

# Methodological issues for the economic evaluation of health interventions: a concise state of the art

GIAN LUCA DI TANNA<sup>(1)</sup>, GIANLUCA BAIO<sup>(2)</sup>

## ABSTRACT

This paper presents a preliminary report of the Italian Society of Medical Statistics and Clinical Epidemiology (SISMEC) working group called SiPrEMAS (Evidence Synthesis and Decision Modelling in Health) collating some topics addressed throughout the first two years of collaboration. It contains a rapid overview of the principal methods used for the economic evaluation of health interventions. Special focus is given to the process of assembling and pooling the available evidence, modeling methods, the analysis of uncertainty (structural and on parameters), cost analysis and cost consequences analysis. This paper intends to stimulate the discussion among different professionals involved in the decision making process at national level, trying to (re)bridge the gap between decision makers and researchers.

*Key words: Systematic reviews; Meta-analysis; Cost effectiveness analysis; Markov model, Analysis of uncertainty; Cost consequence analysis; Economic evaluation; Health technology assessment*

(1) UBC-United BioSource Corporation, London (UK) & Facoltà di Farmacia e Medicina, Sapienza-Università di Roma, Rome Italy

(2) Department of Statistical Science, University College London (UK)

**CORRESPONDING AUTHOR:** Gian Luca Di Tanna; UBC-United BioSource Corporation, UBC 26-28 Hammersmith Grove London W6 7HA UK & Facoltà di Farmacia e Medicina, Sapienza-Università di Roma, Rome Italy. e-mail: GianLuca.DiTanna@unitedbiosource.com

**DOI:** 10.2427/8917

## INTRODUCTION

Health systems around the world are faced with many relevant issues, such as the changes in socio-demographic conditions, the growing expectations of health systems' users and the (rather) uncontrolled proliferation of health technologies.

This situation induces a growing need for an efficient and effective allocation of a limited amount of resources, in a context characterized

by a growing necessity for high quality health care. The public decision maker thus needs a set of methodological tools to support the process of decisions making, which are based on the existing information and sustainable over time.

Nevertheless, different analyses reach the goal of evaluating a growing amount of information and are characterized by a better quality at the price of increasingly higher complexity. Furthermore, a process of

fragmentation of available scientific studies is taking place. Lastly, the emphasis on hypothesis testing (and on model building), the identification of risk factors and the not always accurate description of the variability associated to the results makes difficult - at times impossible - the integration of sources of information, which is an essential step for making adequate decisions [1].

In close analogy with what happened regarding the *Evidence Based Medicine* movement, the explicit use of up-to-date and best proofs of efficacy (which integrate the individual clinical expertise and patients' expectations) of the assistance of patients [2], *Medical Decision Making* is becoming increasingly established as a discipline based on scientific evidence and, at the same time, on (decisional) models built to address the system's choices in an optimal manner.

The evaluations of the clinical efficacy and of the economic impact are some of the pillars on which Health Technology Assessment (HTA) procedures are held. HTA is a multidisciplinary discipline which plays a supporting role to the programming through which health technologies are critically examined under various points of view (clinical, epidemiological, economic, ethical and organizational) [3].

Therefore, HTA has the difficult task of predicting, as most accurately as possible, the consequences (i.e.: the costs and benefits) of specific health policies, paying attention to the medium and long-term direct and indirect effects. In light of this, in order to perform an appropriate HTA, an estimation of the following variables cannot be avoided:

- the probability that a technology will bring a significant health improvement;
- the probability that a technology will have a significant impact on other health policies (e.g.: reduction in inequalities);
- the probability that a technology will have a significant impact on Health System's resources (financial or other) and/or, more generally, on the resources of society as a whole.

Thus, an independent evaluation of the relevant proofs of efficacy of the technologies under study is an integrating part of the above process, specifically through the systematic identification of strengths and weaknesses of the available literature (phase 1). Models for

decision making analysis are subsequently employed, which help in determining and estimate the patients therapeutic (or diagnostic) pathways in a pre-determined time-frame, integrating information on costs and benefits so as to identify the optimal strategy (phase 2).

Essential components of model building are *i)* the quantification of the uncertainty associated to the choice of a particular strategy; and *ii)* the possibility of identifying which of the parameters used in the model have mostly contributed towards such uncertainty in the decision making process.

Consequently, two basic problems exist: The first has as a focus on the ability to build "a bridge between theory and practice" [4], that is enabling the decision makers to interpret and consequently use HTA studies adequately - note however that the different use of economic evaluation at local and national level exist and constitutes a limitation [5]. The second problem is the dissemination/diffusion/sharing of methods relevant to the practitioners. This latter aspect is particularly true in Italy, where the absence of a "reference" institution (such as NICE or the NIHR in the UK), who is the repository and guarantor of the methodology, has been perceived in the past. This issue can be partially overcome with the entrust provided by the "Conferenza Unificata Stato Regioni" (September 2007) with the creation of the National Agency for Regional Health Services (AGENAS) which duty is to support the regions in the development of HTA related activities, and with the creation inside AGENAS of RIHTA - the Italian Network for HTA. On the other hand, the national overview is characterized by a delay in receiving, using and comparing opinions and experiences on the various and multiprofessional HTA methods.

