, as in the case of surveillance. It may be very difficult to separate normal activities from caregiving activities. Preparing and serving meals may be defined a normal activity, but helping someone to eat is not. This problem may be solved by using a structured interview or questionnaire in which caregivers are asked about the type of care provided and loss of time from paid work, unpaid work and leisure time.

If informal care is provided at the expense of paid labour, the worker's labour costs will be used to estimate informal care. The method follows that used for estimating productivity costs. If, however, informal care is provided at the expense of unpaid work and/or leisure time the cost measurement is not straightforward. As for patients, the valuation of unpaid working time by caregivers can be valued by the market wage, that is by using the wage a person would have received for a similar work in the market. An attempt to define how much to value leisure time has been done by Posnett and Jan [39]. If caregivers' time allocated to the provision of caring and nursing of patients involves a displacement of non-working time, the cost can be given by the value (to the individual) of leisure activities forgone, that is by using the net hourly income caregivers would earn had they provided that activity in the market. There is not however much agreement on the valuation of leisure time. Brouwer et al. [38] suggest for instance to value lost leisure time through quality of life and not monetarily. Alternatively, the costs of informal care can be assessed through the replacement approach, by assuming that a professional worker would have been hired had the lay person (i.e. caregiver) not been available.

The last is thought to be the most consistent with the methods used to cost the other cost components since what is important is the type of activity provided to patients more than the nature of time (i.e. leisure or work) displaced.

5.2. Theoretical issues behind COIs

After the first attacks on COI analysis from welfare economists on the ground that COIs have nothing to do with resource allocation and with decision-making since they do not say anything on the worthwhileness of public projects, COI analysis has been deeply debated within health economics and its role as a tool for decision-making is still unclear.

At the end of 1980s, a debate fostered by Shiell et al. [40] on COIs followed the publication of Behrens and Henke's [41] work on the economic costs of illness in the Federal Republic of Germany. Shiell and colleagues' critiques pointed to: (1) the view of cost-of-illness estimates as the economic benefits of treatments, (2) the HCA and (3) the usefulness of COI in the policy context. COI studies have been criticised because of the embedded circularity. As Drummond et al. [42] already noted, Shiell and colleagues pointed out that COI analysis may lead to priority being given to those illnesses which are apparently costly because they already have a large amount of resources devoted to them. If past resource allocation decisions have been made in an irrational manner, subsequent policy decisions perpetuate and amplify the initial mistake. Shiell et al. conclude that although more information is needed on the social impact of illness, the natural history of disease and the effectiveness of different medical intervention, it is not COI analysis that can accomplish these tasks. Health economics – they say – does have more to offer health services planners in terms of ranking of treatment programmes according to their cost/QALY scores while COI studies can only confuse, mask and mislead. However, in spite of that, it can be said that COIs provide a useful framework for the problem to be addressed. Decision-makers become knowledgeable about the actual resources allocation and if they have been allocated in an irrational way for whatever reason, they can readjust them upon knowledge of this.

The HCA and more generally, the whole analysis concerning the economic costs of illness were strongly criticised by welfare economists who denied that individuals' productivity and wages had any relationship to the amount that should be spent to save their lives. Conversely, it was put forward that allocative decisions must be taken on the basis of the individuals' WTP. This led researchers to study what type of relationship exists between COI and the human capital

and WTP approaches. The assumption is that the economic costs of illness might be considered to have a lower bound to the individuals' WTP to reduce the risk of getting ill. Theoretical models suggest that WTP reflects four components: (1) lost wages, (2) additional medical expenses, (3) the dollar-value of the disutility of additional illness and (4) the change in defensive expenditures [43]. In contrast, COI studies measure the lost wages and the additional medical expenses. The amount an individual would be willing to pay is larger than the COI measure as long as the omitted quantities of the WTP measure are positive. WTP also overcomes the problem of incorporating non-monetary information on quality of life into the economic calculus. The quality of life of ill patients may be reduced beyond the restorative capability of current rehabilitation efforts. For that reason, the quantifying of intangible costs remain rather intractable, with the HCA at least. WTP partly solve this problem by providing a comprehensive estimate of the value attached to a reduction in risk of illness. Individuals are asked to reveal their maximum WTP and by doing so, they tend to consider all possible (physical and psychosocial) effects [44] a reduction in health risk may have on to their lives. If a WTP methodology is used, then many intangible costs are implicitly included in the monetary values presented. Stated in this way, COI measure of the benefits of a reduction in morbidity is a lower bound to the theoretically correct WTP measure, although not necessarily a very good approximation of it [45].

