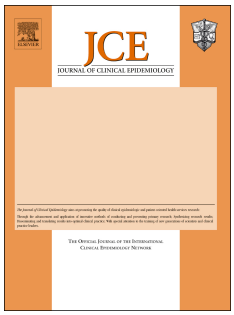

GRADE guidance 23: Considering cost-effectiveness evidence in moving from evidence to health-related recommendations

Feng Xie, Ian Shemilt, Luke Vale, Francis Ruiz, Michael F. Drummond, Joanne Lord, Kirsten H. Herrmann, María Ximena Rojas, Yuan Zhang, Carlos Canelo-Aybar, Pablo Alonso-Coello, Tatyana Shamliyan, Holger J. Schünemann, GRADE Working Group

PII: S0895-4356(23)00200-7

DOI: https://doi.org/10.1016/j.jclinepi.2023.08.001

Reference: JCE 11147

To appear in: *Journal of Clinical Epidemiology*

Received Date: 26 January 2023

Revised Date: 28 June 2023

Accepted Date: 2 August 2023

Please cite this article as: Xie F, Shemilt I, Vale L, Ruiz F, Drummond MF, Lord J, Herrmann KH, Rojas MX, Zhang Y, Canelo-Aybar C, Alonso-Coello P, Shamliyan T, Schünemann HJ, GRADE Working Group, GRADE guidance 23: Considering cost-effectiveness evidence in moving from evidence to health-related recommendations, *Journal of Clinical Epidemiology* (2023), doi: https://doi.org/10.1016/j.jclinepi.2023.08.001.

GRADE guidance 23: Considering cost-effectiveness evidence in moving from evidence to health-related recommendations

Feng Xie[1,2], Ian Shemilt[3,4], Luke Vale[3,5], Francis Ruiz[3,6], Michael F Drummond[3,7], Joanne Lord[3,8], Kirsten H Herrmann[3,9], María Ximena Rojas[10] Yuan Zhang[1], Carlos Canelo-Aybar[10], Pablo Alonso-Coello[10,11], Tatyana Shamliyan[12], Holger J Schünemann[1,13,14], GRADE Working Group

[1]Department of Health Research Methods, Evidence and Impact (formerly Clinical Epidemiology and Biostatistics), McMaster University, Health Sciences Centre, 1280 Main Street West, Hamilton, ON L8S 4K1 Canada

[2]Centre for Health Economics and Policy Analysis (CHEPA), Faculty of Health Sciences, McMaster University

[3]Campbell & Cochrane Economics Methods Group (http://methods.cochrane.org/economics/welcome).

[4]EPPI-Centre, University College London, 10 Woburn Square, London WC1H 0NR United Kingdom.

[5]Health Economics Group, Population Health Sciences Institute, Newcastle University, The Baddiley-Clark Building, Richardson Road, Newcastle upon Tyne NE2 4AX United Kingdom.

[6]Department of Global Health and Development, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT United Kingdom.

[7]Centre for Health Economics, University of York, Alcuin 'A' Block, Heslington, York YO10 5DD United Kingdom.

[8]Southampton Health Technology Assessments Centre, University of Southampton 1st Floor, Epsilon House, Enterprise Road Southampton Science Park, Southampton SO16 7NS United Kingdom.

[9]ResearchGCP, Dunantstr.5, 85521 Ottobrunn, Germany.

[10]Iberoamerican Cochrane Centre, Department of Clinical Epidemiology and Public Health, Biomedical Research Institute Sant Pau (IIB Sant Pau), Barcelona, Spain.

[11]CIBER of Epidemiology and Public Health, CIBERESP, Madrid, Spain.

[12]American College of Physicians, 190 N Independence Mall West Philadelphia, PA, USA, 19106

[13]Department of Medicine, McMaster University, Health Sciences Centre, 1280 Main Street West, Hamilton, ON L8S 4K1 Canada

[14]Department of Biomedical Sciences, Humanitas University, Via Rita Levi Montalcini 4 - 20090 Pieve Emanuele (Milano), Italy

What is new?

Key findings

When using the GRADE Evidence to Decision (EtD) framework, the approach to integration of evidence about cost-effectiveness may vary by available evidence. When there are multiple cost-effectiveness analyses, top-level tabulated summaries of key findings, and quality of the evaluations can be developed and presented as evidence profiles. In the absence of relevant cost-effectiveness evidence, the cost of preventing presents one alternative.

What this adds to what was known?

GRADE EtD frameworks are widely used, but users are challenged about how to integrate cost-effectiveness evidence. This guidance describes practical scenarios for guideline developers and other decision-makers.

What is the implication and what should change now?

Practical information for summarizing and presenting cost-effectiveness evidence is key to its integration in the development of recommendations.

Abbreviations

EtD – Evidence-to-decision
GRADE – Grading of Recommendations Assessment, Development and Evaluation

Keywords

GRADE, Guidelines, Systematic reviews, Cost effectiveness analysis

Abstract

Background

This is the 23rd in a series of articles describing the GRADE approach to grading the certainty of evidence and strength of recommendations for systematic reviews, health technology assessments, and clinical guideline development.

Objectives

We outline how resource utilization and cost-effectiveness analyses (CEAs) are integrated into health-related recommendations, using the GRADE Evidence-to-Decision (EtD) framework.

Methods

Through iterative discussions and refinement, in-person, and online meetings, and through email communication, we developed draft guidance to incorporating economic evidence in the formulation of health-related recommendations. We developed scenarios to operationalize the guidance. We presented a summary of the results to members of the GRADE Economic Evaluation Project Group.

