Methods for Socio-Economic Evaluation of Marketed Medicines Ⅲ: Factors Affecting Methodological Quality and Transferability

SUN Xin¹, LI You-ping¹*, Thomas Alured Faunce², CHENG Gang³

1. Evidence-Based Medicine Center and Clinical Epidemiology Center, West China Hospital, Sichuan University, Chengdu 610041, China; 2. Medical School and College of Law, Australian National University, Australia; 3. Center for Drug Evaluation, State Food and Drug Administration, Beijing, 10061, China

Key Words
Socio-economic evaluation
Cost-effectiveness evaluation
Value for marketed medicines
Methodological quality
Transferability

Abstracts
Methodological quality and transferability will be important issues for the credibility and usefulness of both published studies and administrative methods for evaluating the socio-economic value of marketed medicines in China. This paper critically examines factors commonly contributing to, or inhibiting, the quality and transferability of socio-economic evidence of the value of medicines, with specific reference to the Chinese community. It discusses appropriate approaches to design, performance, and reporting of published economic evaluation studies, as well as guidance on assessment of quality of economic evaluations and recommends two internationally established methods that may be suitable for training in this setting.

In the last decade, socio-economic evaluation (often termed cost-effectiveness evaluation) has been widely used by governments throughout the world as a final stage in the assessment of health care technologies, particularly, prescription pharmaceuticals. Its role is generally viewed as defining the efficiency or community value of marketed medicines after they have been approved on safety and efficacy grounds, thereby informing decision-making in relation to government reimbursement lists and private formularies. In fulfilling their responsibilities in this area, expert evaluators may face difficult situations in weighing the volume and quality of available evidence.

These difficulties may relate to flaws of methodological quality such as the availability and reliability of clinical trial data, the use of surrogate outcomes or inappropriate comparators. Unless, for example, a government provides tax incentives or other encouragements, no valid clinical trial data may be derived comparing a new (allegedly innovative) product head-to-head on safety, quality and cost-effectiveness grounds against generic or other medicines already in use in China. On the other hand, they may also concern the applicability of cost-effectiveness evidence from different geographic areas (or health care settings), producing so-called “transferability” or “generalisability” issues, which may influence whether trial data can be applied to all citizens in China. If manufacturers of pharmaceutical products can readily point to such methodological flaws in such evaluations, this may inhibit the capacity of the Chinese authorities to use such evidence and processes, for example, to negotiate scientifically appropriate reimbursement levels or intellectual property rents.

To assist expert evaluations, important published cost-effectiveness evaluations are frequently included in systematic reviews. Such reviews use a variety of instruments to assess the quality of the included economic evaluations. Jefferson et al., however, having investigated the quality of published systematic reviews of economic evaluations from 1990 to 2001, identified a need of standardization of such assessment instruments. In addition, checklists for assessing transferability of the assessments may differ in items and scales.

The aims of this paper are to discuss, in the Chinese context, possible factors affecting the quality and transferability of socio-economic (or cost-effectiveness) evaluations, and commence the process of developing relevant guidelines.

Factors affecting methodological quality and transferability

General elements of the process of socio-economic (or cost-effectiveness) evaluation are set out in Figure 1. Assessments of the socio-economic or overall community value of pharmaceuticals, usually consist...
Fig The Process of Socio-Economic Evaluation and Price Negotiation for Marketed Medicines

of four consecutive steps: ① design of a clinical study of the relevant medicine, ② measurement and valuation of health outcomes and costs related to its recommended use, ③ presentation and expert analysis of this data and ④ factoring of the economic (or cost-effectiveness) evaluation into a price negotiation with the manufacturer. The definition of these objectives is basic but paramount to the process of conducting economic evaluations. A clear definition of objectives should state what alternative interventions are used to which populations by what measures.

**Design in economic evaluations**

In the initial design of a clinical trial, investigators should clearly define the values perspective of the study and its time horizon (length of follow-up), whilst also appropriately assessing comparative existing interventions (pharmaceutical and non-pharmaceutical) used by the target population, for the studied disease. Each such item of study design has its specific role in affecting the quality and transferability of subsequent economic evaluations based on the resultant data.

