
COLLÈGE DES ÉCONOMISTES
DE LA SANTÉ

# French Guidelines for the Economic Evaluation of Health Care Technologies

September 2004

Special thanks to our Firms Members for their contribution to the translation of this document:

Alcon, Aventis, Bristol-Myers Squibb, GlaxoSmithKline, LEEM, Sanofi-Synthélabo, Servier, Wyeth Lederle

COLLÈGE DES ÉCONOMISTES
DE LA SANTÉ

# FRENCH GUIDELINES FOR THE ECONOMIC EVALUATION OF HEALTH CARE TECHNOLOGIES

## Methodological recommendations

**Work carried out by the Members of the Collège des Économistes de la Santé (the French Health Economists Association)** under the coordination of Émile Lévy and Gérard de Pouvourville

## Participants:

AUQUIER Pascal (University of Marseille)
AURAY Jean-Paul (CNRS – University Lyon 1)
BERDEAUX Gilles (ALCON France)
BÉRESNIAK Ariel (ARES-SERONO Suisse)
BRUN-STRANG Catherine (Novartis France)
CARRÈRE Marie-Odile (GRESAC-University Lyon 1)
CHARVET-PROTAT Suzanne (AFM)
CORCAUD Sylvaine (Servier)
COUDEVILLE Laurent (CRESGE – Lille)
d'ALCHÉ-GAUTIER Marie-José (CHU Caen)
de POUVOURVILLE Gérard (CREGAS-INSERM U537)
DERVAUX Benoît (CRESGE-LABORES, Lille)
DURU Gérard (CNRS – University Lyon 1)
EECKHOUDT Louis (University of Mons - Belgique)
FAGNANI Francis (CEMKA-EVAL)
GABRIEL Sylvie (Sanofi-Synthelabo Groupe)
GUIGNARD Éric (GlaxoSmithKline France)
LAUNOIS Robert (University Paris Nord)
LEBRUN Thérèse (CRESGE – Lille)
LELEU Hervé (CRESGE-LABORES, Lille)
LENNE Xavier (CRESGE – Lille)
LE PEN Claude (University Paris Dauphine)
LÉVY Émile (University Paris Dauphine)
LILLIU Hervé (CLP Santé)
MARISSAL Jean-Pierre (CRESGE – Lille)
MIDY Fabienne (CREDES)
SAILLY Jean-Claude (CRESGE-LABORES, Lille)
SELKE Bernard (CRESGE – Lille)

## Referees:

BOULENGER Stéphanie (CES)
ULMANN Philippe (CES)

**CES – 7 rue de Citeaux – 75012 Paris – France – http://www.ces-asso.org**
**Tel.: 33 (0)1 43 45 75 65 – Fax : 33 (0)1 43 45 75 67 – E-mail : ces@ces-asso.org**

# SUMMARY OF RECOMMENDATIONS

## I) Study viewpoint and perspective

### Recommendation 1

Economic evaluation studies of health care programmes must adopt the widest possible perspective in order to include all the relevant outcomes of each programme studied. The choice of the range of observation must be justified. Given that economic evaluation studies are in many cases carried out from the perspective of a single agent (the Assurance Maladie[1], the hospital, etc.), it is recommended that, wherever possible and where relevant, the impacts of the therapeutic management strategies being studied should also be explicitly analysed from the perspective of the other agents concerned (general practice in the case of hospital studies, other agents in the case of budgetary impact studies on the Assurance Maladie, etc.). Finally, the public health perspective supposes that the consequences of health care policies as regards epidemiology, the long term and inequality of access to health care or the redistribution effects should be clearly explained whenever these aspects are deemed relevant.

## II) Timing of studies over the life of a product

### Recommendation 2

Economic evaluation studies may be carried out at any stage in the lifecycle of a health care strategy. The study questions differ depending on the stage being studied, and also on the methods employed, the degree of uncertainty as to the aspects under consideration and the data available. It is therefore necessary to specify the stage of development of the health care strategy under study. Furthermore, since the data required for an economic study differ from that required for a clinical trial, it is important that an economist be involved at the earliest possible stage in the design of clinical trials to determine whether or not it would be relevant to collect the cost and effectiveness data required for the economic analysis at the same time as the clinical data.

## III) Theoritical framework

### Recommendation 3

Economic evaluation studies must be seen as part of a decision analysis that is able to call on multiple criteria and methodologies adopted from different disciplines. Different theoretical frameworks may therefore be used, on the authors' responsibility, borrowing from methods of optimisation, multicriteria decision analysis and the economics of well-being, i.e. from the fields of management science and economic theory.
In certain cases (cost-consequences analysis), the studies will be limited to a presentation of the various criteria of choice, leaving users to apply their own weighting or to aggregate criteria as they see fit.

## IV) Environment studies

### Recommendation 4

It is recommended that an economic evaluation study should include a preliminary descriptive section which provides an overview of the main information available (descriptive epidemiology and types of management) for the disease concerned in terms of both public health and health expenditures. This section may be as detailed as necessary.

## V) Target population

### Recommendation 5

The target population of the evaluation must be clearly described. In the case of an economic evaluation study associated with a clinical trial, the target population is the trial population. If several studies are used to estimate effectiveness, the choice of target population and the estimates of results must be clearly set out. It may be relevant to carry out an economic evaluation study by sub-group of the target population, in so far as sub-group statistical analysis justifies it.

[1] All French abbreviations, acronyms and definitions are translated and explained in the "abbreviations, acronyms and definitions" section, p. 52.

## VI) Choice of comparator and comparison methods

### Recommendation 6

Where several strategies exist in practice (including non-drug therapies and non-treatment), these should be listed in full, and the strategies studied should be taken from this list (and justified). The therapeutic strategies to be used as comparators will be chosen among those most frequently used (including non-treatment) or newer strategies which may legitimately be deemed likely to become reference strategies in the very near future.

### Recommendation 7

When performing an economic evaluation study, it is the direct comparison between two strategies that presents the greatest scientific relevance. It may, however, be necessary to perform an indirect comparison based on published data, which is acceptable on a provisional basis in the absence of any original studies allowing for direct comparison.

## VII) Types of study

### Recommendation 8

Each evaluation has its own particular scope and limitations. The type of study selected should be clearly stated and justified with respect to the issue addressed, and must be described at the start of the study. The author should also provide his personal definition of the type of study used.

## VIII) Time frame and transferability to other countries

### Recommendation 9

The time frame of an economic evaluation study must be justified in terms of the natural history of the disease and the availability of data.
It is recommended that the time frame chosen should be long enough that all outcomes, both positive and negative, of the treatments used and evaluated be included in the study. This recommendation may be difficult to apply in certain cases, in particular as regards very long-term effects, mainly due to major uncertainties or the impact of new treatments in the intervening period. In such cases, the reasons for the choice of the time frame should be expressly stated, and the consequences of not including certain events in the study should be discussed.

### Recommendation 10

Epidemiological and economic data cannot be transferred from one country to another without prior verification. When the economic evaluation study relies on international data or national data for a country other than France, the authors should demonstrate that these data (clinical, epidemiological or economic) may be transferred with sufficient plausibility to the French context. In the absence of any supporting evidence, the study should mention that the data so transferred are to be used with caution.

## IX) Definition and measurement of costs

### Recommendation 11

The databases of reference used to calculate the cost of a hospital short-stay are those generated by the Programme de Médicalisation des Systèmes d'Information (PMSI) [the medical information system which is the French DRG program], which offer the most precise method of identification of the clinical characteristics and the treatments of the patients. An *ad hoc* observational survey is therefore possible, but the results obtained for a sample of hospitals should be compared against national data to ensure that the results are truly representative.

**Recommendation 12**

The development of the PMSI for *Services de Suite et de Réadaptation* (SSR) [medical follow-up and rehabilitation services] provides useful information in terms of epidemiological knowledge of medium-stay admissions but doesn't allow the valuation of its cost, because of the absence of sufficient data on unit costs. The provisional scale of costs is not considered reliable. It is possible to apply the daily rates established by the CNAMTS for the rest/convalescence, physiotherapy and rehabilitation disciplines to the number of days per disease derived from the PMSI-SSR database, or to carry out *ad hoc* surveys of specialised institutions.

**Recommendation 13**

For the hospital sector, any evaluation on the transportation costs must take into account the distinction between primary (home/hospital) and secondary transportation (between institutions) and how the institutions is funded (lump sum allowance/daily tariff). Where necessary, a survey of the institutions concerned should be carried out in addition to consulting the *Assurance Maladie* information systems.
A study of transportation costs will also need to take into account the nature of the hospital activity (accident & emergency, for example) or the diseases under consideration (chronic diseases in particular).

**Recommendation 14**

Conventional tariffs and the latest data published by the Assurance Maladie reimbursement offices should be used as the basis for measuring the costs of general practice and, if a societal perspective is adopted, the portion of costs not reimbursed and any supplementary fees charged by doctors over the tariff should be included. A new set of specifications is available, the *Classification Commune des Actes Médicaux* (CCAM), which replaces the former *Catalogue des Actes Médicaux* (CdAM) and the *Nomenclature Générale des Actes Professionnels* (NGAP).

**Recommendation 15**

The disease and its treatment may induce direct non-medical costs incurred by patients and their families. These costs may relate either to expenditures for the acquisition of goods, or may represent a cost in time. These direct costs should be measured by *ad hoc* surveys carried out on a rigorous basis, particularly in view of the potential income-effect. Non-working time may be measured in one of three ways: a time total is calculated but is not included in the cost-outcome ratio; the time total is valued at the hourly rate applicable to an equivalent resource on the market; alternatively, as part of a cost-benefit analysis, this time is measured by the patient's willingness to pay. On a subsidiary basis, it will be possible to evaluate the number of working hours the caregiver is required to give up.

**Recommendation 16**

The impact of the disease and its treatment on the functioning of the economy must be taken into account in the economic evaluation of a health care programme, when there is good reason to believe that this impact is significant, in terms either of the number of people concerned or of the lasting of the consequences of the disease in the lives of patients. The number of lost working days is measured by *ad hoc* surveys. The valuation with the human capital approach is simple but unrealistic. The basis for the friction-cost method is more realistic, but requires much more empirical work before it can be applied in France. Neither approach allows evaluating the time lost by persons not in employment. The theory of well-being recommends including the value of lost time in the evaluation of the results of a treatment programme, rather than in the costs. In any event, the inclusion of indirect costs must be specifically analysed and its impact on the results of the evaluation studied in the light of the method adopted.

**Recommendation 17**

The value of human life may be measured in two ways: using the human capital approach, by calculating per capita income discounted over the life expectancy of the French population, or using the theory of well-being approach, using the valuation placed by individuals on the life years gained.

**Recommendation 18**

The evaluation of intangible costs (or benefits) is not justified, given that the human and psychological impact of the treatment may be studied by means of quality of life indicators, measurements of utility or by using contingent evaluation methods (see section 10.4).

## X) Definition and measurement of outcomes

### Recommendation 19

When effectiveness is expressed by a single indicator, this solution should be justified and the aspects of effectiveness not examined should be clearly stated. When several indicators are combined, the methods of combination should be described.
If the indicator expresses a number of events avoided, the period over which this outcome is obtained should be stated.
Effectiveness is to be preferred to efficacy. If, however, efficacy is the sole indicator available, every effort should be made to convert this into effectiveness by means of appropriate adjustments.

### Recommendation 20

Ideally, a health care intervention should be evaluated on the basis of final end-points (expressed either directly or through surrogate end-points). If only intermediate end-points are available, the relationship between these end-points and the final end-points should be estimated, primarily on the basis of published literature.

### Recommendation 21

Evaluating quality of life offers real descriptive and informative potential; carrying out such studies is therefore highly relevant.
In terms of decision-making in public health, evaluating quality of life therefore is a complement to the evaluation of effectiveness.

### Recommendation 22

It is recommended that only instruments that have been subject to a validation process (linguistic, cultural and psychometric) be used to measure quality of life.
Since a quality of life measurement instrument is based on specific conditions of use, on a questionnaire and a scoring function, none of these components can be modified without forfeiting the benefits of the validation process.

### Recommendation 23

In view of the wide variety of tools available for measuring quality of life, it is important to ensure that the tool employed is appropriate to the aims of the study. Since the strengths and weaknesses of generic and specific tools are mutually complementary, it is recommended that the two tools should be combined wherever possible. Otherwise:
. the use of a specific instrument is recommended when the aim of the study is to compare several strategies for a given disease;
. the use of a generic instrument is recommended when the aim of the study is to compare several diseases, methods of health care management, etc., or when no specific instrument is available for the population considered.

### Recommendation 24

The economic approach to quality of life is based on the utility function which consists in assigning a numerical value to each of the states for which a preference is expressed. If A is preferred to B, the utility of A is a number higher than or equal to the utility of B.

The methods most commonly used to identify preferences are the standard gamble, time trade-off (TTO), and visual analogue scale methods. These methods allow for the expression of individual preferences within a strict framework. The use of these methods must therefore take into account their theoretical and empirical characteristics.

A number of theoretical problems arise when it comes to moving from individual preferences to a collective preference. Since no solution has yet been found to these problems, using the results of any collective preference as the sole criterion on which to base public health decisions is not recommended.

**Recommendation 25**

The QALYs approach consists in aggregating into a single dimension the two dimensions that describe the outcomes of a health care programme in terms of life years gained and quality of life.
The proposed aggregation rule poses many problems in terms of both methodology and philosophy. The limited robustness of this approach allows the manipulation of the conclusions of a study.

In view of this, the readers or users of a study setting out outcomes in terms of QALYs should pay close attention to the following rules:

a) It is essential to verify that each study presents the reference standards used to measure quality of life, the measurement method used, and that the specification of the multiplying aggregation function used for the reference standards has been validated.

b) It is equally important to enquire into the origins of the quality of life measurements, and particularly whether these measurements are "psychometric" or derived from techniques of preference identification such as TTO or standard gamble methods.

c) When the QALY versions used are based on utility or expected utility theory, the reader should bear in mind issues such as whether the behavioural assumptions have been verified, the choice of methodology and epistemology, and should seek to identify any bias these choices might induce.

d) It is important in all circumstances to be cautious of any attempts at inter-studies and inter-pathologies comparisons (e.g. league tables). It is important to ensure, for example, that the same systems of reference have been used in all the different cases. Nor is it recommended that the valuation of health states from one study be used for another, without careful validation to ensure that the health states are indeed equivalent.

e) Attention should also be paid to the risk of rogue distribution effects when comparing populations with very different socio-economic characteristics or where preferences for the treatments analysed may be very heterogeneous.

f) In the current state of research, it is not recommended that public health decisions should be based on study results expressed in terms of QALYs unless conditions a) to e) above are verified and/or validated, given the possibility of arriving at divergent results from the same observed data.

**Recommendation 26**

The use of contingent evaluation in the health care field is still at an the experimental stage, which means that the limitations of the technique must be clearly set out and the potential for bias examined.

## XI) Discounting costs and benefits

**Recommendation 27**

In the case of a collective choice, discounting should be applied to:
- costs
- health outcomes, if the method used for measuring benefits does not already take into account the preference for the present.

It is recommended that costs and outcomes should be discounted at the same rate, except if studies suggest that the preference for the present in health outcomes and in costs is different.

For the purposes of international comparisons the rates used should be 0%, 3% and 5%. A sensitivity analysis should be carried out systematically on discount rates in order to test the robustness of the conclusions (including a 0% rate, i.e. the evaluation without discounting).

## XII) Modelling

### Recommendation 28

Modelling covers a series of clearly-defined techniques whose application to the economic evaluation of therapeutic strategies has been validated. It can be used to evaluate both costs and outcomes. A model is a decision aid which can be used to aggregate data from different sources. It is used to systematise decision-making elements, particularly by allowing for long-term estimates, indirect comparisons, evaluation of the budgetary impact of a health care strategy on the management of a disease and the extrapolation of the results from a survey to the general population. The choice of the type of model to be used must be justified and adapted to the issue being studied.

### Recommendation 29

The interpretation of modelling results must remain linked to the assumptions underlying the model's construction (given population, etc.). All the data and assumptions used in the model must therefore be clearly stated, documented and validated. The reliability of the models is subtended by a dual process of internal and external validation. The robustness of results must be tested by a sensitivity analysis. The choice of parameters to be varied must be presented/justified, as must the intervals inside which values are varied.

## XIII) Verifying the robustness of study conclusions

### Recommendation 30

In the case of studies using data from different sources, a sensitivity analysis must be carried out on the variables and assumptions used. These must be clearly identified and justified, as well as the intervals inside which they are varied. In the case of a complex model, an analysis in terms of best case and worst case scenarios may be carried out.

### Recommendation 31

When carrying out stochastic analyses where data on the cost and effectiveness of treatments for each patient are available, confidence intervals on the cost-effectiveness result must be used in order to take sampling errors into account. Preference will be given to the parametric method based on Fieller's theorem and the non-parametric bootstrap approach. The representative nature of the data used will be assessed prior to the analysis. An alternative and more informative approach to take uncertainty into account is to present the acceptability curve of the strategies which represents the probability that the strategy in question is cost-effective for different ceiling values of the cost-effectiveness ratio, ranging from 0 to infinity.

## XIV) Presentation of the results

### Recommendation 32

As a general rule, the results of an economic evaluation study should be presented in such a way as to allow peer examination and review. The report must contain all the elements that will enable an informed reader to understand the methodology adopted, check the sources and verify the relevance and accuracy of calculations. All data, references and calculation procedures used in the study must be available to any interested party on request, if not explicitly included in the study. The only limit to this principle is respect of the confidentiality of data on identifiable individuals. Although it is perfectly legitimate for authors to comment freely on their results in the conclusion or the discussion section, they should be careful to avoid any interpretation of these results that might seem exaggerated in the light of reality.

### Recommendation 33

The publication of the results of economic evaluation study in the field of health care may have important economic or regulatory repercussions. It may influence a public decision, founded on the quality of the scientific study. It is therefore important that the process of publication and dissemination should meet the criteria of integrity applied by the scientific community. It is important, for example, that all the authors have played an effective part in the scientific part of the study. Each signatory is personally responsible for the veracity of the study and for ensuring that the methodological approach followed was indeed as described in the study. Authors' institutional affiliations must be stated and the source(s) of funding, whether public or private, must be disclosed.

### Recommendation 34

Cost-effectiveness (or cost-utility ratios) must be expressed in incremental form so that they may be used as a criterion to assist in decision-making. Variations in the use of resources should be included in the numerator and those influencing health states in the denominator, taking care to avoid double counting.
The components of these ratios (costs and results) should also be presented with their statistical distribution (mean, median, confidence interval, etc.).

## XV) Interpretation of the results

### Recommendation 35

On completion of a study, the distributive aspects of the result should be considered by differentiating it by as many factors as may seem pertinent (socio-demographic factors in particular) and which can raise equity problem by favouring certain groups or agents at the expense of others.

### Recommendation 36

1. If studies lead to recommendations on public policy, it is important to underline that such studies serve merely as decision aids and under no circumstances as imperatives for action.
2. It is important to ensure that the type of recommendation is appropriate to the type of study. Avoid making cost-benefit type judgements (maximising collective well-being) in the context of a cost-effectiveness study, for example.
3. The results of cost-effectiveness studies, particularly those which focus on mutually exclusive programmes (the most common), must emphasise that the choice between the non-dominated options relies to a large extent on a value judgement which the decision-maker may legitimately apply.
4. The value of any reference to external comparators ("*the programme analysed is as cost-effective as another commonly implemented*") is purely indicative rather than prescriptive.

### Recommendation 37

The possibility of developing recognised reference values for cost-effectiveness type studies based on criteria such as cost/life year gained, adjusted or not for considerations of handicap or quality of life, will require major efforts to standardise methods and parameters, particularly for the calculation of costs. In the absence of such rules, great caution should be observed in any comparison of results from different studies. It is recommended that such comparisons should always be limited to the same therapeutic field and that the comparability of results should be justified as fully as possible.

## XVI) Budgetary impact

### Recommendation 38

In addition to economic evaluation studies, which operate at the micro-economic level, it may be useful to extrapolate the results obtained in order to estimate the short and medium-term budgetary impacts for the main agents concerned in case of a generalisation of the treatment. These estimates, which may be as complex or as simple as circumstances require, will have to provide a detailed presentation of the assumptions on which they are based.

# TABLE OF CONTENTS

# INTRODUCTION

Given the growing number of guidelines on good practice in economic evaluation studies in the field of health and the existence in France of recommendations first issued in 1997[2], some explanation is necessary of how the following guidelines, issued in 2003, compare to those of 1997 and to the most recent versions of international guidelines.

In France, the context, the elaboration process, the level of scientific rigour demanded and hence the content has evolved since 1997.

The context, and in particular the institutional context, made it possible to bring together a tripartite commission of 18 members representing the industry, the Public authorities and health economists. Over the course of its year-long debates, the commission was able to review the main methodological problems and formulate responses in a spirit of consensus.

The level of scientific rigour required under these earlier guidelines was the level applied in most countries at the time. Since then, however, that level has risen appreciably in the field of outcome estimation (better-documented figures, statistical methods offering greater refinement in sensitivity analyses and the handling of uncertainty or discounting, more diversified models for the projection of costs and outcomes).
Lastly, content has been enhanced by types of study and issues rarely or never encountered in the mid-1990s; cost-consequences studies have achieved a degree of acceptance (albeit timid), attempts have been made to evaluate the contribution made by volunteers; contingent evaluation has made (some) progress; the problem of the transferability of international studies has, if not solved, at least been raised.

In addition, the number of French health economists has increased substantially over the last 5 years and greater cohesion between them has been achieved through associations like the Collège des Économistes de la Santé (CES)[3].

It is the CES that is responsible for initiating this revision of our recommendations, organising working sub-groups with broad remits on a range of subjects to produce the Guidelines. These are only now being submitted to the industry and the Public authorities for discussion, since we were unable to establish ground rules with these partners along the way that would, in parallel to the Guidelines, have represented a consensus on the main orientations for the economic evaluation of health care and therapeutic strategies.

How do these recommendations compare to those produced by other countries? The CES has made every effort to position its work at the cutting edge of current thinking in health economics and there are many points of convergence with published scientific work and recommendations. We must acknowledge a great debt to our Canadian colleagues, since we used their work as an inspiration for our own: accompanying each recommendation, set out as concisely and clearly as possible, with an explanatory text supported by key references to enable the reader to explore further. Following their example once again, we have sought to adopt the greatest possible neutrality wherever active scientific controversy existed on a certain number of points. The reader will nonetheless find that we have taken a stance on certain important points, where we felt that the apparent consensus of the scientific community was not fully justified. Lastly, as far as the measurement of costs is concerned, we clearly had to take into account the particular ways in which health care is funded in our country, and the specific characteristics of existing databases, in order to put forward appropriate recommendations.

On which points do we find the greatest convergence with work done outside France? On all points relating to the general design of studies, the use of medical data and their extrapolation, decision-making rules deriving from economic evaluation. To cite one example: recommendations 30 and 31 set out in great detail the precautions to be taken for sensitivity analysis of results in the assumptions used in models and to the uncertainty associated with clinical and economic data. There has been considerable progress in methods addressing both these points since the last version of the Guidelines, and the credibility of the studies we carry out in France can only gain should we adopt them. It is of the greatest importance to provide the decision-maker with the fullest possible information on the robustness of the results of economic evaluation studies.

In contrast, the reader will find us more circumspect, or alternatively bolder, on the use of preference elicitation methods or contingent valuation methods to measure the outcomes of health care programmes. We wanted to stress out that these are based on standardised theoretical frameworks which are open to dispute, and that the empirical methods on which they rely, while ingenious, do not as yet justify their use as decisive factors in decision making. Two points seem important to us.

[2] Published in the CES Newsletter of April 1997 and reprinted in the Journal d'Économie Médicale, 1998, T.16, no. 4-5, pp. 353-372.
[3] The Collège des Économistes de la Santé is a so-called learned society, founded in 1989, and gathering about 200 French Health Economists.

We are not entirely convinced (but the debate was quite lively) that the present methods for the measurement of outcomes of health care programmes, based on either utility measures or on the measurement of willingness to pay, are sufficiently robust to serve as a rule for public decision-making. We might have been content to issue the usual caveat and modestly remind our readers that studies are only one factor in decision-making, but we felt it more useful to explain why. In particular, we felt that there was a risk of deviation in the measurement of QALYs, as more and more studies make uncontrolled use of utility values published for other health care programmes and calculated using different methods. Such hybrid methods are no longer able to guarantee the comparability of outcomes from one health care programme to another, thereby undermining the principle of league tables in which programmes are ranked by comparison with one another in ascending or descending order of cost per QALY.

The second point relates to the theory of well-being, which postulates that the situation of all is improved when the aggregation of all individual benefits increases, irrespective of the distribution of these benefits. It matters little if some gain while others gain less or lose, as long as the balance is positive, because the workings of the economy will enable the necessary adjustments to be made. Just because everyone is in better health on average does not mean that each individual is in better health, however, and the democratic debate consists precisely in ensuring that compensations are truly being organised.

Criticism has its uses, art its difficulties. It is time for us to play our part in taking forward methods of economic evaluation of health care programmes. We hope with these guidelines will contribute to the improvement of studies carried out within the French health care system. We must persevere, however, in the task of reflection.

## I) STUDY VIEWPOINT AND PERSPECTIVE

Different costs and outcomes measurement methods exist, depending on the perspective chosen; it is important, therefore, that the viewpoint(s) adopted for a given study be clearly stated: budgetary impact or forecast for the Assurance Maladie [French national health insurance programme], the hospital, etc., or impact on total consumption of medical goods and health care irrespective of the source of funds, and on public health (societal perspective).

However, since the overall aim of economic evaluation studies is to provide decision aids in the field of public health policy, it would be preferable if a "societal" perspective was adopted in any event. This concept of "societal" perspective as yet has no precise definition in France. Depending on circumstances, it may relate to a concept of collective interest arising out of the economic theory of well-being, to a general perspective of public health or to the inclusion of concepts of equity between groups and between generations. It also relates to a wider concept of budgetary impact for society as opposed to a simple measurement of budgetary impact for one agent alone (see Appendix I).

**Recommendation 1**

Economic evaluation studies of health care programmes must adopt the widest possible perspective in order to include all the relevant outcomes of each programme studied. The choice of the range of observation must be justified.
Given that economic evaluation studies are in many cases carried out from the perspective of a single agent (the Assurance Maladie, the hospital, etc.), it is recommended that, wherever possible and where relevant, the impacts of the therapeutic management strategies being studied should also be explicitly analysed from the perspective of the other agents concerned (general practice in the case of hospital studies, other agents in the case of budgetary impact studies on the Assurance Maladie, etc.).
Finally, the public health perspective supposes that the consequences of health care policies as regards epidemiology, the long term and inequality of access to health care or the redistribution effects should be clearly explained whenever these aspects are deemed relevant.

## II) TIMING OF STUDIES OVER THE LIFE OF A PRODUCT

The economic evaluation of a health care strategy may be carried out at any stage in its lifecycle. Each stage may be associated with one or more decisions to be taken, and the method must be chosen according to the decision on which the study is intended to shed light.

During the research and development phases, the economic study may provide the promoter of a new treatment with information on its potential value by identifying target populations, by identifying competing products or health care strategies and by an initial estimation of the treatment's relative value compared to the competition.

At present, many studies are carried out before the marketing of a health care strategy for negotiations with the main health care funding bodies. These *ex ante* studies are generally based on clinical efficacy data derived from experimental and quasi-experimental studies which do not reflect the real future conditions of use of the health care strategy. With increasing frequency, these clinical trials include measurements of the effect of health care strategies on patient quality of life, but they more rarely include any measurement of their utility. The relevance of costs measurement alongside clinical trials is open to debate. It will be important to assess what proportion of the costs is attributable to the experimental protocol itself. In contrast, there are circumstances in which the involvement of an economist at an early stage in the design of the clinical trial can be useful: where the choice of the final relevant outcome is not the same for the clinician and for the economist (short or medium-term mortality, for example), or when it is important to know what has become of patients who have dropped out or for whom treatment has failed in both arms of the trial.

There is growing pressure from funding bodies to obtain studies allowing the evaluation of outcomes of health care strategies under real conditions, once the product is marketed. These studies may be used as grounds for renegotiation of indications for use and price. They can also be used to verify whether the assumptions made in preliminary studies are confirmed or not. Such studies have an advantage over randomised clinical trials in that they are closer to real practices. They also call for methods other than those used in clinical trials.

**Recommendation 2**

Economic evaluation studies may be carried out at any stage in the lifecycle of a health care strategy. The study questions differ depending on the stage being studied, and also on the methods employed, the degree of uncertainty as to the aspects under consideration and the data available. It is therefore necessary to specify the stage of development of the health care strategy under study. Furthermore, since the data required for an economic study differ from that required for a clinical trial, it is important that an economist be involved at the earliest possible stage in the design of clinical trials to determine whether or not it would be relevant to collect the cost and effectiveness data required for the economic analysis at the same time as the clinical data.

## III) THEORETICAL FRAMEWORK

Economic evaluation studies are defined by their joint consideration of medical criteria, based on clinical epidemiology, and cost criteria. However, this association may be made either within the formal framework of the economic theory of well-being and utility, or within the multidisciplinary methodological framework sometimes referred to as decision analysis. In the first instance, the expression of social utility must conform to very strict conditions, often difficult to meet; in the second instance, the primary objective is an operational decision model where we must accept that these conditions are abandoned.

Since the ultimate purpose of economic evaluation studies is to aid the decision-maker in making choices, the second approach, that of the decision-making process, seems the most relevant.

**Recommendation 3**

Economic evaluation studies must be seen as part of a decision analysis that is able to call on multiple criteria and methodologies adopted from different disciplines. Different theoretical frameworks may therefore be used, on the authors' responsibility, borrowing from methods of optimisation, multicriteria decision analysis and the economics of well-being, i.e. from the fields of management science and economic theory.
In certain cases (cost-consequences analysis), the studies will be limited to a presentation of the various criteria of choice, leaving users to apply their own weighting or to aggregate criteria as they see fit.

## IV) ENVIRONMENT STUDIES

Economic evaluation studies of therapeutic strategies may be supplemented by a range of complementary approaches such as the cost of disease and budget forecast. This latter approach is dealt with in section 16.

Cost of disease type studies are carried out ahead of economic evaluation studies. They provide general descriptive elements on the disease concerned by the treatment being studied. They generally summarise the information available on the descriptive epidemiology (incidence, prevalence, natural history, consequences in terms of morbidity, handicap and mortality), the current methods of management within the health care system and the associated costs (direct and indirect). Repercussion analysis may include quality of life studies (see paragraph 10.2).

Cost measurement in these studies raises the same type of difficulties as appear in economic evaluation analysis as such (see Chapter 9: Definition and measurement of costs).