Results

We describe how to estimate the cost of preventing (or achieving) an event to inform assessments of cost-effectiveness of alternative treatments, when there are no published economic evaluations. Evidence profiles and Summary of Findings (SoF) tables based on systematic reviews of CEAs can be created to provide top-level summaries of results and quality of multiple published economic evaluations. We also describe how this information could be integrated in GRADE's EtD frameworks to inform health-related recommendations. Three scenarios representing various levels of available cost-effectiveness evidence were used to illustrate the integration process.

Conclusions

This GRADE guidance provides practical information for presenting cost-effectiveness data, and its integration in the development of health-related recommendations, using the EtD framework.

1. Introduction

This is the 23rd in a series of articles describing GRADE guidance to grading the certainty of evidence and strength of recommendations for authors of systematic reviews and guideline developers. In the 10th article of the series we described challenges with rating certainty of evidence (also known as quality of the evidence or confidence in an effect estimates) for resource use which is included in GRADE evidence profiles and Summary of Findings (SoF) tables (1). The evidence profile and summary of finding table have been validated in multiple studies including randomized trials (2, 3, 4, 5). We also introduced health economic evaluation as a methodological framework for assessing the relative efficiency (cost-effectiveness) of alternative management strategies (interventions) (6).

An economic evaluation may be conducted either as part of an empirical study, such as a clinical trial or registry study, or using decision modelling. Trial-based economic evaluations measure the cost-effectiveness of alternative management strategies amongst trial participants by combining data on clinical effects, resource use, unit costs, and patient reported outcomes (e.g., health state utilities) that are sometimes collected as part of a trial (7). In contrast, model-based economic evaluations compare the cost-effectiveness of alternative management strategies in a specified target patient population, through mathematical modelling techniques. Typically, they do this by synthesising clinical and economic data, gathered from a number of different research and administrative sources, and predicting incremental cost-effectiveness of the alternative strategies (8). Trial- and model-based economic evaluations are complementary approaches, for example, modelling can be used in extrapolating findings beyond the trial follow-up. The 15th article of the series (9), and the subsequent GRADE Working Group papers introducing the GRADE Evidence to Decision (EtD) frameworks (10, 11, 12, 13, 14), highlight resource use and cost-effectiveness as key criteria that may determine the direction and strength of guideline recommendations. Decision-makers may consider resource use and cost-

effectiveness alongside the rest of the factors considered in the EtD frameworks (e.g. equity, and acceptability) (8, 9, 15, 16).

2. **Objective**

This article provides guidance on the incorporation of cost-effectiveness evidence into the GRADE EtD framework when making health-related decisions. Specifically, we present three scenarios and explain how to integrate economic evidence, and how this information may help inform the direction and strength of recommendations.

3. **Methods**

We conducted most of this work, prior to GRADE's creation of project groups and new processes for article approval and processes (17, 18). It was developed using the approach that GRADE pursued for earlier guidance articles, including discussion of the lead authors at GRADE meetings, with input from the GRADE working group. Through iterative discussions and refinement during in-person and online meetings and through email communication, we developed draft guidance to using cost-effectiveness data in the formulation of recommendations and making decisions in 2013 and first submitted for publication in January 2019. Given changes in the leadership of the project group, substantial revisions that addressed the original peer reviewers' comments were completed from 2022 to 2023. This was followed by a new round of peer review. Further revisions based on comments and input received were made by the current authors to finalize this guidance.

We developed three key scenarios that highlight situations that guideline developers and other decision-makers often face, when they use GRADE EtDs and similar frameworks. These scenarios were chosen to reflect various levels of available cost-effectiveness evidence, ranging from none to multiple relevant published economic evaluations. Decision-makers should use the best available evidence to inform recommendations. A systematic literature would be needed to search and identify published evidence on cost effectiveness (19). However, often there might be no relevant evidence on

cost-effectiveness at the time of guideline development. Scenario 1 represented a situation when there is no published cost-effectiveness evidence. In this situation a simplified approach to consider cost and effectiveness may be useful. One such approach is by calculating the cost of preventing an event (COPE). Scenarios 2 and 3 dealt with the situations where there are published CEAs. When there was published CEAs, we developed economic evidence profiles that describe key findings of the CEAs including cost of treatment, incremental cost per patient, incremental effect per patient, incremental cost effectiveness ratio (ICER), uncertainty, and overall certainty rating. For each scenario, we used published reviews or guidelines developed using the GRADE approach to illustrate how to integrate cost-effectiveness evidence into the EtD framework. We described how to summarize cost-effectiveness evidence contained in the economic evidence profile in the relevant sections in the EtD framework and integrate into guideline recommendations.

4. Results

Scenario 1: No published cost-effectiveness evidence

Decision-makers should use the best available evidence to inform recommendations. However, often there might be no relevant evidence on cost-effectiveness at the time of guideline development. Therefore, guideline panels may undertake a structured discussion or a simple 'back-of-the-envelope' calculation. One approach is a limited form of considering cost-effectiveness, such as the COPE.

We refer to the scenario in which the guideline panel could consider whether it is plausible to estimate cost-effectiveness without a cost-effectiveness analysis (CEA), given the evidence on safety and effectiveness of the intervention. For example, if an intervention is very low-cost and improves health (and therefore prevents the need for subsequent treatment) then it would be highly plausible that the intervention would be less costly (i.e., economically dominant).

The above-mentioned situation can occur, but when a decision is not clear, then simple modelling approaches can be useful. The COPE represents the cost of preventing a single adverse outcome or achieving a single beneficial outcome (20). COPE is the product of the number needed to treat (NNT) to prevent a single adverse event, or achieve a single additional success, multiplied by the total cost of treatment for a single individual over a specified time frame. For instance, if one needs to treat 20 patients for 2 years to prevent a single premature death and the cost of treatment for a single patient over 2 years is $10,000, the COPE is 20 X $10,000, or $200,000. Evidently, COPE differs from properly conducted CEAs because it considers only the direct costs of providing the intervention. It does not include consideration of costs associated with the impact of the intervention on downstream use of resources or any other costs.