**Perspective**

The definition of research perspective or values has direct impact on the identification of costs associated with the evaluated medicine. For example, the choice of whether to value opportunity costs for patients on low incomes, or whether the overall aim is primarily to ensure equity of access for all Chinese citizens to essential medicines (rather than limited access by the wealthy to “innovative” medicines) will affect the decision on whether indirect costs are included.

The perspective with which the research is conducted might also influence the applicability of results. For example, results from a perspective which places universal access to essential medicines as a primary social aim, might be not be readily transferable to those trials conducted with a narrower perspective (e.g. to facilitate rapid market entry and government reimbursement for an expensive medicine).

**Time horizon**

The time horizon may affect both the choice of health outcome measures and cost range. The health measures should be not only clinically relevant, but also relate to the duration that investigators are interested in. Surrogate outcome measures (such as mere physiologic change in cholesterol level or in blood pressure), which are applied in short duration, should be extrapolated to endpoints such as five year survival and health-related quality of life (HRQOL) wherever possible. Similarly, the cost for patients will also increase along with the length of time needed for treatment. The length of duration also relates to decision on choice of epidemiological design. For example, trial-based or observational studies are usually limited to medium-term outcomes, while modeling has the ability to expand the length of study time.

**Choice of comparator**

The choice of comparator is an area frequently exploited by pharmaceutical manufacturers who fund and steer the design of clinical trials. In order to make
their new product seem most cost-effective, often the most expensive medicine in the relevant therapeutic class is chosen as a comparator and then administered to trial subjects at just below the recommended therapeutic dose. Alternatively the new product is compared against placebo, making it easier for the manufacturer to claim a therapeutic advantage, but harder to assert the product is “innovative” in any sense which has any context of value to society. More socially appropriate choices of alternative therapy may allow a study to reflect a better measure as possible of opportunity costs for patients of using the new treatment. Theoretically, the competing intervention should be the most cost effective in current practice, regardless of how cheap and easily available it is. In practice, the most widely used alternative should be most appropriate. Placebos should not be used unless the current practice is “doing nothing”.

Target populations
The range of target population in socio-economic evaluation (or cost-effectiveness) may be narrow or wide. Inclusion of a wide spectrum of population (rich and poor, rural and city, old and young, ill and healthy) may make the trial results much more applicable to China as a whole. Relevant factors include demography, epidemiology, disease complications, case mix, and culture/altitudes. These factors might have direct impact on the costs and health outcomes, but the magnitude of impact is likely to vary in a specific setting.

① Demography. Demographic characteristics, such as sex and age, are routine aspects of any well conducted clinical trial and should not influence transferability. In fact a trial which has been designed so that it is not immediately transferable to the entire Chinese population should have a presumption of having to justify why it has been designed in such a restrictive fashion.

② Incidence of diseases. The disease relates chiefly to resource use. For example, in a specific setting where disease incidence is low, the medication costs only account for 40% of the total costs. When the disease incidence increases, the proportion of medication costs may increase from 40% to 50%, since the medicines may be used more intensively than other resources.

③ Case mix. Case mix is used to identify patients that are similar in type and in terms of resources used as measured in patient days of care. However, the criteria of forming the case mix group are usually different across economic studies. Therefore, the effectiveness and costs generated from a specific case mix group are usually not transferable to the setting that applies to different criteria for the case mix group.

④ Culture/traditions. Variations of cultures or traditions usually result in difference of medication behaviors, which are mostly likely cause variations of effectiveness and costs. However, as far as reasonably practicable, clinical trials in China should be designed to predict how cultural differences in all areas of society (rich and poor, rural and city) are likely to change the effectiveness and costs of a studied medicine.

