

SYSTEMATIC REVIEWS OF ECONOMIC EVALUATIONS IN HEALTH TECHNOLOGY ASSESSMENT: A REVIEW OF CHARACTERISTICS AND APPLIED METHODS

Miriam Luhn

Institute for Quality and Efficiency in Health Care (IQWiG)
miriam.luhn@iqwig.de

Barbara Prediger

Institute for Research in Operative Medicine

Edmund A.M. Neugebauer

Brandenburg Medical School Theodor Fontane; Interdisciplinary Centre for Health Services Research

Tim Mathes

Institute for Research in Operative Medicine

Objectives: When making decisions in health care, it is essential to consider economic evidence about an intervention. The objective of this study was to analyze the methods applied for systematic reviews of health economic evaluations (SR-HEs) in HTA and to identify common challenges.

Methods: We manually searched the Web pages of HTA organizations and included HTA-reports published since 2015. Prerequisites for inclusion were the conduct of an SR-HE in at least one electronic database and the use of the English, German, French, or Spanish language. Methodological features were extracted in standardized tables. We prepared descriptive statistical (e.g., median, range) measures to describe the applied methods. Data were synthesized in a structured narrative way.

Results: Eighty-three reports were included in the analysis. We identified inexplicable heterogeneity, particularly concerning literature search strategy, data extraction, assessment of quality, and applicability. Furthermore, process steps were often missing or reported in a nontransparent way. The use of a standardized data extraction form was indicated in one-third of reports (32 percent). Fifty-four percent of authors systematically appraised included studies. In 10 percent of reports, the applicability of included studies was assessed. Involvement of two reviewers was rarely reported for the study selection (43 percent), data extraction (28 percent), and quality assessment (39 percent).

Conclusions: The methods applied for SR-HEs in HTA and their reporting quality are very heterogeneous. Efforts toward a detailed, standardized guidance for the preparation of SR-HEs definitely seem necessary. A general harmonization and improvement of the applied methodology would increase the value of SR-HE for decision makers.

Keywords: Systematic review, Economic evaluation, Health technology assessment, Reimbursement

Continuing pressure on health budgets worldwide makes an efficient resource allocation increasingly crucial. In recent years, particularly the introduction of several high-cost interventions presents enormous challenges regarding accessibility and sustainability of healthcare systems (1;2). Economic considerations are increasingly important for health authorities and their pricing and reimbursement decision-making process regarding new technologies.

Health technology assessment (HTA) frequently uses systematic reviews of health economic evaluations (SR-HEs) (3). These could provide evidence about the cost-effectiveness of an intervention within a limited time period (approximately 3 to 12 months) (4). SR-HEs are valuable (i) to inform the development of an own economic model, (ii) to identify the most relevant study for a particular decision, and (iii) to identify the implicated economic trade-offs (3). Even though, the transferability and generalizability of

results stemming from different jurisdictions represents a major challenge (5).

Jefferson et al. (2002) found that SR-HEs show fundamental methodological flaws, especially regarding their search strategy and the application of an appropriate quality assessment tool (6). Moreover, it was discovered that search methods applied in SR-HEs are not extensive enough and inconsistent with published recommendations (7). Universally accepted methods for SR-HEs do not seem to exist so far: More recent studies focusing on the available methodological guidelines found that the recommendations still vary widely and are partly imprecise (8–11). As a result, the conduct of SR-HEs in HTA may vary widely and methodological shortcomings are apparent. The aim of this study was to (i) provide a detailed overview of the applied methods for SR-HEs in HTA, (ii) identify similarities and differences within the HTA-reports of different agencies, (iii) identify common challenges.

Its findings may support the needed generation of precise and universally accepted methods for SR-HEs in HTA,

We thank Consuela Jacobi-Yniguez for proofreading the article. No funding was received for the present research.

increasing their acceptance and usefulness for medical decision makers.

METHODS

The authors conducted a review of the methods applied in systematic reviews for economic evaluations conducted within HTA-reports.

Search Strategy and Study Selection

We used the publicly accessible member lists of the International Network of Agencies for Health Technology Assessment (INAHTA), of Health Technology Assessment International (HTAi), and the European Network for Health Technology Assessment (EUnetHTA) to identify HTA agencies (available on <http://www.inahta.org/>, <https://www.htai.org/>, and <http://www.eunetha.eu/>, request date June 22, 2016). Between July 2 and August 22, 2016, we searched the Web pages of all 115 member organizations for HTA-reports published since January 2015. We then manually screened the list of published HTA reports of each of the agencies. We used this timeframe to show a current overview of the methods applied for SR-HEs. We also considered reports of joint assessments by EUnetHTA published within that time span. Inclusion criteria were as follows (see Supplementary Table 1 for further explanation): (i) publication type: inclusion of reports that evaluated a health technology to inform medical decision making; (ii) assessment includes a SR-HE: HTA-reports that described economic aspects of the technology under assessment and had conducted a literature review of economic evaluations in at least one database; (iii) language: inclusion of reports written in English, German, French, or Spanish

One reviewer (M.L.) screened all identified full-texts for eligibility. A second reviewer (T.M.) rescreened a random sample of 10 percent excluded articles. In case of uncertainty regarding inclusion/exclusion, a second reviewer was involved (T.M., B.P.). Any disagreements were resolved through discussion or involvement of a third reviewer. In case of frequent and/or substantial disagreements, a verification of all excluded articles was planned. As there were no discrepancies in the calibration exercises, the verification process was not extended.