For example, when guideline developers formulated a recommendation regarding the use of pulmonary rehabilitation for patients with chronic obstructive pulmonary disease (COPD) and recent exacerbation of their disease (21), none of the trials included in the source systematic review reported estimates of resource use or a CEA associated with pulmonary rehabilitation. Table 1 shows NNTs with respect to the outcome 'mortality', derived from corresponding pooled effect size estimates reported in a trustworthy systematic review of pulmonary rehabilitation, when compared with usual care without rehabilitation in adults with COPD (21). The median length of follow-up amongst the 6 RCTs included in estimating the pooled treatment effect for the outcome 'mortality' was 12 months. The inference from NNTs was that 23 COPD patients would need to be treated with an initial program of pulmonary rehabilitation, following a recent exacerbation of their disease, in order to prevent one death during a subsequent 12-month period. The estimated up-front costs of treating one patient using a program of pulmonary rehabilitation were assumed to be $3,000 to $5,000 per patient (for good practices on costing, please refer to Drummond *et al.*, (22)). In this example, there is no need to annualise the costs of treatment because pulmonary rehabilitation is provided as a discrete, time-limited program. The

COPEs are shown in the final column of Table 1. Estimates of the cost of preventing a single death through the use of pulmonary rehabilitation for patients with COPD (with recent exacerbation of their disease) over 12 months range between approximately $70,000 to $116,000 (2010 US Dollars). Of note, with this approach we are not grading the certainty of the COPE for this outcome; rather we draw upon the benefit/harm and cost outcomes, based on the evidence profile published. Such estimates help guideline developers to answer the question how large the resource requirements (costs) are in various settings. Guideline developers should define, ideally, a priori thresholds of acceptable in various countries resource requirements.

[Insert Table 1 about here]

The systematic review also reported a pooled estimate of the effect of pulmonary rehabilitation on the rate of hospital readmission, over a median of 9 months follow-up (21). We calculated that 5 patients would need to be treated to prevent one hospital readmission (21). One implication of this finding is that, if costs associated with avoided hospital readmissions are the only important impact of pulmonary rehabilitation on downstream use of resources, then COPE for mortality based on treatment costs would overestimate the cost of preventing a single death due to the cost saving from avoided readmissions. A further implication is that, given a sufficient number of hospital readmissions avoided, cost-savings from avoided hospitalization in the future may entirely offset up-front treatment costs, and pulmonary rehabilitation would become a cost-saving management strategy compared with usual care.

It is possible to elaborate the basic form of COPE presented above by combining respective NNTs for ‘mortality’ and ‘hospital readmission’, with information about the cost of treatment and the cost of a hospital bed day. By varying the estimated impact on length of stay for readmitted patients it would be possible to estimate the ‘switching point’ at which cost-savings would entirely offset up-front treatment costs. It should be noted that COPE

might be more suitable for clinical events than for other types of outcomes (e.g., quality of life). Another limitation of using COPE is that the number of outcomes (and hence the number of trade-offs between outcomes) considered by guideline panels increases, more sophisticated economic modelling techniques will be needed to inform this process. This more sophisticated modelling can be thought of as a higher level synthesis, as it allows for synthesizing evidence for multiple outcomes (e.g., use of quality-adjusted life year (QALY)) and associated resource use (costs). This modelling activity can be proportionate to the resources available to conduct this work for a decision and the circumstance, with COPE being the fastest and most frugal approach.

Scenario 2: Limited source of cost-effectiveness evidence

Sometimes, evidence on resource use and cost-effectiveness is limited and may come from a single source. In 2020, the World Health Organization (WHO) developed guidelines for the management and care of drug-resistant tuberculosis (23). For multidrug- or rifampicin-resistant tuberculosis (MDR/RR-TB), the guidelines considered whether all-oral bedaquiline-containing shorter regimen of 9–12 months compared with standard care of shorter regimen with injectable agent recommended by WHO safely improve patient outcomes (23). In this guideline, one CEA was identified, critically appraised, and included (24). This study used a Markov cohort model to compare cost-effectiveness of the shorter (6-9 months) all-oral regimen containing bedaquiline, pretomanid and linezolid (BPaL) compared with context-specific standard care for patients with extensively drug resistant tuberculosis (XDR-TB) in South Africa, Georgia, and the Philippines (24). Clinical effectiveness was derived from a single non-randomized study (24). Costs were estimated based on the literature and local consultation and converted to 2018 US dollars (24). The outcome was measured using disability-adjusted life year (DALY) averted, representing the number of years of healthy life made possible by a given intervention. The analysis was conducted with a lifetime horizon from a health sector's perspective. Both

deterministic and probabilistic sensitivity analyses were conducted to assess the uncertainty of CEA (24).

We created an economic evidence profile to summarise the principal findings and the certainty of CEA from the only available evaluation (Table 2). The cost of BPaL varied from $245 in the Philippines to $362 in South Africa, while the cost of standard care ranged from $450 to $624 at the intensive phase and from $88 to $239 at the continuation phase. The saving in total cost and DALYs averted by BPaL compared with standard care ranged from $1410 and 8.73 in South Africa to $4059 and 10.14 in Georgia (that is, in both settings BPaL was more effective and less costly and hence based on these data dominant).