Measurement and valuation
Epidemiological design
In socio-economic evaluations of data obtained from clinical trials, the resultant evidence is usually categorized in a hierarchy of importance. Generally, meta-analysis of double blind, randomized control trials and randomized trials are rated as the first and second best available sources for effectiveness data. Observational or individual case studies and expert opinions can be used, but are considered inferior in evidentiary value. Epidemiological design has little influence on the cost measurement and valuation, since cost could be obtained independently from other resources.

Health outcomes measurement
Health outcomes, which consist of natural health outcomes and health utilities, are measured differently in varying forms of epidemiological design. If the natural health outcomes are obtained from the randomized trials and observational studies, a set of criteria could be used to assess their quality, which is a similar process to appraising the methodological quality for systematic review of clinical evidence. If the meta-analysis is used to assess the quality, a similar criteria checklist could also be applied. Expert opinions should also be justifiable, if there is no other evidence available. In the case of measuring health state utilities, different methods are available, either preference-based valuation or quality-of-life instrument. Investigators should also justify their use of the methods for outcome measurement against broad social criteria such as equity and social justice. In the preference valuation, standard gamble or time trade-off are preferred to use, since the results measured by these two are more consistent than those by visual analogue scale.

① Clinical practice variations. There are usually variations of clinical practice between different health care settings, such as tertiary care versus primary care. The differences in level of care and capacity to perform medical procedures, are likely to result in diverse health outcomes. It is extremely hard to quantify the differences, but the qualitative analysis on the clinical practice in specific setting to the local setting is useful in determining the applicability of existing economic evidence. This will be particularly true of cost-effectiveness analysis of medical devices rather than medicines.

② Compliance. The compliance of patients with physician recommendations for medication usage has the direct impact on the effectiveness. Due to the differences in culture, education, level of private
insurance, medication habits and ethnic origins, compliance of medication varies significantly between geographic areas in China.

3. Interviewees for the health state utilities. The target resources for the health state utilities may range from the general public, health professionals, and patients. This is an acceptable variation in the more socially-oriented Chinese context from the approach adopted in the more privatized United States system as should not affect transferability of data.

Measurement and valuation of costs
The costing process involves measurement of resource use and valuation of unit prices, which are separate in logic order. The methods of measuring resource use are generally categorized into (1) a synthetic approach and (2) prospective data collection. The synthetic approach consists of using secondary data such as administrative data, expert opinions, and retrospective chart reviews. The strength of synthetic approach is the accessibility and readiness of measurement, whereas the data are likely to be incomplete and exposed to risk of bias. Prospective data collection, on the other hand, is conducted either along with the clinical trial or the study on its own. It is superior in providing detailed data on resources use, but inferior in its laborious nature and time-consuming structure. There is no clear-cut point as to decide the appropriateness of using the methods of resources use. The choice of measurement depends on the requirements on the cost details and the efforts needed for the measurement. Regularly, the cost components that have significant contribution to the total should be measured in detail to allow analysis of its impact on the results. Minor cost components are usually obtained through synthetic methods. The data on resources use could be aggregate or disaggregate. However, the measurement of resources use in natural units and reporting separately are more appropriate since this makes the results more transparent.

Five approaches to valuing costs are available: (i) prices derived from national registries; (ii) prices derived from health economics literature and previous research; (iii) standard costs; (iv) tariffs or charges; and (v) calculation of unit costs. The choice of valuation approach is based on the measurement of resources use. The detailed measurement of resources use usually relates to much efforts on valuation of unit prices, and vice versa. There is no simple appropriate method to valuation of costs. Regularly, the cost components that have significant contribution to the total should be measured and valued in detail, while others of less significance are not. To determine the quality of performance in measuring and valuing costs, investigators should decide whether the methods of measuring the resources use and unit prices are appropriate and clearly described.

The measurement of resources use and unit prices is restricted to a specific setting, where many factors might contribute to the variations of costs across different geographic regions. These include:

1. Unit prices. Unit prices have direct and substantial impact on the costs. Usually, the unit prices of different geographic areas vary significantly. Unless the unit prices are reported separately with the quantities of resources use, the costs generated from a specific health care setting and geographic area might not be applicable to the other. These transferability issues can be solved with careful trial design.