**Recommendation 4**

It is recommended that an economic evaluation study should include a preliminary descriptive section which provides an overview of the main information available (descriptive epidemiology and types of management) for the disease concerned in terms of both public health and health expenditures. This section may be as detailed as necessary.

## V) TARGET POPULATION

A health economic evaluation study must specify the population that will potentially benefit from its effects. This population may be described by demographic characteristics (age, sex), by the nature of the disease concerned, by the stage of development or severity of the disease, by the existence of associated diseases, by exposure to certain risk factors, etc. In the case of an economic evaluation study associated with a clinical trial, the target population is generally the population that meets the criteria for inclusion in the trial. This target population will be different from the population that will benefit from the treatment in current practice.

When the effectiveness data are drawn from multiple sources on different populations, an accurate description must be provided of the method used to arrive at a definition of the target population and of the estimates of effectiveness on this population.

An effectiveness study may show differences in results for sub-groups of patient. In order to exploit these differences, however, care must be taken to ensure that the initial design of the study provides adequate statistical control and power. The same reasoning applies to a cost-effectiveness study, which may only be extended to certain sub-groups if the effectiveness for these sub-groups has been demonstrated.

However, an *ex post* multivariate analysis is considered an acceptable method for evaluating the effectiveness of a treatment depending on the characteristics of the patients treated. It will then be a case of estimating the clinical results for each sub-group on the basis of its characteristics.

**Recommendation 5**

The target population of the evaluation must be clearly described. In the case of an economic evaluation study associated with a clinical trial, the target population is the trial population. If several studies are used to estimate effectiveness, the choice of target population and the estimates of results must be clearly set out. It may be relevant to carry out an economic evaluation study by sub-group of the target population, in so far as sub-group statistical analysis justifies it.

# VI) CHOICE OF COMPARATOR AND COMPARISON METHODS

## 6.1 CHOICE OF COMPARATOR

**Recommendation 6**

Where several strategies exist in practice (including non-drug therapies and non-treatment), these should be listed in full, and the strategies studied should be taken from this list (and justified). The therapeutic strategies to be used as comparators will be chosen among those most frequently used (including non-treatment) or newer strategies which may legitimately be deemed likely to become reference strategies in the very near future.

## 6.2 METHODS OF COMPARISON

Economic evaluation is frequently constrained in its choice of comparators by the therapeutic strategies available from clinical trials. These are often either a placebo treatment or standard medical strategies, but they are often old and sometimes obsolete at the time of introduction of the therapeutic strategy being studied.
N.B. The method of comparison using placebo results is often considered equivalent to "non-treatment". This approximation generally results in an under-estimation of the difference in effectiveness between the strategy being studied and the non-treatment, since a placebo effect exists in virtually all diseases. This disadvantage may be accepted in so far as it leads to a demonstration "by default" with regard to the strategy being studied.

**Recommendation 7**

When performing an economic evaluation study, it is the direct comparison between two strategies that presents the greatest scientific relevance. It may, however, be necessary to perform an indirect comparison based on published data, which is acceptable on a provisional basis in the absence of any original studies allowing for direct comparison.

## VII) TYPES OF STUDY

Different types of study may be envisaged: some seek to establish a link – in the form of a ratio, for example – between the costs and the outcomes of a therapeutic strategy (cost-result study), whereas others just present the costs and the various outcomes (or assessment criteria) of a programme without seeking to formally identify a link (cost-consequences study).

These two types of study are also differentiated by the synthetic nature of the first type and the multicriteria aspect of the second (cost-consequences). In this latter instance, the multidimensional nature of the criteria makes any attempt at aggregating them problematical.

In contrast, any cost-result study necessarily implies that a reference strategy has been defined, against which all other strategies will be evaluated.

### 7.1) COST MINIMISATION STUDIES

Cost minimisation studies are used when the strategies being compared differ only in the costs they incur. When two strategies have the same therapeutic efficacy and the same consequences (medical or social for the patient), but different costs, the lowest-cost strategy is sought. Some would say that there is no such thing as a cost minimisation study, arguing that the assumptions they imply (same effectiveness, same utility) are never met in reality.

### 7.2) COST-EFFECTIVENESS STUDIES

Cost-effectiveness studies are used to identify the strategy that will produce the maximum effectiveness for a given cost or, vice versa, to achieve a given objective at the lowest cost. Effectiveness is measured by a clinical outcome indicator or an objective health state indicator. These studies also provide decision-makers with information on the additional effectiveness obtained from an extra cost.

### 7.3) COST-UTILITY STUDIES

Cost-utility studies are a generalisation of cost-effectiveness studies for the result obtained. Their theoretical interest lies in their ability to provide a summary indicator of results. This type of analysis requires a knowledge of patient preferences. Cost-utility analysis is used particularly when impacts on survival/quality of life are important criteria for judging the effects of health care strategies.

### 7.4 COST-BENEFIT STUDIES

Cost-benefit studies make it possible to determine whether a new health care strategy generates a net benefit for society. Cost-benefit analysis can therefore be distinguished from cost-utility analysis in that it implies that all the costs and all the outcomes of the strategy being studied are expressed in monetary terms. However, assigning a monetary value to health care results poses many problems, particularly since these are non-merchantable effects.

### 7.5 COST-CONSEQUENCES STUDIES

Cost-consequences studies seek to draw up an inventory of all the costs incurred by a programme and all the outcomes, positive or negative, to which it might lead. It is therefore in their nature to consider phenomena which are expressed in a highly heterogeneous form, from both a qualitative and a quantitative point of view. This being the case, cost-consequences studies can also incorporate quality of life studies. Cost-consequences studies may also constitute a useful preparatory basis for a future economic evaluation.

When these studies fail to link resources to results by means of an explicit criterion, they are not considered as full-scale economic evaluations. Instead, they are seen as decision aids which leave a wider margin of freedom for the decision-maker than the studies described above. In reality they involve identifying, for a given subject, all the information that is relevant and useful for decision-making, to collect and classify that information according to two aspects of an action: on the one hand, all the costs the action entails, whatever the nature of those costs and how they are measured; and on the other hand, all the outcomes, economic or otherwise, that the programme will bring about (excluding costs, to avoid double counting).

**Recommendation 8**

Each evaluation has its own particular scope and limitations. The type of study selected should be clearly stated and justified with respect to the issue addressed, and must be described at the start of the study. The author should also provide his personal definition of the type of study used.

# VIII) TIME FRAME AND TRANSFERABILITY TO OTHER COUNTRIES

## 8.1) TIME FRAME

Different time frames may be envisaged. A study may seek to measure the impact of the treatment over the patient's entire lifetime (particularly in the case of chronic diseases) or may instead prefer to restrict the study to a more limited time frame.

Although determined primarily by the natural history of the disease, the choice of time frame for the study is often restricted by the data available (particularly data from clinical trials).
But modelling may be used to estimate the long-term consequences both in terms of costs and health outcomes (see section 12).

**Recommendation 9**

The time frame of an economic evaluation study must be justified in terms of the natural history of the disease and the availability of data.
It is recommended that the time frame chosen should be long enough that all outcomes, both positive and negative, of the treatments used and evaluated be included in the study. This recommendation may be difficult to apply in certain cases, in particular as regards very long-term effects, mainly due to major uncertainties or the impact of new treatments in the intervening period. In such cases, the reasons for the choice of the time frame should be expressly stated, and the consequences of not including certain events in the study should be discussed.

## 8.2) TRANSFERABILITY TO OTHER COUNTRIES

In the absence of national data, economic evaluation studies of therapeutic strategies often rely on international data, particularly in the fields of epidemiology and clinical practice.

It is generally accepted that clinical results are transferable from one country to another, or from one group of countries to a single country, as long as the criteria defining the population concerned and the method of health care organisation are respected. Epidemiological data may be subject to greater variability.

Economic data, however, are not generally transferable from one country to another, due to differences between health care management methods, and hence in resource consumption, and to differences in the prices/tariffs of the goods and services used.

**Recommendation 10**

Epidemiological and economic data cannot be transferred from one country to another without prior verification. When the economic evaluation study relies on international data or national data for a country other than France, the authors should demonstrate that these data (clinical, epidemiological or economic) may be transferred with sufficient plausibility to the French context. In the absence of any supporting evidence, the study should mention that the data so transferred are to be used with caution.

# IX) DEFINITION AND MEASUREMENT OF COST

## 9.1) DIRECT MEDICAL COSTS

From a purely logical point of view, the only costs that should be considered as direct costs are those directly attributable to the disease, in respect to its treatment and/or prevention, whether primary or secondary.

In reality, it has become common practice, by a form of tacit understanding, to confine direct costs to medical costs alone, and to exclude the value of the time spent by the patient or relatives which is (wrongly) assigned to indirect costs, even when directly attributable to the disease.

### 9.1.1) Hospitalisation

#### *9.1.1.1) Short-stay*

There are two ways of estimating the costs of hospital stays in the Medical, Surgical and Obstetric (acute care) units of public, private not-for-profit and private commercial institutions. Either carry out an *ad hoc* survey as part of a study on a sample of institutions, by compiling the resources used and by valuing these on the basis of unit costs supplied by the institutions. This method offers the advantage of providing a very precise estimate of the resources used, but the disadvantage of limiting the representativeness of the results to the sample selected. The second way is to use the *Programme de Médicalisation des Systèmes d'Information* (PMSI) databases [the medical information system which is the French DRG program]. There are three databases available, and their conditions of access must be negotiated on a case by case basis.

The first of the three is the national database of *Résumés de Sortie Standardisés* [standardised discharge abstracts], which covers all short-stay hospital admissions in France. It does not include cost data, but can provide data on the length of stay for admissions identified by diagnosis or by medical procedure.

The second is the *Etude Nationale de Coûts* [national cost survey] database, which is built on a sample of public and private not-for-profit institutions. It provides costs per stay broken down into four main categories: medical cost, medico-technical cost, logistic cost and structural cost. It can therefore provide an estimation of the costs of admissions selected on the basis of principal and associated diagnoses, or of medical procedures.

The third is the database developed for private commercial hospitals by BADIMEHP, an offshoot of the *Fédération Hospitalière Privée* [private hospital federation]. It consists of two sub-bases. The first can be used to calculate the charges billed per stay, and is therefore a tariff database. The second uses the same methods as the *Etude Nationale de Coûts* database to calculate the cost per stay.

In economic evaluation studies, it is recommended that the most specific data be used, i.e. the data which allow identifying the cost of a stay for a patient identified by his diagnostic codes and/or the medical procedures which the patient received in the course of the stay. Appendix II presents the technical recommendations on how these data may be used, and in particular how to ensure that they are representative at the national level.

**Recommendation 11**

The databases of reference used to calculate the cost of a hospital short-stay are those generated by the Programme de Médicalisation des Systèmes d'Information (PMSI) [the medical information system which is the French DRG program], which offer the most precise method of identification of the clinical characteristics and the treatments of the patients. An *ad hoc* observational survey is therefore possible, but the results obtained for a sample of hospitals should be compared against national data to ensure that the results are truly representative.

#### *9.1.1.2) Medium and long-stay*

***1) Public hospitals (including PSPH)***

We should begin by saying that the exhaustive information gathered in 1999 on the public sector (including PSPH) allows absolutely no possibility of cross-checking with previous years, given the way in which the *PMSI- Service de Suite et de Réadaptation* (SSR) was designed. The counting of patients and admissions are not performed – or at least not displayed – and the only option available is to attempt to count the number of hospitalisation days (full hospitalisation or day-patient or follow-up hospitalisation, as well as the number of sessions) of patients on grounds of morbid events, through a complex system of data based on the CIM 10, implying the solution to the problem of CIM 9 – CIM 10 code correspondence encountered for short-stay admissions.

SSR data is classified by:

1 – CMC (*catégories majeures cliniques*) [major clinical categories]
2 – GHJ (*groupe homogène de journées*) [homogeneous groups of days]

The data are drawn from the RHAs (resumés hebdomadaires anonymisés) [anonymised weekly abstracts], which are situation data reporting the number of days of hospitalisation per week. But the number of RHAs, by CMC or by GHJ, cannot be used to represent the number of stays in one of these categories.

An unencrypted list of the GHJs shows that they are differentiated on the basis of the main diagnosis, the degree of dependence, the level of health care management required and the complexity of physiotherapy/rehabilitation. For each GHJ, we have global information (number of institutions in which they have been identified, number of days, mean age, sex ratio, post-operative delay, etc.), information on the disease treated (in CIM 10 codes), on the aetiology when this differs from the disease, and on the medical procedures carried out.

As yet, however, there is still no cost database for the medium-stay hospital admission in public hospitals and PSPH. On the other hand, the CNAMTS has broken down the tariffs of 1,073 establishments into 8 main groups of disciplines for the year 2000 by calculating the maximum tariff, the minimum tariff and the standard deviation for each.

***2) Private hospitals***

At present, the only documents produced for SSR care have been methodological guidelines and no figures have been published.

The CNAMTS has nonetheless calculated 6 classes of discipline and occupation rates, average length of stay, number of days, number of beds and average tariffs for 450 institutions (institutions operating under CRAM contract and on a daily rate established by prefectorial order).

**Recommendation 12**

The development of the PMSI for *Services de Suite et de Réadaptation* (SSR) [medical follow-up and rehabilitation services] provides useful information in terms of epidemiological knowledge of medium-stay admissions but doesn't allow the valuation of its cost, because of the absence of sufficient data on unit costs. The provisional scale of costs is not considered reliable. It is possible to apply the daily rates established by the CNAMTS for the rest/convalescence, physiotherapy and rehabilitation disciplines to the number of days per disease derived from the PMSI-SSR database, or to carry out *ad hoc* surveys of specialised institutions.

### *9.1.1.3) Transportation costs*[5]

Transportation costs are monitored by the *Assurance Maladie* and may be extracted by means of the *Système d'Information de l'Assurance Maladie* (SIAM) [Information system of the Health Insurance] requests performed on the information systems of the reimbursement offices. It nonetheless remains difficult to assess the total transportation costs in the hospital sector:

- In institutions operating on a lump-sum allowance (public and private PSPH), since part of the transportation expenses are covered by the allowance as a result of the distinction made between "primary" transportation (home/hospital) and "secondary" transportation (between institutions), any study of transportation costs must involve a survey of the institutions concerned. In principle, the *Etude Nationale des Coûts par Activité Médicale* [national survey of costs per medical activity] can be used to calculate the cost per half-hour of *Service Mobile d'Urgence et de Réanimation* (SMUR) emergency medical transport for the sample of institutions taking part in the survey. Indeed, in the cost records broken down by functional unit, there exists a "SMUR" functional unit that identifies the direct costs (salaries, vehicle depreciation and maintenance, fuel, medical equipment depreciation and maintenance, various medical consumables) in relation to the work unit of "a half-hour of SMUR". Access to the SMUR functional unit files is obtained by special request to the CTIP[6].

- The method of transportation must be taken into account: helicopter, rail or motor vehicle. The first of these (helicopter) is a key component of "emergency" transportation costs, a field in the process of reorganisation (a ministerial working group has been set up to address the subject) and for which many studies exist. In cases where motor vehicles are used, a distinction must be made between the use of a "medical" vehicle (light medical vehicle or ambulance) and the use of a private vehicle. Where a medical vehicle is used, any comparison (e.g. geographical) will need to take into account the levels and the (local) structure of transportation supply (the light medical vehicles/ambulances ratio, for example). For what concerns the use of private vehicles, this may be of particular significance if the aim is to study the costs of

[5] See also Appendix VI
[6] Centre de Traitement de l'Information du PMSI [PMSI information processing centre],University of Paris VI, 12 rue Cuvier, 75005 Paris

transportation for patients with chronic diseases (chronic renal failure, for example), given that expenses are reimbursed on the basis of a national scale.
- Finally, it is important to point out that in hospitalisation there exists what is known as the "nearest institution" rule which may influence the evaluation of transportation costs, part of which may not be covered by the Assurance Maladie.

**Recommendation 13**

For the hospital sector, any evaluation on the transportation costs must take into account the distinction between primary (home/hospital) and secondary transportation (between institutions) and how the institutions is funded (lump sum allowance/daily tariff). Where necessary, a survey of the institutions concerned should be carried out in addition to consulting the *Assurance Maladie* information systems.
A study of transportation costs will also need to take into account the nature of the hospital activity (accident & emergency, for example) or the diseases under consideration (chronic diseases in particular).

### 9.1.2) General practice

When the perspective adopted is that of the Assurance Maladie, evaluating general practice costs poses no particular problem since reimbursement tariffs are available. It is also entirely justifiable to integrate additional payments charged by doctors, when the amount is known, to the estimations of costs not reimbursed by the Assurance Maladie.

When taking a societal perspective, however, it is legitimate to consider whether the reimbursement tariffs constitute satisfactory estimates of the corresponding opportunity costs. We know that these tariffs mainly cover, at least in the case of GP consultations, the remuneration of a medical time. Indeed, the overhead costs generally represent a small fraction of the fees charged. GP consultation reimbursement tariffs are set through negotiations between the health care authorities and representatives of the medical profession, and it is by no means sure that the tariffs applied are at the level they might be in an open market. It is also clear, however, that the social management of health care spending substantially modifies access to health care and leads to an increase in the volume of those seeking treatment. Furthermore, health care professionals adjust consultation times which enable them, to a certain extent, to maintain their hourly rates at acceptable levels. These considerations suggest that attempts to estimate an opportunity cost other than with the reimbursement rates would raise many questions of methodology and feasibility, and that it seems difficult to find any satisfactory solutions.

**Recommendation 14**

Conventional tariffs and the latest data published by the Assurance Maladie reimbursement offices should be used as the basis for measuring the costs of general practice and, if a societal perspective is adopted, the portion of costs not reimbursed and any supplementary fees charged by doctors over the tariff should be included. A new set of specifications is available, the *Classification Commune des Actes Médicaux* (CCAM), which replaces the former *Catalogue des Actes Médicaux* (CdAM) and the *Nomenclature Générale des Actes Professionnels* (NGAP).

## 9.2 DIRECT NON-MEDICAL COSTS

An individual's health state and the treatments he needs have a direct impact on the daily life of the individual and his relatives, translating either into time "lost" or monetary expenditures. Examples of such costs include the time and cost of access to health care services, time spent on childcare while parents, spouses or relatives (elderly) are unavailable, time spent by family visiting during hospitals stays, time spent by home-helps to make up for the inability to do housework, but also the cost of re-fitting a home to cope with a handicap (in the case of rheumatoid arthritis, for example), expenditure on substitute equipment not considered as medical products (such as a wig to conceal hair-loss induced by chemotherapy). The main resource in the first case is time; the second case involves the acquisition of goods.

In the case of the acquisition of goods, there valuation is performed by identifying the corresponding expenditure. This is done by means of ad hoc surveys, which supposes the selection of a representative population. Such work should be subordinated to theoretical assumptions which may be made regarding the scale of these costs relative to others, and regarding the differential impact of one health care programme compared to another in the case of a comparative study.
When the resource used is time, it is also measured by means of *ad hoc* surveys. In theory, we could suppose that these costs will be low and may be considered as transitional adjustment costs in the case of short-term acute care episodes leading to rapid patient recovery (under 1 month). In this case, it could be explicitly assumed that they will not significantly modify either the values of the cost-outcome ratios or the general conclusions. This assumption will need to be justified, however. In contrast, the lengthier the disease and the treatment, and the greater the potential for long-term invalidity, the more pertinent it will be to take such costs into consideration. Caution will be required,

however, in comparing health care programmes, even of short duration, which significantly change patient management, for example from full hospitalisation to day-care hospitalisation. In such instances, it is possible that a solution which is less costly in terms of medical care may translate into a significant transfer of the burden of cost to the family. A full cost review should therefore take such transfers of costs into consideration.

Evaluating these costs in monetary terms is a subject of controversy. In many cases, the time is taken out of leisure time, corresponds to mutual help between family and friends and involves no expenditures. The decision to rely on paid help depends on household income and the value assigned to the use of non-working time. In this instance, it could be argued that the time lost has no opportunity cost for society. This lost time will therefore be counted separately, will not be included in the incremental cost-outcome ratio, and will figure as a supplementary argument in the comparison of two health care programmes.

This reasoning may reach its limits, however, when a large number of patients suffers from a particular disease or when the treatments implemented require a great deal of time, such as in the case of chronic incapacitating diseases (multiple sclerosis, muscular dystrophy, Alzheimer's disease, etc.): if the whole of society were to be so ill as to require constant assistance from those around them, this unpaid time devoted to managing the consequences of the disease would not be available, either for other domestic activities on behalf of those in good health (e.g. bringing up children) or for consumption, and would therefore have an impact on the potential for wealth creation. Once again, caution will be required in comparing health care programmes, even of short duration, which significantly change patient management, for example from full hospitalisation to day-care hospitalisation. In such instances, it is possible that a solution which is less costly in terms of medical care may translate into a significant transfer of the burden of cost to the family. A full cost review should therefore take such transfers of costs into consideration. Once again, it is the expected amount of these costs which should determine whether or not they should be evaluated in monetary terms.

A monetary evaluation (of the cost of child-care or domestic help, for example) can then be derived from the cost of an equivalent resource on the market (replacement cost). This approach reflects the viewpoint of a rational collective decision-maker who chooses the most efficient replacement solution, without taking individual preferences into account. In a cost-benefit approach, it could be argued that patients include these consequences in the measurement of willingness to pay, if they are included in the description of health states. This means, however, that these consequences must be clearly incorporated into the method used to describe the health states on which the measurement of willingness to pay is based.

**Recommendation 15**

The disease and its treatment may induce direct non-medical costs incurred by patients and their families. These costs may relate either to expenditures for the acquisition of goods, or may represent a cost in time. These direct costs should be measured by *ad hoc* surveys carried out on a rigorous basis, particularly in view of the potential income-effect. Non-working time may be measured in one of three ways: a time total is calculated but is not included in the cost-outcome ratio; the time total is valued at the hourly rate applicable to an equivalent resource on the market; alternatively, as part of a cost-benefit analysis, this time is measured by the patient's willingness to pay. On a subsidiary basis, it will be possible to evaluate the number of working hours the caregiver is required to give up.

## 9.3) INDIRECT COSTS

These consist primarily of production, productivity and human life losses.

### 9.3.1) Production losses, productivity losses

A disease and its treatment may have an impact on the productivity of patients in active employment and hence on the production of wealth at a national level.

Loss of productivity is valued by measuring the number of working hours or days lost due to the disease and its management. As with non-working time, this time should be measured by *ad hoc* surveys, but these pose major problems of sampling and correction in the light of the potential bias arising from the professional activity of patients and the mechanisms for attribution of sick leave by the national insurance system (the difference between salaried employees and the self-employed in France, for example).
Three methods co-exist for evaluating these losses in monetary terms. The first, known as the human capital approach, measures the impact of the disease in terms of production losses it entails, by multiplying the number of work-days lost by their production value. This approach makes no distinction between the impact of the disease on the different companies or sectors in which it occurs. The value chosen to measure production by is GDP per head of active population calculated per working day. This approach has been criticised for being unrealistic, however. Indeed, it relies on the assumption of a full employment economy in which a work-day lost has an automatic and

proportional impact on production. It also fails to consider the fact that there is a certain flexibility in the workplace, that absenteeism can be compensated in the short term by an increase in productivity by the employees without increasing total payroll (or even with a reduction in the total payroll if the absent employees are neither replaced nor receive monetary compensation). In view of these limitations, the method has limited acceptability for public decision-makers since it is not based on a realistic identification of resource and financial flows.

The so-called friction-costs[7] school proposes to make up for the lack of realism inherent in the previous approach by empirical study of how employers adjust to employee sick leave. The authors point out that an employee's absence from work does not necessarily lead to a fall in production, since the absence may be offset either by higher productivity among other employees or by the employee catching up on his work following the period of sick leave. There may well be a period, known as the friction or adjustment period, during which the employee will not be replaced. Only subsequently will the employer introduce a replacement and incur the costs of a temporary employment agency and of training the replacement. The tighter the employment market (the higher the demand for staff), the higher these costs will be (and the longer the adjustment period). The basic assumption of this method is that the elasticity of production to working time is not equal to 1. An elasticity of 0.8 is used in Holland, where this approach was first developed. In addition, the method also measures the medium-term impact of sick leave on the employment market, and in particular the impact of disability leave. In France, this would involve studying the medium-term impact of sick leave on social security contributions, with two contradictory effects. In a period of unemployment, an increase in sick leave may eventually improve the employment situation, but it also leads to an increase in social security contributions, hence to an increase in the cost of work or a decrease in workers' net income and hence to a fall in purchasing power. To sum up, the friction-costs approach considers firstly that the loss of production is not exactly proportional to the number of work-days lost; secondly, it offers a macro-economic modelling of the impact of sick leave on the employment market and, hence, on the cost of work and on corporate competitiveness. It therefore requires specific study, constantly updated, in order to be applied.

The third approach is derived from the theory of well-being and applies to cost per QALY studies or cost-benefit studies with the measure of willingness to pay. With this approach, the disadvantages arising from sick leave are measured by including them in the evaluation made by patients of their health state following a given treatment. So-called indirect costs are thus included in the denominator of the cost-results ratio and are not explicitly mentioned.

The first two approaches for the measurement of production and productivity losses make no provision for taking into account unpaid work, in particular domestic work, or the time of those not in active employment such as retired persons, or of those excluded from the job market because of a handicap. A standard pragmatic approach is the evaluation of "lost time" as described above. It could reasonably be assumed that retired people in good health perform domestic or voluntary activities which contribute to the national wealth. It may equally be assumed that they generate consumption externalities by using the purchasing power of their pension, which is either a deferred income or forced savings. A standard pragmatic approach would be to measure lost time, as described for direct non-medical costs. The welfarist approach, on the other hand, allows for inclusion of the evaluation of lost time by patients themselves.

Lastly, developed countries operate mechanisms for offsetting the loss of income due to disease. In France, the Sécurité Sociale makes up part of the loss of income by paying daily allowances, financed from social security contributions. By definition, the expenses corresponding to the daily allowances received are equal to the social security contributions paid. From the viewpoint of economic analysis, these flows do not represent additional expenses but a redistribution of income between agents. They are therefore not accounted for. From the point of view of the Sécurité Sociale, however, a reduction in the level of sick leave generates savings, and it may be relevant to measure them. In the long term, if the number of sick leave falls significantly, social security contributions could decrease, generating additional disposable income to be shared between employers and employees, with a positive impact on national wealth.

**Recommendation 16**

The impact of the disease and its treatment on the functioning of the economy must be taken into account in the economic evaluation of a health care programme, when there is good reason to believe that this impact is significant, in terms either of the number of people concerned or of the lasting of the consequences of the disease in the lives of patients. The number of lost working days is measured by *ad hoc* surveys. The valuation with the human capital approach is simple but unrealistic. The basis for the friction-cost method is more realistic, but requires much more empirical work before it can be applied in France. Neither approach allows evaluating the time lost by persons not in employment. The theory of well-being recommends including the value of lost time in the evaluation of the results of a treatment programme, rather than in the costs. In any event, the inclusion of indirect costs must be specifically analysed and its impact on the results of the evaluation studied in the light of the method adopted.

[7] Koopmanshap M.A. et al, 1995

### 9.3.2) Cost of human life

Economic evaluation of the loss of human life is based not only on the value placed by developed societies on life as such, but also on the established link between longevity and wealth. A society whose population is in a poor state of health, with high early mortality (by comparison with the current standards of developed societies), is a society in which children begin to work at an early age and acquire no educational capital, and in which adults do not save and therefore do not lay the foundations for the accumulation of capital that creates future wealth. There is therefore a link between the human capital accumulated through education and health, and the wealth of a society. In such a model, death deprives society of future wealth. The calculation of a cost per year of life gained ratio from death arises from this link between human capital and wealth. The cost of a lost human life can therefore be measured by the loss of potential income associated with survival. This annual cost may be calculated as the discounted total income over a lifetime which is equal to the life expectancy of the population, divided by this life expectancy. The incomes considered should be gross incomes, after deduction of social security contributions but before tax. This annual cost has the advantage of being the same for all, irrespective of the date at which death occurs, the difference lying solely in the number of life years remaining, which will naturally be lower for the elderly than for the young.

The following point is currently under debate: should not the future health care spending of survivors be deducted from the income expected from survival? The proposition seems logical. And yet, in a universal health insurance system within which intergenerational solidarity operates, it may be assumed that the discounted amount of lifelong contributions is equal to the discounted amount of health care spending. The future health care spending of survivors would only be taken into account if this assumption could not be verified.

If the discounted annual income is divided by the whole population, it may be assumed that this value measures the loss of a human life whether the person was in active employment or not. The conclusion is, therefore, that those not in employment contribute, through the use of their time, to the possibility of wealth production by those in employment.

An alternative to the human capital approach is that derived from the theory of well-being. The benefits of survival are valued by the individual's willingness to pay for an additional year of life, and this valuation incorporates their health state. In this approach, potential future incomes linked to survival should not be included in the cost-benefit analysis, since the value of this income is taken into account in the individual's willingness to pay.

**Recommendation 17**

The value of human life may be measured in two ways: using the human capital approach, by calculating per capita income discounted over the life expectancy of the French population, or using the theory of well-being approach, using the valuation placed by individuals on the life years gained.

## 9.4) INTANGIBLE COSTS AND BENEFITS

Intangible costs reflect the loss of well-being of the patient and its relatives arising from the disease; this loss may represent a human and psychological cost (suffering, pain, loss of life, etc.).

It is universally agreed that these negative factors should be included not in the costs but in the outcomes, either in studies including concepts of quality of life or in cost-benefit studies.

**Recommendation 18**

The evaluation of intangible costs (or benefits) is not justified, given that the human and psychological impact of the treatment may be studied by means of quality of life indicators, measurements of utility or by using contingent evaluation methods (see section 10.4).

# X) DEFINITION AND MEASUREMENT OF OUTCOMES

## 10.1) EFFECTIVENESS

### 10.1.1) Definition of effectiveness: experimental efficacy and effectiveness

Clinical trials measure experimental efficacy, in that their results are obtained from experimental protocols or under conditions of controlled prescription. These results cannot therefore be considered as completely equivalent to a measurement of effectiveness in real conditions of use. And clearly it is effectiveness in real conditions of use that is of interest to the patient, the doctor and the decision-maker.

Working on the basis of the ideal model of theoretical or experimental efficacy, which implies a patient selection process governed by precise criteria of inclusion and exclusion, as well their randomisation and the use of double-blinding, different deviations may arise which will affect the measurement of efficacy. The dosages prescribed may not correspond to standard practice in different countries; patient randomisation and the double-blinding method may not have been rigorously applied; above all, patients compliance is not 100% under real conditions.

Is it possible to overcome this bias? Supposing that it were possible, does this really give a measurement of effectiveness? And does this effectiveness really measure the effectiveness of treatments in a non-selected population from a public health perspective?