[Insert Table 2 here]

The EtD framework summarizes the key design and the main findings of the CEA and the recommendations of the guideline panel (Table 3) (23). The panel considered the cost saving by the 9–12-month all-oral regimen was moderate compared with the short injectable-medication containing regimen. Some key factors related to resource uses were noted, including existing stock of second-line medications during the transition, diagnostic capacity, and drug safety monitoring and management capacity. The magnitude of saving was dependent on the context, where the extent to which higher drug costs were outweighed by the reduced costs of delivering injectable agents and treating recurrent and secondary cases may vary. The certainty of evidence was considered moderate as the estimates on resource use were similar across settings.

For cost-effectiveness, BPaL was dominant compared with standard care across all three settings. The cost-effectiveness results were robust in most sensitivity analyses, except when the drug prices became extremely expensive. Therefore, the panel considered that the cost-effectiveness evidence probably favors the shorter all-oral regimen over the injectable-containing regimens for this patient population.

The panel's final recommendation was that a shorter, all-oral, bedaquiline-containing regimen of 9–12 months' duration be used in eligible patients with confirmed MDR/RR-TB, who have not been exposed to treatment with second-line TB medicines used in this regimen for more than 1 month, and in whom resistance to fluoroquinolones has been excluded. The conditionality of the recommendation was mainly attributed to the very low certainty in the efficacy evidence, and the diagnostic capacity requirement (23).

[Insert Table 3 here]

Scenario 3: Multiple sources of cost-effectiveness evidence

The number of publications on CEAs has been growing over the last decades (25). Guideline developers may have to consider multiple CEAs when formulating recommendations. In a recommendation, from the European Commission Initiative Breast Cancer Guidelines, on whether an organized mammography screening program (MSP) or an opportunistic mammography screening program (OS) should be used for early diagnosis of breast cancer in asymptomatic women (26, 27, 28), three relevant CEAs for the European setting were identified (29, 30, 31). The economic evidence profile (Table 4) shows that in the two analyses conducted in Switzerland, one reported that the MSP was associated with €1473 higher cost but gained 0.02 life year compared with OS (incremental cost-effectiveness ratio = €75,602) (29), while the other found that MSP gained more life years at lower costs (i.e., "dominant") (30). The third CEA also reported that MSP was dominant to OS in terms of cost and life years in Austria (31). Within these studies, the cost-effectiveness findings were robust in most sensitivity analyses (Table 4).

[Insert Table 4 here]

In the EtD framework (Table 5), the guideline panel highlighted the differences in the cost of MSP and OS, as well as incremental cost-effectiveness ratio among the published CEAs (26, 28). Different models (i.e.,

Markov cohort model vs microsimulation model) with the cost and effect data input from different time periods may contribute to the discrepancies. The participation rate of the screening program was also found to affect the cost-effectiveness estimates. There was consensus in the panel that the resources required for implementing the screening program would vary across countries. The panel noted that the MSP was associated with higher benefits, with the incremental cost varying across countries. As a result, the cost-effectiveness probably favors MSP.

The guidelines panel recommended using an organised screening programme with MSP for early detection of breast cancer in asymptomatic women. This was a strong recommendation, with low certainty in the evidence on resource uses. The panel also noted that additional research on the cost-effectiveness of organised screening in different settings was needed (26, 28).

[Insert Table 5 here]

5. **Discussion**

GRADE recognises that evidence for relative efficiency of intervention alternatives can be useful in moving from evidence to recommendations and suggests including this information in GRADE EtDs. Some guideline developers have adapted GRADE evidence profiles and SoF tables to include summaries of the quality and principal findings of CEAs (32). For the purposes of this GRADE guideline article, we created economic evidence profiles that summarize key information about cost-effectiveness, including the cost of intervention, the cost of comparator, the incremental cost, the incremental effect, incremental cost-effectiveness ratio, uncertainty from deterministic or probabilistic sensitivity analysis, and overall certainty of evidence. The economic evidence profile provides relevant information for the EtD process, documentation of the judgments decision makers have to make and informs recommendations.

Strengths and Limitations

Note that the evidence profile does not substitute for a detailed description and appraisal of methods and results of each included CEA. Guidelines often lack appraisals of the certainty of clinical benefits and harms (models inputs) (33) and the certainty of CEA evidence (34). Therefore, before incorporating economic evidence into the EtD framework, we recommend appraisal of reporting and methodologic quality of individual CEA that includes the assessment of certainty of all clinical input variables (6, 35, 36, 37). We also recommend transparent appraisal of the certainty of evidence from all relevant CEAs using GRADE domain of the directness, risk of bias in the body of evidence, heterogeneity in incremental cost per QALY estimates, and the publication bias (6). Countries are increasingly developing their own "reference cases" for economic evaluation to guide the development and interpretation of CEAs for use in decision-making (38, 39). These reference cases would shape the assessment of relevance and quality of economic evidence in specific contexts.

There are challenges in summarising economic evidence in a clear but concise way. The evidence profile examples presented in this article focus on pairwise comparisons. However, economic evaluations often compare multiple treatments. In order to present an incremental cost-effectiveness ratio for each pairwise comparison in a series of economic evidence profiles, each of which accompanies the corresponding GRADE evidence profile, it will not always be appropriate or useful to disaggregate the results of an economic evaluation. In an incremental analysis each option is compared with the next most expensive, non-dominated option. Not all incremental cost-effectiveness ratios are comparable. Many decision makers recommend the use of generic outcome measure (e.g., QALY or DALY derived using very specific methods) in economic evaluations to allow for broad comparisons (40, 41). Nevertheless, there are some other sources of variability that make the incremental cost-effectiveness ratio not comparable such as differences in the perspective of the analysis and the time horizon used, among others.