2. Clinical practice variations. Clinical practice variations have direct impact on both the health outcomes and costs. It is hard to predict the magnitude of the impact on the costs, since there are no quantitatively analytical tools to examine the effect of variations between different settings on the costs. Investigators should be aware of the risk of the variation, and justify whether they are transferable to the local setting.

3. Compliance. The differences of compliance of could cause the variations of costs between different setting. This is mainly because the quantity of resources use differs from the variations of compliance.

4. Healthcare resources (technology availability). The availability of diagnostic and treatment technologies usually cause significant variations between health care setting. Tertiary care is generally equipped with more advanced technologies than primary care, thus resulting in higher costs.

Modeling in economic evaluations
Modeling techniques are used in economic evaluation to simulate the diseases progression within a certain period of time. The reasons for use of modeling in economic evaluation have been explained elsewhere. In the model, effectiveness, costs and transition probabilities are hierarchical, with the economic evaluation have been explained elsewhere. In the model, effectiveness, costs and transition probabilities are hierarchical, with the randomized trials (or meta-analysis of randomized trials) the preferred resource. Costs are available from the regular sources, as described above. To decide the quality of modeling studies, three different aspects should be considered. These include the effectiveness and cost resources, and modeling itself. Effectiveness and probability resources are hierarchical, with the randomized trials (or meta-analysis of randomized trials) the preferred resource. Costs resources are measured and valued according to the regular practice. However, the quality of modeling studies is firstly determined by the model itself.

The way of model itself effecting the quality is the structure that simulates the progress of the diseases. In principle, the structure of the model should be identical to the progress of the diseases, that is, any progress of disease from the intervention, either major or minor, should be simulated in the model to reflect the “real process”. However, practical problems usually occur when even trivial changes of disease are considered. For example, confusion might arise when the models become redundantly complicated; the available
analytical tools may fail to calculate the model if the models incorporate excessive parameters. In practice, the structure of the model should be established in the way that is able to simulate as detailed as possible the major progress of diseases of interest, whereas those changes without clinical and economic significance should not be considered. An important factor that affects the establishment of structure is the objective of economic evaluation. If the investigator could justify the structure of the model, its validity of simulating the disease progress is acceptable. In the presentation of outcomes of the modeling studies, the transparency of simulating the structure should also be important enough, that is, the structure of model should be presented as what the analysis has been carried out.

There is no specific issue of transferability of modeling structure. The factors that may cause problems of transferability in modeling studies are usually derived from the cost and effectiveness measurement and valuation.

Analysis of data, listing on government formulary and price negotiation
Although economic analysis is often generally referred to as cost-effectiveness analysis, the latter is actually a very specific subset of the initial concept. There are two types of questions that relating to analytic techniques that need to be kept in mind in designing a clinical trial: a) are the target interventions likely to be effective in achieving their stated clinical goals, and b) given a limited budget, which is the most efficient treatment option? Cost-benefit analysis is usually used to address the type (a) question, while the cost effectiveness and cost utility analysis are used to address the type (b) question. Appropriate use of analytic techniques relate to the issue of quality of economic evaluations.

Discounting
Discounting is used to adjust the future value of costs and health outcomes to the present value. Without discounting, it is likely that the impact of interventions on the present costs and outcomes is underestimated. Therefore, whether or not use of the discounting rate is directly related to the quality of studies. Although there remain discrepancy on discounting health outcomes, most analysts support the application of discounting to health benefits and costs. On the other hand, the range of discounting rate might differ in studies within different health care setting and geographic areas, and the variations of discounting rates across studies are likely to cause problems in transferring the results to the local setting. Generally, a discount rate of 5% is preferred to use in the economic evaluations across studies, although the rate may be subject to the socioeconomic status of local setting. The discount rate should be justified in the study.