Of a higher number of reports issued by each HTA agency, we randomly selected a sample of ten reports. This was to provide a more balanced overview of the methods applied at the various agencies and to prevent single agencies from being overrepresented in the analysis.

Nonreporting of details of the search strategy for the economic domain resulted in contacting the authors for explanatory information.

As a further source of information and where applicable, we used the recommendations on SR-HEs of the respective agencies' manual. If a report did not present methods in detail but referred to the respective manual, information was

extracted there. We searched the Web pages of all agencies, of which at least one report could be included, for relevant methodological guidance.

Data Extraction and Summary

We extracted the following information in standardized, piloted data extraction forms (Supplementary Table 2): scope of the systematic review, statement of research question and formulated eligibility criteria, literature search strategy, study selection, data extraction, assessment of methodological study quality, assessment of generalizability/transferability/applicability, presentation of cost data, and method for data synthesis.

Using Microsoft Excel 2010, we developed an electronic extraction form. The approach for data abstraction and data presentation was inspired by the publication of Page et al. (12), which provides a similar overview of epidemiology and reporting characteristics of systematic reviews of biomedical research. To ensure the inclusion of all relevant aspects needed to answer our study objective, the extraction form was tested and discussed before use.

A single reviewer (M.L., T.M., or B.P.) extracted data. After extraction of the first reports, a 10 percent random sample ($n = 9$) was verified for accuracy and correctness of data entries by a second reviewer. If necessary, discrepancies were resolved through discussion or third party involvement. In case of frequent and/or substantial disagreements, a verification of 100 percent was intended for this step. After sample extractions and discussion this, however, was not necessary.

We considered assessment of methodological study quality and generalizability/transferability/applicability only as given if a publicly available tool or an internal checklist was used. Moreover, conduct of the assessment needed to be on a study level (as opposed to an evaluation on the synthesized results). We considered data items as extracted if they were: (i) presented in a standardized extraction form or (ii) in the running text consistently across all studies included in the SR-HE.

We also included "rapid reviews" or "rapid HTA-reports" (4). Rapid reviews provide quick information for decisions needed within a short time period (13;14). We identified rapid HTA-reports by their title and the description of the report's objective (4). As the comparability of rapid HTA-reports and full HTA-reports is limited, we decided to analyze them separately. We reported only the main differences between rapid HTA-reports and other included HTA-reports.

Where available, we also extracted recommendations on SR-HEs from the agencies' manuals in the data extraction forms (Supplementary Table 2).

All data were analyzed using Microsoft Excel 2010. For nominal data, we provided numbers and percentages. Additionally, we provided median and ranges for count data.

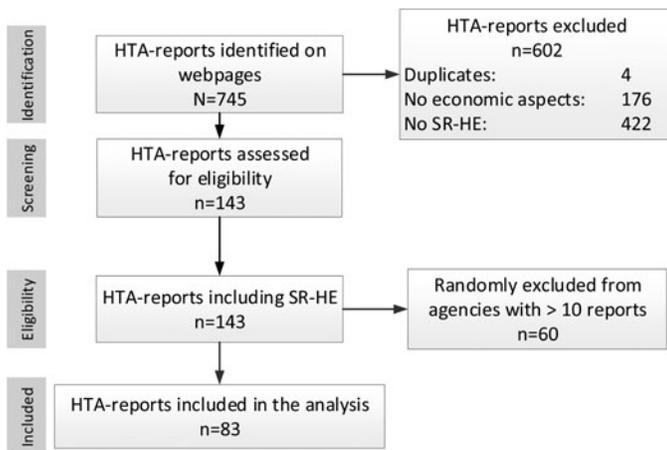


Figure 1. PRISMA flow chart of report selection.

RESULTS

Selection of Health Technology Assessment Reports

Figure 1 illustrates the selection process. Overall, 745 reports were found. After full-text screening, 143 reports were considered eligible.

More than ten reports were available ($n = 80$) from two agencies (HQO and NIHR), leading to the exclusion of sixty randomly selected reports. Finally, eighty-three HTA-reports were included in the analysis (Supplementary Table 3); of these, ten were rapid reports. A presentation of included reports per agency (Supplementary Table 4) and excluded reports, providing reasons for the exclusion (Supplementary Tables 5–7), is given in the supplement.

In nine cases, details of the search strategy for the economic domain were only available on request.