Another challenge arises when clinical or economic findings vary across patient groups. If relative treatment effects for clinical outcomes are inconsistent across different patient sub-groups, the findings for each patient sub-group may be presented in separate rows of a GRADE evidence profile, each accompanied by an economic evidence profile. However, if the relative treatment effects for clinical outcomes are constant across sub-groups, but the economic findings differ across sub-groups (for example, if incremental cost-effectiveness varies with baseline risk), it may be preferable to present a single GRADE evidence profile and an associated economic evidence profile with separate rows for each patient sub-group. Or, if economic findings vary across a continuum of risk, it may be simpler to present a threshold at which treatment becomes cost effective. We now working with the GRADE project group on economic evidence on an integration of these types of evidence profiles and summary of findings tables in GRADE's app GRADEpro to facilitate their development.

Although the number of economic evaluations has been increasing, cost-effectiveness evidence may not always be available for guideline developers, especially for innovative new treatments. Calculations using the COPE could offer a quick assessment on cost-effectiveness of alternative treatment strategies. However, this method should not replace the need for economic modelling when substantial resource use is expected throughout the course of treatment, and the comparison involves trade-offs among multiple important outcomes. Here, best practice would suggest the use of generic measures such as QALY or DALY (39).

There is rarely a simple relationship between research questions, a body of health evidence, an economic evaluation or a set of economic evaluations and a set of recommendations. While standard formats for economic evidence are useful, expressing those in complex evidence networks remains an area of research.

6. Conclusions

Cost-effectiveness evidence has played an increasingly important role in informing health-related recommendations. GRADE evidence profiles and SoF tables can be used to inform a quick assessment of cost-effectiveness of alternative treatments when there are no published economic evaluations. More often than not, there may be multiple relevant economic evaluations for which economic evidence profiles can provide top-level summaries of the quality and results of the evaluations. The evidence summary helps guideline panels to move from evidence to recommendation through the EtD framework. Further work is needed to address the presentational challenges of summarising complex clinical and economic evidence networks in a clear but concise way and for the assessment of certainty in the evidence of modelled evidence (6).

Declaration of Interest

The GRADE system has been developed by the GRADE Working Group. The named authors drafted and revised this article. FX is a member of the EuroQol Group and elected deputy chair of its Executive Committee. Seven authors of this article (IS, LV, FR, MFD, JL, KH, MXR) are all convenors or members of the Campbell and Cochrane Economics Methods Group, a Cochrane Methods Group. FX, IS, LV, YZ, MXR, CCA, PAC and HJS are members of the GRADE working group. The views expressed in this article are those of the authors and not necessarily those of any organization.

Acknowledgement

The authors would like to thank the following members of the Grade Working Group for helpful comments provided on earlier drafts of this article: Erin Graybill (Newcastle University), Massimo Brunetti (Azienda USL, Modena), Gerald Gartlehner (Danube University Krems), Gordon Guyatt (McMaster University) and Susan Norris (World Health Organisation).

REFERENCES

1. Brunetti M, Shemilt I, Pregno S, Vale L, Oxman AD, Lord J, et al. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. J Clin Epidemiol. 2013;66(2):140-50.

2. Rosenbaum SE, Glenton C, Oxman AD. Summary-of-findings tables in Cochrane reviews improved understanding and rapid retrieval of key information. J Clin Epidemiol. 2010;63(6):620-6.

3. Carrasco-Labra A, Brignardello-Petersen R, Santesso N, Neumann I, Mustafa RA, Mbuagbaw L, et al. Improving GRADE evidence tables part 1: a randomized trial shows improved understanding of content in summary of findings tables with a new format. J Clin Epidemiol. 2016;74:7-18.

4. Vandvik PO, Santesso N, Akl EA, You J, Mulla S, Spencer FA, et al. Formatting modifications in GRADE evidence profiles improved guideline panelists comprehension and accessibility to information. A randomized trial. J Clin Epidemiol. 2012;65(7):748-55.

5. Glenton C, Santesso N, Rosenbaum S, Nilsen ES, Rader T, Ciapponi A, et al. Presenting the results of Cochrane Systematic Reviews to a consumer audience: a qualitative study. Med Decis Making. 2010;30(5):566-77.

6. Brozek JL, Canelo-Aybar C, Akl EA, Bowen JM, Bucher J, Chiu WA, et al. GRADE Guidelines 30: the GRADE approach to assessing the certainty of modeled evidence-An overview in the context of health decision-making. J Clin Epidemiol. 2021;129:138-50.

7. Petrou S, Gray A. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. BMJ. 2011;342:d1548.

8. Drummond MF, Sculpher, M.J., Claxton, K., Stoddart, G.L., Torrance, G. W. Economic evaluation using decision analytic modelling. In: Drummond MF, Sculpher, M.J., Claxton, K., Stoddart, G.L., Torrance, G. W., editor. Methods for the Economic Evaluation of Health Care Programmes 4th ed. Oxford: Oxford University Press; 2015.

9. Andrews JC, Schunemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, et al. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. J Clin Epidemiol. 2013;66(7):726-35.

10. Schunemann HJ, Wiercioch W, Brozek J, Etxeandia-Ikobaltzeta I, Mustafa RA, Manja V, et al. GRADE Evidence to Decision (EtD) frameworks for adoption, adaptation, and de novo development of trustworthy recommendations: GRADE-ADOLOPMENT. J Clin Epidemiol. 2017;81:101-10.

11. Neumann I, Brignardello-Petersen R, Wiercioch W, Carrasco-Labra A, Cuello C, Akl E, et al. The GRADE evidence-to-decision framework: a report of its testing and application in 15 international guideline panels. Implement Sci. 2016;11:93.

12. Alonso-Coello P, Oxman AD, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. BMJ. 2016;353:i2089.

13. Alonso-Coello P, Schunemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ. 2016;353:i2016.

14. Schunemann HJ, Mustafa R, Brozek J, Santesso N, Alonso-Coello P, Guyatt G, et al. GRADE Guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health. J Clin Epidemiol. 2016;76:89-98.