Comparison of benefits and costs
In the economic evaluations, it is not enough to merely calculate the ratio of cost-effectiveness ratio for a specific intervention, and compare the ratios of alternative interventions. Within the healthcare context, economic evaluations – cost-effectiveness analysis and cost-utility analysis – are conducted based on the given health budget. Thus, the incremental ratio – extra costs per a unit of benefit gained – should be calculated to examine whether the target intervention is value for money compared to the competing interventions.

Analysis of uncertainty
Socio-economic evaluations are allowed for uncertainty due to variations of many substantially important parameters. The consideration of uncertainty enables users of evidence to decide whether conclusions are robust and meaningful. Failure to perform sensitivity analysis is most likely to affect the quality of economic evidence. The approaches to conducting sensitivity analysis vary from univariate to Monte Carlo simulation analysis, with each weighted different in their strength to investigating the uncertainty. The use of a specific approach to sensitivity analysis should be justified by the investigators. In the meanwhile, the parameters that might result in variations of results should be investigated as many as possible. Sensitivity analysis is an issue that affects the methodological quality and the robustness of conclusions, rather than the transferability.

Other factors affecting methodological quality
There are several other underlying factors that might affect the quality of economic evidence. These include conflict of interest and ethical considerations.

Conflict of interest is potentially an important factor that may bias the results. The investigators who conduct the economic evaluations might have a pre-existing preference to the active pharmaceutical intervention if they have received research funding from or are employed by the manufacturer. Therefore, it is necessary to have a statement by the investigators as to the conflict of interest in the study reports. It is also important that the central government continue to fund clinical trials as does not leave their production entirely up to industry.

Socio-economic evaluations aim to identify the intervention that is the most cost-effective option. Inherently, this does not mean the interventions which costs least is the best option. Usually, the intervention that have better effectiveness are likely to associate with higher costs, and the incremental analysis indicated such interventions are the best options within a specific health budget. However, given the accessibility and affordability of medicines, it is not always that the relatively cost-effective interventions are the right options. Investigators should make it clear whether the cost-effective interventions are affordable and available to as wide a target population as possible in China.
To summarize the possible factors that affect the methodological quality and transferability of economic evaluations, we have tabulated these factors and their magnitude of impact.

**Factoring Socio-economic evaluation into a price negotiation**

Although this is not the major topic of this presentation, socio-economic (or cost-effectiveness) evaluations will only have true value for the Chinese people if they factor into a price negotiation conducted by a central government using its large monopsony buying power to supervise levels of reimbursement over a “list” or formula. If the socio-economic evaluation, for example, shows that the submitted medicine performs better in terms of cost and effectiveness than existing treatments in its therapeutic class then its manufacturer is entitled to a price premium. If not, then a negotiation based on cost minimization can be undertaken. The global benchmark price will be an important consideration in such negotiations.

**Assessment tools for the methodological quality of economic evaluation**

There are now several checklists to assess the quality of economic evaluations. Some are original, whereas others are modified from the original checklists. The range of the quality items vary significantly across the checklists, although several are considered to influence quality in all checklists. Currently, there are no single criteria to assess quality that are used consistently across the world. However, several are widely used in most systematic reviews of economic evidence. These include Drummond-10 checklists and Chiou-19 checklists.

**Drummond-10 checklist**

In the 10-item checklist, three broad questions are evaluated: ① is the economic evaluation likely to be usable, ② how were costs and consequences assessed and compared, and ③ will the results help in purchasing services for local people. Each item is asked to record “yes”, “no” or “can’t tell”. Although the 10-item checklist was not validated originally, the late studies followed showed that this checklist is a useful and quick tool to assess the quality of economic evaluations, which is very helpful for decision makers.

However, the disadvantage of this checklist is that some items are combinations of several quality items. For example, the item 4-b (measured accurately in appropriate units prior to evaluation) comprises of the issue of both effectiveness and costs measurement, which might cause confusions to those who assess the quality. Item 8 includes many issues, which may have been beyond the scope of quality. This causes difficulties in determining the methodological quality.