Moving from these essential requirements, a strong need for a gathering of professionals who could share and discuss methodological issues was felt; this is how a working group SiPrEMAS - Sintesi delle Prove di Efficacia e Modellistica per l'Analisi Decisionale in Sanità was established within the SISMEC - Italian Society of Medical Statistics and Clinical Epidemiology. The primary objective of SiPrEMAS is the creation of a "platform for confrontation" where to share methods for the integration of heterogeneous sources of "evidence" and for the decision-making modelling, and for enhancing in value the

experiences so far acquired. In particular, if the problems to be overcome in the decision making analysis are on clinical, economical, epidemiological and statistical issues, the methodologist (i.e.: the medical statistician, the epidemiologist, the health economist) is required to interact with different HTA team professionals; moreover, she/he is also required not just at the stage of quantitative synthesis of sources or at the stage of choosing the decisional technique or building/validating the model and executing the analyses. Typically, methodology has to be accounted for properly also at the delicate stage of translating such a process in results realistically usable for the decision makers. The objective with the highest priority is that of providing detailed information (supported by evidence), helpful for the decision makers when coming to make decisions on the appropriateness and efficacy of resource allocation [6].

Clearly, the topics to be dealt with are various and all deserving of specific attention. However, in this paper we mainly focus on some aspects of which the group members have already performed applications and deepened their knowledge. In particular, these are:

- Systematic Reviews and Meta-Analysis
- Model Building
- Uncertainty Analysis
- Cost Analysis
- Cost-Consequences Analysis

## SYSTEMATIC REVIEWS AND META-ANALYSIS

The amount of scientific information available in the literature has grown noticeably in the last decades. About one million articles are published every year on 30 000 scientific and divulgative journals, along with 17 000 biomedical textbooks. Currently, the PubMed search engine includes over 19 million citations. The difficulty the health operator has to deal with when she/he attempts to identify, select and critically evaluate information relevant to his/her clinical question is evident. And also the quantitative synthesis of the results of the information retrieved is difficult. Literature reviews were established to address these needs. The revision process has evolved over time: it has moved from narrative reviews, based on studies which selection and evaluation

methodologies were not always made explicit, to the development of systematic reviews including a meta-analytic approach to the evidence synthesis.

The expression “systematic review” is generally referred to the process of literature revision planned according to an explicit and reproducible methodological pathway, specified by means of a protocol aimed at minimizing possible distortions and wrong conclusions due to the loss or omission of important studies conducted on a given topic [7] – a process sometimes referred to as “cherry-picking”.

The fundamental steps in a systematic review that answers a specific clinical question are: the systematic research of bibliographic sources, the evaluation of the quality of eligible studies, the qualitative synthesis of the information and the quantitative synthesis of the results. The latter phase, that it is not always advisable to perform, is usually called “meta-analysis”, and is defined as the statistical combination of results coming from independent studies carried out to answer the same clinical question [8-10].

The evaluation of the quality and variability of the characteristics of the studies is a very important aspect in the process of systematic reviews. Despite the multiple methods for the qualitative evaluation of single studies in the literature, the majority of these are based on the analysis of specific characteristics: the study design, the characteristics in the enrolment of patients under comparison, the methods of measurement and evaluation of outcomes, the characteristics of the patients for the entire period of observation.

The essential element in a systematic review is the evaluation of the variability or heterogeneity of the study characteristics that can have a direct effect on the results. Studies can differ for various clinical and methodological reasons. The clinical heterogeneity relates to differences among studies concerning patients, treatment and setting. The methodological heterogeneity relates to the study design adopted (i.e. observational, experimental), the quality of the study and the type of analysis performed.

The first attempt to control for heterogeneity usually takes place during the planning phase of a systematic review, by defining criteria for the selection of studies based on the type of study design, on the modality of treatment/

intervention and on the type of patients. Such criteria are chosen a priori so as to select studies that are as homogenous as possible. The presence of heterogeneity between studies is then evaluated a posteriori, by means of a statistical test able to assess the evidence against the hypothesis that the variability of results is due to chance, and not to possible clinical and methodological differences among studies. A significant result indicates excess of variability, defined as *statistical heterogeneity*. This variability is attributed to the methodological and clinical differences among the selected studies. In such cases, reviewers tend not to perform meta-analysis and the systematic review will be concluded with the simple description of the characteristics and quality of the studies and of the results obtained.

In many systematic reviews the presence of statistical heterogeneity is tackled by means of subgroup analysis. Studies are grouped according to similarity in the characteristics that are likely to be the cause of the variability in the results. The choice of the characteristics must be specified in the protocol. Results are combined within such subgroups and group-specific estimates of treatment effect are provided only in the case of absence of statistical heterogeneity within the groups.