The concept of effectiveness has more to do with a continuum of situations than the simple efficacy/effectiveness opposition would suggest: it offers a variable geometry, depending on the objective, of the conditions for its measurement and the possibility of correcting certain forms of bias in order to come closer to reality without prejudicing the consistency of the indicator.

**Recommendation 19**

When effectiveness is expressed by a single indicator, this solution should be justified and the aspects of effectiveness not examined should be clearly stated. When several indicators are combined, the methods of combination should be described.
If the indicator expresses a number of events avoided, the period over which this outcome is obtained should be stated.
Effectiveness is to be preferred to efficacy. If, however, efficacy is the sole indicator available, every effort should be made to convert this into effectiveness by means of appropriate adjustments.

### 10.1.2) Effectiveness indicators: final end-point, intermediate end-point and surrogate end-point

Medical evaluation distinguishes final end-points (life expectancy, life saved, functional improvement, clinical events avoided) and surrogate end-points, which are criteria for which it has been demonstrated that their improvement leads to an improvement in health criteria (a reduction in blood pressure or cholesterol level, for example). In the absence of surrogate end-points, intermediate end-points, for which no such link has been proven, will be used.

The use of intermediate end-points in an economic evaluation raises two types of difficulties:

- a difficulty of comparison: only treatments with an impact on the selected intermediate end-point can be compared, which therefore excludes treatments which act on other parameters but which nevertheless have an impact on the occurrence of the final end-points;
- a difficulty in the economic interpretation of a variation in effectiveness on the intermediate end-point; it is difficult to know what a 20% reduction in one of these parameters means in terms of effectiveness.

**Recommendation 20**

Ideally, a health care intervention should be evaluated on the basis of final end-points (expressed either directly or through surrogate end-points). If only intermediate end-points are available, the relationship between these end-points and the final end-points should be estimated, primarily on the basis of published literature.

## 10.2) QUALITY OF LIFE

### 10.2.1) Definition of health-related quality of life

"Quality of life is defined as *individuals' perception* of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations and concerns".[8]

In the field of health care, however, it is preferable to restrict our approach to the health-related quality of life: we only consider the aspects relating to health and patient management.

Interest in measuring health-related quality of life has grown considerably among researchers, health care professionals and decision-makers. It allows for subjective criteria (physical, social, emotional) to be combined with purely medical evaluation criteria (often intermediate). It therefore provides a patient's (or a member of society) assessment of the various consequences of his disease and its management. Quality of life is a legitimate criterion of result in an economic evaluation.

Measuring quality of life looks to a less restrictive vision of health than "an absence of disease", which makes it particularly appropriate in the case of chronic or recurrent diseases.

Two approaches put forward methods integrating quality of life:

- the psychometric approach, which produces quality of life measures using questionnaires (recommendations 22 and 23);
- the economic approach, which produces utility measures using preference revelation methods (recommendations 24 and 25).

Evaluation of quality of life offers a relevant and validated approach for most diseases or treatment processes, whereas preference revelation methods are still the subject of considerable theoretical debates.

**Recommendation 21**

Evaluating quality of life offers real descriptive and informative potential; carrying out such studies is therefore highly relevant.
In terms of decision-making in public health, evaluating quality of life therefore is a complement to the evaluation of effectiveness.

### 10.2.2) Measuring quality of life: the psychometric approach

#### *10.2.2.1) Conditions of use*

A quality of life evaluation tool consists of a medium through which to describe states of health and a calculation procedure allowing associating a qualitative numerical value to a health state. When a tool is chosen, its conditions of use must be respected, in particular: the target population, the disease, the method of administration, the rules for the calculation of the score and the rules for the interpretation of the results.

The growing interest being shown in the measurement of health-related quality of life has given rise to the development of a large number of tools, the quality and documentation of which may vary widely. In-depth bibliographical research[9] on the population or disease concerned must be carried out systematically.

In order to properly assess the quality of a tool, the researcher must have access to the original publications, although these may often be incomplete. In which case, it is strongly recommended that the developers of the instrument be contacted and questioned about the circumstances in which it was developed[10], whether it has already been used, if it can be used freely but, above all, if it has already been tested to demonstrate its validity, reliability and sensitivity.

Validity: a measure is said to be valid when the proportion of bias (systematic measurement error) is low. Internal validity means that the instrument actually measures what it is supposed to measure (content, criterion and construct validity). External validity indicates the extent to which the conclusions of a study may be generalised to other populations.

[8] WHO definition quoted by J. Orley, Paris, 1996
[9] The three main bibliographical databases listing most publications on quality of life are Medline, Pascal and Embase. Key words: Health Quality/Quality of life/disease or sub-population considered.
[10] As a general rule, a questionnaire is more acceptable if it has been generated from interviews with patients than from expert opinions.

Reliability: a measure is said to be reliable when the random measurement error is low: the variability of the measure is then explained by the variability of the criterion. In general, several minimum properties must be satisfied:

- Internal consistency: the property whereby items measuring the same attribute produce strongly correlated scores (Cronbach coefficient alpha).
- Reproducibility: the tool is tested on an identical sample at two different times, to see if the results are stable over time (test/re-test method).
- Inter-observer reliability: questionnaires, which are not self-administered, are tested on an identical sample by two different interviewers, to see if the sample is interviewer-sensitive.

Sensitivity: an instrument's capacity to detect variations in quality of life over a given period.

When a French version of an instrument developed abroad is used, it must have been transposed in cultural as well as linguistic terms. Hence a French-language instrument developed in Quebec may not necessarily be directly adaptable to French culture.

Among the different instruments used are indexes, which allow calculating a global score, and profiles, which give a score for each of the different quality of life dimensions explored, without combining these into a single score. It is essential to comply strictly with the scoring procedure described in the instrument user manual, and in particular no attempt must be made to aggregate a profile.

**Recommendation 22**

It is recommended that only instruments that have been subject to a validation process (linguistic, cultural and psychometric) be used to measure quality of life.
Since a quality of life measurement instrument is based on specific conditions of use, on a questionnaire and a scoring function, none of these components can be modified without forfeiting the benefits of the validation process.

### *10.2.2.2) Measuring quality of life: generic and specific instruments*

Evaluation instruments are generally distinguished by whether they are generic or specific in nature. Generic quality of life questionnaires are instruments developed to correspond to a broad spectrum of treatments or diseases of differing degrees of severity, unlike specific instruments which are specific to a particular disease.

Each approach has its advantages and disadvantages, which the researcher must choose between.

| Approach | Strengths | Weaknesses |
|---|---|---|
| Generic | Detects the incremental effects of a procedure on the different dimensions of life.<br>Comparisons between different health states or populations. | May hide a key factor.<br>Low sensitivity (requires a large sample). |
| Specific | Clinical sensitivity<br>Temporal sensitivity | Depending on the nature of the specific instrument, does not allow for comparison of:<br>- different dimensions;<br>- different populations;<br>- different health states. |

According to Guyatt, 1995

In so far as they have complementary strengths, it is preferable to associate a specific instrument and a generic instrument, for example to study the unexplored aspects of one scale compared to another.

From a practical point of view, however, their systematic association is not always possible, since it may compromise the overall acceptability of the study by making it more cumbersome. The final choice of questionnaire type thus depends directly on the aim of the study.

**Recommendation 23**

In view of the wide variety of tools available for measuring quality of life, it is important to ensure that the tool employed is appropriate to the aims of the study. Since the strengths and weaknesses of generic and specific tools are mutually complementary, it is recommended that the two tools should be combined wherever possible. Otherwise:

. the use of a specific instrument is recommended when the aim of the study is to compare several strategies for a given disease;

. the use of a generic instrument is recommended when the aim of the study is to compare several diseases, methods of health care management, etc., or when no specific instrument is available for the population considered.

## 10.2.3) Utility

In economics, the preferences of an agent are modelled by means of a binary relation between the options proposed to the agent: this "preferred or indifferent to" relation is known as a preference relation. In certain axiomatically defined cases, this preference relation may be represented by a utility function.
In the field of health care, utility[11] therefore expresses the preferences of individuals (patients, the general population) with respect to different health states. The methods developed by health economists may be used to reveal these preferences through individual choices[12].

### *10.2.3.1) Definition of utility*

There are three types of utility:

- **cardinal (or ratio) utility** is supposed to measure absolute intensities of satisfaction (if the utility of A is equal to 3 times the utility of B, state A is "three times better" than state B);
- **interval (or Neumannian) utility** is supposed to measure relative intensities, i.e. satisfaction variations (all that can be said is that the difference between the utility associated with state A and the utility associated with state B is three times greater than the gap between the utilities associated with states B and C respectively);
- **ordinal utility** measures satisfaction in simple order of preference using real numbers.

### *10.2.3.2) Measurement of utility: methods of preference revelation*

***1 – Standard gamble method***
This method involves asking a subject if he prefers contract no. 1: "being certain to live out the rest of his life with a given handicap" or contract no. 2: "undergoing an operation that will enable him to live the rest of his life without handicap and accepting a 1-p risk of death". The subject is then asked to determine the value of p for which contracts 1 and 2 would be equivalent. This value is then taken to be the interval utility associated with the handicap being considered. Three conditions must be accepted for this:

1. the utility of living without a handicap is 1, and the utility of death is 0;
2. the individual's preference for these contracts may be represented by an interval utility function (existence of a Neumannian structure for all the contracts and risk-neutrality in probability);
3. the utility of the handicap, in a context where there is no uncertainty, is equal to the utility of contract no. 1, while this utility is determined in a context of uncertainty (see risk-neutrality in probability).

Since the condition of the individual's risk-neutrality in probability, in particular, is only rarely verified, this compromises the validity of this measurement technique.

***2- Time trade-off***
This involves asking an individual with a life expectancy of t years if he prefers to live out this time t with a handicap, or to have a shorter lifespan but in good health. The aim is to reveal the length of time t* such that: living for a length of time t* in good health is equivalent for the individual to living a length of time t with the handicap being considered.
The t*/t ratio is then taken to be the value of the interval utility attributed to the handicap considered. As in the case of the standard gamble method, the time trade-off method rests on a strict axiomatic structure which requires that the utility function on the pairings (lifetime; health state) should be equal to the lifespan multiplied by the utility of the health state. The stability of preferences over time and the constancy of lifespan substitution rates are also set down.[13]

***3- Visual analogue scale***
This method consists of asking the subject to identify his present health state directly on a visual graduated scale. Although simple to use, it provides only an ordinal scale of measurement.

---

[11] In this context, utility is used in the economic sense of the term. Economic utility is distinct from the concept of medical utility, which assesses the sound basis of a treatment for a patient in a given set of circumstances.
[12] The additive scoring methods of psychometric scales may also be considered as a form of preference revelation of the tool's designer through the choice of items and weighting (a scoring function is, in the strict sense of the term, a utility function). Factorial analysis itself may be interpreted as a form of preference expressed by the author if he chooses to adopt the scale produced by this analysis.
[13] This axiom implies that the percentage of life years the individual is prepared to sacrifice in order to move from one health state to another of better quality is independent of life expectancy.

*4- Systems of classification of health states*

These systems propose to calculate the utility associated with a health state from the utilities attached to the dimensions composing the given health state; these utilities are derived from surveys of the general population using preference revelation methods. The utility so calculated is thus an evaluation of the average preference rating that would be attributed to this health state by a random sample of the general population.
The main methods used are QWB (Quality of Well-being), HUI (Health Utility Index) and Euroqol: validations of French versions of the latter two are proposed. These methods cumulate the assumptions of the techniques used to determine utility scores for the general population (standard gamble or time trade-off) and the assumptions of multi-attribute utility, but they have the advantage of simplicity in that they avoid the need for population surveys.

In addition to the revelation of individual preferences, and its subjection to a strict axiomatic structure, these four approaches all pose the problem of the aggregation of the individual preferences into a collective preference, since no procedure exists to do this. Authors employing them in their work are content merely to calculate utility means or medians. Being unfounded, such approaches therefore cannot be used as a basis for public decision-making.

**Recommendation 24**

The economic approach to quality of life is based on the utility function which consists in assigning a numerical value to each of the states for which a preference is expressed. If A is preferred to B, the utility of A is a number higher than or equal to the utility of B.

The methods most commonly used to identify preferences are the standard gamble, time trade-off (TTO), and visual analogue scale methods. These methods allow for the expression of individual preferences within a strict framework. The use of these methods must therefore take into account their theoretical and empirical characteristics.

A number of theoretical problems arise when it comes to moving from individual preferences to a collective preference. Since no solution has yet been found to these problems, using the results of any collective preference as the sole criterion on which to base public health decisions is not recommended.

## 10.3) QALYs (QUALITY ADJUSTED LIFE YEARS)

The outcome of a therapeutic strategy may be described according to different dimensions: economic, therapeutic effectiveness and impact on quality of life are the dimensions most commonly used. The economic effectiveness of a therapeutic strategy is usually measured by the ratio of the value of the resources employed by a measurement of the outcome achieved. When the outcome is described in more than one dimension, this ratio can only be calculated when it is possible to aggregate these dimensions into a single numeric criterion. The QALYs (Quality Adjusted Life Years) approach offers a procedure which aims to aggregate a criterion of effectiveness, generally life years gained, with a criterion measuring the quality of life that will be experienced during those years.

By definition, and in its simplest form, the specification of the aggregation function proposed by the QALYs method, used to aggregate a gain of t life years in a state of life whose quality is measured by the value q, is equal to the period t multiplied by quality of life q.

This specification has the advantage of allowing for a simple interpretation of the result obtained. If we agree that the value of the "good health" state is equal to 1, then it is easy to see that one year spent in a quality of life measured by q and that q years spent in "good health" lead to the same aggregate result. It can therefore be stated that living one year in a state of life whose quality is measured by the value q is equivalent to living q years in good health. It is customary to say that living one year in a quality of life measured by q produces 1xq = q QALYs.

In its most complex form, the aggregation function used by the QALYs approach to aggregate a sequence of "lengths-quality of life" described by the series S = ($q_1$, ...$q_t$, …$q_s$), where $q_t$ measures quality of life during the year t, is defined by: $Q(S) = \sum_{t=1}^{t=s} q_t /(1+r)^t$ where r is the discount rate for the life years.

### 10.3.1) The pros and cons of QALYs

The QALYs approach is widely used in international literature and in certain countries, such as the UK and Canada, actually constitutes the tool of reference imposed by the authorities for the economic evaluation of health care technologies. Threshold values have been identified. The UK's NICE (National Institute for Clinical Excellence) thus considers that a value of £30,000 per QALY constitutes the limit of acceptability for the introduction of new treatments.

There is a number of factors explaining the international success of QALYs. The approach is presented as a simple and intuitive extension of the number of life years gained which includes a subjective assessment of the quality of the life years saved. It thus takes into account (at least potentially!) the patient's viewpoint. Compared to a cost, a uni-dimensional measurement of the scarcity of the resources deployed for a medical technology, the gain in QALYs is a uni-dimensional measurement of the technology's medical effectiveness. This is a final end-point (as opposed to intermediate end-points) which tends to measure how treatments actually affect the life of patients in both quantitative and qualitative terms. Finally, the QALY claims to be universal. Any new technology, irrespective of its field of application, is in theory likely to lead to a calculation of QALYs in so far as it necessarily affects the quantity and/or the quality of life. The QALY can therefore be used to rank all medical techniques, irrespective of their application, along a single axis of effectiveness (their cost-QALYs gained ratio), thereby orienting public health priorities in a deliberately rational and transparent manner.
And yet, despite these advantages, QALYs have from the outset fuelled heated and even passionate debates within the health economists' community, with some going so far as to dismiss them as "This "sQualyd" nonsense" (according to Roy Carr-Hill).[14]
The debate over QALYs is complex because it brings together different empirical, theoretical and ethical dimensions, all of which merit the closest attention. Certain criticisms, although directed specifically towards QALYs, in fact have a much wider bearing and are aimed not merely at QALYs but at the very philosophy of economic calculation or at general aspects of economic theory sometimes used in the construction of QALYs. This is the case, for example, with the von Neumann-Morgenstern (vNM) utility theory or the "expected utility" theory.
If the philosophical and theoretical foundations underlying the QALY are open to criticism, there is obviously no justification for its existence.

Clearly, using QALYs presupposes that the user subscribes to philosophical and methodological principles which may perfectly and legitimately be contested, whether within the framework of philosophical options or within that of a scientific debate within the world of economic science.

The success of QALYs in the UK is no doubt explained by the quasi-cultural predominance of philosophical and methodological options that are viewed elsewhere (and in France in particular) with greater mistrust.[15]
However convenient it might be, referring the debate over QALYs to an external epistemological reference is an unsatisfactory solution.

It would therefore seem necessary to summarise the terms of the debate briefly and, let us hope, as clearly as possible – with the exception of the issue of discounting, which is dealt with in another recommendation – with the sole aim of avoiding a naïve and potentially dangerous use of QALYs.

## 10.3.2) Methodological problems

### *10.3.2.1) Specification of the aggregation function: multiplicative form, reference standard*

These problems relate first and foremost to the relevance of the specification of the aggregation function used. For the sake of simplicity, we will offer two examples.

The first illustrates the lack of robustness in the specification of the aggregation function by a multiplication.
Suppose that we wish to compare the following two outcomes: "gain 10 life-years to be spent in a state with a quality evaluated as 0.3," and "gain 20 life-years to be spent in a state with a quality evaluated as 0.1".
The first outcome gives 10 x 0.3 = 3 QALYs, the second gives 20 x 0.1 = 2 QALYs, hence "gaining 10 life-years to be spent in a state with a quality evaluated as 0.3," is better in terms of QALYs than "gaining 20 life-years to be spent in a state with a quality evaluated as 0.1".
Now let us change the initial figures for the measurement of quality of life by a simple amplitude change of 0.4; the evaluations of quality of life are now 0.7 and 0.5. The number of QALYs calculated for the first result is then equal to 7 and is equal to 10 for the second result. The ranking produced is the complete opposite of the first, while the outcomes compared are exactly the same!

Clearly, the multiplicative specification is extremely sensitive to linear changes to the reference basis in which the arguments are measured. If the pertinent ranking of outcomes is the first one obtained, and this ranking is to be retained after translation, the specification of the aggregation function must be modified and the function used must be duration x quality – 0.4 x duration, which is not multiplicative. But which is, in fact, the pertinent ranking?
This simple example illustrates the possible manipulation of conclusions provided by QALYs.

[14] Carr Hill R., 1993.
[15] Moatti J.-P, et al., 1995.

The second example illustrates the caution with which the authorities must issue standards.
The UK's NICE considers that a value of £30,000 per QALY constitutes the limit of acceptability for the introduction of new treatments.
Imagine that a treatment costs £270,000 and offers the patient a life expectancy of 10 years in a health state defined as "Substantial limitation of professional activity… and average moral suffering" on the Rosser-Kind index[16].
If we apply this index and an index constructed from a survey of a sample of nurses, this health state scores 0.963, giving a cost per QALY of £28,037.38. The treatment is therefore acceptable by NICE standards.
If, however, we use an index constructed from a survey of a sample of doctors, the same health state scores 0.760, giving a cost per QALY of £35,526.32. The same treatment is therefore no longer acceptable by NICE standards.
Note that for the two indices, the state equivalent to death (scored 0) and the state equivalent to good health (scored 1) are the same.
Common sense would suggest the following conclusion:
Any result or standard presented as a cost per QALY must include the definition of the reference base within which it was calculated and within which the specification of the aggregation function in the form of a product has been validated.

Results presented without definition of the reference base and without specification of the aggregation function in the form of a product should therefore not be used as the basis for public decision-making.
Assuming that the validation is presented, any attempt to compare results from one study presented in cost-per-QALY form with those of another study using the same type of ratio must be carried out according to the same reference base and with the same instrument for measuring quality of life.

### *10.3.2.2) Instruments used for measuring quality of life in QALYs*

#### *1) The Rosser-Kind index*

The source of the quality of life coefficients $q_t$, used to weight each life-year gained, is of fundamental importance. The early work of Allan Williams used weightings derived from a two-dimensional scale established by Rachel Rosser and Paul Kind[17], which assigned a coefficient defined by two reference values (0 for death and 1 for good health) to states of life described from a dual standpoint of handicap and suffering. The coefficients themselves were derived from a psychometric study carried out in the hospital setting and involving doctors, nurses and patients. Since the authors published their sub-scales by type of person interviewed, it is easy to identify the significant differences in the evaluation of the coefficients according to the interviewee's position in the health care system: evaluations carried out using the weightings given by doctors, nurses and patients do not coincide.[18] It is therefore important to be aware that divergent points of view may lead to major distortions in the ranking of health states if the scales used are heterogeneous and derived from different sources. This is an important limitation on the QALY's claim to universal applicability.

#### *2) Standard gamble and time trade-off*

Subsequent and more sophisticated versions of QALYs used methods based on the rational choice theory to derive weightings, in particular the standard gamble or time trade-off (TTO) methods. QALYs are sometimes even identified with these methods. This is, in fact, erroneous since QALYs can perfectly well be calculated using "psychometric weightings". The use of methods based on preference revelation techniques can nonetheless be justified by a concern for compatibility with the principles of the economic theory of well-being and of rational choice. The coefficients produced by these methods are in fact "utilities" as defined in terms of economic theory[19] and QALYs are thus expressed as the discounted sum of the utilities attached to successive health states over the patient's survival, as $Q(S) = \sum_{t=1}^{t=e} u(q_t)/(1+r)^t$ where $u(q_t)$ is the utility attached to the health states in year t.

The standard gamble and TTO methods are not equivalent and may produce different results for the same health states. The standard gamble method is derived directly from the vNM expected utility theory[20]. Its use is therefore dependent on accepting the postulates of that theory. It is based on an axiom of "rational behaviour" in response to risk, the validity of which has been contested on the basis of experimental observation of behaviour in real situation. This discussion is central to theoretical research, with a whole school of thought –particularly authors working on the "rank-dependent" utility of outcomes – seeking to reconcile standardised rules for expressing rationality in response to risk with the reality of how agents behave.

---

[16] Gudex C., Kind. P.
[17] Williams A., 1998.
[18] Taking a state of health defined as follows: 3 years with a slight social handicap and no moral suffering, followed by 2 years with substantial limitation of professional activity and average moral suffering. If this state of health is evaluated using the doctors' coefficients given by the Rosser-Kind index, the result is: vQALYs (doctors) = 3 * 0.981 + 2 * 0.187 = 4.837. Using the coefficients given by nurses, the result for the same state is: QALYs (Nurses) = 3 * 0.992 + 2 * 0.911 = 6.724.
[19] See Appendix VII.
[20] See Appendix VII.

Even if we accept the Neumannian axiom, we must still be careful to ensure that the utilities we compare are indeed derived from the same system of reference. As they are defined to the nearest linear transformation, comparisons of utility are virtually meaningless if the definitions of the states corresponding to 0 and 1 differ[21]. Unlike the standard gamble methods, the TTO method does not involve a choice made under uncertainty[22]. It is also much simpler to implement, and this is no doubt why it is the method recommended by certain advocates of QALYs, particularly the economists at McMaster's University in Canada. It nonetheless assumes acceptance of the principles of rational behaviour in general and contains a certain number of methodological weaknesses. Indeed, this measurement method of the utility of a health state relies on the assumption that for an agent the utility of the pairing t life-years gained in a state of health s is equal to the product of the length of time by the utility v(s) of the state of health s (this being the specification of QALYs), without ever specifying the reference base against which this utility is measured and without having validated the choice of this specification in multiplicative terms.

## 10.3.3) Conceptual problems

### *10.3.3.1) The QALY as utility*

An ultimate conception of the QALYs is to consider them as utilities and to postulate that an agent prefers a state of health S to another, S', if and only if the number of QALYs associated with S is higher than the number associated with S'[23]. This promotion of the QALY to the rank of a utility "in its own right" changes both its nature and its significance. From a sophisticated health state indicator, the QALY becomes a measurement of patient well-being with regard to health. We could go on to conclude that maximising the number of QALYs is thus the ultimate objective of health care systems!

The QALY as utility comes up against two types of difficulties, practical and theoretical. From an empirical point of view, it is possible to find counter-examples demonstrating that a rational individual may perfectly well prefer a state of health which "procures" fewer QALYs than another. Mehrez and Gafni have given two examples[24]. They conclude that QALYS should be rejected as a measure of utility, in favour of Healthy-Years Equivalents (HYEs). We will return to this issue. In theoretical terms, it is a question of what system of axioms agents' preferences should obey in order for them to be correctly represented by QALYs. This highly complex issue was addressed by Pliskin et al. as long ago as 1980[25], and may be summarised as follows. Suppose, for the sake of simplicity, that the health states are described by two parameters, length of life t and quality of life q, which is assumed to be uniform over the entire period considered. Suppose that we are "under uncertain conditions", i.e. that a probability distribution is defined for the various health states. Now let us suppose that an individual is capable of expressing preferences on these health states, that is to say that he has a utility function providing him the well-being associated with each health state U(t,q). This utility function is specific: in fact, it is a vNM type utility function with two parameters, t and q (instead of a single parameter as in the vNM functions considered earlier). These are referred to in the literature as "multi-attribute utility (MAU) functions", the theory of which was advanced by authors such as Keeney, Luce, Raiffa, etc. in the 1960s-70s[26].

The question is then to what extent: 1) such a function exists, 2) it can be expressed in the usual form of QALYs, i.e. to what extent U(t,q) = t.u(q), i.e. the product of a lifetime multiplied by the utility (vNM) of quality of life. Pliskin et al. showed that this result requires a whole series of axioms: that preferences on standard gambles defined on states of health are "Neumannian", that there is a "mutual independence in utility" between the two arguments, that the duration-quality substitution rate is always constant and that the individual is risk-neutral as regards longevity. These axioms are highly restrictive and Duru et al. have recently produced experimental data showing that supposedly rational agents do not necessarily obey them[27].

---

[21] If we say that u is a vNM utility function, we have seen that any utility v defined by v = au + b, where a is a strictly positive real number and b is any real number, is a representation of the same preference. It is therefore simple to demonstrate that for two pairings (t,s) and (t', s'), such that t.u(s) > t'.u(s'), i.e. that the pairing (t,s) is strictly preferred to the pairing (t',s'), b can be found such that: tv(s) < t'.v(s'). The pairing (t',s') is then strictly preferred to the pairing (t,s).

[22] To avoid confusion, certain authors suggest reserving the term "utility" for coefficients derived using the standard gamble methods, and the term "value" for coefficients derived from a choice made "under certainty".

[23] For a more technical presentation of this question, see Le Pen C., Théorie de l'utilité et mesure des états de santé, Analyse et Prévisions, 129-130: 37-54, 1997.

[24] Let us consider two 5-year life horizons as follows: S = (0.8; 0.8;0.8;0.8;0.8) and S' = (0.1;1;1;1;1). In the case of S, the patient suffers a handicap reducing his quality of life by 20% and the number of QALYs (non-discounted is 4). This patient may undergo a very painful operation which will sharply reduce his quality of life for the year of the operation, but which will restore a normal quality of life over the following four years. The number of QALYs associated with this situation is 4.1. Theoretically, if the QALY is a utility, the patient should choose S' which is supposed to procure greater well-being. In fact, it is perfectly possible that an informed and rational patient may choose S to avoid the inconvenience of the painful treatment. On this point, see Mehrez A., Ghafni A., 1989.

[25] Pliskin J. et al., 1980.

[26] Keeny R.L., Raiffa H., 1976.

[27] Duru G. et al., 2002.

### *10.3.3.2) – QALYs and HYEs*

The practical and theoretical difficulties involved in constructing QALYs as utilities have led Canadian authors Birch, Mehrez and Ghafni to put forward an alternative measurement of health states known as Healthy-Years Equivalents (HYEs)[28]. These authors admit the existence of an MAU function on health states U(t,q), but they do not attempt to give it a clearly-defined functional form. On the contrary, they seek rather to summarise the information contained in the function U(.) in a single indicator, the HYE, which they define as follows: if h measures the HYEs associated with the state of health (t,q), h is such that U(h,1) = U(t,q). In other words, HYEs are the number of life-years in good health (q=1) which give the same utility as health state S, for a lifetime t and a uniform quality q. Caution: although there is a strong formal resemblance, this result is not analogous to that of a TTO. Here we are in the theoretical framework of the MAU function under uncertainty and not in that of a "sure" choice. The advantage of this formula, in the eyes of its advocates, is that it does not entail an inversion of preferences. If an agent prefers a state of health (t,q) to another (t',q'), therefore if U(t,q) > U(t',q'), then we will necessarily have h > h', given the properties of the function U. As we have seen, this is not necessarily the case with QALYs.

There are two objections, however: firstly, the empirical measurement of HYEs is difficult and requires a complex double standard gamble mechanism[29] and secondly, the utility function still requires acceptance of the "Neumannian" assumptions which, as we have seen, are not self-evident.

### *10.3.3.3) QALYs and preference aggregation*

A delicate matter – and one often neglected – is the aggregation of QALYs. In empirical studies, the outcome of a treatment is often presented in the form of a total gain in QALYs for a population of patient or, which is the same thing, in the form of an average gain per patient. Can the QALYs of different individuals be aggregated? How legitimate is such an aggregation? Can the QALY for one individual be compared and added to the QALY of another individual? The addition of QALYs seems in theory to pose few problems: it seems quite legitimate to add QALYs, just as it is legitimate to add life-years (two individuals, one surviving 5 years and the other 10 years, "produce" 15 years of survival, giving 7.5 years on average). Given that QALYs include a subjective component, however, it is by no means self-evident that QALYs can be added together. Can the subjectivity of one individual be added to the subjectivity of another? We are familiar with the difficulties of interpersonal comparisons of utility in economics. In practice, the aggregation of QALYs is only legitimate if we admit the additional assumption of a "representative individual" whose preferences are identical to those of the community. This assumption is acceptable, however, if all individuals in the community are identical! This leads us to an important *caveat*: it is dangerous – and not particularly rigorous – to use aggregated QALYs for populations for which we have not first verified that they were homogeneous in respect of their preferences regarding health states.

### *10.3.3.4) The distributive dimension of QALYs*

QALYs also pose a problem of redistribution which has less to do with the QALY as such than with its use in the calculation of cost-effectiveness. The cost per QALY will always be lower, cost for cost, for a younger patient (with a longer life expectancy) than for an elderly patient; it will be lower for a patient in good health (whose quality of life is greater) than for a handicapped patient. Health care programmes aimed at young people in good health will therefore be favoured over programmes at substantially equal costs aimed at populations which are elderly and/or in poor health[30]. Certain authors have suggested to weight the QALYs by a coefficient designed to neutralise these discriminatory effects: the QALY of an elderly or underprivileged individual would be worth more, for example, than the QALY for a young or favoured individual. This option is contrary to what is sometimes known as the "(Allan) Williams principle", which states that "*a QALY is a QALY, no matter who gets it*"[31]. Adopting this distributive, egalitarian ethical principal does not mean overlooking the potentially discriminatory dimension of calculating cost per QALY. It simply reflects, at least in Williams' view, the lack of any clear societal directives on health care priorities[32]. From a certain point of view, the egalitarian ethic derives from the application of the principle of "insufficient reason" for the distribution of the health benefits. It is nonetheless important to bear in mind the possible redistributive effects when comparing programmes relating to populations with different socio-demographic characteristics.