That the HCA has nothing to do with individuals' WTP for risk reduction and that discriminates amongst paid and unpaid workers as well as between workers is a well known issue. However, the WTP approach seeks to avoid the issue of distribution of income through an appeal to the doctrine of the potential Pareto improvement. The problem is that Pareto optimality is biased towards maintenance of the status quo and if the initial distribution of income is not optimal, as assumed by welfare economists, the strict adherence to the potential Pareto improvement criterion perpetuates the unequal distribution of income. Moreover, individuals' WTP is strongly related to their ability to pay. At one extreme this means that rich people are willing to pay more to reduce risks as well as the HCA values time off work because of illness higher for people with high remunerative jobs. This issue could be tackled by using an average national wage instead of age-, gender- or

sector-specific wages. In addition, as Weinstein et al. [46] point out, the WTP approach reveals that people place different values on risk-bearing in different situations and that adherence to this preference-ranking could direct governmental expenditures into curative programmes aimed at acutely ill individuals, to the detriment of potentially more effective preventive programmes. This would not worry the advocates of the WTP approach who says that only individuals' values count [47]. There is however another set of values in society that emphasises the ability and responsibility of society and its members to improve themselves, largely through an educational process that molds values and ultimately choices. The emphasis given by WTP to individuals' responsibilities and its rejection of paternalism denies the fact that individuals' attitudes towards risk are not born with the individual but evolve throughout life on the basis of personal experiences and contact with organisations whose efforts are explicitly devoted to influence those attitudes (see for instance anti-smoking campaigns and other health promotion programmes). WTP values are influenced by governmental programmes and regulations and, hence, are endogenous to the very system they are supposed to guide!

The whole debate on the human capital technique can be approached by saying that it does not aim to provide the economic value of human life [11]. What the HCA does value is a certain component of the cost of disease. Morbidity and mortality, by causing persons to lose time from work and other productive activities, destroy labour, a valuable economic resource. Diseases thus create a loss to society and it is this loss that the HCA attempts to measure. Critiques of the HCA must then be made on the basis of how well it measures this aspect of the burden of diseases and whether is useful. Limits of estimates based on the HCA should be made clear by analysts who have to point out the fact that the estimation of productivity costs strongly refer to the specific composition of the population studied in terms of number of paid and unpaid workers, age, sex, type of job performed and present productivity cost estimation separately in order to leave the decision-maker makes his own considerations. The methodology proposed and tried by Koopmanschap and van Ineveld [36] is quite appealing. However, the FCM is not straightforward in its calculations and requires a lot of work at firm level to measure the length of friction periods by

age, sex and education. In addition, this approach does not solve the problem of discrimination against people outside the labour force that has been put forward by various health economists [48,49] against the HCA. Besides, a lot of difficulties arise with the estimation of production losses for unpaid workers. These data are difficult to retrieve at national level. Additionally, if the use of the HCA would end in an overestimate of the actual losses of production, it would however facilitate comparisons among COI studies (as well as inter-temporal comparisons) as: (i) it is the commonest method used since mid-60s and (ii) it depends less than the FCM on macroeconomic variables.

Lastly, the implicit contention that welfare economics based methods provide a fundamental link between analysis and policy making is not entirely true. The criterion that underlies every CBA is that of Pareto optimality, but Pareto optimality is a very weak basis on which to evaluate public policy: firstly, because it assumes consumers' sovereignty and optimal distribution of income which do not exist, and secondly, because decision-makers are very reluctant to consider benefits from health care programmes in monetary terms. Warner and Luce [50] point out that the perception that CBA should produce policy decisions is a common but misguided one and that in general CBA does not answer a basic policy question. Neither CBA nor cost-effectiveness analysis (CEA) or COI are policy decision-making techniques. It seems there is a largely shared view that at best these techniques can serve as one of several inputs into a complex decision-making process which will always remain a political process [50–52]. In recently published surveys, the influence of economic evaluation studies on decision-making has been assessed [53–55]. In these studies, it is concluded that results of economic evaluation analysis are not widely used in decision-making. Decision makers feel that institutional dimensions, such as difficulties in transferring budgets, and the lack of credibility of studies are among the most important barriers and these apply to all economic technique.