In order to explore some sources of variability, techniques have also been developed, such as the meta-regression that allows to evaluate the association between the investigated measures of effect and specific covariates collected at the study level.

At this point, one must recall that the assessment of the heterogeneity among studies, even though this is an integrating part in the evaluation of random-effects models (different from fixed-effects models for the explicit use of the heterogeneity estimation in the process of pooling), goes beyond the statistical computation and requires a careful analysis particularly, for instance, with regards to the part for which a regression model is not able to account.

Systematic reviews are a valid instrument for the selection of the best published studies, for their qualitative and quantitative evaluation, and are useful in solving controversies when studies provide discordant results. These play an important role when making recommendations on clinical behaviour, in assisting medical staff and patients in the decision-making and in the appropriate management of specific clinical

conditions, and for identifying areas to be addressed by future research.

Furthermore, the meta-analytic results are, especially if of clinical connotation, one of the privileged inputs (particularly because deduced by the available literature) that support decision-making models of economic evaluations.

## MODEL BUILDING FOR THE DECISION-MAKING ANALYSIS

The role of the decision-making model building for economic evaluations of alternative health interventions is commonly recognised and appreciated as a tool to support decision makers involved in public health policies. Well calibrated models are often able to predict realistically direct and indirect effectiveness and costs of the interventions (under observation and, through the quantitative synthesis of proofs of effectiveness, they allow the identification of specific gaps in research and the extension of the time-span, often limited, of the traditional epidemiological studies. Model building approaches that can be adopted to answer a question regarding the economic evaluation of health intervention are various [11-16]. The choice of the type and structure of the model depends primarily on the intervention taken into consideration, and on the nature and availability of the data needed for the parameterization. In the phase of project planning of an economic evaluation, two questions play a particularly relevant role for the identification of the principal types of analysis:

- The first relates to the degree of independence of the individuals; in other words, the possibility of interaction among study subjects that can be due to factors such as infections or limitations in the modality through which treatments can be administered (e.g.: budget limitations or waiting lists). In practice, economic evaluations are largely conducted assuming independence among individuals and the most used techniques in such cases are decision trees and Markov models [17];
- The second relates to the grouping level that is wished to be adopted and that distinguishes cohort models from individual agent models (microsimulation). The first quantifies the proportion of people

with common characteristics and can take into account different patients' attributes/characteristics, subdividing the number of stages or decisional branches according to these. In complex cases where the structure dimensions increase exponentially, the usefulness of this approach is limited. Individual-based models, instead, allow the simulation of the progression of each individual with his/her own characteristics.

### Decision trees

These are the simplest and most used cohort-type models under the hypothesis of independence: all possible patients' pathways through clinically relevant stages are explicitly shown with associated probabilities of event and measures of outcome (in terms of efficacy, cost and net benefit) [17]. The expected value of a particular decision is calculated by summing the products between the probabilities of each tree branch (the proportion of individuals that will follow that path) and the outcome relative to the specific condition ("roll-back" method). Usually, the time-span is short and patients' mortality is not differential among the studied strategies. However, it is difficult to study recurrent events unless the simplicity of the tree is sacrificed. The simulated decision trees constitute a valid alternative for the calculation of the mean value of each decision: this approach generally uses Monte-Carlo simulations, which allow the estimation of the uncertainty in quantitative problems by calculating a series of possible future realizations of the phenomenon under examination. In practice, the number of individuals for each tree branch is simulated and the greatest advantage of such technique derives from the possibility of evaluating the variability in the number of patients that will end-up at the different stages [18].

### Markov models

This type of models is becoming increasingly popular in economic evaluations thanks to its versatility and simplicity in the building process: the greatest advantage is in the ability to model "risks in time" and repeated events. However, a limitation to this technique is that it does not allow for an interaction

between patients (although this can be partially modelled with the software advancements). At any moment, each individual could find him/her-self in one (and only one) of a finite series of stages (or health conditions) assumed to be exhaustive and mutually exclusive. To each stage is attributed a cost and a particular measure of efficacy (or utility) that can be of different types according to the nature of the status (i.e. repetitive, tunnel, absorbent). The cohort is followed-up at regular time-intervals, defined as cycles, until all subjects move to the absorbent stages (i.e. those from which it is not possible to move away, e.g.: death). All possible movements from one stage to another occur according to some probabilities of transition that represent the chance that a patient finds him/her-self in the destination stage at the end of a cycle, given that the same individual was at a particular stage at the start of the time interval: it is relevant to notice how such probabilities only depend on the stage at which the patient is at the beginning of the cycle under consideration. This condition, known as 'Markov assumption', implies that the transition probabilities cannot depend on the time spent by an individual in one stage, or from the particular 'history' experienced by the patient before moving to the considered stage. Such an assumption can be "relaxed" by adopting a series of time-dependent conditional probabilities (Markov processes) [17-18]. For each strategy, the proportion of patients in each hypothesized condition will be calculated cycle by cycle and, based on this distribution, the relative costs met and the benefits gained at each time interval will be obtained. Taking into consideration that a future benefit has less value today and that a future cost at today's prices has also less value, in practice such quantities will have to be "discounted" to obtain to-date values. Markov models for cohorts can be analysed analytically - and by Monte Carlo simulations, sampling a series of values starting from the distribution of hypothesized probabilities.