---

[28] Mehrez A., Ghafni A., 1989.
[29] Mehrez A., Ghafni A., 1991.
[30] There is, however, a discriminatory effect of the same order if a less controversial indicator is used, such as "cost per year of life saved".
[31] Williams A., 1992.
[32] "At present I am sticking uneasily to a QALY is a QALY is a QALY, pending further clarification of just how much extra weight is to be given to the favoured categories over the unfavoured ones," A. Williams, op.cit., 1992.

### *10.3.3.5) QALYs and utilitarianism*

A final point which, if we were to address it in full, would take us far beyond the framework of these recommendations, concerns the ethical debate over the use of QALYs. Is not the use of QALYs associated with a utilitarian approach contrary to the professional code of practice on which medical ethics are based? Is the reduction of a human life to a number not morally inacceptable? Such is the stance taken by certain opponents of QALYs, the best-known of whom is the philosopher John Harris[33]. In his view, rather than resorting to an immoral calculation of QALYs, random selection would be the only procedure compatible with medical ethics if a scarce resource had to be allocated to certain patients rather than to others. But as Allan Williams rightly points out[34], this fundamentalist criticism actually condemns not just the QALY but any approach attempting to establish collective priorities in health care choices. The lottery solution only shifts the problem from choosing which patients are to be treated to choosing the procedure used to decide which patients are to be treated! And few people in society would consider a lottery as a morally "fair" procedure. Finally, it seems that there is a risk of "over-interpreting" QALYs by turning them into a measure of collective well-being in the field of health. If, to paraphrase Williams, the use of QALYs is clearly "consequentialist" (in that it stresses the impact of health care on the health of patients measured in quantity and quality of life), it is not necessarily rooted in a utilitarian philosophy the objective of which is simply to maximise health gains.

Using the QALY as a local indicator for the allocation of health care resources is undoubtedly not incompatible with philosophical options other than utilitarianism for the definition of social justice in health care, options which at the overall level would, for example, stress the equity of the system. If the criticism of QALYs is to be legitimate, it must be directed at what they really are (an indicator for resource allocation built on a fragile theoretical construction and with distributional aspects that are locally debatable), not at what they can't claim to be (a utilitarian measurement of social justice in health care).

**Recommendation 25**

The QALYs approach consists in aggregating into a single dimension the two dimensions that describe the outcomes of a health care programme in terms of life years gained and quality of life.
The proposed aggregation rule poses many problems in terms of both methodology and philosophy. The limited robustness of this approach allows the manipulation of the conclusions of a study.

In view of this, the readers or users of a study setting out outcomes in terms of QALYs should pay close attention to the following rules:

a) It is essential to verify that each study presents the reference standards used to measure quality of life, the measurement method used, and that the specification of the multiplying aggregation function used for the reference standards has been validated.

b) It is equally important to enquire into the origins of the quality of life measurements, and particularly whether these measurements are "psychometric" or derived from techniques of preference identification such as TTO or standard gamble methods.

c) When the QALY versions used are based on utility or expected utility theory, the reader should bear in mind issues such as whether the behavioural assumptions have been verified, the choice of methodology and epistemology, and should seek to identify any bias these choices might induce.

d) It is important in all circumstances to be cautious of any attempts at inter-studies and inter-pathologies comparisons (e.g. league tables). It is important to ensure, for example, that the same systems of reference have been used in all the different cases. Nor is it recommended that the valuation of health states from one study be used for another, without careful validation to ensure that the health states are indeed equivalent.

e) Attention should also be paid to the risk of rogue distribution effects when comparing populations with very different socio-economic characteristics or where preferences for the treatments analysed may be very heterogeneous.

f) In the current state of research, it is not recommended that public health decisions should be based on study results expressed in terms of QALYs unless conditions a) to e) above are verified and/or validated, given the possibility of arriving at divergent results from the same observed data.

[33] Harris J., 1987.
[34] Williams A., 1996.

## 10.4) CONTINGENT VALUATION: WILLINGNESS TO PAY

Contingent valuation is used to determine an individual's willingness to pay for a health-related good, by placing the individual in the hypothetical situation of a health care market. This is the method that is the most consistent with the assumptions of the theory of well-being. By answering precise questions (see below), the individual reveals the maximum amount he is willing to pay in order to obtain a certain benefit. This approach makes it possible to incorporate individual preferences into cost-benefit analyses and makes it possible to express, in monetary terms, an overall judgement on a health state, including in terms of quality of life[11].

Various methods are possible, for example the auction mechanism and the payment card mechanism.
The auction mechanism operates as follows: the individual is offered the opportunity to purchase a good at a price set arbitrarily. If the individual accepts, the price is gradually increased up to the point at which the individual refuses to pay. This is more expensive that the mechanism which involves asking the individual to choose directly between several different amounts (the payment card mechanism), but is probably less biased.

Possible biases may be related to the conception of the scenario setting out the health states, the non-neutrality of the monetary values proposed as reference points and, above all, the unreality of the market situation which the individual is meant to suppose.

The contingent valuation methods in the field of economic evaluation of health care programmes are still at the experimental development stage, which means they cannot be used as tools to help public decision-making.

**Recommendation 26**

The use of contingent evaluation in the health care field is still at an the experimental stage, which means that the limitations of the technique must be clearly set out and the potential for bias examined.

# XI) DISCOUNTING COSTS AND BENEFITS

## 11.1) THE PRINCIPLE OF DISCOUNTING

The outcomes and costs of health care intervention protocols may sometimes spread over time horizons of varying lengths. The question of discounting involves asking what an outcome or cost at date $t_1$ represents, relative to the same outcome (in physical terms or in constant euros) or cost (in physical terms or constant euros) today at $t_0$. If we answer this question by saying that there is an equivalence between these elements irrespective of the date at which they occur, then no discounting is required (or, if preferred, that the discount rate is zero). In economics, particularly when comparing costs or outcomes expressed in monetary terms, it is common practice to discount on the basis that decision-makers prefer present situations (i.e. using a positive discount rate).

Thus an income of S euros at date $t_1$ (annual period immediately following period $t_0$) will have a "present value" PV = $S/(1+a)$ if a is the discount rate. More commonly, if the income S is obtained only k years later (period $t_1$), the present value becomes PV = $S/(1+a)^k$.

Clearly, discounting subsequently makes it easier to compare and aggregate outcomes and effects spread over time. Conversely, aggregating effects or costs of a similar nature without any particular precautions is equivalent to adopting a discount rate of zero.

## 11.2) SHOULD HEALTH CARE OUTCOMES BE DISCOUNTED?

The issue of discounting health care costs and benefits lies at the heart of a sustained debate which has yet to be – or is perhaps never to be – decided, between those economists who argue that costs and benefits (expressed in monetary or physical terms) must be discounted to reflect the preference for the present, and those who consider that it is impossible to discount benefits expressed in non-monetary terms, due to the non-transferability of outcomes in some health care programmes that have irreversible effects (death).

Discounting outcomes is only in fact meaningful if the decision-maker really is in a position to renounce to a situation x at a date t to choose for a situation y at date t + dt. In the health care context, for example, the decision-maker must be in a position to give up the option "an immediate life-year" in favour of "a life-year in five years". Clearly, this is meaningless at the level of the individual.

From a collective point of view, the postulate of transferability may be justified, including in the event of irreversible outcomes. This supposes that the decision-maker is in a position to evaluate the health state of a population from health states of the individuals composing the population. This brings us back to the more general question of the justification for the aggregation of "units of result". This problem of inter-individual comparisons health states or, in the wider context, of utilities, is not specific to discounting but it is coupled with the introduction of the time dimension into the reasoning which may, when choices with long-term repercussions are being made, lead to choices being made between generations (see below). Adopting the principle of discounting outcomes means that the decision-maker is in a position to exchange "a certain quantity of life-years for a population X at a given date t", for "a certain quantity of life-years for a population X' at a given date t'". This approach clearly implies value judgements which must be presented and discussed. This difficulty is particularly true for the evaluation of prevention programs. The discounting of outcomes has a tendency to disadvantage prevention strategies, which call for an investment of resources today in order to obtain some benefits in a remote future. Failure to discount outcomes, however, may produce results which are debatable on grounds of ethics or equity when the time horizon of the economic evaluation is long-term, as it is in the case of studies on the importance of vaccination in eradicating a disease, or when it comes to evaluating the aggregate cost of the consequences of handicap when life expectancy is high.

## 11.3) SHOULD THE SAME DISCOUNT RATE BE APPLIED TO HEALTH CARE COSTS AND OUTCOMES?

If, adopting a collective perspective and accepting the premises of the theory of well-being, we envisage discounting costs and benefits, the question still remains as to what discount rate to apply. Is the rate the same for both the costs and the health outcomes? Three arguments may put forward for justifying the choice of different discount rates:

- the decision-maker's preference for the present differs according to whether he is considering the costs or outcomes of health care;
- the tools used to measure the utility attached to health states take into account the decision-maker's preference for the present. As a general rule, if the benefits are expressed in terms of utility or willingness to pay, the benefits must be discounted at a lower rate than that applied to monetary elements (or even at zero), in so far as preference revelation methods for series of health states already include the preference for the present;
- the preference for health over other economic goods changes over time (see Appendix VIII).

## 11.4) WHAT DISCOUNT RATES SHOULD BE USED?

As far as the choice of discount rate is concerned, economic theory enables us to identify processes for determining the rate i in relation to assumptions as to the rationality of the decision-maker and the economic environment in which the decision is taken. The rate may happen to coincide with the real monetary interest rate, but this is not necessarily the case, particularly when financial markets are imperfect. One possible reference in France is the discount rate used by the authorities when deciding on public investments[35]. For information, Canadian guidelines opt for discounting both costs and benefits (expressed in either monetary or physical terms) at the same rate of 5%, while allowing researchers to carry out a sensitivity analysis on the rate applied to benefits at rates below 5% (0% or 3%). Other international recommendations discount outcomes at rates varying from 1.5% to 6% with a sensitivity analysis including the 0% rate.

**Recommendation 27**

In the case of a collective choice, discounting should be applied to:
- costs
- health outcomes, if the method used for measuring benefits does not already take into account the preference for the present.

It is recommended that costs and outcomes should be discounted at the same rate, except if studies suggest that the preference for the present in health outcomes and in costs is different.

For the purposes of international comparisons the rates used should be 0%, 3% and 5%. A sensitivity analysis should be carried out systematically on discount rates in order to test the robustness of the conclusions (including a 0% rate, i.e. the evaluation without discounting).

[35] For example, see reports by the Commissariat Général du Plan [The French Planning Office], such as: Boiteux M. & Baumstark L. (2001): choix des investissements et coût des nuisances, Rapport pour le Commissariat Général du Plan, La Documentation Française, June 2001, 325 p. Available free of charge in electronic format from the website of La Documentation Française : http://www.ladocumentationfrancaise.fr/brp/notices/014000434.shtml

# XII) MODELLING

A model is an idealised and formalised representation of a real phenomenon. It is an analytical technique capable of simulating the impact of one or more factors (inputs into the model) on the body of expected results (outputs of the model).
All models, including observations governed by protocols (clinical trials), are a representation of reality.

## 12.1) FIELDS OF APPLICATION

Models are widely used in such varied fields as demography, epidemiology or even finance. The choice of model depends on the question posed. For the economic evaluation of therapeutic strategies, models have many fields of application.

- The information provided by randomised clinical trials are often insufficient to judge of the clinical and economic benefits of the strategy under evaluation, in actual use. Modelling helps to overcome certain of these shortcomings, particularly by exploiting other existing sources of data.
- Modelling makes it possible to extrapolate clinical and economic results beyond the time horizon of the clinical trial. It can also be used to transpose the results of an international study to the French context or to provide estimates for hypothetical populations (sub-groups of patients at risk, etc.).
- Modelling makes it possible to link results based on intermediate or surrogate end-points provided by clinical trials to the expected final end-points derived from epidemiological studies; for example, lower blood pressure (surrogate end-point) measured during a clinical trial and the expected reduction in cardiovascular mortality (final end-point) in the population likely to benefit from the treatment.
- Modelling can be used to compare the strategy being studied to therapeutic strategies other than those selected in the clinical trial, offering the opportunity, for example, of comparing one or more therapeutic strategies within a single model, as long as clinical data is available for the population concerned.
- Modelling helps to clarify the consequences of decision-making processes when a new therapeutic strategy is introduced. This is the case, for example, with budgetary impact studies (see chapter 16).

## 12.2) MODELLING TECHNIQUES

There exists many modelling techniques. Those most frequently encountered in the field of health care are decision trees, in particular the Markov models used for chronic diseases requiring long-term evaluation. A distinction is generally drawn in modelling between deterministic and stochastic approaches.

In a deterministic approach, the data used as the basis for modelling consists of estimates, generally drawn from published studies, referred to as "secondary data".

- A decision tree is a graphic representation of the consequences of a therapeutic decision. Probability nodes, or chance nodes, are used to incorporate the frequency of expected events: rates of success, failure, cure, relapse, etc. At the end of this series of probabilities, one or more values may be associated with the final result obtained. This may be an effectiveness measure, a utility and/or a monetary cost.
- Markov models are graphic constructions similar to the decision tree, but they incorporate the concept of time in the form of a cycle. This is integrated in the form of a recursive process known as a Markov chain. The process is defined by five elements: a set of dates, states, actions, transition probabilities and "rewards". By convention, an absorbent state must be included: a patient entering this state cannot leave it. Schematically, the Markov model is used to simulate the path taken by individuals through a finite number of states and to thus accumulate the "rewards" that arise from taking this path: life-years saved, for example, and QALYs but also costs. Without going into the mathematical detail of this type of model, their fundamental property is that they are based on "no memory" processes. Thus the number of patients present in state A at time t depends solely on the number of patients present in state A at time t-1 and the probability of entering (or leaving) state A at time t. This simplification represents a serious limitation when modelling diseases with consequences which, for a given patient, may evolve over time (for example, when it comes to incorporating the date of the diagnosis). It is then possible to incorporate a “memory” into the model, and the result is known as a semi-Markov process. For both decision trees and Markov models, the issue of the inclusion of uncertainty tainting results remains a major methodological problem.
- The Monte Carlo method involves simulating a theoretical cohort of patients, each of whom follows a path among the possible consequences of the therapeutic action. Each patient's path is conditioned by the probabilities associated with the chance nodes. Ideally, the distribution functions of the probability nodes should be included in the model, as should the possible correlations between the probabilities of the different chance nodes.

The stochastic approach differs from the deterministic approach in that it is based on the observation of individual or raw data. The data may be derived from a clinical trial, from hospital data, or from medical consumption files. The modelling techniques employed are based on inferential estimates. Uncertainty is approached in a totally different manner, through observation of the statistical distribution of model variables.

**Recommendation 28**

Modelling covers a series of clearly-defined techniques whose application to the economic evaluation of therapeutic strategies has been validated. It can be used to evaluate both costs and outcomes.

A model is a decision aid which can be used to aggregate data from different sources. It is used to systematise decision-making elements, particularly by allowing for long-term estimates, indirect comparisons, evaluation of the budgetary impact of a health care strategy on the management of a disease and the extrapolation of the results from a survey to the general population.

The choice of the type of model to be used must be justified and adapted to the issue being studied.

## 12.3) CONDITIONS OF USE

The data and assumptions used in developing the model must be stated, referenced and listed in tables. For each of the alternatives studied in the model, the results for costs and benefits will be presented separately.

Each resource unit will first listed in a summary table in terms of level and value (in euros) before being aggregated. Direct costs will be shown separately from indirect costs. Discounted and non-discounted results will also be presented separately.

The reliability of the models is subtended by a dual process of internal and external validation. Internal validation is based on the relevance of the assumptions and of the data which must be very well documented (published and/or referenced data) and on the consistency of the structure of the model. External validity is based on how appropriate the dynamics simulated in the model are close to those observed in real life.

Because of the many assumptions used in the construction of models, it is important to use sensitivity analysis to test the robustness of the evaluation results with regard to variations in the key parameters used in the model. The sensitivity analysis may be univariate or multivariate. The choice of parameters to be varied must be justified, as must the interval through which variables are varied from the baseline.
If discounting is applied, the robustness of conclusions to the application of the various discount rates (including a zero rate) may also be tested.

**Recommendation 29**

The interpretation of modelling results must remain linked to the assumptions underlying the model's construction (given population, etc.). All the data and assumptions used in the model must therefore be clearly stated, documented and validated.
The reliability of the models is subtended by a dual process of internal and external validation.
The robustness of results must be tested by a sensitivity analysis. The choice of parameters to be varied must be presented/justified, as must the intervals inside which values are varied.

# XIII) VERIFYING THE ROBUSTNESS OF STUDY CONCLUSIONS

Economic evaluation is an approach performed under conditions of uncertainty. Indeed, it is this which makes decision-making so difficult.

The origin of the uncertainty differs according to whether the model is based on assumptions of the value of certain parameters (e.g. the discount rate) arising from different sources (known as deterministic analysis) or whether a cost-outcome calculation is performed using individual data on effectiveness and/or resource consumption derived, for example, from a clinical trial (known as stochastic analysis). In the first instance, the uncertainty relates primarily to the assumptions and in the second to sampling. These uncertainties may, of course, be cumulative, for example in the case of a probabilistic model based on data from a clinical trial. They should nonetheless be dealt with separately.

## 13.1) SENSITIVITY ANALYSIS OF MODELS

The procedure most commonly used to evaluate the impact of assumption uncertainty on the final decision is a sensitivity analysis and the observation of its repercussions on choice of the strategies.

Sensitivity analysis consists of three stages:

- the identification of uncertain assumptions and parameters for which sensitivity analysis is required;
- the choice of a plausible assumption of variation of uncertain factors;
- the presentation of the results of the sensitivity analysis (the value of the result against that of the parameters varied).

A distinction is made between univariate and multivariate analysis, and also between first order and second order analysis.

In a univariate sensitivity analysis, each uncertain parameter is varied separately while the others remain at their baseline value. The goal is to study the influence of the parameter on the final result in such a way as to determine whether the result is sensitive to variations in the parameter. The greater the sensitivity of the results, the greater the caution required in interpreting the results. Successive univariate sensitivity analyses on all the uncertain parameters – as long as each is indeed independent of the others – may be enough to assess the robustness of the results of the model.

A multivariate sensitivity analysis simultaneously varies two (or more) parameters in order to study the combined effect of these parameters on the results of the analysis. The greater the number of parameters, however, the harder it becomes to represent the results. To overcome this difficulty, the multivariate analyses may be presented in the form of scenarios. In this way it is possible to explore the implications for results of different states, these latter being defined by the pivot values of the parameters. Extreme scenarios, corresponding to the best-case and worst-case situations, for example, can thus be considered. A sensitivity analysis in terms of threshold could then be envisaged. This consists of estimating threshold values for parameters, above or below which the conclusions of the analysis change. This analysis is particularly useful when the baseline value of a parameter is unknown.

Another choice to be made is that of the mode of variation of parameters. A first possibility is to take different plausible values (for example, 0%, 3%, 5% and 7% for the discount rate) and to observe the impact on the result. A second possibility is to introduce, not different values for a parameter, but instead the law of probability that this parameter probably obeys (for example a normal distribution with mean m and standard deviation s).

Second-order Monte Carlo simulation techniques will then be used to obtain the distribution of the results, which will then depend on the type of distribution followed by the parameters. Several distributions, corresponding to several parameters, may thus be studied simultaneously; the Monte Carlo simulation then giving the results in a probabilistic form (for example, a mean and a standard deviation for the cost-effectiveness ratio). It will then be possible to study the properties of the statistical distribution of the results and construct a "confidence interval". In addition, this probabilistic sensitivity analysis enables us to estimate the relative probability of certain situations (e.g. the probability of strict dominance of one strategy over another).

**Recommendation 30**

In the case of studies using data from different sources, a sensitivity analysis must be carried out on the variables and assumptions used. These must be clearly identified and justified, as well as the intervals inside which the are varied.

In the case of a complex model, an analysis in terms of best case and worst case scenarios may be carried out.

## 13.2) TREATMENT OF UNCERTAINTY IN STOCHASTIC ANALYSES

When a "sampling uncertainty" exists, i.e. when the economic calculation uses individual data from a clinical trial or an observational study rather than determining results on the basis of the mean of the different variables, it is preferable to use all the information contained in the individual data.

Recent literature in the field of health economics proposes several methods for estimating the confidence interval of the cost-outcome ratio in these circumstances:

- parametric approaches: the "box" method, the Taylor series method, the confidence ellipse method, the Fieller method;
- non-parametric bootstrap approach with its different estimates of the confidence interval (normal approximation, confidence interval by percentile, percentile interval, corrected and/or accelerated).

These methods produce very different intervals. An examination of their statistical properties suggests that they may only be used subject to certain assumptions specific to each method. Several authors have tested the properties of the estimators produced by these methods, using Monte Carlo simulations: the parametric method based on Fieller's theorem and the non-parametric basic bootstrap approach produce the best results in terms of coverage probability (probability that the parameter of the population will be inside the confidence interval).

It is frequently interesting to plot the acceptability curve, i.e. the curve representing the probability that the strategy being studied is cost-effective according to different ceiling values of the cost-outcome ratio. Note that this curve does not suffer from the problems that may be encountered with confidence intervals, i.e. the interpretation of a negative cost-outcome ratio. Another advantage is that it instantly highlights the probability that the strategy being studied is not cost-effective.

**Recommendation 31**

When carrying out stochastic analyses where data on the cost and effectiveness of treatments for each patient are available, confidence intervals on the cost-effectiveness result must be used in order to take sampling errors into account. Preference will be given to the parametric method based on Fieller's theorem and the non-parametric bootstrap approach. The representative nature of the data used will be assessed prior to the analysis. An alternative and more informative approach to take uncertainty into account is to present the acceptability curve of the strategies which represents the probability that the strategy in question is cost-effective for different ceiling values of the cost-effectiveness ratio, ranging from 0 to infinity.

# XIV) PRESENTATION OF THE RESULTS

## 14.1) A STANDARD FORMAT FOR GREATER TRANSPARENCY

As far as possible, a study report should provide the reader with information on the following points:
- the issues at stake and the reasons for the study;
- the type of study chosen and the reasons for that choice;
- the population concerned;
- the reference medical situation (comparator(s), for example);
- the medical data used to evaluate health impacts;
- the origin of the economic data;
- the different types of costs taken into account;
- the perspective adopted for the calculation of costs;
- the method of dealing with time (discounting);
- the method of dealing with uncertainty (sensitivity analysis);
- the consistency or inconsistency of results with other studies on the same subject;
- factors which must be taken into consideration in order to assess the scope of the study (and particularly its limitations in terms of time, population, information, etc.);
- the dates at which the study was carried out;
- the period for which the results are valid;
- relevant bibliographical references;
- the identity of any source(s) of funds;
- compliance with current regulations regarding the confidentiality of data collected, data held on computer, and with current national and European regulations on biomedical research.

**Recommendation 32**

As a general rule, the results of an economic evaluation study should be presented in such a way as to allow peer examination and review. The report must contain all the elements that will enable an informed reader to understand the methodology adopted, check the sources and verify the relevance and accuracy of calculations. All data, references and calculation procedures used in the study must be available to any interested party on request, if not explicitly included in the study. The only limit to this principle is respect of the confidentiality of data on identifiable individuals. Although it is perfectly legitimate for authors to comment freely on their results in the conclusion or the discussion section, they should be careful to avoid any interpretation of these results that might seem exaggerated in the light of reality.

## 14.2) SCIENTIFIC INTEGRITY AND CONFLICTS OF INTEREST

**Recommendation 33**

The publication of the results of economic evaluation study in the field of health care may have important economic or regulatory repercussions. It may influence a public decision, founded on the quality of the scientific study. It is therefore important that the process of publication and dissemination should meet the criteria of integrity applied by the scientific community. It is important, for example, that all the authors have played an effective part in the scientific part of the study. Each signatory is personally responsible for the veracity of the study and for ensuring that the methodological approach followed was indeed as described in the study. Authors' institutional affiliations must be stated and the source(s) of funding, whether public or private, must be disclosed.

## 14.3) RATIOS

Cost-effectiveness (or cost-utility) ratios are intended to express in summary form the results of a comparative evaluation of different treatments. This evaluation implies that the ratios be presented in incremental form, with the difference in costs between two treatments being divided by the difference in effect produced by the same treatments. It is important that the content of both numerator and denominator be set out clearly, in particular to avoid double counting (i.e. the same phenomenon appearing in both numerator and denominator). With this end in view, for example, when the ratio expresses the cost of a life-year saved, the author will avoid valuing life-years lost under the heading of costs. For the interpretation of these ratios, readers are strongly recommended to read the comments in paragraph 15.2.

**Recommendation 34**

Cost-effectiveness (or cost-utility ratios) must be expressed in incremental form so that they may be used as a criterion to assist in decision-making. Variations in the use of resources should be included in the numerator and those influencing health states in the denominator, taking care to avoid double counting. The components of these ratios (costs and results) should also be presented with their statistical distribution (mean, median, confidence interval, etc.).

# XV) INTERPRETATION OF THE RESULTS

## 15.1) DISTRIBUTION ASPECTS AND PROBLEMS OF EQUITY

The result of a cost-effectiveness or cost-utility study should never take the form of a single figure, such as the cost of a life-year saved, which results from a mean ratio or marginal ratio.
- If we take the example of hypertension or high blood cholesterol, there are clearly as many "costs" per life-year saved as there are categories of individual differentiated by age, level of hypertension or cholesterol, age of onset, sex, etc. This differentiation, which requires the result of a preventive health care campaign to be expressed by several series of ratios, arises from the wide diversity in the individuals making up society.
- This diversity emerges once again when, after carrying out an evaluation study from the point of view of society, we look at how an overall social gain is distributed among its components ("economic agents" for example, to use the terminology of the Comptabilité Nationale [National Accounts]).

Using the following example of vaccination against an acute disease, we see that the overall gain of 2,000 for society translates, among other consequences, into a loss of 600 for the Assurance Maladie, while the other three agents gain.

*Example: vaccination of the population aged over 59 against an acute disease*

| | Society | Assurance Maladie | Patients | Productive sector (business) | Health care providers |
|---|---|---|---|---|---|
| Cost of vaccination | 0 | - 1,260 | - 540 | 0 | +1,800 |
| Cost of care | 0 | + 560 | + 240 | 0 | - 800 |
| avoided | +2,000 | +100 | + 400 | + 1,500 | 0 |
| Indirect costs avoided | | | | | |
| Total | + 2,000 | - 600 | + 100 | + 1,500 | + 1000 |

+ = gains
- = losses

Similarly, the decision-maker will need to take the sociological differentiation into account in the outcome of a programme (or campaign).

Is it equitable to concentrate the benefits of the programme on a small number of individuals, or should we choose another programme which, for an equal cost-benefit ratio, will benefit a greater number of people (the benefit per individual being consequently reduced)?

The question may be illustrated by the following example:

| | Programme A | Programme B |
|---|---|---|
| Cost | 2,500,000 F | 2,500,000 F |
| Effectiveness | 100 | 100 |
| C/E | 25,000 F | 25,000 F |
| Number of individuals who benefit from the program | 10 | 50 |
| Number of life-years gained per individual | 10 | 2 |

**Recommendation 35**

On completion of a study, the distributive aspects of the result should be considered by differentiating it by as many factors as may seem pertinent (socio-demographic factors in particular) and which can raise equity problem by favouring certain groups or agents at the expense of others.

## 15.2) THEORETICAL PRINCIPLES OF A DECISION BASED ON AN ECONOMIC EVALUATION STUDY

How in practice are public policy recommendations formulated on the basis of economic evaluation of health care?

1. In the first instance, it is important to remember that the quantitative results of the economic evaluation cannot, in themselves, mechanically determine public decision, for three main reasons:

a) The economic evaluation is never complete. However sophisticated, it cannot claim to be exhaustive in taking into account all the cost and benefit elements of each of the alternatives analysed. Besides, the economic measure cannot be reduced to a simple mechanical operation of identifying and valuing accounting items, even if this is how it may sometimes appear. It is a scientific approach, which therefore makes reasoned methodological choices that must be clearly identified, discussed and brought to the attention of the decision-maker – hence these guidelines. The decision-maker is the ultimate judge of their relevance.

b) Distributional aspects are rarely taken into account in the economic measure, which is more interested in the allocation dimension of public policy. The "right" decision, that which is recommended by the results of the evaluation, may well benefit certain agents in society and penalise others. In general, it is the role of the political powers to arbitrate between these divergent interests, at least in representative democracies. These powers may well decide that one Franc of spending by the Assurance Maladie is not equivalent to a Franc of income of a health care provider. The calculation may, in certain instances, introduce weighting systems to take into account for redistributive aspects. It is, however, the political powers which have the ultimate legitimacy to validate this calculation.

c) Finally, the evaluations are generally performed without reference to the institutional and political context, even if a particular "viewpoint" is adopted for the calculation of costs. At no point is the economic evaluation in a position to transcend the institutions of public decision-making. On the contrary, it is structurally designed to serve them, whether the studies have been commissioned by those institutions or are produced by third parties in order to influence their decisions. Economic evaluation is not an institution: it is merely an instrument designed to shed light on the logic behind choices, to bring greater rationality to often complex institutional workings.

Despite these reserves, it is nonetheless important to clarify the actual meaning of the recommendations that may be made on the basis of cost-outcome studies.

2. The first comment relates to the type of analysis. Cost-benefit analysis, which measures costs and benefits in monetary terms, is the only form of analysis which can be used to make recommendations that are wholly unambiguous in their interpretation. A health care action is recommended, subject to the above reserves, if "*the total social benefits exceed the total social costs by a sufficient percentage*"[67]. As everyone is aware, however, this cost-benefit criterion remains extremely delicate to apply in practice, particularly in view of the uncertainty associated with the economic evaluation of the effects of health care (life-years gained, for example) and, more generally, of effects which cannot be evaluated in monetary terms on a market (time, pain, etc.).

3. The cost-effectiveness criterion, whose result is expressed in the form of a ratio such as "*cost per life-year saved*", avoids these problems. In return, it is more difficult to interpret in that it is purely a technical ratio measuring the "productivity" of a health care programme and, as such, implies nothing as regards the political and social "desirability" of the programme. This point must be stressed: the cost-effectiveness ratio is not an indicator of collective preference, but an indicator of the technical performance obtained from a good, a policy or an institution. The community may "prefer" to choose the most technically efficient programme, but this is in no way an obligation. It would not be difficult to find examples of a community "preferring" a health care programme that costs 12 million euros but which yields only 1,000 life-years (therefore 12,000 euros per life-year), over an alternative, more cost-effective which would cost 120,000 euros for a benefit of 15 life-years (therefore a ratio of 8,000 euros per life-year). This choice is not necessarily irrational. It is simply the result of the fact that the decision-maker may find that the absolute benefit of the most profitable solution is too small[68].