6. Do COIs still have a role in health economics?

COI studies can have an important role in health economics as a decision-making tool. A major error

has been that of continuing to consider COI analysis as a sort of CBA. Put it in this way, the majority of the efforts made by health economists in the last 30 years have been directed either to demonstrate how far COI studies are from the well accepted body of welfare economics or to find relationships between COI and the WTP techniques that could give credibility to COIs [56]. It is here believed that, although these efforts are of paramount importance in aiding a better understanding of the main shortcomings of the two approaches, COI studies could gain relevance if they are considered from another perspective.

COI analysis is different from any other economic evaluation analysis because it basically does not compare costs with outcomes. It is a descriptive study whose main objectives are:

1. to assess the economic burden of illness to society. This would give information about the amount of scarce resources consumed because of illness and, along with epidemiological data on morbidity and mortality may help ranking diseases according to global burden,
2. to identify the main cost components and their incidence over total costs. This would help health policy makers defining and/or limiting: (i) any cost containment policies to those cost components that weigh heavily on total costs and (ii) controlling for the actual implementation of previous health policies. In the case of schizophrenia for instance, it emerged that the de-institutionalisation process that has occurred in Italy since 1978 had not always been substituted for by effective community care. In some geographical areas this resulted in a shift of costs from the Italian NHS to the families who had to bridge the gap brought about by the ineffective and inefficient application of the national mental health policy [57],
3. to identify the actual clinical management of illness at a national level. This would help policy makers and managers to analyse the production function used to combine inputs and/or intermediate services to deliver the final output that could range from a single product, such as hospitalisation to an entire therapeutic pattern which encompasses multiple medical services. Inefficient and/or ineffective functions can then be put forward and may represent the base for the re-engineering of the whole

process in case of inefficiencies, and for the re-assessment of the clinical strategy as it happens for the evidence-based medicine. Clinical guidelines may for instance be one of the final products in this case, that is whenever the identification of the clinical management of illness is judged ineffective or too diverse in the same country,

4. to explain the variability of costs. In this case, statistical analysis can be performed to check whether there is a relationship between costs variability and variables, such as those related to the illness (e.g. severity), to the patient (e.g. demographic variables) or to the health care providers (e.g. teaching hospitals versus district hospitals). These results would help managers to feed the planning process with more accurate information as to the future provision of services, since knowing the cost drivers that explain – at least partially – the consumption patterns of services can be very useful when planning the provision of health care services. In the COI study on schizophrenia [57], the authors found that nearly 40% of direct costs variability could be explained by type of accommodation and the presence of extra pyramidal symptoms (EPS). Patients with EPS consumed 2.4 times more health care resources than those without EPS. This finding can help decision makers forecasting the future consumption of health care services (e.g. hospital care, outpatient care, drugs, ...) when combined with the characteristics of the local patients group.

In order to reach these goals, COIs need however to be bottom-up and the top-down approach has to be definitely abandoned.

7. Conclusion

COI was the first economic evaluation technique used in the health field. The principal aim was to measure the economic burden of illness to society. COIs have been traditionally retrospective prevalence-based studies conducted by allocating national health care expenditures to different broad disease categories. The informative power of such an economic study relied upon the implicit assumption that the economic costs of illness would represent the economic benefits of a health intervention had it eradicated the illness.

Its usefulness as a decision-making tool has however been questioned since its inception. New types of COI studies have recently emerged and are now gaining importance due to the unrealistic assumption of an all-of-nothing approach which embedded the traditional COIs. These relate to COIs conducted in the field by a bottom-up approach, which is by estimating the quantities of services used by the illness by going prospectively or retrospectively through the patient's records. If the COI study is incidence-based, the informative power for decision-makers is also more relevant as it would entail the identification of the patients' clinical management pathways, or the production function, currently used in a country/region to diagnose and treat the illness.

Finally, COI can be a good economic tool to inform decision makers if it is considered from another perspective. COI is a descriptive study that can provide information to support the political process as well as the management functions at different levels of the health care organisations. To do that the design of the study must be innovative, capable of measuring the true cost to society, to identify the different subjects who bear the costs and to explain cost variability. In a study by Koopmanschap [6], it has been found that out of 30 COI studies randomly sampled from the then recent literature no attempt has been made to explain past trends in disease costs, thus leaving the study's findings rather static and difficult to use in the policy context.

Observational studies (incidence- or prevalence-based, prospectively or retrospectively designed) with a bottom-up approach and with the measurement of costs based upon accounting principles (i.e. micro- or gross-costing approach) can fulfil these goals and are therefore put forward as the studies that can interpret the decision makers needs better than the traditional ones.