The assumption of homogeneity of the classical Markov models (that is, the observations are dependent exclusively from the last preceding observed time-point) is, however, not sufficient for modelling many phenomena adequately. Such an assumption is for instance restrictive in the presence of persistent, irregular, with regime changes, phenomena.



On the other hand, relaxing such assumption directly drives to extremely complex models, in which the number of parameters to estimate is excessive. Two valid alternatives [19] are Markov latent models and mixture transition distribution (MTD) models.

### Individual sampling models

Instead of identifying the flow of patients at the aggregated level, single patients whose progression over time can be influenced by their specific characteristics and by the particular paths that they have been following until that moment can be tracked [11]. Usually, with this type of models, the number of individual histories simulated is sufficiently large to provide relying point-estimations with the relative measures of associated variability. Decision tree models simulated at the individual level, through which patients on various tree branches are simulated keeping track of the different personal histories, and the individual markovian models, which constitute the most used technique nowadays, belong to this family. Effectively, the latter models are those seen in the previous paragraph, analysed through the simulation of single individuals at each time interval, that allow to model eventual comorbidities and that allow to introduce particular conditions for the transition from one stage to another. In some cases, these models also allow to take into consideration the time to the following event instead of using pre-fixed cycles.

### System dynamics models

Among models that allow interaction between individuals, dynamic models constitute the cohort approach that permits to model the status of the system over time and, eventually, to reproduce its variations after a particular event has happened [11]. A classic example of the use of this type of models can be found in epidemiological models for infectious diseases. In such cases, the higher the level of infection, the higher the risk of contracting the infection. However, at the same time a decrease in the number of subjects at risk occurs, with a rate of response of the system that necessarily varies also according to present constrictions (economic

and/or instrumental availability). Models that use ordinary differential equations (ODE) systems for describing the rate of change of variables in the system belong to this family of models. The dynamics of the change in both cases are deterministic and fractions of individuals can be found in the stages under consideration. These two assumptions can be overcome by the markovian models introduced before.

### Individual event history models

Defined by Koopman [20] in the attempt to unify the modellistic approach used in the epidemiological and demographic fields, this type of models allow to use parameters and rate/probabilities different for each individual so as to accurately consider the heterogeneous characteristics of the population under study, to introduce in a dynamic manner eventual limitations of resources of the system and to take into account the “unique” history of each patient [21].

### Simulation with discrete events

Among the modelling techniques, these probably represent the most flexible ones, since they allow to study over time the trend of single individuals undergoing different processes and different types of event that can condition their characteristics and outcomes [22, 23]. Furthermore, it is possible to introduce a “tail” structure that permits to simulate an interaction between individuals due to particular constrictions (e.g.: the problem of the waiting lists). The status of the system depends on individuals at risk and on their characteristics, on the list of events that can happen at a particular moment in time or that are in any case programmed to happen: each event is determined separately and its happening induces the re-definition of events that could follow.

From this overview a fundamental question naturally rises: what makes a model a “good” model? First of all it must be taken into account that the model building (and its results) must be an easily readable and interpretable tool for the decision makers. Simple (a model, after all, is just a simplified representation of the reality), not simplistic, adapted to the objectives

for which it was built and (as suggested by the Consensus Statement for the decisional model building in HTA in 2000 [24]) it must be characterized by transparency, clearly defined objectives, internal and external consistency, parsimony, reproducibility, inferential validity and interpretability.

The detailed exploration of the uncertainties associated with the model, that is, the necessary presence of an analysis of sensitivity for the evaluation of the robustness of the results deserves a separate debate. This will be discussed in the following paragraph.

## UNCERTAINTY ANALYSIS

A fundamental distinction in the analysis of decision-making models for economic evaluations is that between *variability* and *uncertainty*. The two terms are not considered synonyms since they refer to different phenomena. Variability refers to the natural variations that can be observed in the value of the variables measured on the study sample. When intended as such, individual variability is an element for which it is not possible to account and it is not associated with the amount of available evidence (even though this is obviously in inverse relation to the sample size [27]).

On the contrary, uncertainty is associated with the imprecise knowledge of the population parameters (e.g.: costs and clinical effects) of a specific intervention on the study population. In particular, it is possible to distinguish between the uncertainty associated with parameters, defined as the precision with which parameters are estimated (extrapolation of the values of reference parameters from the sample to the general population) and the uncertainty due to the choice of methods of analysis [28] or structural uncertainty [29].