In addition, further efforts have been contributed to develop a scoring system based on the Drummond-10 checklist. Three methods have been developed to assess the impact of different scoring methods on the validity. It was found that to score the economic evaluation using equally weighted item is good at differentiate the low-quality studies, whereas not adequate to differentiate the good-quality papers. A second scoring system, which uses a hierarchy of effectiveness evidence and introduces an item on transferability, decreases the points of study which have been score the highest. The third scoring system has multiply each item, rather than using an additive score, is good at distinguishing good quality studies, but blunt in differentiating poor quality studies.

**Chiou-16 checklist**

Chiou et al developed and validated a scoring system for grading the quality of cost-effectiveness and cost-utility studies. This scoring system comprised of a 16-item checklist, and each item was weighted by points, ranging from 1-9. The total score was 100, and those studies scoring over 75 were high in quality.

This checklist was developed by several steps. A committee comprised of health economists firstly pooled the quality items from the 14 published checklists and guidelines, and then added new items if they agreed with the their importance. When consensus was achieved, the checklist was sent out to value the weight of each item and quantify the scores against the quality by a 2-phase survey of health economists. Spearman rho and Wilcoxon tests were used to examine the validity and analysis of covariance. This checklist is the currently best validated instruments that scores the quality of economic evaluations.

**Cases study**

To improve the understanding on assessment of methodological quality, we exemplified by appraising the methodological quality from a published economic evaluation. We have abstracted the main methodological items, as below.

① From the health care sector perspective, this paper used modeling to investigate whether the introduction of lamivudine in the Australian market was cost-effective in patients with HBeAg positive chronic hepatitis B in Australia.

② The patients included in this study were characterized of HBeAg positive, elevated ALT level, no progression to cirrhosis, and native to interferon-α. A major assumption was made that 70% of patients were male, and they started the management at the average age of 30.

③ Two antiviral strategies, each representing the availability of antiviral medicines in the Australian market, and one “do nothing” strategy, indicating the baseline results were considered in the study.
## Table 1 The possible items influencing the quality and transferability of economic evaluations

<table>
<thead>
<tr>
<th>Study items</th>
<th>Items for Quality</th>
<th>Best practice for conducting economic evaluations</th>
<th>The possible factors and ways of affecting transferability</th>
<th>Possible efforts and adjustment used to improve the transferability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design items</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>√</td>
<td>Clear description. Societal perspective preferred, but subject to the objective of study. Reasons for choice of perspective should be justified.</td>
<td>Perspective: Influence the identification of costs</td>
<td>Disaggregate resources use and unit prices</td>
</tr>
<tr>
<td>Time horizon</td>
<td>√</td>
<td>The duration of interest should be long enough to capture effectiveness and costs.</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Interventions</td>
<td>√</td>
<td>Ideally, the most cost-effective options currently available are the competing interventions. In practice, the most widely used option is the most appropriate competing intervention.</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Populations</td>
<td>√</td>
<td>Ideally, those that have the indications and are diagnosed of no contraindications are the population. In practice, more specific populations should be justified.</td>
<td>Demographic characteristics: influence costs and effectiveness</td>
<td>Few is available to improve the transferability.</td>
</tr>
<tr>
<td>Analytic techniques</td>
<td>√</td>
<td>Clear description of the analytic techniques. The reasons to chose analytic techniques should be justified.</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td><strong>Measurement and valuation items</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epidemiological design</td>
<td>√</td>
<td>Pragmatic trials are the ideal standards, but impractical in most cases. No single choice is right. The choice of designs should be justified.</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Health outcomes measurement</td>
<td>√</td>
<td>Justify the use of the epidemiological design. Clear and adequate description of methods to measure the effectiveness. Standard gamble and time trade-off methods are preferred to value utilities. The use of disease-specific or general quality-of-life instruments should be justified.</td>
<td>Clinical practice variations: cause difference in diagnostic and treatment outcomes; Compliance: result in significant change of effectiveness for different population using the same medication. Different target populations for health utility valuation: have different viewpoints on health utilizes, thus produce the possibly different results.</td>
<td>Few is available to improve the transferability. Few is available to improve the transferability. Few is available to improve the transferability.</td>
</tr>
<tr>
<td><strong>Analysis and presentation items</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discounting</td>
<td>√</td>
<td>The use of rate of discount should be justified.</td>
<td>Discount rate: influence the costs, and possibly effectiveness if health outcomes are discounted</td>
<td>Re-calculated the discounted costs and outcomes</td>
</tr>
<tr>
<td>Incremental analysis</td>
<td>√</td>
<td>Incremental cost-effectiveness ratio should be calculated and presented and the judgment is needed to assess the cost-effective option.</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Sensitivity analysis</td>
<td>√</td>
<td>Sensitivity analysis should always be performed to identify the factors affecting the robustness. The methods used should be justified.</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Conflict of interest</td>
<td>√</td>
<td>Conflict should be indicated whenever possible.</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Ethical considerations</td>
<td>√</td>
<td>Affordability, accessibility of medicines, in addition to the cost-effectiveness should also be discussed.</td>
<td>None</td>
<td></td>
</tr>
</tbody>
</table>
1 The model structure was developed under the Australian clinical practice setting, and assumptions were made and clarified where appropriate. The structure of model was reported and clear.