4. Allocating resources according to the criterion of cost-effectiveness achieves maximum potential benefit under a given budget constraint. Any reallocation of budget from one technique to another which is more cost-effective increases the health benefits for the same cost. In practice, however, these possible re-allocations are local: it is a matter, for example, of choosing the most cost-effective treatment for a given disease and for a given budget. In many cases, the choice is even limited to two options (a health care strategy and its comparator). It is much rarer for

[67] Mishan E.J. 1976.

[68] In formal terms, if we consider the cost of alternative health care programmes characterised by their cost, C, and their effectiveness in terms of lives saved, V, and if the decision-maker's preferences are expressed by the utility function U(C,V) which depends positively on U and negatively on C, the maximisation of this function over all the possible programmes does not necessarily lead to the choice of the most cost-effective programme. The choice of this programme implies a highly particular form of this utility function.

analyses to address more general re-allocations, such as may occur between two medical fields, or even between the health sector and other social sectors such as education, environment, etc.

5. As far as the cost-effectiveness ratio itself is concerned, several situations must be considered depending on the nature of the projects and the severity of the budget constraint. We will make the distinction between the case where several projects may be implemented independently from one another (for example, defining a breast-cancer screening plan involving several measures) and the case where the problem is to choose one programme from a number of mutually-exclusive options (for example, the choice between two treatments for the same disease).

In both instances, we will base our arguments below on a "local allocation of resources" as defined above: we will envisage the choices within a predetermined budget for a given objective, without considering other possible uses for this budget. If our object, for example, is to identify the most effective actions for breast-cancer screening, we will not concern ourselves with whether or not this budget could be more effectively allocated to screening for colorectal cancer or for any other public health problem, or indeed to any other public choice outside the realm of health.

6. Where the budget allows for the financing of several programmes contributing towards the same objective, the recommendation involves ranking programmes in descending order of cost-effectiveness ratio and financing those which fall within the budget envelope, starting with the most cost-effective. It can easily be demonstrated that no other method of choice brings a greater collective benefit. In particular, and by a process which is not intuitive for anyone unfamiliar with economic reasoning, the result is superior than would have been obtained by selecting the most effective campaigns in absolute terms, without consideration of cost.

7. Matters become more complex when the choice is to be made between two or more mutually-exclusive programmes. Common sense would then dictate choosing first the programme that is most cost-effective, considered incrementally against a baseline option (possibly the status quo). The solution is not always so simple, however. There may, for example, be a very great difference between the cost and effectiveness of different programmes. Let us return to the example used in §3 and assume that the community has decided to spend the budget of 12 million euros on a health care policy objective. We are in a "local" situation and the choice is between two programmes: one [A1: 12 million euros – 1,000 life-years], the second [A2: 120,000 euros – 15 life-years]. It is not logical, in this instance, to choose A2, even though it is in fact the more cost-effective, because it leaves 11.8 million euros unused for a deficit of 985 life-years. In contrast, it would be logical to rule out the choice of a third programme of the type [A3: 12 million euros – 800 life-years]. It would be equally logical to rule out an option [A4: 250,000 euros – 15 life-years]. Both programmes are effectively strictly dominated by A1 and A2 respectively.

The procedure to be followed, which is an application of the concept of incremental ratio, is as follows:
a. Identify all programmes able to contribute towards the objective established;
b. Eliminate from the analysis any programme dominated by at least one of the other programmes[69];
c. Of the non-dominated programmes (which form the "boundary"), select the one which provides the greatest benefit for the budget allocated.

This procedure does not guarantee that the programme chosen will necessarily be the one with the highest mean cost-effectiveness ratio. It does, however, guarantee that it will produce the maximum health effects within the framework of the budget allocated to the health care policy objective being pursued[70].

**Recommendation 36**

5. If studies lead to recommendations on public policy, it is important to underline that such studies serve merely as decision aids and under no circumstances as imperatives for action.
6. It is important to ensure that the type of recommendation is appropriate to the type of study. Avoid making cost-benefit type judgements (maximising collective well-being) in the context of a cost-effectiveness study, for example.
7. The results of cost-effectiveness studies, particularly those which focus on mutually exclusive programmes (the most common), must emphasise that the choice between the non-dominated options relies to a large extent on a value judgement which the decision-maker may legitimately apply.
8. The value of any reference to external comparators ("*the programme analysed is as cost-effective as another commonly implemented*") is purely indicative rather than prescriptive.

[69] There are two dominance criteria: "strong" dominance when a programme is more costly for the same effectiveness or less effective for the same cost than another programme, and "weak" dominance when a programme is more costly for the same effectiveness or less effective for the same cost than a linear combination of two other programmes.
[70] Our example leads us to prefer programme A1 to A2, although A2 is more cost-effective.

## 15.3) INTERPRETATION OF COST-EFFECTIVENESS AND COST-UTILITY STUDIES: "LEAGUE TABLES"

One of the possible uses of cost-effectiveness studies is to constitute a decision aid in matters of public health policy, acceptance of a treatment for reimbursement by the Assurance Maladie or recommendations on good clinical practice. The results obtained in terms of effectiveness are expressed, in the case of cost-effectiveness studies or more often cost-utility studies, in terms of ratios relating the incremental cost to the incremental benefit associated with the comparison of an innovative procedure against a baseline situation.

Different situations may present themselves, however, as to how these results should be interpreted. This depends, in fact, on the nature of the benefit indicators used.
- When these indicators are clinical or biological criteria specific to a disease (e.g. the number of patients reaching a certain threshold for a biological parameter such as total cholesterol or blood pressure, an intermediate clinical criterion such as MMSE in Alzheimer's disease or a Lequesne score in arthritis), the result obtained may only be compared and hence interpreted in the light of the result that would have been reached by another study using the same methodology in the same therapeutic field.
- When the benefit indicator used is more general and universal in nature, such as the gain in survival, survival adjusted for disability (e.g. the DALY) or QALYs, it becomes possible to examine the effectiveness results obtained in the light of data of the same type already published. Anglo-Saxon economists employ the term "league tables" (Tables used in the UK to present to present the classification of the football clubs in the different divisions) to designate the summary tables built up by compiling the cost-effectiveness ratios quoted in the various published studies.

The first difficulty that arises with the comparison of cost-effectiveness ratios between studies is the lack of detail about the definition and exact content of the policies studied and of the comparators in the results often quoted in league tables. Another and more important difficulty is the lack of comparability between the methods and assumptions used in the various studies. As a way of offsetting these problems, it would be useful if these baseline values were established on the basis of a selection of results offering all the necessary guarantees of homogeneity of methods and assumptions (e.g. same unit costs, relating to the same period and the same health care system, same discount rates, etc.).

Some authors have tried to homogenize *a posteriori* published results in order to arrive at consistent data[71]. Such compilation suggests that the orders of magnitude for economic acceptability thresholds of preventive and therapeutic procedures vary enormously according to whether we are working from the standpoint of treatment or prevention, from that of safety or of environmental risk.

Some international institutions involved in development have suggested that a medical procedure could be considered effective for ratios of a cost per life-year in good health gained that is less than the value of per capita GDP[72]. The corresponding value for European countries would be of 25,000 euros, which in fact corresponds to around half of the threshold values often used by convention in published works based on criteria of this type. The order of magnitude of the threshold often used is around 50,000 euros and it is interesting to note that this value more or less corresponds to the annual cost of outpatient haemodialysis, a costly medical technology vital to the survival of patients suffering from chronic renal failure, but also to the order of magnitude of GDP per number of employments in France.

To sum up, it would probably be useful for such work to be carried out for a future version of these recommendations, in order to produce baseline values in the context of the French health care system. One of the necessary preconditions would be to arrive to a standardisation of methods. A first stage in this standardisation would be, for example, to compile a collection of standard unit costs and explicit rules for calculation (discounting, fixed and variable costs, time horizon, etc.).

Lastly we will note that the above considerations as regards the limits to possible comparisons between the results of different studies apply also, and to an even greater extent, to cost per QALY type ratios. The limits of interpretation of QALYs explained earlier also apply here.

**Recommendation 37**

The possibility of developing recognised reference values for cost-effectiveness type studies based on criteria such as cost/life year gained, adjusted or not for considerations of handicap or quality of life, will require major efforts to standardise methods and parameters, particularly for the calculation of costs. In the absence of such rules, great caution should be observed in any comparison of results from different studies. It is recommended that such comparisons should always be limited to the same therapeutic field and that the comparability of results should be justified as fully as possible.

[71] Tengs et al., 1995.
[72] Bulletin of the WHO, 2000, Vol. 78, no. 2.

## XVI) BUDGETARY IMPACT

Estimating the short and medium term (2 to 3 years) budgetary consequences for the different agents in the health care system (Assurance Maladie, complementary health insurance, households, public hospitals, general and private practitioners, etc.) of the implementation of a new treatment is an important secondary objective of economic evaluations. This type of analysis naturally follows on from the previous types of study but may also be carried out independently.

It is also distinguished in many respects from those other types of study by the methodological choices it supposes:

1. a systematic inclusion of tariffs and actual financial flows (direct and variable costs; indirect costs represented, for example, by the payment of a daily allowance);
2. a short to medium term perspective (2 to 3 years) which assumes that the tariff, regulatory and organisational environment of health care remains constant;
3. a perspective of generalisation of the treatment (i.e. to the entire health care system) including the substitution effects with the current therapies, forecasts of market share in a competitive situation, and costs avoided in the short and medium term given the effectiveness of the treatment;
4. inclusion of the specific characteristics of the populations actually receiving the treatment being studied (target population, population reached, defined patient sub-groups, etc.).

**Recommendation 38**

In addition to economic evaluation studies, which operate at the micro-economic level, it may be useful to extrapolate the results obtained in order to estimate the short and medium-term budgetary impacts for the main agents concerned in case of a generalisation of the treatment. These estimates, which may be as complex or as simple as circumstances require, will have to provide a detailed presentation of the assumptions on which they are based.

# ABBREVIATIONS, ACRONYMS AND DEFINITIONS

## A

**ARH** (Agence Régionales d'Hospitalisation): Regional hospital boards
**Assurance Maladie**: French national health insurance programme

## B

**BADIMEHP** (Bureau de l'Assurance Qualité et de l'Information Médico-Économique de l'Hospitalisation privée): Quality insurance and of the medico-economic information of private hospitals

## C

**CCAM** (Classification Commune des Actes Médicaux): Common classification of medical procedures
**CDAM** (Catalogue des Actes Médicaux): Medical Procedure Catalogue
**CEPS** (Comité Économique des Produits de Santé): Medical products Pricing committee
**CMC** (Catégories majeures cliniques): Major Clinical Categories
**CMD** (Catégories Majeures de Diagnostiques): Major diagnostic categories
**CNAMTS** (Caisse Nationale d'Assurance Maladie des Travailleurs Salariés): National sickness fund for salaried workers
**Comptabilité Nationale**: National Accounts
**Comptes de la Santé**: National Health Accounts

## D

**DGI** (Direction Générale des Impôts): The French tax department
**Direction des Hôpitaux**: Hospital administration Directorate
**Direction des Hôpitaux et de l'Organisation des Soins**: Hospitals and Health care organisation Directorate

## E

**EPHMRA**: European Pharmaceutical Marketing Research Association
**Étude Nationale de Coûts**: National cost survey
**Étude Nationale des Coûts par Activité Médicale**: National cost survey per medical activity

## F

**Fédération Hospitalière Privée**: Private hospital federation
**FSE** (Feuille de soins électronique): Electronic claim form
**FSO** (Frais de salle d'opération): Operating room expenses

## G

**GHJ** (Groupe homogène de journées): Homogeneous groups of days
**GHM** (Groupe homogène de Malades): Homogeneous groups of patients

## I

**ICR** (Indice de Coût Relatif): Relative cost index
**ISA** (Indice Synthétique d'Activité): Casemix index

## N

**NABM** (Nomenclature des Actes de Biologie Médicale): Lab Test Procedure Nomenclature
**NGAP** (Nomenclature Générale des Actes Professionnels): General Nomenclature of medical procedures

## O

**ONDAM** (Objectif national des dépenses d'assurance maladie): National Objective for health insurance expenditures

## P

**PERNNS** (Pôle d'Expertise et de Référence Nationale des Nomenclatures de Santé): National Reference Centre for Medical and Health Classifications
**PMSI** (Programme de Médicalisation des Systèmes d'Information): the French DRG program
**PSPH** (Participant au Service Public Hospitalier): Private not-for profit hospitals incorporated into the public hospital service.

## R

**RHA** (resumés hebdomadaires anonymisés): Anonymised weekly abstracts
**RSS** (Résumés de Sortie Standardisés): Standardised discharge abstracts
**RUM** (Résumés d'Unité Médicale): Hospital Ward Discharge Abstract

## S

**Sécurité Sociale** : Social Security
**SIAM** (Système d'Information de l'Assurance Maladie): Information system of the Health Insurance
**SMUR** (Service Mobile d'Urgence et de Réanimation): Emergency and Resuscitation Mobile Service
**SNIR** (Système National Inter-Régimes) : A system that collects and aggregates data about the activities of health care professionals in the private and public sectors at a national level
**SSR** (Service de Suite et de Réadaptation): Rehabilitation services

## T

**TPJ** (Tarif des Prestations Journalières): Per Diem Rate

## U

**UCANSS** (Union des Caisses Nationales de Sécurité Sociale): National union of Social Security Funds

# *TECHNICAL APPENDICES*[73]

[73] The information contained in these appendices was written up in 2002 and 2003 and may therefore be subject to change after publication of this document.

## APPENDIX I: Concepts and measurements of costs according to the perspective adopted

The cost of a good does not exist. Cost is not an intrinsic characteristic of an object or of a product as temperature might be for air or water. It is a calculation based on a theory and conventions. A cost may only be defined if we make a choice as regards the point of view adopted, the scope of the perimeter to be explored, the object of the study and the time horizon to be adopted.

1 – CHOICE OF VIEWPOINT

Identifying the costs to be taken into account in economic evaluations is largely determined by the areas of responsibility of those commissioning the studies and/or for whom they are carried out. In principle, these fall into five categories: households, the producers of medical goods and services (hospitals, general and private practitioners, manufacturers), purchasers (*Agence Régionales d'Hospitalisation* (ARH), the Assurance Maladie, the *Comité Economique des Produits de Santé* (CEPS)), health care system regulators and society seen as an established political entity.

HOUSEHOLDS

To examine the repercussions of the treatment on household budgets, we calculate the out-of-pocket medical costs in the health care sector, such as actual or de facto co-payments. We will then add to this any expenditure not directly related to the treatment but which frequently accompanies treatment and is not reimbursed by social security. This category includes primarily: the accommodation expenses for long-stay hospitalisation which families have to pay for while the social security system pays for all costs associated with treatment, all the equipment costs incurred by families when they wish patients to be cared for at home, plus the cost of non-medical transportation to visit a family member or close friend in hospital. Although these are not monetary expenditures as such, any evaluation of the burden that falls on family and close friends of the patient as a result of the treatment must also take into account the potentialities sacrificed: absence from work in order to look after a family member or close friend, loss of income for the patient due to the outbreak of the disease. All of these costs make up the household cost.

HEALTH CARE PROVIDERS

*Hospitals*

Hospital expenditure may be classified under four headings:
- the accounting classification is based on a distinction between direct and indirect expenses,
- the PMSI classification distinguishes between medical expenditure, logistics expenditure and structural expenditure,
- the economic classification distinguishes variable costs from fixed costs,
- lastly, expenditures may be grouped according to the purpose of the study. The content of their estimation will be very different depending on whether the costs are being calculated for the purposes of decision-making, of the management of departmental budget differences or for the financing of the hospital structures.

These classifications are closely interlinked. Both direct and indirect expenses include variable and fixed costs. Analysis of the financial repercussions of a decision records new expenditures which arise when the decision is taken, compared to the expenditures that would have been observed had the decision not been taken. Such repercussions affect both direct and indirect costs, variable and fixed expenses. Table 1 gives a better illustration of the different aspects from which hospital spending may be analysed.

Table 1

## BREAKDOWN OF HOSPITAL COSTS

AIM OF THE STUDY

### DIRECT COST

**1. Medical expenditures**

*1.1 Variable medical expenditures*

- Consumables, individualised or otherwise
- Medico-technical procedures

*1.2 Fixed medical expenditures*

- Salaries (medical, nursing, auxiliary)
- Medical equipment (depreciation and maintenance)

**2. Logistics expenditures**

*2.1 Variable logistics expenses*

- Laundry
- Catering

**SUB-TOTAL 1**: DIRECT COST = 1.1 + 1.2 + 2.1

**Choice of a strategy**

**Expenditure control**

**Tarific ation**

### INDIRECT COST

*2.2 Fixed logistics expenses*

- Other (admin., IT, maintenance)
- Administrative staff

**SUB-TOTAL 2:** QUASI COMPLETE COST = 1.1 + 1.2 + 2.1 + 2.2

**3. Structural expenditures**

- Finance costs
- Depreciation

**SUB-TOTAL 3**: INDIRECT COST
(Fixed logistics expenses + structural expenses)

### COMPLETE COST

DIRECT COST (SUB-TOTAL 1) + INDIRECT COST (SUB-TOTAL 3)

QUASI COMPLETE COST: SUB-TOTAL 2 + Structural expenses

The framework of costs to be taken into account in any economic evaluation must correspond to the budgetary concerns of the party embarking on a health care project. Hospital doctors and pharmacists are faced with both medical and economic choices, particularly at a time when innovative and effective technologies are arriving on the market but are considered costly. Adopting their standpoint, we will concern ourselves only with the costs directly linked to the implementation of these new protocols, irrespective of the administrative sections which record the costs (costs of a therapeutic innovation). Finance managers cannot ignore the repercussions of these clinical choices on departmental budgets (costs directly associated with the running of the department or incurred as a result of its activity). At this level, however, it would be inappropriate to include overheads and structural costs since neither doctors nor pharmacists can be held responsible for the evolution of these costs. As for hospital managers, they always pay great attention to balancing the hospital's books (complete costs) to the positive impact to be gained from faster patient turnover in terms of ISA points when one technology is used in preference to another, and from income, which may be described as attenuating in that it partially offsets the extra costs associated with the innovation.

Under no circumstances should the *Tarif des Prestations Journalières* (TPJ) [Per Diem Rate] be used: this cost valuation unit is purely for accounting purposes. It simply enables funding from the global budget to be distributed to the different social security schemes on the basis of the number and length of stays in the hospital by their contributors. These are crude tariffs covering all departments, all patients and all conditions without distinction, within the 8 medical pricing categories to which they have been assigned. They should not be used in health economic evaluation studies. Cost/GHM is the only approach which reflects the real cost of treatment.

Three avenues of approach are therefore proposed for analysing the structure of hospital costs. Initially we will isolate all the financial repercussions of the introduction of a new therapeutic protocol for the treatment of a disease. Next we will combine these expenses with the fixed costs of clinical departments and the overheads for the institution which are included in the GHM corresponding to the indication, in order to evaluate the real cost of patient management. Lastly, the consequences of the therapeutic innovation for the institution in terms of ISA points will be quantified.

- *Costs of the implementation of a therapeutic protocol*

The analysis of the financial repercussions of a new therapeutic protocol records the new expenditures which arise after adoption of the new approach, compared to the expenditure that would have been incurred had the new treatment not been adopted. Expenses not directly linked to the introduction of this protocol should not affect the decision. It is on this point that the evaluation of the cost of the implementation of a new technology differs fundamentally from the surveillance of department budgets or the search for the "right" price that balances the hospital's accounts. The cost is used as a management tool and not as an instrument for control or pricing. It allows us to estimate whether the results obtained justify the scale of the efforts deployed.

The cost which can be directly related to the implementation of a particular treatment is made up solely of expenses which can be assigned to the treatment without special agreement or calculation. This definition makes no assumption as to how far the scope of the innovation extends upstream or downstream of the department in which it appeared, nor of the laws of variation of expenditures (variable cost and fixed cost), nor of the time horizon adopted (immediate consequences of the treatment administered or remote repercussions due to complications).

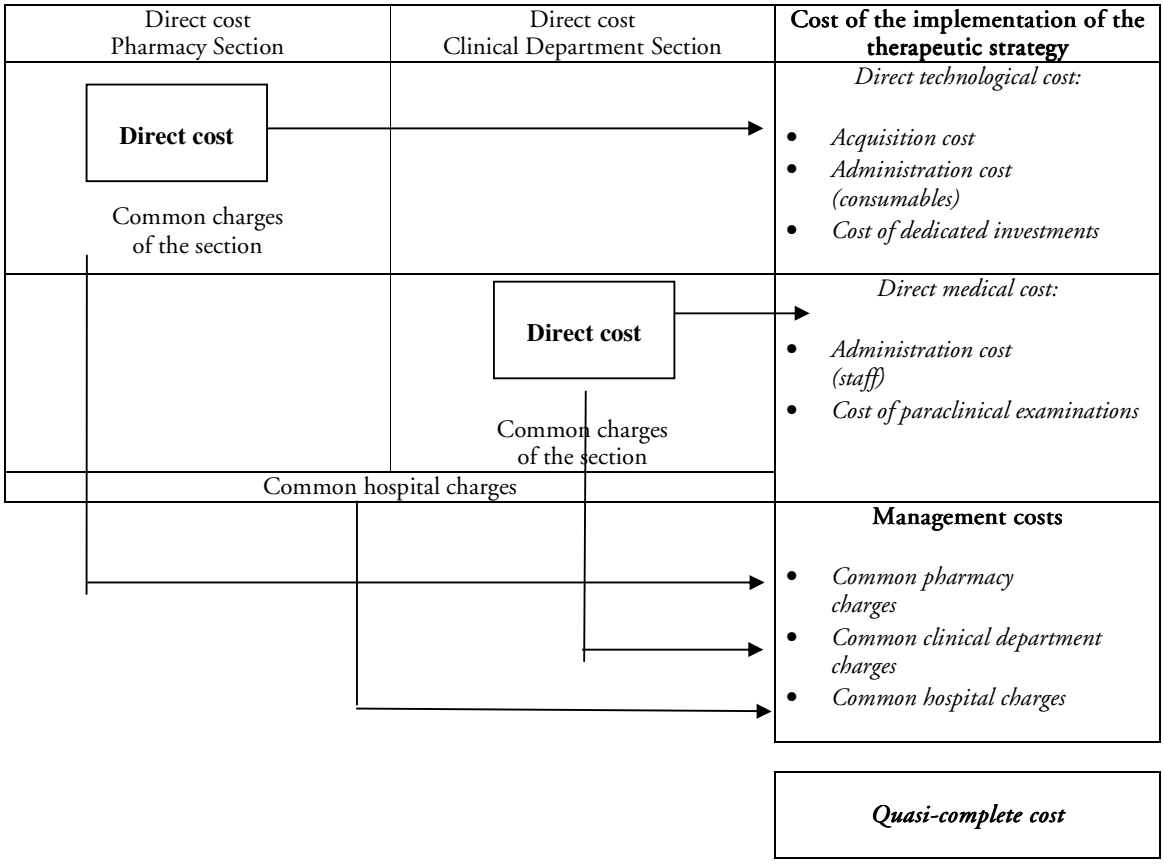
The term "directly" may, however, be understood in two different ways: in one, reference is to the variability of the expenses, in the other it is the possibility of identifying these expenses individually that is the focus of attention, which makes it possible to include dedicated equipment. In general, micro-economic studies focus on the variable costs related to investigations specific to the technology implemented, to the drugs prescribed and the consumables required for their administration. The calculations are generally based on standard costs as regards both quantities consumed and unit costs. The standard quantities used for medical/technical procedures are those taken from observation logs or defined by expert opinion. In the case of drugs, the dosages used are generally those recommended in the summary of product characteristics. Unit costs must be calculated on a first-line basis from *Indice des Coûts Relatifs* (ICR) [relative cost index], from the average cost of code letters on the national costs survey and from NGAP coefficients. Products and consumables for which the prices were negotiated are extracted directly from the department's accounting data. If personnel time is specified, its monetary equivalent must be deducted from the corresponding GHM headings to eliminate double counting.

- *Cost of patient management in an appropriate structure*

The National cost survey can be used to evaluate hospital stays associated with the treatment of various diseases. The cost categories defining the components at the level of each GHM may be aggregated according to their content into each of the PMSI headings in table 1. At present, a GHM is not specific to a disease, however. Given that a certain number of pathologies are associated with a GHM and that a single disease can be found in several GHMs, the GHM is a composite index of diseases. As an index, the GHM also represents an average of the lengths of stay associated with certain pathologies. The technique known as the "GHM amenagé" [adjusted GHM] can be used to correct the first bias. One or more of the 7 headings relating to variable medical costs will therefore be excluded from the 19 cost components analysed in the PMSI study, to be replaced by the real costs that can be directly associated with the implementation of a particular protocol. The real opportunity cost of initiatives adopted, i.e. the value of what cannot be done because of doing what is done, may be incorporated into the analysis instead of the mean baseline values. The second bias, however, is more difficult to neutralise. A new allocation of the expenses included in each GHM according to length of stay would mean totally reconstructing the GHMs, which is precisely what the proposed method seeks to avoid.

The total hospital cost is equal to the sum of the substitute cost of implementation of the therapeutic strategy and of management in a hospital structure. Capital expenditures are excluded from the cost, since the amount is linked more to the past history of the hospital than to the management abilities of its managers. For this reason, the ISA points "scale" will be referred to in the rest of the text as the quasi-complete cost.

*Figure 1: breakdown of direct and indirect hospital costs by budget allocation type*




- *Total cost of attenuating income*

When a department's occupancy rate is close to 100%, the adoption of a treatment which reduces the number and duration of complications makes it possible to increase patient turnover rates and admit new patients, which in turn leads to further GHMs being recorded. The management of the new entrants entails extra variable costs, but it increases the number of ISA points generated by the hospital. The monetary equivalent of these points is greater than the variable costs associated with the new entrants. This margin on variable costs helps to finance the hospital's fixed costs. The amount it represents must be deducted from the additional cost of using the new treatment or from its full cost price in order to calculate the net medical cost associated with its use.

Conversely, treatments which entail longer stays due to complications block beds and reduce turnover rates. In this instance, the real opportunity cost of the situation must be calculated. When a complication arises, the hospital loses money: it can no longer include in its theoretical budget the monetary equivalent of ISA points which would have been associated with new entrants. At the same time, though, the hospital also gains because it does not incur the variable costs engendered by the arrival of new patients. The shortfall is the net balance between the sums which the hospital would have received and the expenses it would have incurred in treating an additional case. This amount defines the hidden cost of complications. To obtain the total cost of treatment, these indirect opportunity costs must be added to the direct costs of treatment or to the complete costs, calculated as explained above.

*General and private practice*

The CCAM to be introduced in 2004 will bring about substantial changes for private practitioners (generalists and specialists). First and foremost, the CCAM will allow for procedures to be coded by type, as had been planned under a previous convention between the CNAM and GPs. This will make it possible to analyse the activities of practitioners of the various medical disciplines. In addition, the ranking of the classifications has been based on the following factors: stress, mental effort, technical skill, duration and work. Its resulting application to tariff-setting, even if all the terms and transitional measures are not yet in place, will lead to a redistribution of income between the disciplines on the basis of more logical criteria.

The system of tariffs will be based on the use of codes and will no longer use the key letters and coefficients currently in use.

A code = a designation = a unit price

The unit price of a procedure will correspond to the fee paid to the health care professional and will be the outcome of a tariff negotiation.

Fee = (W*Fc) + W* (Gg + S)

(1) (2)

(1) Price of work: physical and intellectual resources

- W: point of medical work corresponding to the sum of the physical and intellectual resources deployed by the practitioner. It includes stress, technical skill, mental effort and duration of the procedure, of which it is a complex synthesis.
- Fc: A single monetary conversion factor, identical for all specialities.

(2) Practice cost: professional charges

- W: point of medical work.
- Cg: after calculating the mean cost of general charges per point for a discipline, general expenses are allocated to the different procedures proportionally to the work they require.
- S: any excess costs associated with specific practices not carried out by all practitioners. The calculation of excess costs includes only charges not included in general expenses.

(See Appendix IV for a detailed presentation of methods for estimating the burden of charges in private practice.)

HEALTH CARE PURCHASERS

The choice of expenditure items in a study should be dictated not so much by abstract reference to a principle of general interest which no administrative institution may claim to represent given its designated remit, as by the real existence of a forum for negotiation in which real purchasers and vendors of health care may come together. From this point of view, it would seem wise to concern ourselves only with changes in the consumption of medical care and goods, excluding any loss of production, by adopting from the outset the perspective of the health care system. But even though a Sécurité Sociale department does exist within the health ministry and is responsible for steering the system as a whole, it must be said that the department's managers have no responsibility as purchasers. The real negotiating power lay for a long time with the Assurance Maladie but it does so no longer after the redefinition of the respective roles of the state and the Sécurité Sociale in the management of the French health system. The existence of a bilateral monopoly or a monopsony in the purchase of health insurance, as our health care system has often been described, is a concept devoid of content that masks the reality of how the system operates. At present there are quite definitely three purchasers: the Agences Régionales d'Hospitalisation (ARH) which set hospital budgets, the Assurance Maladie, which negotiates the conditions of general and private practice with practitioners, and the CEPS, which is responsible for setting drug prices.

*Table 2: Grouping of expenses according to viewpoint adopted*

| | | | ***Viewpoint adopted*** | | |
|---|---|---|---|---|---|
| **Costs to collate** | **Families** | **Health care professionals** | **Purchasers CEPS/ARH/AM** | **Health care system** | **Society** |
| Medical expenses | Non-reimbursed | Linked to treatment | Reimbursed | Yes | Yes |
| Non-medical expenses | Yes | No | No | No | Yes |
| Potentialities sacrificed | Yes | No | No | No | Yes |
| Impact on economic activity | Loss of income not made up | No | No | No | No, if QALYs |
| Welfare state cost | Substitute income | No | Daily allowance | No | Management costs |

- *The Agence Régionale d'Hospitalisation (ARH)*

From the standpoint of the director of the ARH, the demonstration must be made in terms of ISA points per GHM. Thanks to the introduction of "medicalised" information by the hospitals, which now present their activity in the form of GHMs, it is possible to offer a calculation of the global budget which takes into account the range of cases treated. With this information, the regulator is able to examine whether the actual budget is very different from the theoretical budget. If this is so, the regulator can agree with the hospital on a plan for the gradual alignment of its budget with the reference budget. Unit prices per GHM are global and do not need to show their breakdown, as was previously the case at hospital level when using the Enquête Nationale de Coûts to calculate the net margin on variable cost. This information would be useless to the regulator, who would have no idea of how to use it. The calculations made for the ARHs need only ISA points without the need to use the results of the Enquête Nationale de Coûts.

