

# VU Research Portal

## **Criteria list for assessment of methodological quality of economic evaluations:consensus on health economic criteria.**

Evers, S.M.; Goossens, M.E.J.B.; de Vet, H.C.W.; van Tulder, M.; Ament, A.

### ***published in***

International Journal of Technology Assessment in Health Care  
2005

### ***document version***

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

### ***citation for published version (APA)***

Evers, S. M., Goossens, M. E. J. B., de Vet, H. C. W., van Tulder, M., & Ament, A. (2005). Criteria list for assessment of methodological quality of economic evaluations:consensus on health economic criteria. *International Journal of Technology Assessment in Health Care*, 21, 240-5.

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

### **E-mail address:**

[vuresearchportal.ub@vu.nl](mailto:vuresearchportal.ub@vu.nl)

# Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria

**Silvia Evers, Mariëlle Goossens**

*Maastricht University, Institute for Rehabilitation Research*

**Henrica de Vet, Maurits van Tulder**

*VU University Medical Centre*

**André Ament**

*Maastricht University*

**Objectives:** The aim of the Consensus on Health Economic Criteria (CHEC) project is to develop a criteria list for assessment of the methodological quality of economic evaluations in systematic reviews. The criteria list resulting from this CHEC project should be regarded as a minimum standard.

**Methods:** The criteria list has been developed using a Delphi method. Three Delphi rounds were needed to reach consensus. Twenty-three international experts participated in the Delphi panel.

**Results:** The Delphi panel achieved consensus over a generic core set of items for the quality assessment of economic evaluations. Each item of the CHEC-list was formulated as a question that can be answered by yes or no. To standardize the interpretation of the list and facilitate its use, the project team also provided an operationalization of the criteria list items.

**Conclusions:** There was consensus among a group of international experts regarding a core set of items that can be used to assess the quality of economic evaluations in systematic reviews. Using this checklist will make future systematic reviews of economic evaluations more transparent, informative, and comparable. Consequently, researchers and policy-makers might use these systematic reviews more easily. The CHEC-list can be downloaded freely from <http://www.beoz.unimaas.nl/chec/>.

**Keywords:** Review, Systematic, Cost and cost analysis, Delphi technique, Quality assurance, Health care, Evidence based medicine

Health-care professionals, consumers, researchers, and policy-makers can be overwhelmed by the sometimes un-

manageably large number of trials and economic evaluations of health care interventions. Systematic reviews of these studies can help in making well-informed decisions on which intervention to adopt. For maximum usefulness, systematic reviews of economic evaluations should be transparent, that is, all relevant methodological information from the included studies should be described in a systematic way. However, there is no generally accepted criteria list for reviewing economic evaluations, which may be because most of the criteria lists are created single-handed. The aim of the Consensus on

The authors thank the following persons for their participation in the Delphi panel: D. Banta, M. Buxton\*, D. Coyle\*, C. Donaldson, M. Drummond\*, A. Elixhauser\*, B. Jönsson, E. Jonsson\*, K. Kesteloot, B. Luce, D. Menon, M. Mugford\*, E. Nord, B. O' Brien († February 13, 2004), J. Rovira, L. Russell, F. Rutten\*, G. Simon, J. Sisk, R. Taylor, G. Torrance, A. Towse, and L. Vale. We thank J. van Emmerik for his technological assistance with the current study and E. Brounts for the literature review. \*These members were also part of the Task Force Group.

Health Economic Criteria (CHEC) project was to develop a generally accepted criteria list, which should be regarded as a minimum standard.

The CHEC-list focuses only on the methodological quality of economic evaluation aspects, as existing criteria lists focus already on the methodological quality of more general aspects of effectiveness studies (20;23). The CHEC-list was developed for systematic reviews of full economic evaluations based on effectiveness studies (cohort studies, case-control studies, randomized controlled trials). The focus on economic evaluations alongside trials is due to practical considerations, because other methodological criteria are relevant when using other designs, for example, modeling studies or scenario analyses (27). The criteria list has been developed using a Delphi method. A similar method was used in development of an instrument to assess the methodological quality of randomized controlled studies (30).

## METHODS

The study started with the creation of a large item pool, followed by reduction by a Delphi method. In this study, three Delphi rounds were sufficient to reach consensus (defined as general agreement on a substantial majority).

The project team was responsible for the construction and reduction of the item pool, the selection of participants of the Delphi panel, the construction of the Delphi questionnaires, the analysis of the response, and the formulation of the feedback.