15. Sculpher MJ, Pang FS, Manca A, Drummond MF, Golder S, Urdahl H, et al. Generalisability in economic evaluation studies in healthcare: a review and case studies. Health Technol Assess. 2004;8(49):iii-iv, 1-192.

16. Piggott T, Baldeh T, Dietl B, Wiercoch W, Nieuwlaat R, Santesso N, et al. Standardized wording to improve efficiency and clarity of GRADE EtD frameworks in health guidelines. J Clin Epidemiol. 2022;146:106-22.

17. Schunemann HJ, Brennan S, Davoli M, Mustafa RA, Akl EA, Meerpohl JJ, et al. The development methods of official GRADE articles and requirements for claiming the use of GRADE – a statement by the GRADE Guidance Group. J Clin Epidemiol. accepted for publication pending revisions.

18. McGowan J, Akl EA, Coello PA, Brennan S, Dahm P, Davoli M, et al. Update on the JCE GRADE series and other GRADE article types. J Clin Epidemiol. 2021;140:163-4.

19. Aluko PG, E.; Craig, D.; Henderson, C.; Drummond, M.; Wilson, E.C.F.; Robalino, S.; Vale, L. Economic Evidence. In: Higgins JPTT, J.; Chandler, J.; Cumpston, M.; Li, T.; Page, M.J.; Welch, V.A., editor. Cochrane Handbook for Systematic Reviews of Interventions Version 6.3 ed: Cochrane Collaboration; 2022.

20. Maharaj R. Adding cost to number needed to treat: the COPE statistic. Evid Based Med. 2007;12(4):101-2.

21. Puhan MA, Gimeno-Santos E, Cates CJ, Troosters T. Pulmonary rehabilitation following exacerbations of chronic obstructive pulmonary disease. Cochrane Database Syst Rev. 2016;12:CD005305.

22. Drummond MS, M.J.; Claxton, K.; Stoddart, G.L.; Torrance, G.W. Methods for the Economic Evaluation of Health Care Programmes. Fourth Edition ed. Oxford: Oxford University Press; 2015.

23. WHO consolidated guidelines on tuberculosis: Module 4: treatment - drug-resistant tuberculosis treatment. Geneva: World Health Organization; 2020.

24. Gomez GB, Siapka M, Conradie F, Ndjeka N, Garfin AMC, Lomtadze N, et al. Cost-effectiveness of bedaquiline, pretomanid and linezolid for treatment of extensively drug-resistant tuberculosis in South Africa, Georgia and the Philippines. BMJ Open. 2021;11(12):e051521.

25. Neumann PJ, Fang CH, Cohen JT. 30 years of pharmaceutical cost-utility analyses: growth, diversity and methodological improvement. Pharmacoeconomics. 2009;27(10):861-72.

26. European Commission Initiative on Breast Cancer (ECIBC): European guidelines on breast cancer screening and diagnosis. Italy: European Commissio; 2019.

27. Schunemann HJ, Lerda D, Dimitrova N, Alonso-Coello P, Grawingholt A, Quinn C, et al. Methods for Development of the European Commission

Initiative on Breast Cancer Guidelines: Recommendations in the Era of Guideline Transparency. Ann Intern Med. 2019;171(4):273-80.

28. Schunemann HJ, Lerda D, Quinn C, Follmann M, Alonso-Coello P, Rossi PG, et al. Breast Cancer Screening and Diagnosis: A Synopsis of the European Breast Guidelines. Ann Intern Med. 2020;172(1):46-56.

29. Neeser K, Szucs T, Bulliard JL, Bachmann G, Schramm W. Cost-effectiveness analysis of a quality-controlled mammography screening program from the Swiss statutory health-care perspective: quantitative assessment of the most influential factors. Value Health. 2007;10(1):42-53.

30. de Gelder R, Bulliard JL, de Wolf C, Fracheboud J, Draisma G, Schopper D, et al. Cost-effectiveness of opportunistic versus organised mammography screening in Switzerland. Eur J Cancer. 2009;45(1):127-38.

31. Schiller-Fruehwirth I, Jahn B, Einzinger P, Zauner G, Urach C, Siebert U. The Long-Term Effectiveness and Cost Effectiveness of Organized versus Opportunistic Screening for Breast Cancer in Austria. Value Health. 2017;20(8):1048-57.

32. Agache I, Song Y, Alonso-Coello P, Vogel Y, Rocha C, Sola I, et al. Efficacy and safety of treatment with biologicals for severe chronic rhinosinusitis with nasal polyps: A systematic review for the EAACI guidelines. Allergy. 2021;76(8):2337-53.

33. Mensch A, Beck T, Civello D, Kunigkeit C, Lachmann N, Stock S, et al. Applying GRADE Criteria to Clinical Inputs to Cost-Effectiveness Modeling Studies. Pharmacoeconomics. 2018;36(8):987-94.

34. Riva JJ, Bhatt M, Brunarski DJ, Busse JW, Martins CC, Xie F, et al. Guidelines that use the GRADE approach often fail to provide complete economic information for recommendations: A systematic survey. J Clin Epidemiol. 2021;136:203-15.

35. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Statement: Updated Reporting Guidance for Health Economic Evaluations. Clin Ther. 2022;44(2):158-68.

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(7052):275-83.

37. Kim DD, Do LA, Synnott PG, Lavelle TA, Prosser LA, Wong JB, et al. Developing Criteria for Health Economic Quality Evaluation Tool. Value Health. 2023.