2 The outcome measures were life-years gained, HBeAg seroconversion, development of cirrhosis. The data for transition probabilities were derived from pooled results of clinical trials, published studies, and expert opinions if the trials were not available. The 1-year and life-time results were obtained from analysis of modeling.

3 The unit costs of resource use were valued and obtained from the national registries and databases. However, units of resource use were not well reported, though some recorded the data.

4 A discounting rate of 5% was used to the cost data, but not to health outcomes data. Incremental analysis was conducted to compare three alternative strategies, both to costs and health outcomes. Sensitivity analyses were performed by using a range of alternative assumptions. However, the methods for the sensitivity analyses were not explained in the study.

5 In the discussion section, this paper discussed the weakness and strength of the results, and its application to the local decision. However, it did not presented where the study funding was obtained.

By using the abstracted data, we thus filled out the two checklists (appendix). The results of both checklists indicated the quality was high. However, from the checklists, the assessment of quality differed between the checklists. The Drummond checklist is short, and combined several quality items, which made the appraisal more challenging to the assessors. The Chiou’s checklist is easy to follow, which disaggregates the quality items. Several items between Drummond checklist (item 3, 10) and Chiou’s checklist (item 4, 7, 12, 13, 16) differed. However, the quality methods for posing questions, incremental analysis, sensitivity analysis, measurement and valuation of health outcomes and costs were consistent in these two checklists. The major difference between the checklists was that the Chiou’s checklist detailed the items on modeling, and the Drummond checklist considered the applicability of results.

To assess the methodological quality, it might be more appropriate to consider the Chiou’s checklist, which might be more easy to handle and inform the methodology strength.

References


### Appendix Methodological quality of published paper by Drummond 10-item checklist

<table>
<thead>
<tr>
<th>Question point</th>
<th>Yes/no/cannot tell</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Was a well-defined question posed?</td>
<td>Yes</td>
</tr>
<tr>
<td>2. Was a comprehensive description of the competing alternatives given?</td>
<td>Yes</td>
</tr>
<tr>
<td>3. Does the paper provide evidence that the program would be effective</td>
<td>Yes</td>
</tr>
<tr>
<td>4. Were all important and relevant resource use and health outcome consequences for each alternative:</td>
<td></td>
</tr>
<tr>
<td>a) identified?</td>
<td>Yes</td>
</tr>
<tr>
<td>b) measured accurately in appropriate units prior to evaluation?</td>
<td>Yes</td>
</tr>
<tr>
<td>c) valued credibly?</td>
<td>Yes</td>
</tr>
<tr>
<td>5. Were resource use and health outcome consequences adjusted for different times?</td>
<td>No</td>
</tr>
<tr>
<td>6. Was an incremental analysis of the consequences and costs of alternatives performed?</td>
<td>Yes</td>
</tr>
<tr>
<td>7. Was an adequate sensitivity analysis performed?</td>
<td>Yes</td>
</tr>
<tr>
<td>8. Did the presentation and discussion of the results include enough of the issues that are required to inform a purchasing decision?</td>
<td>Yes</td>
</tr>
<tr>
<td>9. Were the conclusions of the evaluation justified by the evidence presented?</td>
<td>Yes</td>
</tr>
<tr>
<td>10. Can the results be applied to the local population?</td>
<td>No</td>
</tr>
</tbody>
</table>