Manuals providing methodological guidance for the conduct of SR-HEs were available for five agencies (CADTH, HIQA, HIS, KCE, and LBI) (13;15–18). Furthermore, the HTA Core Model (19) developed as part of EUnetHTA Joint Action 2 (JA2), contains applicable recommendations and was, therefore, also considered. Overall, the identified six manuals relate to 25/83 (30 percent) of included reports.

Methods of Systematic Reviews of Economic Evaluations

The characteristics of the analyzed SR-HEs are presented in Tables 1 and 2. More detailed versions are available in the supplement (Supplementary Tables 8 and 9).

Scope of the Systematic Review. Sixty percent of the reports included a SR-HE, but no primary cost-effectiveness analysis (CEA). The rest included both. For all rapid reports, only a SR-HE was conducted (Table 1).

Research Question and Eligibility Criteria. The research question/objective of the SR-HE and at least one criterion for study selection

was reported in the majority of reports. The specification of all components of PICOS (population, intervention(s), comparator(s), outcomes, study design) was found in approximately one-third of cases, but only in one rapid report. In one-third of rapid reports, no eligibility criteria were mentioned at all.

Of the reports specifying the economic study types included, one-third explicitly restricted the inclusion to full economic evaluations (20). In most cases, other economic study types (e.g., cost-consequences analyses, budget impact analyses) were also included.

Literature Search. Review authors searched a median of four electronic databases (range, 1–14). Most reviewers searched general medical databases (including MEDLINE and EMBASE). Economic databases and HTA/systematic review databases were each considered in approximately 60 percent of reports. Of those, the National Health System Economic Evaluation Database (NHS EED) and the HTA Database were consulted most often. In rapid reports, both categories of databases were searched less often than in full HTA-reports. A full list of databases that were searched in the analyzed SR-HEs and the frequency of use is presented in the supplement (Supplementary Table 8).

Economic terms were used in more than half of the literature searches. The majority of authors designed own Boolean search strings with economic terms instead of applying published filters developed for the detection of economic studies. The Scottish Intercollegiate Guidelines Network (SIGN) filter for MEDLINE (21) was used most often.

In most reports, at least one limit was applied to the literature search. The median number of search limits was 3 (range, 0–6). More than half of the authors restricted the inclusion of publications to certain languages, with mostly including only English publications.

Additional sources (e.g., online search, reference lists) were searched in the majority of reports.

Study Selection. In 43 percent of reports, at least two reviewers selected eligible studies, mostly by screening all retrieved articles independently. A duplicate study selection was more prevalent in full HTA-reports than in rapid reports.

In seventy-two percent of reports, the study selection was described and/or illustrated in a flow chart. This was rarely the case in rapid reports.

In twelve reports (including one rapid report), no eligible economic studies could be identified in the literature search. Therefore, only seventy-one reports with applicable study results were considered in the following sections.

Data Extraction. One-third of HTA-reports (including two rapid reports) indicated the use of a standardized data extraction

Table 1. Methods of Systematic Reviews of Economic Evaluations in HTA: Scope and Literature Search

Category	Characteristics	Full HTA (<i>n</i> = 73)	Rapid HTA (<i>n</i> = 10)
Scope of systematic review	Exclusively systematic review	40 (55%)	10 (100%)
	SR accompanied by primary CEA	25 (34%)	0
	SR to inform primary CEA	8 (11%)	0
Eligibility criteria	Not specified	3 (4%)	3 (30%)
	Population	58 (79%)	3 (30%)
	Intervention(s)	65 (89%)	5 (50%)
	Comparator(s)	50 (68%)	2 (20%)
	Outcomes	49 (67%)	2 (20%)
	Study design	51 (70%)	6 (60%)
	<i>Specification of economic study types considered</i>	52 (71%)	5 (50%)
	PICO or PICOS	38 (52%)	1 (10%)
	Country/setting	14 (19%)	0
	Publication type	23 (32%)	1 (10%)
	Publication date	10 (14%)	1 (10%)
	Language	4 (5%)	2 (20%)
	Specified requirements for cost data	10 (14%)	0
No. of databases searched		5 (1–14)	4 (1–9)
Information sources	Databases		
	General (medical) databases	69 (95%)	10 (100%)
	Cochrane Library	9 (12%)	2 (20%)
	Economic databases	48 (66%)	3 (30%)
	HTA/SR databases	51 (70%)	2 (20%)
	Specific databases	21 (29%)	6 (60%)
	Citation index (SSCI/SCI)	8 (11%)	0
	<i>Other sources searched</i>	62 (85%)	6 (60%)
	Online search	36 (49%)	5 (50%)
	Hand search in selected journals	2 (3%)	0
	Reference lists	46 (63%)	1 (10%)
	Contact experts or corresponding authors	9 (12%)	0
	Contact manufacturers	4 (5%)	0
Economic search terms/filters	Economic search terms/filters applied	40 (55%)	5 (50%)
	Published economic filter	5 (7%)	3 (30%)
	Own Boolean search string with economic terms	35 (48%)	2 (20%)
Search limits	<i>No. of limitations applied</i>	3 (0–6)	3 (0–6)
	At least one limitation applied	59 (81%)	9 (90%)
	Publication date	37 (51%)	6 (60%)
	Update of previews SR	8 (11%)	0
	Language	39 (53%)	7 (70%)
	English only	26 (36%)	7 (70%)
	English + selected other	13 (18%)	0
	Human studies	32 (44%)	5 (50%)
	Publication type	21 (29%)	4 (40%)
	Study type	7 (10%)	4 (40%)