The presence of uncertainty does not have an impact only from a computational point of view, but it has also direct consequences on the decision process when a choice is made on health technologies on the basis of the evidence produced by cost-effectiveness studies. In fact, the choice of a technology bears on the possibility to acquire more knowledge on this, with the aim of reducing uncertainty. A preliminary estimation of the amount of information available on the

population parameters and of its impact, therefore, becomes a key factor in a rational and competent decisional process.

The most commonly investigated form of uncertainty is that associated with the model parameters. The methodological guidelines [30] recommend to perform a probabilistic sensitivity analysis (PSA). The aim is to propagate the uncertainty associated with the population parameters to the entire model that is at the basis of the study, in order to determine how and to what extent this could have an impact of the output (decision process) [31]. PSA requires the specification of a probability distribution for each parameter (possibly taking into consideration also the correlation levels that exist among the various parameters); the following step typically involves a simulative approach (e.g.: by means of Markov Chain Monte Carlo methods), with the aim to evaluate the variability in terms of costs and benefits. As a matter of fact, the parameters are considered as random variables by choosing an appropriate probability distribution for all of them, which represents the state of the science for each one. An essential step is, thus, the choice of the probability distribution, which must take into account the nature of the parameter, of the methodology used to estimate the parameter and of the decisional context. Sometimes these aspects are conjugated to subjective evaluations, but the procedure implies that such assumptions be made explicit by the experimenter, and therefore opens the choice to criticism and revisions [32].

Even though it is not always possible to encode a distribution as “standard”, the nature of the parameters involved very often allows for the use of a series of candidates that typically produce reasonable results: for instance, if the reference parameters represent the probability of occurrence of a specific event, it is possible to use the Beta distribution; analogously, in order to model the costs associated to a specific program it is possible to use a Gamma or Log-Normal distribution (which allow to take into account the typical asymmetry of these variables).

To test the impact of the uncertainty of a single parameter, one-way sensitivity analysis is used. The model will prove sensitive to the parameter when the choice on which the technology is cost-effective varies after the one-way sensitivity analysis. It could however

happen that the decision is not sensitive to the uncertainty on single parameters, whereas it does vary after a multivariate sensitivity analysis.

With a focus to cost-effectiveness studies, the primary output of the model is represented by the ICER (Incremental Cost-Effectiveness Ratio). It becomes important to present the results of the sensitivity analysis in terms of the ICER. The results of the analysis can be pictorially presented with the use of scatter plots, confidence ellipses or cost-effectiveness acceptability curves (CEAC). The solution preferred by the methodological guidelines is the CEAC, which represents the probability of each technology included in the study to be cost-effective, according to some cut-off values. The cut-off values for the willingness to pay are plotted on the x-axis, while the y-axis reports the probability to be cost-effective.

An additional summary of the sensitivity analysis is the expected value of perfect information (EVPI), which combines the probability that the actual decision (made with the available information) is correct, with the consequences associated with the possibility that, in light of new evidence becoming available, it turned to be actually reversed. (33).

This mode of reasoning is naturally Bayesian and it is not incidental that the uncertainty associated with the estimation of the input parameters is propagated to the model outputs, with the possibility to analyse posterior distributions that can be used again to inform the decision process, with the possibility to study the probability that an intervention is cost-effective and with the eventuality that extra information be requested to inform the decision and, thus, evaluate its value.

## COST ANALYSIS

The distribution of the costs is typically characterized by a noticeable asymmetry and kurtosis, due to the presence of extreme values (that is, subjects who make use of a greater amount of resources) and due to the heteroscedasticity and multimodality. In purely statistical terms, it follows that the measures of location that best describe a distribution are the median and the interquartile range.

In the majority of cases, however, the public decision maker is interested in the mean value of the costs and of the benefits of an

intervention. These values are in fact important because when multiplied by the number of interested subjects, they permit to measure the total amount of resources needed.

As a consequence, the guidelines for the economic evaluation of health interventions point towards the comparison of the mean values of various alternatives [34]. It becomes therefore necessary to obtain an estimation of the average costs, by means of univariate and multivariate methods of analysis, taking into account the non-normality that characterizes the distribution.

The problem of the estimation of a mean cost is not easy to be solved, even with hundreds of observations [35]. In fact, non-parametric methods, such as those based on the normality of the sample mean or bootstrap, can turn out to be poorly efficient and, in any case, they assume a high sample size.

On the other hand, the use of parametric methods is made more complex by the fact that a large asymmetry and kurtosis of these data imply that the mean cost for the population strongly depends on the tail of the distribution; a consequence of this is that different parametric models, that adapt to the data equally well, can lead to very different results.

Regarding this issue, it must be noticed that when the interest of the decision-maker is address towards the average costs, the presence of outliers is of specific interest in some cases. On the one hand, in fact, it can be useful to identify the covariates that explain the presence of extreme observations; however, on the other hand, it can also be useful to estimate robust mean values with respect to the extreme values of the distribution (e.g.: for insurance reasons or in order to set standard tariffs).