### Construction and Reduction of the Item Pool

For the development of the initial item pool, items were selected from various sources, and several search strategies were used to identify the relevant literature in the field of economic evaluations. First, a Medline search was performed for the period 1990 to 2000 using MESH headings "cost and cost analysis," "meta-analysis or review literature." Additionally, Psyclit and Econlit were screened for the same period using keywords "review or meta-analysis" and "economic or cost" in titles. Additional articles were identified by searching the Cochrane Library (2000, issue 3) and the National Health Service EED database using the terms "cost or economic." Finally, handbooks on economic evaluations were monitored and a request to submit additional guidelines was presented to the HealthEcon Discussion List.

Three members of the team (S.E., M.G., and A.A.) developed a classification scheme, which included the various domains of economic evaluations (e.g., economic study question, economic study design, economic identification, outcome valuation) under which all items were ordered. The classification scheme consisted of nineteen categories. Details of the methods have been published elsewhere (2). The Delphi participants were given the opportunity in the first

Delphi round to add additional categories and items they thought were missing.

### Selection of Participants

Twenty-three international experts participated in our Delphi panel (see acknowledgments). We first formed a Task Force Group consisting of seven international experts in the field of economic evaluations. This Task Force Group assisted the project team with the composition of the final Delphi panel. Participants for the Delphi panel were selected if they were authors of guidelines or criteria lists or if they had special expertise in (systematic reviews of) economic evaluations. The project team made a first selection, keeping a balance between various countries and research groups. The inclusion of experts from different research settings all over the world was an explicit goal. The final Delphi panel was approved by the Task Force Group.

### Procedure/Analysis

During the entire Delphi procedure, structured questions were combined with open questions to facilitate the Delphi process. For example, in the Delphi-1 questionnaire, participants were asked to indicate which categories should be incorporated in a minimum set of the criteria list and give their arguments. Participants were always asked for other suggestions. In all the Delphi rounds, participants were asked to complete the questionnaire "bearing in mind the aim of the project." After each Delphi round, feedback was given to inform the participants of decisions, opinions, and arguments of the other participants. The project team decided, on the basis of the majority of the answers and the arguments used, which categories and items should appear in the next Delphi round. Overall, a category or item was included based on an arbitrary cutoff point (if half of the participant plus one agreed on its inclusion).

The decisions of the project team were also presented and justified in feedback reports (Delphi questionnaires, feedback reports, and the final CHEC-list can be downloaded from [www.beoz.unimaas.nl/chech/](http://www.beoz.unimaas.nl/chech/)).

### Delphi Rounds

In the first Delphi round, participants were asked to indicate which categories and subsequently which items should be included in the CHEC-list. In the second Delphi round, the participants received the results and comments of the panel together with the decisions of the project team. In this second Delphi round, participants were asked, "to select one item of each category that should be included in the final CHEC-list." Based on the analysis of this second round, the project team phrased a concept version of the CHEC-list. In the third Delphi round, participants were given a final opportunity to suggest modifications to the concept version of the CHEC-list.

## Pilot

To pilot the applicability of the CHEC-list, two articles on economic evaluations were reviewed by all authors.

## RESULTS

### Participants

Twenty-six experts were invited to participate in the Delphi panel, of whom two did not accept the invitation and one who originally agreed to participate did not respond to our questionnaires. A total of twenty-three members completed all three Delphi rounds.

### Delphi-1 Round

A total of twenty-five guidelines were initially identified (1; 3–22;24–26;28;29). Five of these guidelines were excluded after reading, because they were inadequate for our goal (3;5;14;22;28). Furthermore, some guidelines were published more than once (4;7;10;11;24–26). This finding left a sample of fifteen guidelines available, resulting in an initial item pool of 218 items. The item pool was restructured and reduced using the classification scheme. We eliminated items on general methodology (covered by criteria lists of effectiveness studies), and skipped (almost) identical formulations (see Ament et al. [2]). This strategy reduced the initial item pool to 128 items spread over nineteen categories, which were used in the first Delphi round. In the first Delphi round, all nineteen categories were considered essential for the CHEC-list, that is, the lowest approval (83 percent) was on the category “Ethics and distributional effects.” The items to be included in each category, however, did show some major differences (agreement varied between 9 percent and 91 percent). We used a majority agreement as an arbitrary cutoff point to include an item within a category. If none of the items reached the cutoff point within a category, the project team decided on the most appropriate item based on the arguments of the participants in the first Delphi round.