CEA, cost-effectiveness analysis; HTA, health technology assessment; PICO, population, intervention(s), comparator(s), outcomes; PICOS, population, intervention(s), comparator(s), outcomes, study design; SCI, Sciences Citation Index; SR, systematic review; SSCI, Social Sciences Citation Index.

Table 2. Methods of Systematic Reviews of Economic Evaluations in HTA: Data Extraction, Critical Appraisal, and Synthesis

Category	Characteristics	HTA-reports (<i>n</i> = 71) ^a	
Data extraction			
Data extraction method	Use of standardized extraction form	23 (32%)	
Reviewers involved in data extraction	Duplicate data extraction	20 (28%)	
	All data extracted by one reviewer	9 (13%)	
	Not reported	42 (59%)	
	Assessment of methodological study quality		
Assessment tool	Assessment of methodological study quality on study level	38 (54%)	
	Drummond checklist (22)	13 (34%)	
	ICER Integrated Evidence Rating (46)	9 (24%)	
	CHEERS statement (Husereau et al. 2013 [31])	3 (8%)	
	Philips et al. 2004 (24)	3 (8%)	
	CHEC-list (Evers et al. 2005 [47])	2 (5%)	
	Critical Appraisal Skills Programme Checklist (48)	2 (5%)	
	Questionnaire to Assess Relevance and Credibility of Modeling Studies for Informing Health Care Decision Making (Jaime Caro et al. 2014 [25])	1 (3%)	
	Guide to the Methods of Technology Appraisal - reference checklist for economic evaluations (NICE 2013 [23])	1 (3%)	
	Cochrane 2008 (Handbook for Systematic Reviews of Interventions - Incorporating Economics Evidence [28])	1 (3%)	
	Internal checklist based on published tools	8 (21%)	
	Reviewers involved in quality assessment	Duplicate quality assessment	15 (39%)
		Included studies critically appraised by one reviewer	4 (11%)
		Not reported	22 (58%)
Assessment of generalizability/transferability/applicability			
Assessment tool	Assessment of generalizability/transferability/applicability	7 (10%)	
	Questionnaire to Assess Relevance and Credibility of Modeling Studies for Informing Health Care Decision Making (Jaime Caro et al. 2014 [25])	2 (29%)	
	EUnetHTA Adaptation Toolkit (26)	1 (14%)	
	Internal checklist based on published tools or own criteria	4 (57%)	
	Reviewers involved in generalizability/transferability/applicability assessment	Duplicate assessment of generalizability/transferability/applicability	4 (57%)
Not reported		3 (43%)	
Presentation of cost data			
Presentation of cost data	As reported	56 (79%)	
	Inflated	11 (15%)	
	Currency converted	23 (32%)	
	Unclear	6 (8%)	
Data synthesis			
Data synthesis	Narrative synthesis	53 (75%)	
	Permutation matrix	1 (1%)	
	n.a. – only one study included	13 (18%)	
	No synthesis	4 (6%)	

EUnetHTA, European Network for Health Technology Assessment; HTA, health technology assessment; ICER, Institute for Clinical and Economic Review; n.a., not applicable.

^a For twelve of the included eighty-three HTA-reports, we could not analyze the extracted information as the authors could not identify eligible economic studies in the literature search.

form. The remainder did not contain any information regarding the method of data extraction at all.

In twenty reports, at least two reviewers were involved in the data extraction. In most cases, one reviewer extracted all data, while a second reviewer verified data entries. In the two rapid reports applying a data extraction form, two reviewers performed the data extraction independently.

Extracted items are presented in the supplement (Supplementary Table 10). In four reports, we could not identify the extracted information, as it was neither presented in a results table nor in the running text. Therefore, only sixty-seven reports were available for the analysis of extracted data items. All components of PICO were extracted in 39 percent of reports. Information regarding data sources for costs, clinical data, utility data, discounting, or analyses of uncertainty conducted was available in less than half of the reports. One-third of reports was missing the study design or incremental results of the included studies.

Assessment of Methodological Study Quality. Authors of approximately half of the HTA-reports and of one-third of rapid reports explicitly stated an assessment of quality of included economic studies with a tool or an internal checklist based on published tools. In every other report, either the Drummond checklist (22) or an adaptation was used. Table 2 presents further tools applied.