In observational studies, in particular for the estimation of the costs of some pathology, the problem of zero costs also emerges. A portion of the population affected by some pathology or with an interest in a specific program can, in fact, not have access to medical care facilities, resulting in null costs. In general, such a group represents a sub-population with characteristics different to those of the sub-population with positive costs. When the proportion of zero costs cannot be neglected, it is necessary to conduct the analysis so as to evaluate the different effects of the covariates on the probability of generating zero or positive costs.

Both in the experimental and observational



study settings, an additional problem that makes the analysis of costs even more complex is the presence of right-censored observations [36]. In fact, the conduction of a continuative follow-up, that monitors the totality of the medical care provided to patients until the time of interest, often turns out to be difficult or impossible. In particular, an administrative censoring happens when the duration of the clinical trial is not adequately long with respect to the time of accumulation of the costs for the medical care. The event of death can also be interpreted as a censoring of the observable costs, when the estimation of the medical care costs within a specific time frame (not to the occurrence of the event) is of interest. Due to the dependence between the amount of accumulation of the costs and the censoring, the latter cannot be treated as non-informative. Lastly, in the analysis of the costs accumulation process the problem of left-truncation of the costs of the medical care, provided to a patient before the period of observation of the study, can emerge.

The diffusion of national and international multicentre studies makes it necessary to perform the analysis of the average costs, taking into account the effect due to the clustering of the observations [37]. The medical care models, the clinical pathways and the organizational and structural characteristics of the different realities, in fact, determine a variability that would remain unexplained if the data were analysed as independent observations.

In general, at the time of the analysis it is therefore necessary to carefully evaluate the characteristics of the empirical distribution (of the data) and to choose the most appropriate methodological approach so as to avoid erroneous conclusions [38, 39].

## COST-CONSEQUENCE ANALYSIS

The number of economic evaluations of health interventions in Italy has, with no doubt, increased in the recent years. Actually, it must be highlighted how some studies are still characterized by low methodological quality, such as the exclusive use of tariffs and/or expert opinions, but also the use of surrogate end-points and the lack of a probabilistic sensibility analysis, in order to evaluate the robustness of the conclusions and to reduce the

margin of uncertainty [40].

It is therefore not surprising that also the debate on the potential utility and on the real use of such evaluations by the public decision-makers has become more intense. A recent Italian study on a sample of highly selected health managers has highlighted how economic evaluations nowadays do not have a primary role in the decision-making process [41]. The probability that some information be taken into account by the decision-makers is directly proportional to the utility perceived on the specific context for reducing the uncertainty when making choices on the evaluation of future decisions.

In this context it is therefore the responsibility of the evaluator to pay attention to the content and formal aspects: in other words, the most adequate methodology for generating the best usable information – able to reduce the margin of uncertainty – and to communicate the results in an easily understandable manner, must be chosen. Nowadays the information available to the public decision-maker is usually represented by an incremental ratio of cost-effectiveness that shows whether the economic investment required for the new health program is justified by a greater clinical efficacy than that of the alternative reference program.

Although this is a compulsory step when setting the priorities in the allocation of the resources, the institutional decision-maker necessitates of additional information such as the health impact on the territorial community of reference, but most of all s/he needs an estimation of the budget needed for the implementation of the intervention in the local reality. In fact, when two interventions have the same cost-effectiveness incremental ratio they can still have a very different economic and health impact, even if the most common measures of spread of the disease change, such as the prevalence and the incidence. It is the duty of the public decision-maker not just to assign the priorities in the allocation of the resources, based on the cost-effectiveness analysis, but also to evaluate, by means of a budget impact analysis (BIA), the financial consequences that the new technology will have on the future expenditure, in order to verify its sustainability over time. Briefly, BIA defines the possible scenarios according to which the new intervention substitutes the

standard treatment in the target population. More precisely, the possible rates of penetration of a new drug are taken into consideration and, consequently, its positioning in the market is considered. Coherently with the prospective of the financing body (region or country), BIA takes into account only the direct health costs [42] and, at least on a theoretical basis, it could be hoped also for the inclusion of the secondary costs, that is, the costs not directly associated to the pathology of interest, such as the use of the off-label.

BIA must provide the decision-maker with an exhaustive picture of the real, and not theoretical, costs that will have to be sustained and/or that will no more have to be sustained, with reference to the locally defined costs of single health interventions. The importance of BIA is clearly evident when considering the rigid budget limitations that constrain the decision-makers. Such limitations act in the short term and financial-economic evaluations for a single year are, therefore, necessary in order to provide the decision-maker with a valid instrument for the planning of his/her activity. However, it must be remembered that the public decisions should, globally, be for the medium-term (2-5 years), because one of the aims of public administrations, specifically the regions, is to reach a balance over time between income and expenses.