References

- [1] Bracco A. Gli studi di cost-of-illness in Italia: una review della letteratura. *Farmeconomia e Percorsi Terapeutici* 2001;2(4):247–59.
- [2] Byford S. Cost-of-illness studies. *BMJ* 2000;320:1335.
- [3] Bloom BS, Bruno DJ, Maman DY, Jayadevappa R. Usefulness of US cost-of-illness studies in healthcare decision-making. *Pharmacoeconomics* 2001;19(2):207–13.
- [4] Rice DP. Estimating the cost-of-illness. Washington, DC: US Department of Health, Education, and Welfare, Public Health Service; 1966.
- [5] Hodgson T, Meiners M. Cost-of-illness methodology: a guide to current practices and procedures. *Milbank Memorial Fund Quarterly/Health and Society* 1982;60(3):429–62.
- [6] Koopmanschap MA. Cost-of-illness studies. Useful for health policy? *Pharmacoeconomics* 1998;14(2):143–8.
- [7] Rice DP. Cost-of-illness studies: what is good about them? *Injury Prevention* 2000;6:177–9.
- [8] Gold MR, Siegel JE, Russel LB, Weinstein MC. Cost-effectiveness in health and medicine. New York: Oxford University Press; 1996.
- [9] Hartunian NS, Smart CN, Thompson MS. The incidence and economic costs of cancer, motor vehicle injuries, coronary artery disease and stroke: a comparative analysis. *American Journal of Public Health* 1980;70:1249–60.
- [10] Rice DP. Cost-of-illness studies: fact or fiction? *Lancet* 1994;344:1519–20.
- [11] Tarricone R. *Valutazioni economiche e management in Sanità*. Milan: McGraw-Hill; 2004.
- [12] Cooper B, Rice DP. The economic cost-of-illness revisited. *Social Security Bulletin* 1976;39(2):21–36.
- [13] Drummond MF. Cost-of-illness studies: a major headache? *Pharmacoeconomics* 1992;2(1):1–4.
- [14] Mishan EJ. Evaluation of life and limb. *Journal of Political Economy* 1971;79:687–705.
- [15] O'Brien B, Stoddart G, Torrance G. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 1997.
- [16] Roberts JA. Managing markets. *Journal of Public Health Medicine* 1993;15:305–10.
- [17] Allen C, Beecham J. Costing services: ideals and reality. In: Netten A, Beecham J, editors. *Costing community care*. Ashgate; 1993.
- [18] McCrone P, Thornicroft G, Phelan M, Holloway F, Wykes T, Johnson S. Utilisation and costs of community mental health services. *British Journal of Psychiatry* 1998;173:391–8.
- [19] Gerzeli S, Tarricone R, Zolo P, Colangelo I, Busca MR, Gandolfo C. The economic burden of stroke in Italy. The EcLIPSE study. *Neurological Science* 2005;26(2):72–80.
- [20] Schulman K, Burke J, Drummond M, Davies L, Carlsson P, Gruger J, et al. Resource costing for multinational neurological clinical trials: methods and results. *Health Economics* 1998;7:629–38.
- [21] Koopmanschap MA, Touw KCR, Rutten FFH. Analysis of cost and cost-effectiveness in multinational trials. *Health Policy* 2001;58:175–86.
- [22] Raikou M, Briggs A, Gray A, McGuire A. Centre specific or average unit costs in multi-centre studies? Some theory and simulation. *Health Economics* 2000;9:191–8.
- [23] Wordsworth S, Ludbrook A, Caskey F, Macleod A. Collecting unit cost data in multicentre studies creating comparable methods. *European Journal of Health Economics* 2005;6(1):38–44.
- [24] Petty W. *Political arithmetic, or a discourse concerning the extent and value of lands, people, buildings, etc.* London: Robert Caluel; 1699.