### Delphi-2 Round

On the basis of the results of the Delphi-1, the project team created a feedback report and a Delphi-2 questionnaire, in which all items that met the above criteria were presented to the participants. All twenty-three participants completed and returned the Delphi-2 questionnaire. In the analysis within each category, the item with the highest percentage of agreement was chosen. An important observation from the Delphi-2 round was that a large number of the participants indicated that they preferred items to be selected that give insight into the quality of the study performed, rather than into how the study is performed (e.g., “Is the economic study design appropriate to the stated objective?” instead of “What is the form of design used?”). The project team selected and rephrased items and developed a concept version of the CHEC-list, which was included in the Delphi-3 question-

naire. In addition, guidelines were developed on how to fill out the CHEC-list, which gave an explanation of the meaning of each item.

### Delphi-3 Round

The Delphi-3 questionnaire was also completed and returned by all twenty-three participants. The majority of the participants accepted the concept CHEC-list. To give a taste of the final discussion, we present some examples of the remarks of the panel members. Some participants had specific comments regarding the category “Presentation of results.” Three items in this category were included in the Delphi-2 questionnaire, that is, “Are the methods and analysis displayed in a clear and transparent manner?,” “Do the conclusions follow from the data reported?,” and “Are the assumptions and limitations of the study discussed?.” The agreement of these items was 39.1 percent, 69.6 percent, and 26.1 percent, respectively. We, therefore, had selected item 2 for our final CHEC-list (see Table 1). In the second and the third Delphi round, three Delphi members suggested that all three items should be retained in the criteria list. The project team reconsidered this and decided, as it was only suggested by the minority to keep to their original decision based on the opinion of the majority of the panel members.

In the guideline of the CHEC-list, regarding the category “perspective,” the project team had stated the following “If the study is performed from a societal perspective tick ‘yes,’ as all relevant costs and consequences of an intervention and disease are taken into account, if possible. Other narrower perspectives will only include certain components. The authors should motivate why a narrower perspective is valid.” Based on this statement, one participant remarked that “there is no single appropriate perspective. What is appropriate depends upon the decision/policy-maker. Thus, the answer to this item, if answerable at all, will be use-context specific and not intrinsic to the published study.” Although we agree to a certain extent with the remarks made by this Delphi member, it is an overall limitation of any criteria list to design a general criteria list, in which all items are equally relevant for all studies. Based on the discussion in the project team, we decided to not alter the guidelines and the CHEC-list, as the majority of the Delphi panel agreed with the suggested phrasing.

Regarding the category “independence of the investigator” one participant also remarked that “A study should not be ‘marked down’ because there are acknowledged potential conflicts. If we do that, no studies published by the pharmaceutical industry or by organizations working for them would be acceptable.” Based on this, the project team did not change the item, but in the explanation, we tried to overcome this difficulty by stating that “If an external agency finances the study, a statement should explicitly be given about who finances the study to guarantee transparency in the relationship between the sponsor and the researcher. Whenever a

**Table 1.** Final CHEC-List after Three Delphi Rounds

Items finally included in the CHEC-list As decided upon in the Delphi 3	Agreement	
	On inclusion of category Delphi 1	On inclusion of this item Delphi 2
1. Is the study population clearly described?	100%	65.2%
2. Are competing alternatives clearly described?	100%	69.6%
3. Is a well-defined research question posed in answerable form?	91%	100% <sup>a</sup>
4. Is the economic study design appropriate to the stated objective?	91%	— <sup>b</sup>
5. Is the chosen time horizon appropriate to include relevant costs and consequences?	100%	69.6%
6. Is the actual perspective chosen appropriate?	100%	— <sup>b</sup>
7. Are all important and relevant costs for each alternative identified?	100%	82.6% <sup>3</sup>
8. Are all costs measured appropriately in physical units?	96%	73.3% <sup>3</sup>
9. Are costs valued appropriately?	96%	— <sup>b</sup>
10. Are all important and relevant outcomes for each alternative identified?	100%	100% <sup>a</sup>
11. Are all outcomes measured appropriately?	96%	— <sup>b</sup>
12. Are outcomes valued appropriately?	96%	— <sup>b</sup>
13. Is an incremental analysis of costs and outcomes of alternatives performed?	87%	100% <sup>a</sup>
14. Are all future costs and outcomes discounted appropriately?	100%	— <sup>b</sup>
15. Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?	100%	— <sup>b</sup>
16. Do the conclusions follow from the data reported?	100%	69.6%
17. Does the study discuss the generalizability of the results to other settings and patient/client groups?	91%	78.3%
18. Does the article indicate that there is no potential conflict of interest of study researcher(s) and funder(s)?	91%	56.5% <sup>c</sup>
19. Are ethical and distributional issues discussed appropriately?	83%	100% <sup>a</sup>

<sup>a</sup> This was the only item left in the Delphi-2 questionnaire.