At the same time, the epidemiological indicators, useful for the evaluation of the financial impact, must refer to the local context of the spread of the disease and to the measures of spread of local epidemiological factors that do not recommend the use of the drug; furthermore, it is useful to evaluate the compliance to the treatment, the probability of adverse events that induce the discontinuation of the intervention and the change in the probability of the disease occurrence due to the introduction of the intervention.

A decision-maker that serves the community is obviously not only interested in the economic impact of an intervention, but also in the health impact that this health strategy will have on

the governed population. The evaluation will also have to be based on the local indicators of disease spread and penetration of the new strategy, to which data on the efficacy of the new therapy, derived from clinical trials (such as the increase in the disease free survival or the mortality reduction) will have to be applied.

In summary, in full compliance with the principles of the cost-consequence analysis [43], these data will be provided to the decision-maker in a disaggregated manner, therefore, providing easily interpretable absolute values, useful for making decisions regarding the population of interest.

The present work is a short overview of the most commonly used methods for economic evaluations which goes in the direction to elicit guidelines [44] for the economic evaluation of health interventions, as advocated by the Italian Association of Health Economics - AIES. This is an attempt to link the various field operators with the aim to share knowledge and expertise on the principal methods, to allow the decision-makers to participate to the advancements in this field and strengthen the collaboration with researchers of different disciplines. Given the increasing demand of well-conducted studies to inform decision makers we strongly advocate the start-up of educational initiative at various levels, and the SiPrEMAS has this as one of the main goals of its activities.

**Members of the SiPrEMAS working group (Coordinated by Gian Luca Di Tanna and Gianluca Baio):** *Gianluca Baio, Nicola Bartolomeo, Rino Bellocco, Paola Berchiolla, Marta Blangiardo, Alessandra Buja, Caterina Conigliani, Valentina Dall'Armi, Roberto D'Amico, Chiara De Waure, Rossella Di Bidino, Marta Di Nicola, Gian Luca Di Tanna, Alessio Farcomeni, Giovanni Fattore, Alberto Ferrando, Cipriani Filippo, Debora Formisano, Carmen Furno, Simone Gerzeli, Dario Gregori, Davide Integlia, Marco Marchetti, Antonello Maruotti, Eva Pagano, Francesca Petrucci, Salvatore Scorzaface, Lorenza Scotti, Andrea Tancredi, Paolo Trerotoli, Annarita Vestri.*

## References

- [1] Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999; 18(3): 341-64
- [2] Straus SE, Scott Richardson W, Glasziou P, Brian Haynes R. *Evidence Based Medicine 3rd Ed.* Elsevier Limited, Oxford, United Kingdom. 2005
- [3] Cicchetti A. (a cura di). *Introduzione all'Health Technology Assessment. Principi, organizzazione e metodi per il governo dell'innovazione.* Il Pensiero Scientifico Editore, Roma, 2008
- [4] Battista RN. Expanding the scientific basis of health technology assessment: a research agenda for the next decade. *Int J Technol Assess Health Care* 2006; 22(3): 275-80
- [5] Williams I, McIver S, Moore D, Bryan S. The use of economic evaluations in NHS decision-making: a review and empirical investigation. *Health Technol Assess* 2008; 12(7): iii, ix-x, 1-175
- [6] Erntoft S. Pharmaceutical priority setting and the use of health economic evaluations: a systematic literature review. *Value Health* 2011; 14(4): 587-99
- [7] Egger M, Smith GD, Altman DG. *Systematic Review in Health Care. Meta-Analysis in Context* 2001: BMJ books
- [8] Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2* [updated September 2009]. The Cochrane Collaboration, 2009. Available from [www.cochrane-handbook.org](http://www.cochrane-handbook.org)
- [9] Sutton AJ, Abrams KR, Jones DR, et al. *Methods for meta-analysis in medical research.* John Wiley and Sons, Ltd. Publications, Chichester, UK, 2004
- [10] Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to Meta-analysis.* John Wiley and Sons, Ltd. Publications, Chichester, UK, 2009
- [11] Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ* 2006; 15(12): 1295-310
- [12] Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. *J Health Serv Res Policy* 2004; 9(2): 110-8
- [13] Di Tanna GL. La modellistica per l'analisi decisionale in sanità: pane per i denti degli statistici? Abstract book, IV biennial meeting of the Italian Society of Medical Statistics and Clinical Epidemiology, Palermo, September 2007
- [14] Drummond M, Weatherly H, Ferguson B. Economic evaluation of health interventions. *BMJ* 2008; 337: a1204
- [15] Drummond M, Sculpher MJ, Torrance GW, et al. *Methods for the economic evaluation of health care programmes.* Oxford: OUP; 2005
- [16] Gray AM, Briggs A, Clarke P. New BMJ policy on economic evaluations. Economic evaluations should be judged on scientific merit. *BMJ* 2003; 326: 445
- [17] Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Med Decis Making* 1993; 13(4): 322-38
- [18] Hunink M, Glasziou P, et al. *Decision making in health and medicine. Integrating evidence and values.* Cambridge University Press, Cambridge 2004
- [19] Bartolucci F, Farcomeni A. A multivariate extension of the dynamic logit model for longitudinal data based on a latent Markov heterogeneity structure. *Journal of the American Statistical Association* 2009; 104: 816-31
- [20] Koopman JS, Jacquez G, Chick SE. Integrating discrete and continuous population modeling strategies in population health and aging: strengthening the dialog between epidemiology and demography. In *Annals of the New York Academy of Sciences*, Weinstein M, Hermalin A, Stoto MA (eds), vol.954. New York Academy of Science: New York 2001: 268-294
- [21] Weinstein MC. Recent developments in decision analytic modelling for economic evaluation. *Pharmacoeconomics* 2006; 24(11): 1043-53
- [22] Karnon J, Brown J. Selecting a decision model for economic evaluation: a case study and review. *Health Care Manag Sci* 1998; 1(2): 133-140
- [23] Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation *Health Econ* 2003; 12: 837-48
- [24] Decision analytic modelling in the economic evaluation of health technologies. A consensus statement. Consensus Conference on Guidelines on Economic Modelling in Health Technology Assessment. *Pharmacoeconomics* 2000; 17(5): 443-4
- [25] McCabe C. Guidance on good practice in cost effectiveness modeling: is more needed? *Med Decis Making* 2007; 27: 350-1
- [26] Karnon J, Brennan A, Akehurst RL. A critique and impact analysis of decision modeling assumptions. *Med Decis Making* 2007; 27:491-9
- [27] Briggs A. Handling uncertainty in economic evaluation. *British Medical Journal* 1999; 319: 120
- [28] Briggs A, Goeree R, Blackhouse G, O'Brien BJ. Probabilistic analysis of cost-effectiveness models: choosing between treatment strategies for gastroesophageal reflux disease. *Medical Decision Making* 2002; 22: 290-308
- [29] National Institute for Clinical Excellence -NICE. Briefing paper for methods review workshop on exploring uncertainty. Working paper. [www.nice.co.uk](http://www.nice.co.uk)