In 43 percent of full HTA-reports, at least two reviewers were involved in the critical appraisal. In most cases, two reviewers assessed all included studies independently. None of the rapid reports indicated duplicate quality assessment.

Assessment of Generalizability/Transferability/Applicability. In seven reports (including one rapid report), the generalizability/transferability/applicability of included studies was systematically assessed with a published tool or an internal checklist. Two independent assessors were involved in four reports. In all cases, aspects evaluated included (i) the reliability of results based on input data and methods applied, (ii) relevant differences of conditions in the study and the national context regarding health care provision and sociodemographic/epidemiologic factors, and (iii) differences between the methods applied in the study and the national methodological requirements. Most frequently, an internal checklist based on published tools (22–24) or own criteria was applied. Further checklists used for this purpose comprise the Questionnaire developed by Jaime Caro et al. (25) or the EUnetHTA Adaptation Toolkit (26).

Presentation of Cost Data. Approximately every other report presented cost data as originally reported in the included studies. In approximately one-quarter, cost data were presented both as reported and converted to local currency. In six reports, costs were additionally inflated by means of consumer price indices.

Data Synthesis. Two-thirds of the reports contain a narrative synthesis of data. In one exceptional case, incremental effectiveness and incremental cost of included studies were synthesized in a permutation matrix. In every fifth report, only one study was included in the SR-HE, and consequently a synthesis of results was not applicable.

Tables providing full extractions of the included reports are available on reasonable request.

DISCUSSION

The methodologies applied for SR-HEs in HTA vary widely in all process steps. Furthermore, process steps like data extraction, critical appraisal, or assessment of applicability were frequently not performed at all, not reported, or reported nontransparently. It can be said that methodological improvements in the methods applied for SR-HE in HTA and their reporting quality are needed and possible. These are clear prerequisite to increase the usefulness of SR-HEs for policy decision making.

Tight timeframes for the production of HTA-reports and an expected limited value of information regarding the application of more sophisticated methods could explain our findings. Heterogeneity in methodological approaches may partly be justified by type of review product (e.g., rapid reports, full HTA-reports), different scopes (e.g., determine cost-effectiveness, identify drivers of cost-effectiveness, inform own CEA), or different contexts (e.g., organizational structure for decision making). Moreover, the relative importance of economics for decision making varies by country (27). As a result, the economic aspects in HTA-reports might have different objectives (e.g., economics considered as main information versus only as ancillary information) and consequently different information requirements (e.g., detailed economic information versus rough overview). The differences in relative importance and objectives are also a barrier to complete standardization.

Last but not least, also organizational aspects like unequal availability of personnel and time resources or access to databases might prevent a complete standardization of all process steps. However, we also detected methodological differences in comparable review products, jurisdictions, healthcare systems (e.g., publicly funded) and decision-making contexts. Unjustified heterogeneity and deviations from general established standards for preparing and reporting of systematic reviews (28;29) particularly concerned literature search strategy, methods for data extraction, methods for methodological quality assessment, and assessment of applicability of findings. It seems that these process steps as well as the quality of reporting could especially benefit from standardized and detailed guidance. The relevance of SR-HEs in HTA should not be underestimated, taking into consideration that sixty percent of the analyzed reports included only a SR-HE but no primary CEA. This further emphasizes the need to conduct SR-HE uniformly.

The methods do not seem to adhere to the recommendations of publicly available method manuals (13;15–19). Examples for this deviation include the recommendation to specify at least the population, intervention, and comparator as inclusion criteria for the study selection (13;15–18) or the use of a standardized data extraction form (15–19). Another example is the recommendation to systematically assess the quality of included economic evaluations, which is stated in all six available manuals (13;15–19). However, we found that several reports did not follow these requirements.

Research Question and Eligibility Criteria

We found that authors specified different inclusion criteria. In many cases, a broader inclusion of studies in the economic domain might be justified. It might be reasonable, for example, not to specify the control too strictly, because different interventions are standard of care. However, the definition of all PICOS components is important to ensure homogenous and transparent study selection. If review authors did not restrict selection with regard to one or more components, this should explicitly be stated.

Literature Search

While general medical databases were considered in almost all reports, a search in at least one economic or HTA/systematic review database was frequently missing. Our findings are in line with a recently published study by Wood et al. (7). A sole search in general databases, however, bears the risk that not all relevant economic evaluations are identified (30). A possible explanation for our findings, aside from time constraints, might be that the most frequently recommended database NHS EED has not been updated since the end of 2014 and is, therefore, not useable for more recent interventions. In addition, HEED is no longer available.

In half of the analyzed reports, no economic filters were applied. In most of these cases, a joint literature search was conducted for the economic domain and other domains. When the bibliographic search is not restricted to study types, however, a large amount of retrieved articles means that more time for screening can be expected. Moreover, a joint search for the economic and further domains requires that reviewers are familiar with the domains' specific vocabularies and are capable, for example, to identify the relevant economic study types. A more efficient way of literature retrieval might be to conduct separate literature searches and study selections for single domains, combining search terms related to population/indication, intervention, and domain-specific filters.