38. Wilkinson T, Sculpher MJ, Claxton K, Revill P, Briggs A, Cairns JA, et al. The International Decision Support Initiative Reference Case for Economic Evaluation: An Aid to Thought. Value Health. 2016;19(8):921-8.

39. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine. JAMA. 2016;316(10):1093-103.

40. National Institute for Health and Care Excellence. Guide to the Methods of Technology Appraisal. 2013.

41. Canadian Agency for Drugs and Technologies in Health. Guidelines for the Economic Evaluation of Health Technologies: Canada. 4th Edition ed2017.

Tables

Table 1. Cost of preventing a death (COPE): Pulmonary rehabilitation vs. usual care

Outcome	Risk with control	Risk with rehabilitation	NNT	Cost	COPE
Mortality	150 per 1000	107 per 1000	23	$3,000	$69,000
Mortality	150 per 1000	107 per 1000	23	$4,000	$92,000
Mortality	150 per 1000	107 per 1000	23	$5,000	$115,000

NNT=number needed to treat was calculated based on absolute differences in mortality at 12 months after pulmonary rehabilitee versus usual care in the review (21).
COPE=NNT*Cost (e.g., 23X$3,000=$69,000)

Table 2. Economic evidence profile: BPaL vs standard of care for XDR-TB

Study	*BPal*	*Standard of care*	*Incremental cost per patient*	*Incremental effect per patient*	*ICER*	*Uncertainty*	*Certainty*
	Cost/month	*Cost/month*					
One cost-effectiveness analysis[1]	South Africa $362	South Africa $624 (intensive phase) $239 (continuation phase)	−$1,410	8.73 DALYs averted	SoC dominated	Across all three settings, the finding that BPal was cost saving and more effective was robust in deterministic and probabilistic sensitivity analyses. The only exception was when the drug prices for BPaL became extremely expensive.	⊕⊕⊕◯ Moderate
	Georgia $245	Georgia $450 (intensive phase) $89 (continuation phase)	−$4,059	10.14 DALYs averted	SoC dominated		
	The Philippines $245	The Philippines $455 (intensive phase) $88 (continuation phase)	−$3,864	9.93 DALYs averted	SoC dominated		

BPaL: Bedaquiline, pretomanid and linezolid. XDR-TB: extensively drug-resistant tuberculosis; Soc: standard of care. DALY: disability-adjusted life year.

1. Gomez GB, Siapka M, Conradie F, et al. Cost-effectiveness of bedaquiline, pretomanid and linezolid for treatment of extensively drug-resistant tuberculosis in South Africa, Georgia, and the Philippines. BMJ Open 2021;11:e051521.

Table 3. Evidence to Decision Summary: resources required and cost-effectiveness all-oral shorter regimen containing bedaquiline vs standard care for multidrug- or rifampicin-resistant tuberculosis[1]

Resources required How large are the resource requirements (costs)?	
JUDGEMENT	**RESEARCH EVIDENCE**
● moderate savings	Compared with the injectable-containing shorter regimen, the 9–12-month all-oral regimen is projected to be both cost-saving and more effective than the injectable-containing shorter regimen under nearly all modelled conditions. The projected cost savings average US$ 1000 (2019) in South Africa and depend primarily on the extent to which higher drug costs are outweighed by the reduced costs of delivering injectable agents, and treating recurrent and secondary cases (about US$ 2000 per patient is saved in settings with two-times-higher health care costs, versus < US$ 100 saved in settings with high drug costs but low health care costs).
Certainty of evidence of required resources What is the certainty of the evidence of resource requirements (costs)?	
JUDGEMENT	**RESEARCH EVIDENCE**
● moderate	The cost savings relative to a short injectable-containing regimen are robust, except at extremes with drug costs, health care costs or the cost of bedaquiline and other companion drugs.
Cost effectiveness Does the cost-effectiveness of the intervention favor the intervention or the comparison?	
JUDGEMENT	**RESEARCH EVIDENCE**
● Probably favors the intervention	A cost–effectiveness model was developed, incorporating estimated health system costs with drug procurement, health care delivery, adverse events, retreatments, secondary cases, and morbidity and mortality associated with TB mortality and recurrence, treatment duration, drug toxicity and TB transmission (ref). Costs and cost–effectiveness were evaluated in South Africa, with sensitivity analyses representing a range of drug and health care costs, high and low HIV co-prevalence, 95% CIs for estimates of relative efficacy derived from statistical analysis of patient cohort data, and uncertainty in the values of parameters representing the natural history of TB. Compared with the injectable-containing shorter regimen, the 9–12-month all-oral regimen is projected to be both cost saving and more effective than the injectable-containing shorter regimen under nearly all modelled conditions. The projected cost savings average US$ 1000 (2019) in South Africa and depend primarily on the extent to which higher drug costs are outweighed by reduced costs of delivering injectable agents, and treating recurrent and secondary cases (about US$ 2000 saved per patient is saved in settings with two-times-higher health care costs, versus < US$ 100 saved in settings with high drug costs but low health care costs). In a scenario where the all-oral regimen was no longer cost-saving because bedaquiline prices were fourfold higher than current pricing in South Africa or via the GDF, the incremental cost–effectiveness ratio of the all-oral regimen was US$ 400 per disability adjusted life-year averted (range: US$ 100–US$ 900 across 95% CI for relative regimen efficacy).

[1]WHO consolidated guidelines on tuberculosis: Module 4: treatment - drug-resistant tuberculosis treatment. Geneva: World Health Organization; 2020.