<sup>b</sup> Based on the discussion, this item was rephrased and agreed upon in the Delphi-3 round.

<sup>c</sup> Agreement was high, but based on the discussion, the item was somewhat rephrased and agreed upon in the Delphi-3 round. CHEC, Consensus on Health Economic Criteria.

potential conflict of interest is possible, a declaration should be given of ‘competing interest’.”

**Pilot**

The pilot test of the CHEC-list showed a strong agreement among the assessments of the members of the project team when reviewing two articles. No items were changed or adapted.

**CONCLUSION**

This study is the first in which a broadly accepted criteria list for economic evaluations was developed based on a Delphi consensus procedure. In a consensus procedure, the choice of participants is crucial. In the selection of this Delphi panel, the project group tried to achieve a broad representation of experts on quality assessment of economic evaluations. In the Delphi-2 round, it became clear that the majority of the participants wanted the CHEC-list to include items that give insight into the quality of economic evaluations rather than into how the economic evaluation is performed. As a result, most of the items included are now subjective judgments, not simple statements of “fact.” In its practical use, this may challenge the inter-rater vari-

ability. The project team, therefore, suggests using two or more reviewers when performing a systematic review and conducting a pilot test. Criteria lists, such as the CHEC, can be criticized because of their potential rigidity, which might prohibit further development of methodology. This criticism is only valid if criteria are used injudiciously. To prevent the CHEC-list from being methodologically rigid, the project group emphasizes that the CHEC-list is a minimum set and would like to stimulate researchers to add additional items appropriate for the specific subject under study, if relevant.

Finally, there are many examples of poor reporting of economic evaluations. It is often difficult to conclude what actually happened in the study and this affects the methodological quality assessment. Part of the problem is because journals only accept a limited number of words in an article, thus making extensive explanations almost impossible. A solution might be to contact the authors of the original study, asking them for a more detailed description of the study design. An alternative would be to require the production of a standard technical report for every economic evaluation study. This technical report should be available to everyone, for example on the Internet. Such a technical report, describing the economic design, and details of the study, could also be used in systematic reviews.

Notwithstanding these limitations, we believe that, by creating a criteria list consisting of a minimum set of items, this CHEC-list makes it possible for future systematic reviews of economic evaluations to become more transparent, informative, and comparable.

### Policy Implications

There was consensus among a group of international experts regarding a core set of items that can be used to assess the quality of economic evaluations in systematic reviews. The project team also provided a guideline to the criteria list items to standardize its interpretation and facilitate its use. Using this criteria list will make future systematic reviews of economic evaluations more transparent, informative, and comparable. Consequently, the CHEC-list can help researchers and policy-makers to interpret the results of systematic reviews more easily and can be of help by translating these results into policy implications.

### CONTACT INFORMATION

**Silvia Evers**, PhD (S.Evers@beoz.unimaas.nl), Senior HTA Researcher, Department of Health Organization Policy and Economics, **Mariëlle Goossens**, PhD (m.goossens@dep.unimaas.nl), Assistant Professor, Department of Medical, Clinical, and Experimental Psychology, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

**Henrica de Vet**, PhD (hcv.devvet@vumc.nl), Professor, **Maurits van Tulder**, PhD (mw.vantulder@vumc.nl), Assistant Professor, Institute for Research in Extramural Medicine, VU University Medical Centre, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands

**André Ament**, PhD (a.ament@beoz.unimaas.nl), Associate Professor, Department of Health Organization Policy and Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