Data Extraction

Only every third report indicated the use of a standardized data extraction form. This, along with the absence of a second independent reviewer, might lead to inconsistent or incorrect data

extraction (28). In most cases, the reporting of included studies is scarce. Particularly the lack of information on key methodological features (e.g., perspective) and insufficient information on results is critical, as their reporting is required for proper interpretation of results and their applicability. We suggest using standardized templates for data extraction and reporting based on the items included in available reporting guidelines for health economic evaluations (e.g., Husereau et al. [31]; Drummond, and Jefferson, 1996 [22]). The templates could then be adapted to the respective purpose because the relevance of extracted information depends on the review question, defined eligibility criteria, and the scope of the SR-HE.

Assessment of Methodological Study Quality

Authors of every second report evaluated included economic evaluations with a formalized quality assessment tool. Assessing the methodological quality, however, is crucial to appraise the internal validity and consequently the value of results for decision makers (28;29;31;32). The results of a critical appraisal of reviewed economic evaluations could either be used to identify the most credible evaluation available (3) or to decide whether it is necessary to conduct an own CEA. It should be noted that most of the available assessment tools were developed to support and assess quality of reporting of economic evaluations (e.g., references 22;24;31). An in-depth analysis of the methodological quality of modeling studies also involves a critical appraisal of primary data sources (for costs, clinical data, and utility data) included in the analysis. It is possible that systematic quality assessments were often not performed because they require extensive time and experience.

Assessment of Generalizability/Transferability/Applicability

In the analyzed reports, generalizability/transferability/applicability often could not be distinguished. It was, therefore, not possible to consider them separately (33). Although SR-HEs can provide a range of useful information, there are some local factors, especially prices, the disease-specific incidence/prevalence, and differences in the health care provision. These factors need to be considered as they could heavily influence the results of the evaluations (5;34). The systematic assessment of generalizability/transferability/applicability of results was reported even less often than the assessment of study quality. In these cases, the relevance of results obtained from differing jurisdictions within the national context and setting remains unclear, thereby reducing the value of the SR-HE as a decision base (28;33–35).

Another way of addressing problems of transferability is to restrict the selection of studies to those from one's own country or within the relevant setting. This is in line with our findings, that none of the reports defining country/setting as inclusion criterion assessed generalizability/transferability/applicability

of included studies. This approach is particularly appropriate, if the existence of a sufficient number of applicable economic evaluations can be assumed.

Duplication of Study Selection, Data Extraction, and Assessment of Included Studies

A duplication of study selection, data extraction, and assessment of quality and applicability was often not performed. These process steps are susceptible to errors or misjudgments and often require subjective decisions (36). Even though the involvement of two or more reviewers requires greater effort, it increases objectivity and minimizes errors or misjudgments (34;37;38). If, due to personnel or time constraints, a complete duplication of all steps is not possible, abbreviated methods can be applied that involve less expenditure but also reduce the error rate. An efficient way to duplicate study selection is the method of liberal acceleration, in which one reviewer screens all identified studies with a second reviewer rescreening only excluded studies (37). An alternative to an independent data extraction and critical appraisal by two reviewers might be one reviewer extracting and assessing all data with a second reviewer verifying critical data entries only.

Data Synthesis

Generally, results were synthesized narratively. These syntheses were often unstructured and do not allow an interpretation of results of studies included in the SR-HE, or conclusions on the cost-effectiveness of the intervention(s) under assessment. Methodological guidance is available for the structured narrative synthesis in systematic reviews of clinical studies (39). Similar approaches seem applicable for SR-HEs. Even if a meta-analytical approach for data synthesis might not be feasible in most economic evaluations (3), a graphical presentation such as a permutation matrix, hierarchical decision matrix, or cost-effectiveness plane would create a common framework for included studies, thereby enhancing interpretability and reflection (40;41).

The graphical tools provide a convenient overview of the cost-effectiveness of the whole body of evidence. Another advantage of graphical displays is that they are easier to understand than a narrative synthesis, which is helpful for decision makers not familiar with economic evaluations. In the case of detailed reporting of methods and results of the economic evaluations, the available evidence can also be pooled. This can be done by means of meta-analyses of resource use and outcomes separately, or the economic evaluation can be adapted to the own context (42). Such approaches, however, seem only useful if a high quality economic evaluation is identified and most of the results are applicable to the context under consideration.

Methods Applied in Rapid Reports

We found that rapid reports skipped quality relevant aspects even more frequently. These aspects concern the consideration

of economic and HTA/systematic review databases, duplication and presentation of the study selection process, and assessment of study quality. We detected no relevant differences between rapid and full HTA-reports concerning the use of economic filters and other search limits, presentation of extracted data items, and assessment of applicability. The significance of these findings is limited though, because we included only ten rapid reports in our analysis.