- [30] National Institute for Clinical Excellence - NICE. Guide to the methods of health technology appraisal. London: NICE 2008
- [31] Baio G. Metodi statistici per la valutazione economica delle tecnologie sanitarie. Aracne Editore (Roma), 2010
- [32] Briggs AH. Statistical approaches to handling uncertainty in health economic evaluation. *European Journal of Gastroenterology & Hepatology* 2004; 16(8): 551-61
- [33] Ginnelly L et al. Using Value of Information Analysis to Inform Publicly Funded Research Priorities. *Applied Health Economics and Health Policy*, Wolters Kluwer Health, 2005 Adis, vol. 4(1), pages 37-46
- [34] Luce BR, O'Hagan A. A primer on Bayesian Statistics in Health Economics and Outcomes Research, 2003
- [35] Mihaylova B, Briggs A, O'Hagan A, Thompson SG. Review of statistical methods for analysing healthcare resources and costs, applicable to clinical trial data. Manuscript, 2008
- [36] O'Hagan A, Stevens JW. On estimators of medical costs with censored data. *J Health Econ* 2004; 23(3): 615-25
- [37] Thompson SG, Nixon RM, Grieve R. Addressing the issues that arise in analysing multicentre cost data, with application to a multinational study. *J Health Econ* 2006; 25(6): 1015-28
- [38] Mihaylova B, Briggs A, O'Hagan A, Thompson SG. Review of statistical methods for analysing healthcare resources and costs. *Health Econ* 2011; 20(8): 897-916
- [39] Gregori D, Petrinco M, Bo S, et al. Regression models for analyzing costs and their determinants in health care: an introductory review. *Int J Qual Health Care* 2011; 23(3): 331-41
- [40] Cornago D, Li Bassi L, De Compadri P, Garattini L. Pharmacoeconomic studies in Italy: a critical review of the literature *The European Journal of Health Economics* 2007; 8(2): 89-95
- [41] Fattore G, Torbica A. La valutazione economica in sanità: i decision maker utilizzano gli studi? *Economia e politica del farmaco* 2007; 9: 137-52
- [42] Mauskopf JA, et al. Principles of good practice for budget impact analysis: report of the ISPOR Task Force on good research practices - budget impact analysis. *Value Health*. 2007; 10(5): 336-47
- [43] Mauskopf JA, Paul JE, Grant DM, Stergachis A. The role of cost-consequence analysis in healthcare decision-making. *Pharmacoeconomics* 1998; 13(3): 277-88
- [44] Associazione Italiana di Economia Sanitaria. Proposta di linee guida per la valutazione economica degli interventi sanitari. *Politiche Sanitarie* 10, 2, Aprile-Giugno 2009