### REFERENCES

- Adams ME, McCall NT, Gray DT, Orza MJ, Chalmers TC. Economic analysis in randomized control trials. *Med Care*. 1992;30:231-243.
- Ament AJHA, Evers SMAA, Goossens MEJB, de Vet HCW, van Tulder MW. Criteria list for conducting systematic reviews based on economic evaluation studies: CHEC. In: Donaldson C, Mugford M, Vale L, eds. *Evidence-based health economics: From effectiveness to efficiency in health care*. Chapt 8. London: BMJ Publishing Group; 2002:99-113.
- Blackmore CC, Magid DJ. Methodologic evaluation of the radiology cost-effectiveness literature. *Radiology*. 1997;203:87-91.
- Bradley CA, Iskedjian M, Lanctôt KL, et al. Quality assessment of economic evaluations in selected pharmacy, medical, and health economic journals. *Ann Pharmacother*. 1995;29:681-689.
- Buxton MA. *The Canadian experience: Step change or gradual evolution?* London: Office of Health Economics; 1997.
- Clemens K, Townsend R, Luscombe F, et al. Methodological and conduct principles for pharmacoeconomic research. *Pharmacoeconomics*. 1995;8:169-174.
- Department of Clinical Epidemiology and Biostatistics of the McMaster University Health Science Centre. How to read clinical journals: VII. To understand an economic evaluation (part B). *Can Med Assoc J*. 1984;130:1542-1549.
- Detsky AS. Guidelines for economic analysis of pharmaceutical products. A draft document for Ontario and Canada. *Pharmacoeconomics*. 1993;3:354-361.
- Drummond M, Brandt A, Luce B, Rovira J. Standardizing methodologies for economic evaluation in health care. *Intl J Technol Assess Health Care*. 1993;9:26-36.
- Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ economic evaluation working party. *BMJ*. 1996;313:275-283.
- Drummond MF, O'Brien BJ, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press; 1997.
- Drummond MF, Richardson WS, O'Brien BJ, Levine M, Heyland D. User's guides to the medical literature. XIII. How to use an article on economic analysis of clinical practice. *JAMA*. 1997;277:1552-1556.
- Eisenberg JM. Clinical economics. A guide tot economic analysis of clinical practices. *JAMA*. 1989;262:2879-2886.
- Evers SMAA, van Wijk AS, Ament AJHA. *Economic evaluations of mental health care interventions. A review*. Maastricht: Department of Health Economics; 1994:39.
- Ganiats TG, Wong AF. Evaluation of cost-effectiveness research. A survey of recent publications. *Fam Med*. 1991;23:457-461.
- Gerard K. Cost-utility in practice: A policy maker's guide to the state of the art. *Health Policy*. 1992;21:249-279.
- Gibson GA. *Use of the guidelines to evaluate and interpret pharmacoeconomic literature*. Pharmacoeconomics and outcomes. Kansas City, MO: American College of Clinical Pharmacy; 1996:325-364.
- Gold MR, Russell LB, Siegel JE, Weinstein MC. *Cost-effectiveness in health and medicine*. Oxford: Oxford University Press; 1996.
- Haycox A, Drummond M, Walley T. Pharmacoeconomics: Integrating economic evaluation into clinical trials. *Br J Pharmacol*. 1997;43:559-562.
- Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary. *Control Clin Trials*. 1996;17:1-12.
- Lee JT, Sanchez LA. Interpretation of "cost-effective" and soundness of economic evaluations in the pharmacy literature. *Am J Hosp Pharm*. 1991;48:2622-2627.
- Mason J, Drummond M. Reporting guidelines for economic studies. *Health Econ*. 1995;4:85-94.
- Moher D, Jadad AR, Nichol G, et al. Assessing the quality of randomised controlled trials: An annotated bibliography of scales and checklists. *Control Clin Trials*. 1995;16:62-73.
- Sacristán JA, Soto J, Galende I. Evaluation of pharmacoeconomic studies: Utilization of a checklist. *Ann Pharmacother*. 1993;27:1126-1133.
- Sanchez LA. Evaluating the quality of published pharmacoeconomic evaluations. *Hosp Pharm*. 1995;30:146-148.

26. Sanchez LA. Applied pharmacoeconomics: Evaluation and use of pharmacoeconomic data from the literature. *Am J Health-System Pharm.* 1999;56:1630-1638.
27. Soto J. Health economic evaluations using decision analytic modelling: Principles and practices—utilization of a checklist to their development and appraisal. *Intl J Technol Assess Health Care.* 2002;18:94-111.
28. Task Force on Principles for Economic analysis of health care technology, Economic analysis of health care technology. A report on principles. *Ann Intern Med.* 1995;123:61-70.
29. Udvarhelyi S, Colditz GA, Rai A, Epstein AM. Cost-effectiveness and cost-benefit analysis in the medical literature. *Ann Intern Med.* 1992;116:238-244.
30. Verhagen AP, Vet de HCW, de Bie RA, et al. The Delphi List: A criteria for a list for quality assessment of randomized clinical trials for conducting systematic reviews developments by Delphi consensus. *J Clin Epidemiol.* 1998;51:1235-1241.