Limitations

Several definitions exist for HTA (e.g., references 43–45). The decision of whether a publication can be considered a HTA-report or not was, therefore, challenging in some cases. As a pragmatic approach, we decided to include all reports that fulfilled the criteria of evaluating a health technology to inform medical decision making, as these criteria were included in all definitions.

As only a small number of systematic reviews are conducted within the economic domain in HTA and we aimed to provide a wide overview of the methods, we applied a rather broad definition for SR-HE as stated in the methods section. This could possibly lead to the inclusion of reports for which the authors did not explicitly intend to conduct a systematic review for the economic domain.

We did not extract all data in duplicate, which might have resulted in some errors. Likewise, the study selection process was not completely duplicated. Furthermore, by restricting inclusion to more recent reports, we analyzed only a sample of all HTA-reports available.

In conclusion, the methods applied for SR-HEs in HTA and their reporting quality are very heterogeneous. This is especially true for the data extraction as well as assessment of methodological quality and of applicability of the results to the own setting and the context of the HTA-report. Particularly, the applicability to the own setting and context, and the consideration of limited applicability in the analysis seems a challenge for the preparation of SR-HEs. Efforts toward a detailed, standardized guidance for the preparation process of SR-HEs (28) definitely seem necessary. These guidelines could contribute to the harmonization and improvement of the applied methodology and would thus increase the value of SR-HE for decision makers. The main challenge is to find a balance between standardization, level of detail, and the special requirements of different jurisdictions.

SUPPLEMENTARY MATERIAL

Supplementary Table 1:

<https://doi.org/10.1017/S0266462318000624>

Supplementary Table 2:

<https://doi.org/10.1017/S0266462318000624>

Supplementary Table 3:

<https://doi.org/10.1017/S0266462318000624>

Supplementary Table 4:
<https://doi.org/10.1017/S0266462318000624>
 Supplementary Table 5:
<https://doi.org/10.1017/S0266462318000624>
 Supplementary Table 6:
<https://doi.org/10.1017/S0266462318000624>
 Supplementary Table 7:
<https://doi.org/10.1017/S0266462318000624>
 Supplementary Table 8:
<https://doi.org/10.1017/S0266462318000624>
 Supplementary Table 9:
<https://doi.org/10.1017/S0266462318000624>
 Supplementary Table 10:
<https://doi.org/10.1017/S0266462318000624>

CONFLICTS OF INTEREST

The authors have nothing to disclose.