- [25] Denison E. The sources of economic growth in the united states and the alternatives before us. New York: Committee for Economic Development; 1980.
- [26] Mushkin SJ. Health as an investment. *Journal of Political Economy* 1962;(Suppl.):129–57.
- [27] Brody W. Economic value of a housewife. Research and statistics note no. 9. Washington: US Department of Health, Education, and Welfare; 1975.
- [28] Gronau R. The measurement of output of the nonmarket sector: the evaluation of housewives' time. The measurement of economic and social performance. Washington: National Bureau of Economic Research; 1973.
- [29] Murphy M. The value of nonmarket household production: opportunity cost versus market cost estimates. *Review of Income and Wealth* 1978;24:243–55.
- [30] Smith K, Wright K. Cost of mental illness in Britain. *Health Policy* 1996;35:61–73.
- [31] Lindgren B. Costs of illness in Sweden 1964–1975. Lund: Liber; 1981.
- [32] Williams AH. Economics of coronary artery bypass grafting. *BMJ* 1985;291:326–9.
- [33] Drummond MF, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford Medical Publications; 1987.
- [34] Gerard K, Donaldson C, Maynard AK. The cost of diabetes. *Diabetic Medicine* 1989;6:164–70.
- [35] Levin LA, Jönsson B. Cost-effectiveness of thrombolysis—a randomized study of intravenous rt-PA in suspected myocardial infarction. *European Heart Journal* 1992;13:2–8.
- [36] Koopmanschap MA, van Ineveld BM. Towards a new approach for estimating indirect costs of disease. *Social Science and Medicine* 1992;34(9):1005–10.
- [37] Hutubessy RCW, van Tulder MW, Vondeling H, Bouter LM. Indirect costs of back pain in the Netherlands: a comparison of the human capital method with the friction cost method. *Pain* 1999;80(1–2):201–7.
- [38] Brouwer WBF, van Exel NJA, Koopmanschap MA, Rutten FFH. The valuation of informal care in economic appraisal. *International Journal of Technology Assessment in Health Care* 1999;15(1):147–60.
- [39] Posnett J, Jan S. Indirect cost in economic evaluation: the opportunity cost of unpaid inputs. *Health Economics* 1996;5(1):13–23.
- [40] Shiell A, Gerard K, Donaldson C. Cost-of-illness studies: an aid to decision-making? *Health Policy* 1987;8:317–23.
- [41] Behrens C, Henke K-D. Cost-of-illness: no aid to decision-making? Reply to Shiell et al. *Health Policy* 1988;10:137–41.
- [42] Drummond MF, Ludbrook A, Lowson K, Steele A. Studies in economic appraisal in health care, vol. 2. Oxford: Oxford University Press; 1986.
- [43] Berger M, Blomquist G, Kenkel D, Tolley G. Framework for valuing health risks. In: Tolley G, et al., editors. *Valuing health for policy*. Chicago: The University of Chicago Press; 1994.
- [44] Olsen JA, Smith RD. Theory versus practice: a review of 'willingness-to-pay' in health and health care. *Health Economics* 2001;10(1):39–52.
- [45] Kenkel D. Cost-of-illness approach. In: Tolley G, Kenkel D, Fabian R, editors. *Valuing health for policy*. Chicago: The University of Chicago Press; 1994.
- [46] Weinstein M, Shepard D, Pliskin J. The economic value of changing mortality probabilities: a decision-theoretic approach. *Quarterly Journal of Economics* 1980;94(2):374–96.
- [47] Mishan EJ. Cost-benefit analysis. London: Routledge; 1988.
- [48] Jönsson B. Cost-benefit analysis in public health and medical care. Lund: Liber; 1976.
- [49] Johannesson M, Jönsson B. Economic evaluation in health care: Is there a role for cost-benefit analysis? *Health Policy* 1991;17:1–23.
- [50] Warner KE, Luce BR. Cost-benefit and cost-effectiveness analysis in health care: principles, practice, and potential. Ann Arbor, MI: Health Administration Press; 1982.
- [51] Hodgson T. Cost-of-illness studies: no aid to decision-making? Comments on the second opinion by Shiell et al. *Health Policy* 1989;11:57–60.
- [52] Tolley G, Kenkel D, Fabian R. *Valuing health for policy*. Chicago: The University of Chicago Press; 1994.
- [53] Hoffmann C, von der Schulenburg J-MG. The influence of economic evaluation studies on decision-making. a European survey. *Health Policy* 2000;52:179–92.
- [54] Drummond M, Brown R, Fendrick AM, Fullerton P, Neumann P, Taylor R, et al. Use of pharmacoeconomics information—report of the ISPOR task force on use of pharmacoeconomic/health economic information in health-care decision-making. *Value in Health* 2003;6(4):407–16.
- [55] Maiwen J, Talitha F, Werner BF. Decision makers' views on health care objectives and budget constraints: results from a pilot study. *Health Policy* 2004;70:33–48.
- [56] Glied S. Estimating the indirect cost-of-illness: an assessment of the forgone earnings approach. *American Journal of Public Health* 1996;86(12):1723–8.
- [57] Tarricone R, Gerzeli S, Montanelli R, Frattura L, Percudani M, Racagni G. Direct and indirect costs of Schizophrenia in community psychiatric services in Italy. The GISIES study *Health Policy* 2000;51:1–18.