Total 10/12
Methodological quality of published paper by Chiou’s Grading System

<table>
<thead>
<tr>
<th>Questions</th>
<th>Point</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Was the study objective presented in a clear, specific, and measurable manner?</td>
<td>7</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>2. Were the perspective of the analysis and reasons for its selection stated?</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>3. Were variable estimates used in the analysis from the best available source?</td>
<td>8</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>4. If estimates came from a subgroup analysis, were the groups prespecified at the beginning of the study?</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5. Was uncertainty handled by: 1) statistical analysis to address random events; 2) sensitivity analysis to cover a range of assumptions?</td>
<td>9</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6. Was incremental analysis performed between alternatives for resources and costs?</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7. Was the methodology for data abstraction (including the value of health states and other benefits) stated? b</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>8. Did the analytic horizon allow time for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted and justification given for the discount rate?</td>
<td>7</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>9. Was the measurement of costs appropriate and the methodology for the estimation of quantities and unit costs clearly described?</td>
<td>8</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>10. Were the primary outcome measure(s) for the economic evaluation clearly stated and were the major short-term, long-term, and negative outcomes included?</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>11. Were health outcome measures/scales valid and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?</td>
<td>7</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>12. Were the economic model (including structure), study methods and analysis, and the components of the numerator and denominator displayed in a clear transparent manner?</td>
<td>8</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>13. Were the choice of economic model, main assumptions and limitations of the study stated and justified?</td>
<td>7</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>14. Did the author(s) explicitly discuss direction and magnitude of potential biases?</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>15. Were the conclusions/recommendations of the study justified and based on the study results?</td>
<td>8</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>16. Was there a statement disclosing the source of funding for the study?</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Total points</td>
<td>100</td>
<td>92</td>
<td></td>
</tr>
</tbody>
</table>

Note: The highest possible score is 100, and the lowest is 0. Studies with a score exceeding 75 points are of high quality. A Adequate sensitivity analysis includes 2-way analysis and beyond (e.g. Monte Carlo analysis). Therefore, sensitivity analyses limited to 1-way results are not adequate to receive points for item 5. b At a minimum, studies should describe clearly the databases searched, key words used, dates queried, or prioritization scheme for study types.
上市后药物经济学评价方法 III：影响方法学质量和结果可适用性的因素

孙鑫1 李幼平1* Thomas Alured Faunce2 程刚3

1. 四川大学华西医院循证医学与临床流行病学研究中心（成都 610041）
2. 澳大利亚国立大学，医学院和法律学院
3. 国家食品与药品监督管理局药物评价中心（北京 100061）

摘 要

方法学质量和结果可适用性是评价已发表药物经济学评价研究可靠性和适用性的重要方面，也是影响我国上市后药物社会经济价值评价的因素。本文探讨了经济学评价研究中影响方法学质量和结果可适用性的因素及其方式，并特别强调了在我国条件下的影响因素。本文讨论了经济学评价合理的设计、实施和分析报道方法，并介绍了现有评价经济学研究方法学质量量表，两种适用于我国的评价量表。

关键词 社会经济学评价；成本-效果评价；上市后药物经济价值；方法学质量；结果外在适用性

* 通讯作者 Email: yzymlab@hotmail.com