- *The Assurance Maladie*

Adopting the viewpoint of the Assurance Maladie, we will identify only those expenses it is required to cover, either in the form of benefits in kind by way of reimbursements of sums advanced by households to general/private practitioners, after deduction of the co-payments, or in the form of benefits in cash paid to contributors on sick leave.

For outpatient care, medical consultations (key letters C, V, CS), paramedical procedures (AMC-AMK) and radiological and investigative procedures are evaluated on the basis of the NGAP listings edited by the *Union des Caisses Nationales de Sécurité Sociale* (UCANSS) [national union of Social Security Funds). Biomedical procedures associated with the disease are evaluated on the basis of the *Nomenclature des Actes de Biologie Médicale* (NABM) listings also produced by the UCANSS. Any hydrotherapy (spa) treatments evaluated are those which do not involve hospital admission. Hydrotherapy treatments requiring hospital admission may be considered as hospital admissions. The CCAM is set to replace the NGAP on 1 January 2004 for technical procedures. The NGAP will continue to apply for the time being to consultations and to biomedical, paramedical and midwifery procedures.

The Vidal-SEMP database can be used to evaluate the cost of drug prescriptions. Vidal-SEMP is a company which publishes an exhaustive compendium in the form of fact sheets on all pharmaceutical products available in pharmacies and hospitals. Full pharmaco-therapeutic and economic data on these products is available on computer disk (Sempexdata), which makes it possible to use the data in ad hoc programmes. The information is updated monthly. Products are classified according to the EPHMRA classification. A 7-figure code, the CIP code, provides precise identification of each presentation of a pharmaceutical product. From the standpoint of the Assurance Maladie, all drug prescriptions musts be valued on the basis of retail prices.

- *The Comité Economique des Produits de Santé (CEPS)*

For the CEPS, it is the manufacturer's price which should be used to evaluate drug consumption. Reliable cost calculation requires accurate collection of prescription data. Data collection will also include the CIP code, a unique identifier for the pharmaceutical product prescribed. The evaluation also requires accurate knowledge of the dosage in terms of units taken per day, and of the total length of the prescription period. The number of any prescription renewals must also be known. In the event of a subsequent consultation and a renewal of the prescription prior to the previous prescription expiry date, an item must be specified to indicate whether or not the previous drug treatment is maintained. Drugs administered at frequencies prescribed as "on request" or "as needed" must be evaluated on the basis of a frequency agreed with the doctors involved in the study.

THE HEALTH CARE SYSTEM

The estimation of expenses for the health sector as understood by the Comptes de la Santé [national health accounts] covers both the expenses approved for reimbursement by the Sécurité Sociale and reimbursed after deduction of the amount which the contributor is left to pay (ticket modérateur), and the expenses left for families to pay either because these expenses are not reimbursed (the basic ticket modérateur) or because they are not recognised (a de facto ticket modérateur, arising from an excess fee charge or self-medication). If we consider not the origin of the funding but the nature of the operations funded, this covers expenses relating to hospitalisation, outpatient care, spending on drugs and medical transport costs.

SOCIETY

It is generally agreed that adopting the perspective of society transcends institutional budget divergences by making it possible to aggregate all expenses whatever their nature (direct or indirect) and the source of their funding (Securité Sociale, households, state and local authorities). Added to health care costs are the cost of accommodation in non-medical institutions, the cost of domestic equipment needed for patient care in the home, the value of salaries lost by

members of the family due to caring for the patient, and the reductions in remuneration directly linked to the constraints imposed by the treatment of the patient himself. The broadest definition of the cost of care to society even includes the value of production lost due to illness when this is accompanied by temporary or permanent absence from the job market. Certain guidelines recommend including direct and indirect non-medical costs in the estimation of the socio-economic cost of treatment. These proposals pose problems that are both practical and conceptual, however. These data are not always collected during trials and, in the absence of recent information of indirect origin, cannot be included in the analysis. From the theoretical standpoint, there is a high risk of ending up with double counting in both the numerator and denominator of the cost-effectiveness ratio. The consequences obtained are very often adjusted according to the repercussions of the disease and its treatment on quality of life; however, most scales which explore quality of life take into account the repercussions of the disease on social roles and activities on daily living. Valuing these absences from the workplace in terms of cost at the numerator would be equivalent to counting them twice.

When an evaluation is being made from the viewpoint of society, social transfers should under no circumstances be taken into consideration. The heading of social transfers aggregates all cash payments relating to sick leave, disability and death. It is traditionally accepted that such sums do not benefit the medico-social sector directly. They are deducted from certain contributors and paid to others and are used to finance spending that feeds into general economic activity. They are therefore simple transfer operations which in no way correspond to any specific allocation of resources to the health sector.

## 2 – SPECIFICATION OF SPENDING ITEMS

The scope of definition of the cost from a given perspective depends on the cost object we have chosen to study and the role we intend it to play.

The relationship between cost and quantities requires us to specify the parameter by which we classify the expenditure. In the case of the Assurance Maladie, for example, we could consider that the cost object is confined to benefits in kind or, alternatively, that it extends to all the expenses the paying centres is required to meet in respect of cash payments.

The content of the cost will also differ according to the uses to which it is put. Why do we want to calculate it? Do we want to use it as a management criterion, a means of control or a pricing instrument?

In the first instance, the analysis of the financial repercussions of a decision must record the new expenses which arise when the decision is taken, compared to the expenses that would otherwise have been incurred had the decision not been taken. We will consider only those items affected by the decision and exclude all those which reflect the financial consequences of past choices which cannot be reversed, and we will exclude items that merely record pricing decisions mechanically without taking the reality of expenses into account. What really characterises the adoption of society's viewpoint, however, is not so much the scope of the area explored (health spending, family spending, production losses) as the introduction of reasoning based on opportunity cost. Tariffs and expenditures are partially abandoned in favour of attempts to measure the value of the real efforts made by society in favour of health. The quality of fairness becomes an essential metrological property, which means we must use the "adjusted GHM" technique in order to take into account the reality of the costs involved in the indication being studied.

In the second instance, dividing the undertaking into cost centres enables us to define, for each cost centre, which costs are controllable and which costs the centre manager is able to influence and against which his performance can be assessed. Adopting the responsibility centre approach puts greater emphasis on control than on forecasting. The key properties required are accuracy and sensitivity.

Working from the pricing standpoint, finally, we will take into consideration all the expenses incurred in relation to the activity being studied (complete cost or quasi-complete cost).

In the end, there are as many types of cost as there are possible applications. To be specific about the nature of the cost object, we must always be careful to stipulate where the cursor is positioned in the list of budget items analysed from a given viewpoint, and why.

## 3 – DEFINITION OF A TIME FRAME

The consequences of a decision may be felt indefinitely. We should therefore set a time horizon beyond which we will not consider the effects produced. The forecasting limits must be defined in the cost study.

*Figure 2: Possible impacts according to the time horizon chosen*

| Impact for the Sécurité Sociale (complete cost) | Impact for the hospital (direct cost) | Impact for the department (variable cost) | Impact for society |
|---|---|---|---|
| Direct charges | Variable medical and logistics charges relating to the running of the department | Variable medical and logistics charges relating to the treatment | Short term |
| | Department common charges | | Medium term |
| Indirect charges | | | Very long term |

In the short term, in view of existing rigidities, the only potential savings that can legitimately be expected from a reduction in the length of hospital stays are related solely to the variable expenses associated with patient management.

In the medium term, we might hope that the redeployment of medical and paramedical staff within the department or between departments will bring savings in direct costs. The most favourable scenario would be the complete close-down of the department.

Lastly, incorporating overhead costs or structural costs avoided into the potential savings would mean that the introduction of the new technology would enable the entire hospital to be demolished. Such a solution is clearly possible, as was shown in the case of the Laennec and Broussais hospitals coupled with the creation of the Georges Pompidou European hospital. It has to be recognised, however, that such economies will only merge over the very long term.

Any analysis of the savings to be expected from a choice always depends on the time frame on which the analysis is based. This time horizon must always be specified. In the short term, in view of existing rigidities, the only potential savings that can legitimately be expected from a reduction in the length of hospital stays relate only to the variable expenses of the GHM associated with patient treatment. We should confine ourselves exclusively to marginal costs. In the long term, we will take into account all the expenses recorded for the GHM concerned. The reasoning should be based on complete costs on the basis of the mean cost per stay.

The consequence of this recommendation is that it becomes impossible to calculate the cost of a stay by the mechanical application of value in ISA points. Only by a detailed analysis of the components of the GHMs as they appear in the Etude Nationale de Coûts can we carry out a refined analysis of the savings expected from a reduction in the length of hospitalisation. This distinction between the successive time horizons that may be adopted would enable us to modulate the calculation of costs according to the delays required for the re-allocation of resources and people. Each time horizon chosen could then be associated with a different "cost label" in budgetary forecasting models.

## 4 – SELECTING AN INFORMATION COLLECTION METHOD

The cost analysis may be performed retrospectively, prospectively or on the basis of standard costs. The first method requires us to refer back to the patient's medical record, the second requires us to set up a prospective study with an observation log-book or collection on Internet. The standard costs method is the most economical in terms of time spent: it supposes that there is a quantity of standard health cares per patient which may be valued at a determined unit cost. We may consider that the GHMs of the national costs scale also correspond to standard costs, even if their breakdown into quantity and unit cost is not easy to an external observer, a situation perceived by foreign reviewers as a breach of the sacrosanct rule of the necessary distinction between quantities consumed and unit prices used.

## 5 – CONCLUSION

We clearly see that the classic questions : "A cost for whom? The cost of what? A cost to do what? Costs measured when?" are merely a simplification for teaching purposes. The questions are interrelated and require careful consideration and synchronous responses. The definition of the scope of operations to be considered depends in the

final analysis on the cost object chosen from a particular perspective, on the purpose of the study and on the time horizon adopted.

The economist is not an accountant. He is interested as much in the value of the service rendered to the patient as in the scale of the resources deployed to treat the patient. The importance he places on financial factors stems not, as is too often believed, from the desire to control spending, but above all from the concern to save the greatest number of lives. The real cost of a treatment is not the amount of expenditure it incurs but the scale of the potentialities sacrificed for other patients that it is not possible to treat. This is why we cannot be content with studying the respective costs of different treatments without also considering their clinical effectiveness. It is the whole body of outcomes achieved that must be set against the cost in order to determine whether or not a treatment is justified. When competing needs exert growing pressure on limited resources, it is the concept of effectiveness that must prevail.

# APPENDIX II: Programme de Médicalisation du Système d'Information (PMSI)

1 – THE PMSI

1.1) Background of the PMSI

A new approach to hospital management was introduced in the USA in the late 1960s by the team headed by Professor Fetter at the University of Yale. The constant rise in health care expenditures and the lack of relevant indicators for measuring hospital activity led the Fetter team to develop a system of classification known as the DRGs (Diagnostics Related Groups). The groups are defined to represent an identical level of consumption and to be clinically coherent. The classification system was established by analysing data from a sample of 1,800,000 hospital records. 333 homogeneous groups of hospital stays were initially isolated. The current 13th revision of the DRGs contains over 500 groups. The system is used as part of the Medicare prospective payment system for the elderly and disabled.

The DRG classification has been adapted to many European countries, including France where the PMSI was introduced in 1982. After several years of pilot studies, the circular of 24 July 1989 generalised the use of PMSI to all hospitals and set up a structure to manage it, the Medical Information Department. The requirement to collect data on medical activity was reinforced by the hospitals act of 31 July 1991 (Article L 710-6) which made it an obligation for health institutions to analyse their activity and to set up information systems that take into account the diseases and management methods in order to improve knowledge of activity and costs. In 1994, a regional pilot study was set up in Languedoc-Roussillon to use the PMSI to adjust budgetary bases between hospitals. The circular of 10 April 1995 then extended the experimental scheme to all regions of France.

The private sector was not left out in these developments: the decree of 22 July 1996 extended the PMSI to all private health care establishments. The standard information collected is identical to the public sector, except for the addition of the beneficiary's social security number and ranking. These two items are used to follow the patient through the various institutions he may attend, in order to provide a better knowledge of the health care channels. This "chaining" of hospital stays will be all the more relevant in that it will also be applied to public hospitals in the years ahead.

In 1998, the PMSI system was extended to SSR care with a weighted daily classification. The activity of these sectors is characterised by lengthy hospital stays (averaging 35 days). A weekly summary is generated and the patient is classified during his stay into different GHJs according to the diseases and the treatments followed.

A classification for psychiatry is in preparation with a distinction between full hospitalisation and outpatient activity. The Psychiatric PMSI is due to come into service on 1 October 2001.

1.2) Description of *Résumés d'Unité Médicale* (RUM) [Hospital ward discharge abstracts]

Each short-stay clinical department produces an RUM, which contains a limited amount of administrative and "medicalised" information and is compiled retrospectively at the end of the patient's stay.

As of 1 January 2000, the medical information relates to the principal diagnosis, the associated diagnosis, significant associated diagnoses, documentary diagnoses and procedures.
- *The* ***principal diagnosis*** is considered by the doctor treating the patient, at the end of the stay in a medical unit, as having mobilised the bulk of the medical and nursing effort.
- *The* ***related diagnosis*** clarifies the pathological context, essentially when the diagnosis is not in itself a condition.
- ***Significant associated diagnoses*** represent the diagnoses, symptoms or other reasons for significant recourse to treatment or consumption of resources.
- ***Associated documentary diagnoses*** (optional) cover antecedents cured and diseases not justifying treatment.

Diagnoses are coded according to the 10th revision of the International Classification of Diseases (ICD), ICD 10.

Procedures may be classifying or non-classifying. A procedure is classifying when it is likely to modify the classification of the RSS in a GHM, in other words if it is predictive in terms of a particular treatment, discriminating in terms of consumption of resources. Otherwise, it is non-classifying. It may be operating in the sense of the GHM classification because performed in the operating or non-operating block.

Procedures contained in the RUM are coded according to the CdAM (see Appendix III).

The RUMs are forwarded to the DIM which groups the RUMs and classifies them into GHMs using a programmed grouping algorithm.

1.3) Description of Groupes Homogènes de Malades (GHM)

A GHM is defined as follows: Groups are identified in terms of homogeneity on two levels, medical and economic. Each group contains patients with an identical level of consumption of resources and similar medical characteristics.

*Example: in Cardiology, we distinguish for the same disease: GHM 178 myocardial infarction with cardiovascular complications, patient discharged; GHM 179 acute myocardial infarction without cardiovascular complications; and GHM 180 myocardial infarctions followed by death.*

The classification of the Groupes Homogènes de Malades consists of 560 GHMs broken down into 27 *Catégories Majeures de Diagnostiques* (CMD) [major diagnostic categories] (equivalent to medical specialties). It allows an exhaustive and unique classification. Any hospital stay is necessarily classified under one GHM, and one only.

The main variables determining the GHM classification of the stay are: the principal diagnosis, the existence of a classifying procedure , associate co-morbidity or complication, age.

This classification is regularly updated, and has been constantly enriched over the succeeding versions: creation of GHMs specific to major diseases (transplantations, AIDS, multiple trauma), introduction of lists of severe co-morbidity (septicaemia, pneumonia, endocarditis, etc.) to designate stays which have a high economic impact. Version 6 of the classification is used for the data for the year 2000. The associated diagnosis does not as yet enter into the GHM classification, but version 7 scheduled for 2002 should introduce a radical development including a new method of grouping stays according to the "progressive strip-off" (*Effeuillage Progressif* in French) method, the objective being to qualify hospital stays more precisely in medical terms.

The use of the PMSI for the purposes of budget allocation made it necessary to set up a national costs per stay database which can be used to calculate reference costs per GHM.

1.4) National Cost Survey

With the circular of 28 February 1992, the Direction des Hôpitaux (Hospitals administration) decided to set up a national database of costs per stay based on a sample of public hospitals and volunteer PSPH establishments.

The calculation of costs per stay is performed on the basis of three major components: medical costs, medico-technical costs, structural costs and logistics costs.

***Medical costs*** include:
- expenses directly associated with the patient's stay: blood, prostheses and implants, expensive drugs, expensive disposable equipment, procedures carried out outside the hospital;
- the expenses of each department the patient has stayed in: salaries of medical staff, salaries of nursing staff, salaries of other staff, depreciation and maintenance expenses for medical equipment in these units, medical logistics expenses. These are allocated on a daily basis.

***Medico-technical costs*** include expenses relating to procedures using the hospital's technical support centre, broken down into specific units of work: the ICR.

***Structural costs*** include financial costs, amortisation, property costs and tax excluding staff.

***Logistics costs*** include laundry, catering, general management and other logistics. These last two costs are allocated to activities proportionally to the amount of direct charges and then broken down in patient stays according to the number of days.

1.5) Construction of the *Etude Nationale des Coûts par Activité Médicale* [national cost survey per medical activity]

The scale of cost per GHM is built up from the average cost observed in the database for each GHM.

By convention, GHM 540 (delivery without complications) is given a weighting of 1000 ISA points (Indice Synthétique d'Activité) [casemix index] and all other GHMs are positioned in relation to it.

On the 2000 scale, the values of the GHMs ranged from 105 points for a radiotherapy session to 30,538 for a liver transplantation. The national scale of costs is updated every year to allow for developments in GHM classification, treatment procedures and costs.

## 2- USING THE PMSI TO CALCULATE THE COST OF HOSPITALISATION IN THE PUBLIC SECTOR AND PRIVATE SECTOR UNDER CONTRACT (PSPH)

Several methods, set out here in ascending order of accuracy, may be used to calculate the cost of a hospital stay defined by the principal diagnosis of the disease treated and/or by the procedure or the surgical procedure(s) performed during the stay. We will first review some basic data on the calculation of the number of ISA points per GHM.

The number of ISA points assigned to a GHM is derived from an ongoing survey of volunteer hospitals, which provide both medical information and cost data taken from their internal cost accounting system, using a uniform method established by the *Direction des Hôpitaux et de l'Organisation des Soins* (Hospitals and Health care organisation administration). This survey is known as the Etude Nationale des Coûts par Activité Médicale [national cost survey per medical activity][81]. As a first approximation, we can consider that the number of ISA points allocated to a GHM reflects the relative average cost of each group calculated on the basis of data provided by the volunteer establishments. The national value of the ISA point in francs or euros is calculated each year by dividing spending on short-stay hospital treatment in the public sector by the number of ISA points corresponding to all the stays made over the year in the public hospital and PSPH sector.

The PMSI website (www.le-pmsi.fr) offers unlimited access to the following information. Firstly, online access to the national database of all short stays in public and private sector hospitals (since 1997, with a two year lag on the current year). There is no access, however, to an extraction of a sample of stays, but to tables corresponding to the results of standard requests. It is possible to find, for example, the GHMs in which stays associated with a particular principal diagnosis or particular procedure are classified, or the statistics for a GHM (population, average length of stay, standard deviation, min-max, etc.) by status and type of institution, and access to regional statistics is also available. Secondly, two important tables are published every year and can be downloaded: the scale of ISA points per GHM and the table showing average costs per GHM obtained from the sample of establishments taking part in the *Etude Nationale des Coûts par Activité Médicale*.

**1st method**

The study population for which we wish to calculate the complete cost of a hospital stay is defined either by a series of principal diagnoses coded by the ICD 9th or 10th edition, or by a series of procedures coded (up to 2003) from the CdAM.

We can then initially carry out an online interrogation of the national PMSI database Chapter MCO, sub-chapter Statistiques. Online access allows consulting the national database of all short hospital stays in both public and private sector (since 1997). For our purposes, useful interrogations can be made from a number of entry points: principal diagnosis, classifying procedure, GHM. Statistics in greater or lesser degrees of detail are available for each of these entry points.

The most elementary method consists of interrogating the database to identify the GHMs which contain the pertinent diagnoses or procedures for the population being studied. Let us take the example of an interrogation with a list of diagnoses. We obtain a list of GHMs in which these diagnoses are present as the principal diagnosis, with the relative part played by these diagnoses in all the hospital stays grouped in each GHM. The following table is an extract from the hospitalisation database in the public sector for 1999, with the entry-point diagnosis "Gastro-oesophageal reflux with oesophagitis".

*Public database-Year 99*
*DIAGNOSIS K210 - Gastro-oesophageal reflux with oesophagitis*

| GHM | | Population | % |
|---|---|---|---|
| 213 | Major interventions on the small intestine and colon, age over 69 and/or CMA (Associated complications and morbidities) | 5 | 0.064 |
| 214 | Major interventions on the small intestine and colon, age under 70, no CMA | 12 | 0.153 |
| 215 | interventions on the oesophagus, stomach and duodenum, age over 69 and/or CMA | 281 | 3.576 |
| 216 | interventions on the oesophagus, stomach and duodenum, age 18 to 69, no CMA | 1080 | 13.74 |

[81] Guide Méthodologique de comptabilité analytique hospitalière, Tome 1- Tome 2- Tome 3- Sommaire- le calcul des coûts par activité- le calcul des coûts par séjour. ISBN 2-11-092510-8, Ministère des Affaires Sociales, 1997.
Échelle Nationale des coûts rélatifs par Groupes Homogènes de Malades, ISBN 2-11-120270-3, Ministère des Affaires Sociales, 1997.

| | | | |
|---|---|---|---|
| 217 | interventions on the oesophagus, stomach and duodenum, age under 18 | 151 | 1.921 |
| 226 | interventions to repair hernia or rupture except for inguinal and crural hernia, age over 69 and/or CMA | 4 | 0.051 |
| 232 | interventions on rectum and anus other than rectal resection, age under 70, no CMA | 3 | 0.038 |
| 233 | Other interventions on the digestive tract, age over 69 and/or CMA | 5 | 0.064 |
| 234 | Other interventions on the digestive tract, age under 70, no CMA | 15 | 0.191 |
| 256 | Gastro-enteritis and various diseases of the digestive tract, age over 69 and/or CMA | 1436 | 18.27 |
| 257 | Gastro-enteritis and various diseases of the digestive tract, age 18 to 69, no CMA | 1169 | 14.87 |
| 258 | Gastro-enteritis and various diseases of the digestive tract, age under 18 | 1143 | 14.54 |
| 262 | Conditions of the digestive tract with CMAs | 178 | 2.265 |
| 706 | @ Conditions of the digestive tract: outpatient with other operative procedure | 9 | 0.115 |
| 806 | @ Conditions of the digestive tract: outpatient with no operative procedure | 2353 | 29.94 |
| 890 | @ Immediate transfer | 8 | 0.102 |
| | Other GHMs | 7 | 0.089 |
| | TOTAL | 7859 | 100 |

The valuation of the cost of a stay for gastro-oesophageal reflux with oesophagitis may therefore be made, as a first approximation, by associating each GHM with its number of ISA points and by calculating the number of ISA points weighted by the relative percentage of each GHM (last column of the table). The weighted number of ISA points thus obtained is multiplied by the latest known value of the ISA point at national level, which can also be obtained from the PMSI site.

This method has the advantage of simplicity, but it may be time-consuming to carry out if there are several diagnoses to search for, since the database must be interrogated sequentially. In addition, it reintroduces an heterogeneity in the costs that the GHMs are supposed to overcome, by mixing the average costs of different groups. The method can certainly be refined by separating the "surgical" GHMs and "medical" GHMs, but this is still very basic. It is therefore recommended that this method should not be used except to calculate a cost for the hospital treatment of one disease. It is not specific enough, on the other hand, to be used for a cost-effectiveness study. Note also that the online interrogation function on the PMSI website does not allow cross-referenced Diagnosis/Procedure queries.

It is possible to arrive at a very rough estimate of a confidence interval using this method, using published data by GHM on the PMSI website. Indeed, we know the average length of stay and its standard deviation for each GHM. We must then assume that 1) there is a close relation between the average length of stay of a GHM and its number of ISA points, 2) that the variance in the length of stay is the same for any hospital stay sub-sample for a given GHM.

**2nd method**

This is based on a sample of stays taken from the *Etude Nationale des Coûts par Activité Médicale*. The sample may be obtained on written request to the *Direction des Hôpitaux et de l'Organisation des Soins*. The limits to this demand are as follows:
- no identification of the hospital;
- for patients aged over 90 or under 1 year, the exact age is not given but just the corresponding age category;
- stays terminating in death are not included.

We can then request an extraction according to several criteria, diagnoses or procedures, in order to be as close as possible to the population being studied. We obtain a cost statistic on a sample from which we can estimate a confidence interval. Two adjustments must nonetheless be made.

Firstly, the sample from the National Cost Survey is not representative. A certain number of adjustments are made in order to calculate the number of ISA points per GHM, but when we carry out an extraction we obtain raw data giving no identification of the hospital, so we cannot make any of these adjustments. We must therefore begin by calculating a value for the ISA point, working from the reference costs per GHM file. To do this, we simply calculate the total number of ISA points in the Etude de Coûts and divide them by the total budget for all the hospitals in the study, which in turn is obtained by multiplying the population of each GHM by its mean cost. We can then compare the cost per ISA point of the study against the national figure, which is also published on the PMSI website. If the two are significantly different, we can apply a correction factor to the costs for our sample, equal to the ratio between the national cost per ISA point and the National Cost Survey cost per ISA point.

Secondly, the ISA point does not cover "structural" costs (finance costs, depreciation and maintenance of property), and these will therefore have to be eliminated from the earlier adjustment and added back in subsequently. There is therefore an uncorrected sampling bias for these cost items, but it has little impact on the overall cost.

**3rd method**

In the 2nd method, we assumed a perfect match between the target population of the economic study and the National Cost Survey sample. This assumption is rarely perfectly verified. We may therefore proceed as follows:

We can model the expected cost of a hospital stay for the target population, based on sample data. This is equivalent, for example, to using a regression model where the dependent variable is cost per stay for the sample, and the explicative variables are the characteristics of the target population which are known and which distinguish the target population from other patients in the sample (for example, age, sex, associated diagnoses, procedures performed). We then have a model which allows us to adjust the sample data to the characteristics of the target population.

Other uses of the National Cost Survey and the National Database

When carrying out an economic study, we sometimes have access to direct observations of hospital resource consumption, in terms of the nature of the medico-technical procedures performed, laboratory procedures and the number of days of hospitalisation. Unless we carry out *ad hoc* surveys, however, we rarely have access to the unit costs of these complementary procedures and investigations. When such *ad hoc* surveys are carried out at on a small number of sites, we cannot guarantee that the results are truly representative. This aspect is, of course, relatively minor in calculating the overall direction of a result, since what will really counts is consumption by volume measured between two treatments and valued according to identical unit costs. Even so, the possibly heterogeneous nature of the unit costs makes it difficult to compare different studies, and hence difficult to compare incremental cost-effectiveness ratios. It is possible, with PMSI data, to calculate unit costs per investigation or procedure which are close to representative at the national level, and which could therefore be applied to all studies.

Indeed, the National Cost Survey and the reference costs provide the following information for each GHM:

| | | |
|---|---|---|
| Lab B | Number of laboratory Bs | Laboratory expenses |
| Block ICR | Number of block ICRs | Operating block expenses |
| Radiology ICR | Average number of medical imaging ICRs | Imaging expenses |
| Anaesthesia ICR | Average number of anaesthesia ICRs | Anaesthesia expenses |
| Diagnostic investigation ICR | Average number of diagnostic investigation ICRs | Diagnostic investigation expenses |
| Dialysis ICR | Average number of dialysis ICRs | Dialysis expenses |
| No. of radiotherapy procedures | Number of radiotherapy procedures | Radiotherapy expenses |

We are therefore able to calculate a mean cost per B or ICR across all the GHMs. This mean cost is weighted by the population of each GHM, as obtained from the PMSI national database.

Let B be the mean number of B per stay in GHM i; let Di be the mean expenditures per stay on laboratory procedures for the GHM i; let Ni be the population in number of stays in GHM i, obtained from the PMSI national database. Thus the mean cost of B obtained from the Etude de Coût is equal to:

$\Sigma_i$ Ni Di / $\Sigma_i$ Ni Bi.

From this calculation, if we know the nature of the complementary investigations prescribed in the course of a hospital stay, and the associated number of Bs thanks to the Nomenclature Générale des Actes de Biologie Médicale, we can calculate the cost of these investigations. We must bear in mind, however, that this cost covers only the direct operating charges of the laboratories concerned, i.e. salaries, medical consumables and the depreciation of equipment. The salary item may be distinguished from the other items in the calculation of this cost.

The principle of the calculation is the same for the ICR for surgery, anaesthesia, medical imaging or diagnostic investigations. The reference catalogue in these cases, however, is the CdAM, which is to be replaced in 2003 by the CCAM.

Finally, we also have a database from which we can calculate the cost of procedures performed outside the public hospital. A file can be downloaded from the PMSI website containing the values for key-letters of procedures performed externally. This file can be used to transform the external activity of public hospitals, measured using the key-letters of the Nomenclature Générale des Actes Professionnels (Cs, B, K, Z and others) into ISA point equivalents. In 2000, for example, a CS (consultation with a specialised doctor or a specialised dentist) was quoted as 11.971 ISA points; if the national value of the ISA point was 13 FF, the cost of the CS was 154.92 FF.

## 3 – USE OF THE PMSI TO CALCULATE THE COST OF HOSPITALISATION IN THE PRIVATE SECTOR

The commercial private hospital sector poses a specific problem in calculating the cost of a hospital stay, which is that of medical fees. The following data is available at present. From the PMSI website we can download a file giving a scale in ISA points for GHMs in the private hospital sector. This scale is constructed from form 615, the form used to bill patients for their treatment in the course of each stay. The form does not include medical fees, however, which means it is not comparable with the public scale. To estimate a cost per GHM, we must therefore identify the main procedures associated with it and charge them on the basis of the value of the NGAP key-letters. This is relatively feasible for the "surgical" GHMs but more difficult for the others, for which we will not know the details of the medico-technical procedures prescribed. A highly approximate estimation consists of charging a specialist consultation (Cs) for each day of hospitalisation. Finally, the value of the private sector ISA point is available on request from the BADIMEHP (www.adimehp.com).

This association, an offshoot of the commercial private hospital federation, has very recently produced a cost study using a similar methodology to that of the public sector study. 30 private hospitals provided a database of some 200,000 stays and 28,000 sessions for 1998, making it possible to estimate mean costs per GHM, this time including medical fees. This database can therefore now be used to carry out the same type of calculation as those described above for the public sector.

The PMSI national database can also be used to calculate, GHM per GHM, the relative market shares between public sector and PSPH, and between public sector and commercial private sector. It is therefore possible to calculate a weighted public/private cost.

## APPENDIX III: CdAM and ICR

2.1 - THE CATALOGUE DES ACTES MEDICAUX (CdAM)

The procedures contained in the RUM are coded using the Catalogue des Actes Médicaux (CdAM). First published in 1985, the CdAM was developed by committees of medical experts coordinated by the Direction des Hôpitaux to meet two objectives:
- to identify procedures carried out during the patient's stay;
- to measure the consumption of human and material resources required to perform those procedures, using an ICR.