Table 4. Economic evidence profile: organized mammography screening vs an opportunistic or non-organized mammography screening program for breast cancer in asymptotic women

Study	*Organized mammography screening*	*Opportunistic or non-organized mammography screening program*	*Incremental cost per woman*	*Incremental effect per woman*	*ICER*	*Uncertainty*	*Certainty*
	Cost/woman	*Cost/woman*					
Three CEAs[1,2,3]	Switzerland[1] €4149	Switzerland[1] €2676	€1473	0.02 life-year	€75,602 per life year gained	ICER was sensitive to the mortality and incidence of breast cancer, cost of biopsy and the sensitivity of MSP, as well as to the participant rate. The distribution of ICER ranging from $45,000 to $135,000, with highest probability of occurrence at $75,000.	⊕⊕⊕◯ Moderate
	Switzerland[2] €1630	Switzerland[2] €1637	-€7	0.0171 life-year 0.0159 QALY	Dominant	Different participation rates for the screening program were compared. In sensitivity analyses with different levels of false negative rate, the cost effectiveness finding was consistent across the scenarios.	
	Austria[3] €1667	Austria[3] €1678	-€11	0.009 life-year	Dominant	MSP remained dominant in most deterministic sensitivity analyses. In probabilistic sensitivity analysis, the probability of MSP being cost effective was 70% at the threshold of €50,000.	

CEA: cost effectiveness analysis; ICER: incremental cost-effectiveness ratio; MSP: organized mammography screening program; QALY: quality-adjusted life year

1. Neeser K, Szucs T, Bulliard JL, et al. Cost-effectiveness analysis of a quality-controlled mammography screening program from the Swiss statutory health-care perspective: quantitative assessment of the most influential factors. Value Health; 2007.
2. de Gelder R, Bulliard JL,de Wolf C, et al. Cost-effectiveness of opportunistic versus organised mammography screening in Switzerland. Eur J Cancer; 2009.
3. Schiller-Fruehwirth I, Jahn B, Einzinger P, et al. The Long-Term Effectiveness and Cost Effectiveness of Organised versus Opportunistic Screening for Breast Cancer in Austria. Value Health; 2017.

Table 5. Evidence to Decision Summary: organized mammography screening vs an opportunistic or non-organized mammography screening program for breast cancer in asymptotic women[1]

Resources required How large are the resource requirements (costs)?	
JUDGEMENT	**RESEARCH EVIDENCE**
● Varies	The GDG reviewed evidence incorporated from three research studies (16, 7, 17), notes that the costs were lower for organised screening and higher in one study. The GDG discussed that the differences in the total costs of opportunistic vs organised screening may be related to differences in the year value of costs (one study uses 2004 cost value (16) and the other one uses 2007 cost value (7), or it may also be related to the model inputs or type of modelling used (Markov modelling vs microsimulation). For this reason the quality of the evidence was downgraded to low. There was uncertainty in the results because of indirectness, information from two studies come from the same canton in Switzerland (Voud) and may not be able to be extrapolated (16, 7). The GDG noted that radiologist costs may be higher for opportunistic screening. The GDG also notes that in organised screening there may be additional administration costs, however, the cost per examination may be lower and would vary by country. The GDG notes that health related costs may be higher if additional tests beyond mammogram are ordered as a result of opportunistic screening. The GDG also notes that the definition of organised screening may vary from country to country in Europe. The GDG agreed by consensus that the resources required would vary.
Certainty of evidence of required resources What is the certainty of the evidence of resource requirements (costs)?	
JUDGEMENT	**RESEARCH EVIDENCE**
● Low	The quality is low due to imprecision and indirectness. Two studies reported organised screening as a dominant strategy, that is more effective and less costly and the other one did not. The research was further downgraded for imprecision and indirectness. The GDG suggests that local data may be available in their own languages, and not published, to inform cost-effectiveness evidence. The GDG also notes that grey literature may also inform cost-effectiveness decision-making locally.
Cost effectiveness Does the cost-effectiveness of the intervention favor the intervention or the comparison?	
JUDGEMENT	**RESEARCH EVIDENCE**
● Probably favors the intervention	The GDG reviewed evidence incorporated from three research studies (16, 7, 17). Two studies demonstrated that organised screening was dominant; in the Neeser (2007) study the ICER was 75,602 Euros per life year gained. As one may expect that health benefits are higher in organised screening, costs in relation to these benefits will probably be favoured. Costs may differ, but even if they are higher or lower, in most cases, organised screening will be cost-effective. The GDG agreed by consensus that the cost-effectiveness would probably favour the intervention in most settings.

[1]European Commission Initiative on Breast Cancer (ECIBC): European guidelines on breast cancer screening and diagnosis. Italy: European Commissio; 2019.

What is new?

Key findings

When using the GRADE Evidence to Decision (EtD) framework, the approach to integration of evidence about cost-effectiveness may vary by available evidence. When there are multiple cost-effectiveness analyses, top-level tabulated summaries of key findings, and quality of the evaluations can be developed and presented as evidence profiles. In the absence of relevant cost-effectiveness evidence, the cost of preventing presents one alternative.

What this adds to what was known?

GRADE EtD frameworks are widely used, but users are challenged about how to integrate cost-effectiveness evidence. This guidance describes practical scenarios for guideline developers and other decision-makers.

What is the implication and what should change now?

Practical information for summarizing and presenting cost-effectiveness evidence is key to its integration in the development of recommendations.

Declaration of Interest

The GRADE system has been developed by the GRADE Working Group. The named authors drafted and revised this article. FX is a member of the EuroQol Group and elected deputy chair of its Executive Committee. Seven authors of this article (IS, LV, FR, MFD, JL, KH, MXR) are all convenors or members of the Campbell and Cochrane Economics Methods Group, a Cochrane Methods Group. FX, IS, LV, YZ, MXR, CCA, PAC and HJS are members of the GRADE working group. The views expressed in this article are those of the authors and not necessarily those of any organization.