The CdAM is currently organised into five fields, each corresponding to a medico-technical field identified by a letter of the Greek alphabet:
- Alpha: diagnostic and therapeutic procedures,
- Beta: anaesthesiology procedures,
- Gamma: medical imaging procedures,
- Mu: radiotherapy procedures,
- Omega: intensive care procedures.

2.2 – THE INDICE DE COUT RELATIF (ICR)

The ICR or relative cost index was produced during the development of the hospital cost accounting model used for the Etude Nationale de Coûts. It is an accounting unit of work which can be used to allocate the charges for medico-technical procedures between hospital stays.

The ICR expresses the level of use of human and material resources.

It is a relative index: if one procedure carries an ICR of 100 and another of 50, this means that the first consumes on average twice as many human and material resources as the second.

ICRs were determined under standard optimal conditions qualified as "best practice" by doctors of different disciplines, nursing managers and nurses, specialists in accounting and biomedical equipment in public and private hospitals. An ICR is in fact the sum of three sub-indexes:
- an Indice d'Activité Médicale (IAM) or medical activity index which reflects the mobilisation of doctors: their number, qualification and the duration of the procedure;
- an Indice d'Activité Soignante (IAS) or nursing activity index reflecting the same factors for non-medical nursing staff;
- an Indice de Consommation de Ressources Matérielles (IRCM) or material resources consumption index, calculated on the resources (equipment, consumables, etc.) needed for the procedure.

The Alpha field was the first for which ICRs were calculated. The IAMs were determined in the following stages:
- selection of reference procedures, known as benchmark procedures: these are procedures performed frequently and well standardised in technical terms;
- identification of the resources needed, according to IAM components, to perform the benchmark procedures;
- determination of the IAM of other procedures on the basis of the benchmark procedures.

Once the IAM for each procedure was calculated, the IAS and ICRM were added to it. Calculating the overall ICR then consists of a monetary valuation which takes into account each category of staff and equipment used. The three sub-indices evaluated in francs were then added together and their sum divided by 25 to produce easily usable figures.

The maintenance of the CdAM is the responsibility of the Pôle d'Expertise et de Référence Nationale des Nomenclatures de Santé (PERNNS), but the catalogue is now in the final years of its existence before being replaced by the CdAM.

# APPENDIX IV: Guidelines for the evaluation of doctors' professional expenses when setting tariffs for technical procedures

## 1 – GENERAL PRINCIPLES

A new catalogue of all technical procedures in the medical and surgical specialties has been in preparation since 1996: the CCAM.

Its objectives are twofold:
(a) to construct a list of coded headings, common to both public and private sector, that will replace those of the Catalogue des Actes Médicaux (CdAM) and the Nomenclature Générale des Actes Professionnels) (NGAP);
(b) to restructure medical fees on a global and synchronous basis in order to produce a tariff classification that is coherent and "neutral" (i.e. free of major distortions in relation to the reality of medical work and the resources used for the performance of each procedure).

The "Pôle nomenclature" [nomenclature department] of the CNAMTS has developed a methodology for the purpose, adapted from the work of the Harvard public health department in Boston[82], and based on a ranking of procedures according to the resources used by the practitioner. These are of two types: firstly the medical work carried out by the practitioner, second the professional expenses incurred by the practitioner himself, known as the "costs of practice".

The medical work is evaluated by panels of experts from scientific bodies, using a ranking method which assigns each procedure a relative value (or score) expressed in "work points" and combining indicators of duration, stress, technical skill and mental effort. This ranking is carried out in two stages: firstly within each discipline studied independently from the others, and then between disciplines in order to identify inter-disciplinary cross-over points or equivalences that will lead to a single scale of medical work.

The cost of practice is estimated directly in francs, on the basis of information gathered from the Direction Générale des Impôts (DGI) [the French tax department], from practitioners and from manufacturers. The evaluation method is described in this publication.
The work score is expressed as a number of points. To convert this into a monetary value, a coefficient must be calculated, known as the monetary conversion factor (Fc), which will then act as a unique key-letter for medical work. The conversion factor is evaluated as follows.

The starting point is the total budget envelope allocated to the remuneration of technical procedures by all specialties. This envelope is by definition equal to the sum of the remuneration of medical work plus the cost of practice. The total remuneration of medical work is thus equal to the total budget envelope allocated to technical procedures less the expenses corresponding to technical procedures.

If we know the total number of work points carried out, calculating the conversion factor is easy: it is equal to the total remuneration of medical work divided by the number of work points.

The formula for Fc is therefore:

$$Fc = \frac{[\text{Total envelope} - \text{Total expenses}] \text{ of technical procedures}}{\text{Sum of work points for technical procedures}}$$

The fee for the procedure is then estimated as the sum of the price of work (i.e. work score obtained from the ranking, multiplied by the monetary conversion factor) plus the cost of practice:

*Procedure fee = [Work score x Fc] + Cost of practice for the procedure*

The whole approach requires to know the frequency of technical procedures. For surgical procedures, this will be evaluated on the basis of data from the private PMSI, and for outpatient procedures from a specific survey of practitioners carried out by the Echelon National du Service Médical of the CNAMTS[83].

All work on evaluating medical work and the cost of practice will form the basis for tariff negotiations with professional associations, which will be held prior to any setting of tariffs.

[82] Hsiao W.C., Braun P,, Dunn D., Becker R. "Resource-based Relative Values: an overview", JAMA, 1988, 16: 2347-2443.
[83] This survey covers a sample of 50,000 procedures representative of the technical medical procedures carried out in the general/private outpatient sector, identified by CCAM headings.

A transition phase will allow for the gradual adaptation of the present tariffs to the target tariffs.

Once the CCAM is in place, a maintenance structure will be set up to monitor both developments concerning procedures already registered on the CCAM and the introduction of new procedures or techniques after validation by the ANAES.

## 2 – EVALUATING THE COST OF PRACTICE

The cost of practice covers the professional expenses incurred by the practitioner himself (personnel costs, rental charges, finance costs, etc.). It is estimated in francs per procedure for inclusion in the fee.

### 2.1 The context

*2.1.1) Fee-for-service*

The tariff restructuring of the catalogue is based on *fee-for-service*.

*2.1.2) A budget envelope to be allocated*

Since the total amount of the budget envelope allocated to fees is fixed (deducted from the ONDAM), the estimation of the cost of practice automatically affects the estimation of the monetary conversion factor: any over-estimation of the cost of practice for a discipline leads to an under-estimation of the monetary conversion factor (and vice versa) and introduces a structural distortion of fees between disciplines: the "extra" given to one discipline is taken from the fees for another discipline.

*2.1.3) A single tariff per procedure*

The tariff for a procedure will theoretically be the same irrespective of the region, the specialty and the setting (individual practice, group practice or private establishment). In the case of procedures common to several specialties, the cost of practice will be evaluated for each and then compared: if it differs significantly from one specialty to another, further studies will be required to identify the reasons for the difference. Where necessary, a compromise will need to be reached in order to maintain the principle of the single tariff.

*2.1.4) The restructuring of fees is carried out separately from that of the system of hospital remuneration*

The two restructurings are being managed by different agencies. In addition, the scenario for the reform of the hospital remuneration system has not yet been established. A clear identification of practitioners' real expenses corresponding to the current fee system was therefore essential. Further developments may be possible subsequently, but these will require the agreement of all parties concerned.

*2.1.5) A scope limited to technical procedures*

The CCAM concerns, for the moment, only part of the field of medical and dental activity, that of technical procedures. It does not include consultations. We will therefore estimate the portion of professional expenses corresponding to technical procedures (see §3.2).

### 2.2 The method

The option adopted consists of using observed expenses, with no standards-based theoretical judgement (see §3.1).

To take into account the fact that the rate and structure of professional expenses differ from one specialty to another, expenses are evaluated by discipline or group of disciplines. The evaluation method nonetheless calls for a cross-sector approach, i.e. one that is comparable for the collection of information and the analysis of the structure of expenses for each discipline: an evaluation based on data whose overall consistency cannot be guaranteed could lead to an over-estimation or under-estimation of the cost of practice for a particular discipline, which could lead to a biased estimation of the monetary conversion factor (see §0).

Two separate cases are distinguished in the evaluation of the costs of practice: the general case and the case of heavy equipment. In the general case, the tariff for the procedure does not depend on the volume of activity; in the second case, the remuneration of expenses per procedure varies on either side of a benchmark activity.

*2.2.1) The general case*

**General expenses**
For many procedures, the professional expenses concerned are essentially "general" expenses not specific to any particular activity within a single discipline (staff, rent, fees, etc.), or which relate to a specific activity practised by all doctors within the same specialty, such that it is possible to envisage distributing these expenses across all procedures proportionally to the number of work points.

**Excess costs**
Nevertheless certain specific procedures, not performed by all practitioners, may require specific and substantial investment by those who perform them (equipment, staff, etc.). It is preferable to allocate the specific excess cost to the corresponding procedure, when that excess cost is significant.

*Examples of procedures with and without excess costs:*
*in gastro-enterology, digestive endoscopy procedures will not be considered as excess-costs procedures since they are performed by all gastro-enterologists; in dermatology, procedures requiring the use of a laser will generate an excess cost since not all dermatologists are able to provide this technique.*

For each discipline, the general expenses are equal to the total expenses for all technical procedures (excluding consultations) minus the sum of excess costs and technical lump sum payments. The calculation of specific excess costs therefore only includes expenses not included in general expenses, and the excess cost is allocated directly to the corresponding procedure.

General expenses relating to the technical procedures performed by a discipline are allocated pro rata to the work points associated with each procedure. An amount of general expenses per work point ($Cg_D$) is calculated for each discipline by dividing the total amount of general expenses for the discipline by the total number of work points carried out by the discipline:

$$Cg_D = \frac{\text{General expenses for technical procedures within the discipline}}{\text{Sum of work points for technical procedures within the discipline}}$$

- *Cost of procedure practice without excess costs*
  The expenses for these procedures are remunerated proportionally to the medical work, after evaluating the total mean cost per work point of the general expenses. The cost of practice for procedure i in discipline D is therefore: $CP_{iD} = W_i + Cg_D$
  *where*
  *$W_i$ represents the medical work required to perform the procedure (this depends on the procedure, but not on the specialty: a single procedure performed in two specialties will have only one work score),*
  *$Cg_D$ represents the mean cost per work point of general (or usual) expenses for the discipline: it may differ from one discipline to another.*

- *Cost of procedure practice with excess costs*
  By definition, excess costs apply to procedures not performed by all practitioners and which require a specific and substantial investment by those that do perform them.
  The excess cost specific to each of the procedures concerned is assigned to its procedure. The cost of practice for procedure i in discipline D is therefore: $CP_{iD} = W_i + Cg_D + S_{iD}$
  *where $S_{iD}$ represents the excess cost specific to procedure i in discipline D.*

  For a procedure performed by several specialties, the cost of practice may differ from one specialty to another. A choice will therefore have to be made (see §1.3) to establish the amount of the cost of practice to be included in the tariff for the procedure.
  $H_i = W_i \times Fc + CP_i$
  *where Fc represents the monetary conversion factor (monetary value of a work point), which is identical whatever the specialty; and where $CP_i$ is the cost of practice which is the same for each procedure.*

*2.2.2) Special case of heavy equipment*

Heavy equipment (scanner, MRI, radiotherapy equipment, nuclear medical camera, digital angiography, etc.) is particularly costly equipment which can only be acquired and operated with regional or national authorisation. Such equipment is characterised by high fixed costs, with the total average cost of use decreasing significantly as the level of activity rises. Since the levels of activity are more often heterogeneous, taking these costs into account independently

of the activity may prove inflationary, as the return per procedure increases once the breakeven level of activity is passed[84]. Analysis of the structure of total expenses (fixed and variable) for such equipment[2] has produced a preference for a tariff system partially indexed on activity, such as is already applied to scanners and MRI equipment and which is envisaged for extension to all heavy equipment.

The remuneration of the medical work is dissociated from the operating costs.

The fee remunerates the medical work and is equal to: $H_i = W_i \times Fc$

Operating costs are remunerated by:

- a full-rate technical tariff up to a benchmark level of activity (which corresponds to the minimum level of activity we wish to define for each type of equipment): this is equal to the total mean cost[85] relative to this activity.
- a reduced technical tariff above this level of activity: this is equal to the mean cost of an extra procedure[86] above the benchmark level of activity.

## 2.3 Data required for the evaluation

The approach consists of five stages: an estimation of the total amount of expenses per discipline or group of disciplines, an evaluation of the amount of total expenses relating to technical procedures, an evaluation of each excess cost and then of the total amounts allocated to all excess-cost procedures, an evaluation of the technical lump-sum payments for heavy equipment and, finally, an evaluation of the amount of general expenses per work point.

### *2.3.1 Evaluation of total expenses*

Amounts of total expenses are evaluated by discipline or by group of disciplines on the basis of data supplied by the DGI. These data were selected on the grounds of their exhaustiveness and cross-sector compilation (perfectly comparable from one discipline to another): it meets the twofold constraints of cross-sector applicability and disconnection from the field of fees and that of the remuneration of institutions (see §2 and §1.4).

The data, sourced from tax declaration form 2035 (non-commercial profits), provides details of the expenses, under the headings available, for all professionals completing tax declarations in each specialty of the DGI structure. The series provided since 1993 and the inclusion of recent tax changes will make it possible to estimate the total expenses at the time the restructuring is applied. For certain disciplines, the 2035 form does not reflect total professional expenses, in particular where much of the activity is carried on through companies. In order to have access to all the expenses relating to the medical exercise of the discipline, it will therefore be necessary to adjust these data on the basis of complementary studies, particularly in the field of radiology. In addition, these data do not make it possible to identify expenses by convention sector. It will, however, be possible to recalculate all personal social security contributions as if all practitioners were subject to sector 1 (fee fixed by the Caisse Nationale d'assurance maladie), using individual data from the Système National Inter-Régimes (SNIR) and rules for the calculation of contributions relating to practice in sector 1.

### *2.3.2) Evaluation of expenses relating to technical procedures only*

The amount of expenses relating to consultations will be estimated using three independent approaches (see Questions & Answers no. 5), which will be compared for accuracy and results.
The total expenses relating to technical procedures will be estimated by the difference between total expenses (see §3.1) and expenses relating to consultations.

### *2.3.3) Evaluation of excess costs*

The excess cost of a procedure corresponds to the additional direct costs entailed by the practice of the technique concerned.

Direct expenses are estimated on the basis of data collected from manufacturers and practitioners.

---

[84] The breakeven level of activity is the level at which income and expense are equal.
[2] Aliès-Patin A. "Tarification des équipements lourds. Une allocation forfaitaire adaptée à la réalité des coûts de fonctionnement." J. Radiol. 1990 t7, no. 10, pp 561-569.
[85] The total mean cost for a given level of activity is the total cost corresponding to this activity (sum of total fixed costs and total variable costs) measured against the activity considered.
[86] The mean cost of an extra procedure between two levels of activity is equal to the extra cost divided by the extra activity between these two levels of activity.

Initially, a small number of practitioners in each discipline select the procedures or techniques which, in their view, carry an excess cost (procedures not performed by all practitioners of the discipline and requiring a significant specific investment).

A questionnaire is then sent to manufacturers identified as manufacturing the equipment concerned, and to practitioners. The questionnaire is designed to gather information on the price of the equipment and the total global direct cost necessary for its operation (maintenance costs, consumables). The direct excess per procedure using the technique is estimated on the basis of an average activity compatible with national frequency data.

*2.3.4) Evaluation of technical lump sum payments*

The expenses evaluated include all annual operating costs for the equipment concerned: cost of the equipment, cost of facilities, non-medical staff, consumables, management costs, tax, etc. They are divided into fixed and variable expenses.

Two tariffs are estimated: one, a full-rate tariff applicable up to a benchmark level of activity, the other a reduced rate applicable above this level of activity (see §2.2). The activities primarily concerned are radiotherapy, nuclear medicine, scanner, MRI and operative cardio-vascular radiology.
Data is collected by means of surveys of private and public sites.

*2.3.5) Evaluation of general expenses*

The amount of general expenses for technical procedures is evaluated for a discipline or group of disciplines by deducting expenses relating to excess costs and technical lump sum payments from the total expenses relating to technical procedures.

## Questions & Answers

**1. Why use current observed costs and not standardised costs?**
A "standardised" option would have involved building operating accounts for “typical” technical platforms defined by best practice. While this method may appear more satisfactory in principle, it would have involved the participation of many different parties (doctors, establishments, accreditation bodies) and would thus have been extremely burdensome to set up. More importantly, however, this is an option which encourages overstatement and may therefore, within a limited budget framework, create an imbalance favouring certain disciplines at the expense of others.

**2) Why not allocate the direct cost of the basic equipment found in a doctor’s office to the corresponding procedures?**
A cost accounting study generally distinguishes between two types of costs: direct and indirect. Direct costs are those which can be unambiguously attributed to an activity or product, while indirect costs derive from an allocation by distribution coefficient.

If we take the example of a hospital, the salaries of staff working in the medical department or the cost of drugs distributed in the department can be directly allocated to the patients hospitalised in the department. The allocation of a proportion of the hospital manager's salary, on the other hand, can only result from an agreed calculation in the form of a distribution coefficient, the economic relevance of which is obviously open to discussion.

Direct costs may be fixed or variable: in the above example, drugs represent a variable expense, i.e. an extra patient means extra cost. In contrast, salaries are fixed costs and the admission of one additional patient more or less does not cause them to vary (although obviously these expenses are only fixed up to a certain threshold: if the number of patients increases significantly, more staff will need to be recruited).

If we transpose these general concepts to the professional practice of the general/private practitioner, and attempt to allocate the expenses across the various procedures, it is clear that the bulk of these costs cannot be attributed specifically to such and such a type of procedure, but is allocated across the activity as a whole. This applied to building, furniture, secretarial services, management costs and most purchases of consumables.

In contrast, for equipment which is assigned exclusively to the performance of certain procedures, the question arises as to whether or not to evaluate their direct cost. The option adopted is not to use this type of evaluation for equipment constituting basic equipment for all practitioners of a specialty.

We could, for example, seek to isolate the cost of mammography within a radiology practice and assign it solely to mammography procedures. This method presents a risk, however.

Since the cost of the mammography equipment is made up of fixed costs, the unit cost of a procedure depends directly on the volume of procedures performed. We then find ourselves in a classic situation whereby, if we include in the remuneration of a technical procedure a flat amount for equipment depreciation, maintenance, etc. - an amount which does not vary however much the equipment is used - we create an economic incentive to increase the volume of procedures.

To avoid this distortion, it is preferable to add together all the usual expenses for the profession, including those relating to medical equipment, and allocate them over all procedures proportionally to work points. We lose something in the accuracy of cost attribution, but we avoid creating incentives for certain types of procedure. In addition, by not multiplying our cost calculations by type of procedure, we also avoid multiplying the approximations and uncertainties inherent in this type of calculation.

**3) Why evaluate direct costs for "excess cost" equipment and heavy equipment?**
"Excess cost" procedures are procedures which require specific costly equipment and which are not commonly practised by all practitioners of the discipline. This definition implies that the costs specific to the performance of these procedures cannot under any circumstances be allocated over all the procedures performed in the discipline. Such an allocation would lead to an inadequate remuneration of "excess cost" procedures, thereby penalising practitioners who perform them and thus creating a disincentive.

In such cases, it is necessary to evaluate the direct costs associated with the equipment used for such procedures. But how, then, do we avoid the incentive to increase the number of procedures, as described above? There are two solutions:

- introducing a modulation of the remuneration *ex ante* according to volume, which is the solution applied in the case of technical lump-sup payments (see §2.2);
- introducing a feed-back system, based on observation of practice, which enables the remuneration per procedure to be adjusted when the number of procedures per practitioner increases. It will be important to ensure a regular check (at intervals of less than one year) on excess-cost procedures in order to modify the value of the excess cost as necessary in the light of real activity observed.

**4) Why use work as the key for allocation of general expenses?**
Medical work is a cross-sector allocation key, comparable from one discipline to another, that reflects medical activity in its entirety and whose consistency across all the disciplines is guaranteed because of the inter-disciplinary ranking stage.

**5) How do we evaluate expenses relating to consultations?**
A number of approaches will be examined:

- a "semi-standardised" approach from which it will be possible to reconstitute the proportion of activity assigned to consultations, and hence the appropriate proportion of expenses, from a standard number of consultations per hour for highly clinical disciplines, and from an operating account compatible with the DGI structure of expenses;
- a "DGI" approach which involves evaluating the cost of a consultation in highly clinical disciplines (GPs, paediatricians) by dividing DGI expenses (less the proportion of expenses corresponding to the limited number of procedures performed by the discipline) by the total number of consultations in the discipline;
- an approach which allocates the general expenses of the discipline pro rata to the fees of each group of procedures in the discipline: consultations and technical procedures.

**6) Why not review the remuneration of certain procedures carried out in hospitals?**
The separation between fees and the remuneration of hospitals is a major constraint. Envisaging a change to the setting of tariffs for certain procedures is not out of the question, however:
- procedures which can be carried out in the doctor's office or in an institution (endoscopy, for example) and which in the latter case generate FSO or FSE paid to the hospital in addition to the fee;
- operative cardio-vascular radiology procedures for which all costs could be globalised by means of a technical lump-sum payment.

A modification of this type requires consensus between all parties concerned (state, Assurance Maladie, representatives of practitioners and private establishments) and will call for an evaluation of expenses and accurate re-allocation to practitioners or to establishments in accordance with the options adopted.

EXAMPLE OF AN ESTIMATION OF FEES

The following example is theoretical and the data are entirely fictitious.

**1) Data**

Let us suppose that:

- there are 2 specialties, Sa and Sb, each of which performs 5 different procedures;
- the total budget envelope allocated to these specialties is 3,000 MF, of which 2,550 MF for technical procedures;
- the number of consultations provided by Sa is 2,400,000 and by Sb 600,000;
- the total expenses for Sa estimated on the basis of DGI data are 650 MF;
- the total expenses for Sb estimated on the basis of DGI data are 850 MF;
- the amount of expenses per consultation is 55 F;
- procedures a1, b2 and b4 generate an excess cost.

The frequency of procedures, the work scores and the excess cost per procedure are given below:

| Specialty Sa | | | | | Specialty Sb | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Procedure | Number of procedures | Work score | Number of work points | Excess cost FF/proce-dure | | Number of procedures | Work score | Number of work points | Excess cost FF/proced ure |
| A1 | 8 000 | 600 | 4 800 000 | 250 | b1 | 120 000 | 300 | 36 000 000 | 0 |
| A2 | 70 000 | 300 | 21 000 000 | 0 | b2 | 90 000 | 190 | 17 100 000 | 140 |
| A3 | 200 000 | 130 | 26 000 000 | 0 | b3 | 900 000 | 50 | 45 000 000 | 0 |
| A4 | 250 000 | 90 | 22 500 000 | 0 | b4 | 60 000 | 35 | 2 100 000 | 90 |
| A5 | 600 000 | 60 | 36 000 000 | 0 | b5 | 1 500 000 | 20 | 30 000 000 | 0 |
| TOTAL | 1 128 000 | - | 110 300 000 | - | TOTAL | 2 670 000 | - | 130 200 000 | - |

**2) Estimation of expenses related to technical procedures**

The amount of expenses relating to technical procedures in each specialty are estimated by deducting the amount of expenses for consultations from the total expenses.

Expenses corresponding to consultations come to:
for Sa: 132 MF (2,400,000 consultations x 55 F),
for Sb: 33 MF (600,000 consultations x 55 F).

Expenses relating to technical procedures are therefore:
for Sa: 518 MF (650 - 132),
for Sb: 817 MF (850 – 33).

The total amount of excess costs is:
for Sa: 2 MF (8,000 procedures x 250 F),
for Sb: 18 MF (90,000 procedures x 140 F + 60,000 procedures x 90 F).

The amount of general expenses for technical procedures is therefore:
for Sa: 516 MF (512 – 8),
for Sb: 799 MF (817 – 18).

**3) Estimation of the monetary conversion factor and the cost of general expenses per work point (for technical procedures)**

The amount allocated to remuneration of the medical work for these two specialties is equal to the budget envelope allocated to technical procedures, less the cost of practice relating to technical procedures, i.e. 1,215 MF (2,550 – [518 + 817].

The monetary conversion factor (price of the work point) is therefore equal to the amount allocated to work (1,215 MF) divided by the total quantity of work (110,300,000 + 130,200,000 = 240,500,000 work points), i.e. 5.05 F.

**Cost of general expenses per work point**
The amount allocated to the remuneration of general expenses for each of the specialties is equal to the amount of expenses divided by the quantity of work in the speciality:

Cg is equal to 4.68 F (516 MF divided by 110,300,000 work points)
Cgb is equal to 6.14 F (799 MF divided by 130,200,000 work points).

**4) Fees**

Reminder: fees are the sum of the price of work, the price of general expenses and of any excess cost. The formula for fees is: W x Fc + x Cg + S.

The monetary conversion factor is 5.05 F, the "general" cost of practice (excluding excess costs) is 4.68 F per work point for Sa and 6.14 F per work point for Sb, procedure a1 generated a specific excess cost of 250 F, procedure b2 an excess cost of 140 F and procedure b4 an excess cost of 90 F. The resulting fees are shown in the table below:

| Specialty Sa | | | | | | Specialty Sb | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Procedure | Work score | Price of work (1) | Price of general costs (2) | Excess cost FF/procedure | Fees | Procedure | Work score | Price of work (1) | Price of general costs (2) | Excess cost FF/procedure | Fees |
| a1 | 600 | 3031 | 2807 | 250 | 6088 | b1 | 300 | 1516 | 1841 | 0 | 3357 |
| a2 | 300 | 1516 | 1403 | 0 | 2919 | b2 | 190 | 960 | 1166 | 140 | 2266 |
| a3 | 130 | 657 | 608 | 0 | 1265 | b3 | 50 | 253 | 307 | 0 | 560 |
| a4 | 90 | 455 | 421 | 0 | 876 | b4 | 35 | 177 | 215 | 90 | 482 |
| a5 | 60 | 303 | 281 | 0 | 584 | b5 | 20 | 101 | 123 | 0 | 224 |

(1) Work score x 5.05 F (= monetary conversion factor)
(2) Work score x 4.68 F (= general expenses per Sa work point)
(3) Work score x 6.14 F (= general expenses per Sb work point).

## APPENDIX V: Order of 29 July 1998 concerning the collection, processing and transmission of data on medical activity

Ministry of Employment and Solidarity

Order of 29 July 1998 concerning the collection and processing of data on medical activity, as referred to in article L.710-6 of the code of public health, by public health establishments and private establishments funded by global allowance as referred to in article L.710-16-1 of the said code of public health, and of the transmission, as referred to in article L.710-7 of the code of public health, to the regional hospital agencies and to the state of information arising from such processing.

The Minister for Employment and Solidarity
The Secretary of State for Health

Pursuant to the code of public health, notably articles L.710-2, L.710-6, L.710-7, R.710-5-1 to R.710-5-1 1, R.712-57 to R.712-59;

Pursuant to the social security code, notably article L.162-30-1;

Pursuant to the act 78-17 of 6 January 1978 on data protection;

Pursuant to decree no. 78-774 of 17 July 1978 as amended by decrees 78-1823 of 28 December 1978 and 79-421 of 30 May 1979;

Pursuant to decree 95-1000 of 6 September 1995 on the code of medical ethics;

Pursuant to the opinion of the information systems commission on health care establishments of 24 June 1998;

Pursuant to the opinion of the national council of the order of medicine of 3 July 1998;

Pursuant to the opinion of the data protection commission of 16 July 1998;

Hereby decree:
Article 1 – I – For the purposes of carrying out medico-economic analysis of the treatment activity they perform, health care establishments funded by global allowance and listed in article 710-6-1 of the code of public health will introduce computer processing of the following named patient data: Résumés Hebdomadaires Standardisés (RHS) [standardized weekly abstracts], including sections identifying patients and tracking their movements in medical units, for all patients treated in medical units providing follow-up or rehabilitation treatment (SSR) under full hospital admission, admission by week, day or night, and all outpatient treatment. The definition of medical units is a matter for each health care establishment to determine.

II – The implementation of such computerised processing must be preceded by submissions for approval made by the health care establishments themselves to the Commission Nationale de l'Informatique et des Libertés (CNIL) [data protection agency].

III – The health care establishments will take all necessary steps to enable patients to exercise their right of access and rectification as laid down in articles 34 and 40 of the abovementioned data protection act of 6 January 1978, by addressing the medical officer in charge of medical information through the agency of the practitioner having created their file.

IV- After being anonymised, some of the information contained in the RHS is communicated, under the condition set out in article 6 below, to the DRASS on behalf of the Agences Régionales de l'Hospitalisation (ARH). This information is communicated in the form of Résumés Hedbomadaires Anonymes (RHA) and of half-yearly series of RHA (SSRHA), as described in article 5.

Article 2. – I – In each SSR medical unit, the information categories recorded in the RHS are as follows:
1) Information on patient identification:
- patient's date of birth,
- patient's sex,
- patient's admission number,
- postcode for patient's place of residence,
- SSR stay number: an identifier corresponding to the entire stay in the SSR medical units of the health care establishment;

2) Other mandatory information:
- number of the health care establishment in the Fichier National des Etablissements de Santé et de Soins national file (FINESS),
- number of the medical unit,
- type of activity describing the type of patient treatment,
- week number,
- days of patient presence,
- date of last surgical procedure (if applicable),
- ultimate objective of treatment,
- principal symptom,
- etiological condition (if different from principal symptom),
- significant associated diagnosis/diagnoses (if applicable),
- medical procedures on a closed list,
- scoring of patient autonomy in dressing, moving and walking, feeding, continence, behaviour, relational capacity;
- use of wheelchair,
- weekly therapist time (mechanical, sensory/motor, neuropsychological, cardio-respiratory, nutritional, uro-sphincteral) spent on rehabilitation, fitting equipment, assessments, physiotherapy, balneotherapy, and group rehabilitation,
- solely for patients fully hospitalised or hospitalised by the week:
    - date and method of patient admission to the SSR medical unit,
    - date and method of patient discharge from the SSR medical unit,
    - in the event of admission on transfer, provenance of patient,
    - in the event of discharge on transfer, destination of patient.

This information must correspond to the content of the medical record.

II – As an exception to the above, if the patient has been treated anonymously the identity information is restricted to year of birth, sex, admission number and patient SSR stay number. U1.

III – Under the terms of the rules approved by the establishment's medical committee or council, and after consulting the CNIL, other information, particularly of a medical nature, may also be recorded by a medical unit or by the establishment.

Article 3. – 1 – Several RHSs may be produced successively during the course of the patient's stay, each being classified in a Catégorie Clinique Majeure (CMC) and in a Groupe Homogène de Journées (GHJ).

II – The RHSs are grouped at the level of the legal entity for public sector establishments and at establishment level for those in the private sector except, where applicable, for structures which are authorised by the ARH to collect and group RHSs by establishment or hospital group.

III – Morbidity variables (ultimate objective of treatment, main symptom, etiological condition and significant associated diagnoses) are coded according to the World Health Organisation (WHO) International Classification of Diseases, supplemented where necessary by extensions published in the *Bulletin Officiel du Ministère chargé de la Santé* [Official Bulletin of the Ministry of Health]. Rehabilitation activities are recorded according to the catalogue of rehabilitation activities published in the *Bulletin Officiel*. Dependence variables are coded according to a dependence scale developed specifically for the compilation of RHSs, set out in guidelines for the production of RHSs published in the *Bulletin Officiel*. Medical procedures are recorded and coded according to a closed list of procedures published in the same guidelines.

VI – Guidelines published in the *Bulletin Officiel* set out the methods for production, coding and grouping of RHSs. The CMCs and GHJs are listed and described according to the classification set out in the *Bulletin Officiel*.

Article 4. – 1 – In accordance with the provisions set out in articles R.710-5-1 to R. 710-5-1 1 of the code of public health, the medical officer in charge of medical information for the health care establishment is responsible for setting up, on the basis of information supplied to him, a file for RHSs, for the implementation of their grouping into GHJs, and for the processing of this data.

In accordance with the conditions agreed on consultation of the establishment's medical commission or medical council, this officer is responsible for circulating the information to the management of the health care establishment and to the president of the medical committee or council of the establishment, as well as to the practitioners having provided the treatment, under conditions which guarantee the confidentiality of data and the anonymity of patients. He is informed of the purposes of the information processing which he is required to perform and participates in the interpretation of its results.

II – The minimum retention period for the half-yearly files of grouped RHSs is four years. Files of grouped RHSs retained for control and inspection purposes in the establishment may contain only data on hospital stays where the date of the Monday of the observation week is more than 5 years prior to the date of 31 December of the current year.

Article 5 – The grouped RHS files are used, under the responsibility of the medical officer referred to in article 4, to create RHA files and half-yearly series of RHA files (SSRHA). The SSRHA contain information which supplements that provided by the RHAs and gives an overview of the progress of a stay over the half-year. The RHA and the SSRHA are generated by software proprietary to the state and contain;

- no patient admission number,
- no SSR stay number (replaced by a sequential stay number) or medical unit number (only the number of medical units attended in the course of the half-year),
- no date of birth, just age in years (or in days for infants under one year old) and calculated as of the Monday of the observation week, nor postcode which is replaced by a geographic code assigned from a list approved at national level in agreement with the CNIL,
- no admission or discharge dates, in the case of full hospitalisation or hospitalisation by the week: these are replaced in the RHA by specification of a week of start of SSR stay (yes/no), of a week of end of SSR stay (yes/no) and the length of SSR stay,
- no week number, replaced in the RHA by a sequential week number and by the month and year,
- no date of last surgical procedure, replaced by reference to a previous surgical procedure in the last 3 months (yes/no),
- no separate dependency scores for dressing, moving and walking, feeding, continence: replaced by a physical dependency score,
- no separate dependency scores for behaviour or relational capacity: replaced by a behavioural and relational dependency score,
- no reference to the use of a wheelchair.

II – The director and the president of the establishment's medical committee or council are provided with the statistics, aggregated by medical unit and for the entire structure or, on request, with the RHA files, SSRHA files or other files containing individual patient data previously anonymised.

Article 6. – I – For each half-year period, the health care establishment forwards the RHA and SSRHA files to the DRASS on behalf of the ARH. These files are produced by the latest version of the RHA and SSRHA generator referred to in article 5 above, and are transmitted by means of a magnetic medium approved by the government departments responsible for processing the files, no later than three months from the end of the half-year concerned.
II – Within a month of receipt, the information is forwarded to the ministries of health and social security.
III – Within each health care establishment, the medical officer in charge of medical information is responsible for making a back-up copy of the grouped RHS file from which the RHA and SSRHA files are produced, and for retention of the copy.

Article 7 – Medical inspectors of public health and medical advisors to Assurance Maladie bodies have right of access, via the medical officer in charge of medical information, to the RHS files. As part of the data validation procedures, the medical officer advises the practitioners concerned prior to any comparison of an RHS against a medical record.

Article 8 – The director of hospitals at the ministry of employment and solidarity is responsible for the execution of this decree, which will be published in the Official Journal of the French Republic.

Given in Paris on 29 July 1998

For and on behalf of the minister and secretary of state and by delegation
The Director of Hospitals

Edouard Couty

## APPENDIX VI: Transportation costs

National health insurance bodies cover the transportation costs for the insured and dependents requiring transportation in order to receive the treatment and undergo the examinations appropriate to their state of health.

The medico-administrative conditions for the coverage of costs of medical transport (ambulance, light medical vehicle) and non-medical transport (taxi, public transport, personal vehicle) are restrictive and laid down in regulations.

The coverage of transportation costs is subject to the insured presenting a medical prescription attesting that the patient's state of health justifies the use of the prescribed means of transport. Reimbursement is made on the basis of the least costly means of transport compatible with the patient's condition. The amount varies according to the distance and the means of transport used. The reimbursement rate ranges from 65% to 100% (100% for beneficiaries of the Alsace-Moselle local insurance scheme). Transportation relating to a long-term illness (radiotherapy, chemotherapy, dialysis, etc.) is 100% covered as long as the prescription stipulates that the transport is related to the disease. If the patient's health state requires a series of transports, the reimbursement of expenses incurred is subject to two conditions:

1. The medical prescription must consist of at least 4 journeys within a period of two months,
2. Each one-way journey must be over 50 km.

The travel costs of an accompanying person may be reimbursed when the journey takes place on public transport (train, bus, etc.) and when the condition or age of the patient requires the presence of a third party.
An indemnity for loss of income may be paid when the journey by the accompanying person entails an interruption to their normal work.

When a private car is used, for transport eligible for reimbursement by the Assurance Maladie, the reimbursement allowance is calculated against a scale established by ministerial decree (1/7/99). For journeys of up to 200 km, the rate is:

- 0.20 EUR for vehicles of a fiscal horsepower of 5 CV and under,
- 0.24 EUR for vehicles of a fiscal horsepower of 6 and 6 CV,
- 0.27 EUR for vehicles of a fiscal horsepower of 8 CV and over.

Source: ministerial decree of 01/07/99, CNAMTS website, www.cnamts.fr/ass/remb/somremb.htm

In addition, the Convention Nationale [national agreement] established to govern relations between private ambulance companies and the Assurance Maladie payment centres stipulates that (endorsement no. 2 published in the *Journal Officiel* of 15 June 2002):

For ambulances:

- The tariff per kilometre for ambulances is increased to 2 EUR,
- A unified value of four departmental flat-rate tariffs is increased to 46.5 EUR,
- A unified value of urban flat-rate tariffs for zones B, C and D is increased to 52 EUR (this flat-rate tariff includes the first 5 kilometres),
- A revaluation of the zone A urban flat-rate tariff is increased to 54 EUR,
- A supplement of 20.43 EUR may be paid for emergency transport carried out at the request of a dispatching doctor (centre 15) or a medical emergency service (SAMU),
- A supplement of 10.21 EUR may be paid for the transport of premature babies or if an incubator is used,
- A supplement of 20.43 EUR may be paid for each journey when the patient is transported to or collected from an airport or railway station.

For light medical vehicles:

- The rate per kilometre is 0.78 EUR (0.79 in Corsica),
- The departmental flat rate carries the following maximum values: Zone A: 11.81 EUR, Zone B: 11.47 EUR, Zone C: 10.75 EUR, Zone D: 10.21 EUR.

Source: endorsement no. 2 to the national convention of private transporters, *Journal Officiel* of 15 June 2002/*Bulletin Officiel* no. 2002-24.

N.B. The classification of departments used as the basis for tariffs applicable to approved transporters is available on the following website: www.sante.gouv.fr

## APPENDIX VII: Utility and expected utility

1 - UTLITY

Let (x1, x2, x3, … xn) be elements of a series from which an agent with preferences for these elements is required to make a choice (preferring x2 to x1, for example).

We call utility a number u(x1) attached to the element x1 such that u(xi) is greater than or equal to u(xj) if and only if the agent prefers xi to xj (or is indifferent between xi and xj). The utility coefficients thus give a quantitative representation of the agent's preferences.

The purpose of utility theory is to explicit the conditions which these preferences must meet in order for us to express them in such figures.

Two conditions are essential: (1) the condition of totality whereby the agent is able to express a preference (or an indifference) between any pairing of elements in the series chosen; (2) the condition of transitivity whereby if the agent prefers xi to xj and xj to xl, then he prefers xi to xl. If these conditions (plus several others of a more technical nature) are met, we demonstrate that the preferences may be represented by utilities. This result is of more theoretical than practical interest. We demonstrate that there is in fact an infinity of sets of coefficients that may give a valid representation of the same system of preferences. As there is no reason to choose any one over any other, the practical scope of the result is very limited. In this respect, the vNM expected utility theory is more precise (see below).

We will note that the two conditions above are axioms presented without demonstration. Strictly, we should verify their pertinence before applying the theory. There are two ways of avoiding this digression. One is the standardised manner according to which any rational individual should behave according to the axioms. The second is the positive manner according to which individuals in general behave more or less according to the axioms.

2 – EXPECTED UTILITY

The expected utility theory attempts to formalise the conditions of choice under uncertainty (or, more exactly, under probabilised risk). It considers that the elements x1…xn between which the agent's choice must be made (see above) are "random perspectives" or "gambles". Each element of choice (a surgical treatment, for example, which may – or may not - be preferred to a medical treatment) is associated with a list of possible outcomes plus the probability of their occurrence: for example, the operation has a 98% chance of succeeding and a 2% probability of leading to severe complications.

We suppose that the agent is not interested in the event as such (the operation) but solely in its outcome (to be cured or to suffer from complications). We also suppose that the agent expresses preferences (utilities) as regards the outcomes (he would prefer to be cured, for example, rather than to suffer complications) and that he is aware of the probabilities.

The theory of expected utility says that the utility of the event is equal to the utility of each of its outcomes weighted by the probability of their occurrence. We can thus compare the utility of surgery to that of medical treatment, as long as we know the utility of knowledge. An important result is that the utilities so calculated are defined to the nearest linear transformation (affine). In means that we can calculate utilities for all sorts of events as long as we establish by convention the utility of two reference states. If, by convention, we say that the utility of a cure is worth 1 and the utility of complications is worth 0, then the utility of surgery is worth 0.8 (i.e. 80% x 1 + 20% x 0).

The theory attracts criticism on two levels. The first, and most fundamental, relates to the highly burdensome and complex axiomatic structure which is imposed on the agent in order to arrive at the result. This structure has been disputed from the outset, notably by Maurice Allais in France. The second and less fundamental level relates to the reference states, the definition of which determines the results. There is no point comparing utilities which have been calculated using different references (see following note).

## APPENDIX VIII: Discounting to present value

We are envisaging the application of a health care programme which may be considered as a marginal transformation of the economy. The application of this programme has impacts over two consecutive years, designated 0 and 1, in terms of costs on the one hand, and of health outcomes on the other.

If we take: **dh** (differential of h), the quantity of health outcomes generated by the programme (for example, h is a number of life-years gained);
**P**, the collective marginal willingness to pay for the health outcome (aggregate of individual marginal willingness to pay);
**Cm**, the marginal cost of production of the programme,
**i**, the real annual discount rate of the monetary unit (all monetary values are expressed in constant currency of year 0),

We can express the decision criterion in favour of the realisation of the programme as follows:

$$P_0 dh_0 - Cm_0 dh_0 + (P_1 dh_1 - Cm_1 dh_1) / (1 + i) > 0 \qquad [1]$$

If the preference for health over other economic goods alters over time, we may take:
P1 = P0 (1 + r) where r is the annual rate of growth in collective marginal willingness to pay for the health outcome. In our developed countries, it is generally accepted that r is positive.

Relating this to [1], we obtain:

$$P_0 [dh_0 + dh_1 (1 + r) / (1 + i)] > Cm_0 dh_0 + (Cm_1 dh_1 / (1 + i)]$$

The second side of the equation is equal to the discounted cost of the programme, which we can write more simply as dC.

The first side of the equation may be written as $P_0 dH$ where $dH = [(dh_0 + dh_1 / ((1 + i) / (1 + r))]$

The criteria then becomes: $P_0 > dC / dH$

In the second side we recognise the cost-effectiveness ratio of the programme, in which the cost of year 1 has been discounted at discount rate i and the health outcome in year 1 at the rate i' such that:

$$1 + i' = (1 + i) / (1 + r).$$

There are three cases to distinguish:
**if r = 0,** P is stable over time, i = i' and the two rates are equal
**if r < 0**, P diminishes over time, i > i' and the discount rate on the health outcome is higher than on the monetary unit,
**if r > 0**, P increases over time, i < i' and the discount rate on the health outcome is lower than on the monetary unit,

Even if P increases over time, this change is probably slow and may justify us in considering that the two discount rates are equal when working over relatively short periods.

# BIBLIOGRAPHY

## I) Study viewpoint and perspective

Recommendation n° 1


- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Methods for the Economic evaluation of Health care Programmes*. 2nd ed. Oxford : Oxford University Press, 1997, p. 17-25.
- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Méthodes d'Évaluation Économique des Programmes de Santé*. 2ème ed. Paris : Economica, 1998, p. 22-30.


## II) Timing of studies over the life of a product

Recommendation n° 2


- LAUNOIS R. L'évaluation socio-économique une discipline entre recherche clinique et études observationnelles. *Bulletin du cancer*, 2003, Vol. 90, n°1, pp.97-104.


## IV) Environment studies

Recommendation n° 4


- BERESNIAK A, DURU G. *Approche des coûts par maladie*, in : Économie de la Santé. Paris : Masson, 2001, pp.78-80.


## VI) Choice of comparator and comparison methods

Recommendation n° 6 and 7


- GOLD Marthe R., SIEGEL Joanna E., RUSSELL Louise B., et al. *Cost-Effectiveness in Health and Medicine*. Oxford : Oxford University Press, 1996, p.63-66.


## VII) Types of Study

Recommendation n° 8


- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Methods for the Economic evaluation of Health care Programmes*. 2nd ed. Oxford : Oxford University Press, 1997, p. 11-17.
- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Méthodes d'Évaluation Économique des Programmes de Santé*. 2ème ed. Paris : Economica, 1998, p.16-22.


## VIII) Time frame and transferability to other countries

Recommendation n° 9 and 10


- DRUMMOND MF, BLOOM BS, CARRIN G, HILMAN AL, HUTCHINGS HC, KNILL-JONES RP, DE POUVOURVILLE G, TORFS K. Issues in the cross-national assessment of health technology. *International Journal of Technology Assessment in Health Care,* 1992, Vol. 8, n°4, pp. 671-82.
- DRUMMOND Michael, PANG Francis. *Transferability of economic evaluation results.* In : DRUMMOND Michael, MCGUIRE Alistair. Economic Evaluation in Health Care – Merging theory with practice. Oxford : Oxford University Press, 2002.
- GREINER W., SCHÖFFSKI O., SCHULEMBURG J.M. The transferability of international economic health-economic results to national study questions. *HEPAC,* 2000, Vol. 1, n°2, pp. 94-102.
- SPÄTH HM, CARRÈRE MO, FERVERS B, PHILIP T. Analysis of the eligibility of published economic evaluations for transfer to a given health care system. Methodological approach and application to the French Health Care System. *Health Policy*, 1999, Vol. 49, n° 3, pp.161-177.

## IX) Definition and measurement of cost

Recommendation n° 11 to 18


- DRUMMOND Michael, MCGUIRE Alistair. Economic Evaluation in Health Care – Merging theory with practice. Oxford : Oxford University Press, 2002, pp. 82-86
- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Methods for the Economic evaluation of Health care Programmes*. 2nd ed. Oxford : Oxford University Press, 1997, p. 55.
- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Méthodes d'Évaluation Économique des Programmes de Santé*. 2ème ed. Paris : Economica, 1998, p.62.
- DRUMMOND Michael, PANG Francis. *Transferability of economic evaluation results.* In : DRUMMOND Michael, MCGUIRE Alistair. Economic Evaluation in Health Care – Merging theory with practice. Oxford : Oxford University Press, 2002.
- GARATTINI Livio et al. How do Italian Pharmacoeconomists Evaluate Indirect Costs. *Value in Health,* 2000, Vol. 3, n°4, pp. 270-276.
- KOOPMANSHAP MA, RUTTEN F, VAN INEVELD BM, VAN ROIJEN L. T. The friction cost method for measuring the indirect costs of disease. *Journal of Health Economics*, 1995, Vol.14, n° 2, pp. 171-189.
- LAUNOIS R. Un coût, des coûts, quels coûts ? *Journal d'Économie Médicale* 1999, T. 17, n° 1, pp. 77-82.
- MIDY F, GRIGNON M. La notion de coût dans les évaluations médico-économiques. *Journal d'Économie Médicale,* 2002, Vol. 20, n° 2, pp. 105-117.


## X) Definition and measurement of outcomes

Recommendation n° 19 and 20


- JOHANNESSON M, JONSSON B, KARLSSON G. Outcome measurement in economic evaluation. *Health economics,* 1996, Vol. 5, n° 4, pp. 279-296.


Recommendation n° 21 to 24


- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Methods for the Economic evaluation of Health care Programmes*. 2nd ed. Oxford : Oxford University Press, 1997, p. 11-17.
- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Méthodes d'Évaluation Économique des Programmes de Santé*. 2ème ed. Paris : Economica, 1998, p.165-185.
- FAYERS P M, MACHIN D. *Quality of life : Assessment, Analysis and Interpretation*. Wiley and Sons, 2000, pp. 416.
- LAUNOIS R. La prise en compte des préférences des patients dans les choix de santé individuels et collectifs. *Revue d'Épidémiologie et de Santé Publique,* 1994, Vol. 42, n° 3, pp. 246-262.
- VAN ROIJEN L, ESSINK-BOT ML, KOOPMANSCHAP MA, BONSEL G, RUTTEN F. Labor and health status in economic evaluation of health care. The Health and Labour Questionnaire. *International Journal of Technology Assessment* in Health Care, 1996, Vol. 12, n° 3, pp.405-415.
- EQ5D : www.euroqol.org
- HUI 2 & 3 : www.healthutilities.com
- Questionnaires de qualité de vie en général : www.qolid.org


Recommendation n° 25


- BERESNIAK A., DURU G. *Economie de la Santé*. 5ème Ed, Paris : Masson, 2001, pp. 126-131 (Collection "Abrégés Connaissances et pratique").
- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Methods for the Economic evaluation of Health care Programmes*. 2nd ed. Oxford : Oxford University Press, 1997, p. 165-185.
- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Méthodes d'Évaluation Économique des Programmes de Santé*. 2ème ed. Paris : Economica, 1998, p.180-195.
- DURU G. *Utilité*. In : JESSUA C., LABROUSSE C., VITRY D., et al. Dictionnaire des Sciences économiques, Paris : PUF, 2001, 1096 p.
- DURU G, AURAY J.P., BERESNIAK A., LAMURE M., PAINE A., NICOLOYANNIS N. Limitations of the Methods Used for calculating Quality-Adjusted-Life-Years Value. *Pharmacoeconomics,* 2002, Vol. 20, n° 7, pp. 463-473.
- FAYERS P M, MACHIN D. *Quality of life: Assessment, Analysis and Interpretation*. Wiley and Sons, 2000, 416 p.
- GOLD Marthe R., SIEGEL Joanna E., RUSSELL Louise B., et al. *Cost-Effectiveness in Health and Medicine*. Oxford : Oxford University Press, 1996, p. 84-94.
- HARRIS J. QALYfying the value of a life, *Journal of Medical Ethics*, 1987, Vol. 13, n° 3, pp. 117-123.
- KEENEY R.L., RAIFFA H. *Decisions with Multiple Objectives.* Cambridge : Cambridge University Press, 1996 (1ère edition, 1976).
- MEHREZ A., GAFNI A. Quality adjusted life years, utility theory and healthy-years equivalents, *Medical Decision Making*, 1989, Vol.9, n° 2, pp. 142-149.
- MEHREZ A., GAFNI A. Healthy years equivalent: how to measure them using the standard gamble approach, *Medical Decision Making*, 1991, Vol.11, n° 2, pp. 142-149.

- MOATTI J.-P., AUQUIER P., LE COROLLER A.-G., MACQUART-MOULIN G. QALYs or not QALYs : that is the question ? *Revue d'Epidémiologie et de Santé Publique*, 1995, Vol.43, pp. 573-583.
- PLISKIN J., SHEPARD D.S., WEINSTEIN M.C. Utility function for life years and health status. *Ops Res*, 1980, Vol.28, pp. 206-224.
- WILLIAMS A. Cost-effectiveness, is it ethical? *Journal of Medical Ethics*, 1992, Vol.18, n° 1, pp. 7-11.
- WILLIAMS A. QALYs and Ethics : A Health Economist's Perspective. *Social Science and Medicine*, 1996, Vol.43, n° 12, pp. 1795-1804.

Recommendation n° 26

- GREEN P., CARMONE F., WIND Y. Subjective evaluation models and conjoint measurement. *Behavioural Science,* 1972, Vol. 17, n° 3, pp. 288-299.
- LUCE D., TUKEY J. Simultaneous conjoint measurement: a new type of fundamental measurement. *Journal of mathematical psychology*, 1964, Vol. 1, pp. 1-27.
- MAAS A., STALPERS L. Assessing utilities by means of conjoint measurement: an application in medical decision analysis. *Medical Decision Making*. 1992, Vol. 12, n° 4, pp. 288-297.
- MARKHAM F.W., DIAMOND J.J., HERMANSEN C.L. The use of conjoint analysis to study patient satisfaction. *Evaluation & the Health Professions,* 1999, Vol. 22, n° 3, pp. 371-378.
- RATCLIFFE J. The use of conjoint analysis to elicit willingness-to-pay values. Proceed with caution? *International Journal of Technology Assessment in Health Care,* 2000, Vol. 16, n° 1, pp. 270-275.
- RYAN M. A role for conjoint analysis in technology assessment in health care? *International Journal of Technology Assessment in Health Care*, 1999, Vol. 15, n° 3, pp. 443-457.
- RYAN M., MCINTOSH E., SHACKLEY P. Methodological issues in the application of conjoint analysis in health care. *Health Economics*, 1998, Vol. 7, n° 4, pp. 373-378.
- RYAN M., SCOTT D.A., REEVES C., BATE A., VAN TEIJLINGEN E.R., RUSSELL E.M., NAPPER M., ROBB C.M. Eliciting public preferences for healthcare: a systematic review of techniques. *Health Technology Assessment*, 2001, Vol. 5, n° 5, pp. 1-186.

## XI) Discounting Costs and Benefits

Recommendation n° 27

- CAIRNS J.B. Discounting and health benefits: another perspective. *Health Economics*, 1992, Vol. 1, n° 1, pp. 76-79.
- DRUMMOND M.F., JEFFERSON TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. *British Medical Journal*, 1996, Vol. 313, n° 7052, pp. 275-283.
- DRUMMOND Michael, MCGUIRE Alistair. Economic Evaluation in Health Care – Merging theory with practice. Oxford : Oxford University Press, 2002, pp. 236-56.
- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Methods for the Economic evaluation of Health care Programmes*. 2nd ed. Oxford : Oxford University Press, 1997, pp. 107-109.
- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Méthodes d'Évaluation Économique des Programmes de Santé*. 2ème ed. Paris : Economica, 1998, p. 114-116.
- GAFNI A. Time in health: can we measure individuals' "pure time preferences"? *Medical Decision Making*, 1995, Vol. 15, n° 1, pp. 31-37.
- GOLD Marthe R., SIEGEL Joanna E., RUSSELL Louise B., et al. *Cost-Effectiveness in Health and Medicine*. Oxford : Oxford University Press, 1996, pp. 214-46.
- HILMAN A.L., KIM M.S. Economic decision-making in healthcare : a standard approach to discounting health outcomes. *PharmacoEconomics*. 1995, Vol. 7, n° 3, pp. 198-205.
- KATZ D.A., WELCH H.G. Discounting in cost-effectiveness analysis of healthcare programmes. *PharmacoEconomics*, 1993, Vol. 3, n° 4, pp. 276-285.
- KEELER E.B., CRETIN S., Discounting of life-saving and other non-monetary effects. *Management Science*, 1993, Vol. 29, n° 3, pp. 300-306.
- MacKEIGAN L.D., LARSON L.B., DRAUGALIS J.R., et al. Time preference for health gains versus health losses. *PharmacoEconomics*, 1993, Vol. 3, n° 5, pp. 374-386.
- OLSEN J.A. On what basis should health be discounted? *Journal of Health Economics,* 1993, Vol. 12, n° 1, pp. 39-53.
- PARSONAGE M, NEUBURGER H. Discounting and health benefits. *Health Economics*. 1992, Vol. 1, n° 1, pp. 71-79.

## XII) Modelling

Recommendation n° 28 and 29

- BECK J.R., PAUKER S.G. The Markov process in medical prognosis. *Medical Decision Making*, 1983, Vol. 3, n° 4, pp. 419-458.
- BECK J.R. Markov models of natural history. *Journal of Clinical Epidemiology,* 1988, Vol. 41, n° 7, pp. 619-21.

- CHIANG C.L. *An introduction to Stochastic Processes and Their Applications*. New York : Robert E. Krieger Publishing Co., Inc, 1980.
- COX D.R., MILLER H.D. *The Theory of Stochastic Processes*. London : Chapman & Hall, 1965.
- GRIMMETT G.R., STIRZAKER D.R. *Probability and Random Processes*. 2nd ed. Oxford : Oxford Science Publications, 1992.
- KEELER E. *Decision trees and Markov models in cost-effectiveness research*. In: Valuing health care : Costs, Benefits, and Effectiveness of Pharmaceuticals and Other Medical Technologies. Cambridge : Cambridge University Press, 1995, pp. 185 – 205.
- LAUNOIS R., GROUTSCHE J.J., MEGNIGBETO A., LE LAY K., PORTAFAX C. L'apport indispensable de l'épidémiologie clinique aux modèles de Markov. *Journal d'Economie Médicale,* 1999, Vol. 17, n° 5, pp. 343-361.
- NORRIS J.R. *Markov Chains*. Cambridge : Cambridge University Press, 1997.
- PUTERMAN M.L. *Markov decision processes : Discrete Stochastic Dynamic Programming*. New York : John Wiley & Sons Ltd - A Wiley series – Interscience publication, 1994.
- SONNENBERG F.A. *Théorie de la prise de décision et de la stratégie médicale*. In: LAUNOIS R., REGNIER F., editors. Décision thérapeutique et qualité de vie. Paris : John Libbey Eurotext, 1992, pp 27-54.
- SONNENBERG F., BECK J. Markov model in medical decision making. A practical guide. *Medical Decision Making,* 1993, Vol. 13, n° 1, pp. 322-338.
- TIJMS, H. C. *Stochastic models* : An Algorithmic Approach. Chichester : John Wiley & Sons Ltd, 1994.

## XIII) Verifying the robustness of study conclusions

Recommendation n° 30 and 31

- BRIGGS AH, MONEY CZ, WONDERLING DE. Constructing confidence intervals for cost-effectiveness ratios : an evaluation of parametric and non-parametric techniques using Monte-Carlo simulations. *Statistics in Medicine.* 1999, Vol. 18, n° 23, pp. 3245-3262.
- BRIGGS AH, WONDERLING DE, MONEY CZ. Pulling cost-effectiveness-analysis up by its bootstraps : a non-parametric approach to confidence interval estimation. *Health Economics.* 1997, Vol. 6, n° 4, pp. 327-340.
- CHAUDHARY MA, STEARNS SC. Estimating confidence for cost-effectiveness ratios : an example from a randomized trial. *Statistics in Medicine.* 1996, Vol. 15, n° 13, pp. 1447-1458.
- DOUBILET P., BEGG C.B., WEINSTEIN M.C., BRAUN P., MCNEIL B.J. Probabilistic sensitivity analysis using Monte Carlo simulation. *Medical Decision Making,* 1985, Vol 5, n° 2, pp. 155-177.
- HEITJAN DF, MOSKOWITZ AJ, WHANG W. Problems with interval estimates of the incremental cost-effectiveness ratio. *Medical Decision Making,* 1999, Vol. 19, n° 1, pp. 9-15.
- POLSKY D, GLICK HA, WILLKE R, SCHULMAN K. Confidence intervals for cost-effectiveness ratios: a comparison of four methods. *Health Economics,* 1997, Vol. 6, n° 3, pp. 243-252.
- VAN HOUT BA, GORDON G.S. AND RUTTEN, FF. Costs, effects and C/E-ratios alongside a clinical trial. *Health Economics,* 1994, Vol. 3, n° 5, pp. 309-319.
- WILLAN A, O'BRIEN BJ. Confidence intervals for cost-effectiveness ratios : an application of Fieller's theorem. *Health Economics,* 1996, Vol. 5, n° 4, pp. 297-305.

## XIV) Presentation of the results

Recommendation n° 32 to 34

- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Methods for the Economic evaluation of Health care Programmes*. 2nd ed. Oxford : Oxford University Press, 1997, p. 265.
- DRUMMOND Michael F., O'BRIEN Bernie, STODDART Greg L. et al. *Méthodes d'Évaluation Économique des Programmes de Santé*. 2ème ed. Paris : Economica, 1998, p. 285.
- GOLD Marthe R., SIEGEL Joanna E., RUSSELL Louise B., et al. *Cost-Effectiveness in Health and Medicine*. Oxford : Oxford University Press, 1996, p.276.

## XV) Interpretation of the results

Recommendation n° 35 to 37

- MISHAN E.J. The Use of Compensating and Equivalent Variations in Cost-Benefit Analysis. *Economica*, 1976, vol. 43, issue 17, pages 185-97
- MASON J., DRUMMOND M., TORRANCE G. Some guidelines on the use of cost effectiveness league tables. *British Medical Journal, 1993*, Vol. 306, n° 6877, pp. 570-572.
- TENGS et al. Five-hundred life-saving interventions and their cost-effectiveness. *Risk Analysis,* 1995, Vol. 15, n° 3, pp. 369-390.
- Bulletin of the WHO, 2000, Vol 78 n°2.


COLLÈGE DES ÉCONOMISTES
DE LA SANTÉ

CES – 7 rue de Citeaux – 75012 Paris – FRANCE.
Tél: +33 (0)1 43 45 75 65 – Fax : +33 (0)1 43 45 75 67

E-mail : CES2@wanadoo.fr

Website: http://www.ces-asso.org