REFERENCES

- European Commission. Inception impact assessment - Strengthening of the EU cooperation on health technology assessment (HTA). Published September 14, 2016. http://ec.europa.eu/smart-regulation/roadmaps/docs/2016_sante_144_health_technology_assessments_en.pdf (accessed November 28, 2016).
- OECD. *Fiscal sustainability of health systems: Bridging health and finance perspectives*. Paris: OECD Publishing; 2015.
- Anderson R. Systematic reviews of economic evaluations: Utility or futility? *Health Econ*. 2010;19:350-364.
- Merlin T, Tamblyn D, Ellery B. What's in a name? Developing definitions for common health technology assessment product types of the International Network of Agencies for Health Technology Assessment (inahta). *Int J Technol Assess Health Care*. 2014;30:430-437.
- Drummond M, Barbieri M, Cook J, et al. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. *Value Health*. 2009;12:409-418.
- Jefferson T, Demicheli V, Vale L. Quality of systematic reviews of economic evaluations in health care. *JAMA*. 2002;287:2809-2812.
- Wood H, Arber M, Glanville JM. Systematic reviews of economic evaluations: How extensive are their searches? *Int J Technol Assess Health Care*. 2017;33:25-31.
- Mathes T, Walgenbach M, Antoine SL, et al. Methods for systematic reviews of health economic evaluations: A systematic review, comparison, and synthesis of method literature. *Med Decis Making*. 2014;34:826-840.
- Thielen FW, Van Mastrigt G, Burgers LT, et al. How to prepare a systematic review of economic evaluations for clinical practice guidelines: Database selection and search strategy development (part 2/3). *Expert Rev Pharmacoecon Outcomes Res*. 2016;16:705-721.
- van Mastrigt GA, Hiligsmann M, Arts JJ, et al. How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: A five-step approach (part 1/3). *Expert Rev Pharmacoecon Outcomes Res*. 2016;16:689-704.
- Wijnen B, Van Mastrigt G, Redekop WK, et al. How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: Data extraction, risk of bias, and transferability (part 3/3). *Expert Rev Pharmacoecon Outcomes Res*. 2016;16:723-732.
- Page MJ, Shamseer L, Altman DG, et al. Epidemiology and reporting characteristics of systematic reviews of biomedical research: A cross-sectional study. *PLoS Med*. 2016;13:e1002028.
- Healthcare Improvement Scotland. Standard operating procedure for production of evidence notes. Published September, 2012. <http://www.healthcareimprovementscotland.org/his/idoc.aspx?docid=a3ecc845-8fcd-44bf-b274-f8c14253c4be&version=-1> (accessed June 23, 2016).
- Khangura S, Polisen J, Clifford TJ, et al. Rapid review: An emerging approach to evidence synthesis in health technology assessment. *Int J Technol Assess Health Care*. 2014;30:20-27.
- Belgian Health Care Knowledge Centre. Guidelines for pharmacoeconomic evaluations in Belgium. Published April 22, 2008. https://kce.fgov.be/sites/default/files/page_documents/d20081027327.pdf (accessed June 21, 2016).
- Canadian Agency for Drugs and Technologies in Health. Guidelines for authors of CADTH health technology assessment reports. Published May 16, 2003. https://www.cadth.ca/sites/default/files/pdf/cadth_Authors_Guidelines.doc (accessed June 23, 2016).
- Health Information and Quality Authority. Guidelines for the retrieval and interpretation of economic evaluations of health technologies in Ireland. Published September 30, 2014. <https://www.hiqa.ie/publications/guidelines-retrieval-and-interpretation-economic-evaluations-health-technologies-ireland> (accessed August 31, 2016).
- Ludwig Botzmann Institut. (Internes) Manual Abläufe und Methoden. Teil 2. Published March, 2007. http://hta.lbg.ac.at/uploads/tableTool/UIICmsPage/gallery/InternesManual_2.Aufl..pdf (accessed September 16, 2016).
- EUNETHTA Joint Action 2. Work Package 8. H TA Core Model version 3.0. Published January 25, 2016. www.htacoremodel.info/BrowseModel.aspx (accessed September 15, 2016).
- Drummond MF, Sculpher MJ, Claxton K, et al. *Methods for the economic evaluation of health care programmes Vol. 4*. Oxford: Oxford University Press; 2015.
- Scottish Intercollegiate Guidelines Network. Search filters. <http://www.sign.ac.uk/search-filters.html> (accessed January 17, 2017).
- Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. *BMJ*. 1996;313:275-283.
- National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. Published April 4, 2013. <https://www.nice.org.uk/process/pmg9/chapter/foreword> (accessed January 17, 2017).
- Philips Z, Ginnelly L, Sculpher M, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess*. 2004;8:iii-iv,ix-xi,1-158.
- Jaime Caro J, Eddy DM, Kan H, et al. Questionnaire to assess relevance and credibility of modeling studies for informing health care decision making: An ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health*. 2014;17:174-182.
- EUNETHTA. HTA adaptation toolkit. Work package 5 (Version 5). Published October, 2011. https://www.eunetha.eu/wp-content/uploads/2011/01/EUNETHTA_adptation_toolkit_2011_version_5.pdf (accessed November 28, 2016).
- Franken M, Heintz E, Gerber-Grote A, Raftery J. Health economics as rhetoric: The limited impact of health economics on funding decisions in four European countries. *Value Health*. 2016;19:951-956.
- Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions - Version 5.1.0. Published March, 2011. <http://handbook.cochrane.org/> (accessed June 10, 2016).
- Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Ann Intern Med*. 2009;151:264-269,w64.
- Alton V, Eckerlund I, Norlund A. Health economic evaluations: How to find them. *Int J Technol Assess Health Care*. 2006;22:512-517.

31. Husereau D, Drummond M, Petrou S, et al. Consolidated health economic evaluation reporting standards (CHEERS) statement. *BMJ*. 2013;346:f1049.
32. Sackett DL, Rosenberg WM, Gray JA, et al. Evidence based medicine: What it is and what it isn't. *BMJ*. 1996;312:71-72.
33. Burchett H, Umoquit M, Dobrow M. How do we know when research from one setting can be useful in another? A review of external validity, applicability and transferability frameworks. *J Health Serv Res Policy*. 2011;16:238-244.
34. Barbieri M, Drummond M, Willke R, et al. Variability of cost-effectiveness estimates for pharmaceuticals in Western Europe: Lessons for inferring generalizability. *Value Health*. 2005;8:10-23.
35. Burford B, Lewin S, Welch V, et al. Assessing the applicability of findings in systematic reviews of complex interventions can enhance the utility of reviews for decision making. *J Clin Epidemiol*. 2013;66:1251-1261.
36. Centre for Reviews and Dissemination. CRD's guidance for undertaking reviews in health care. published January, 2009. http://www.york.ac.uk/media/crd/Systematic_Reviews.pdf (accessed June 23, 2016).
37. Shemilt I, Khan N, Park S, Thomas J. Use of cost-effectiveness analysis to compare the efficiency of study identification methods in systematic reviews. *Syst Rev*. 2016;5:140.
38. Buscemi N, Hartling L, Vandermeer B, et al. Single data extraction generated more errors than double data extraction in systematic reviews. *J Clin Epidemiol*. 2006;59:697-703.
39. Rodgers M, Sowden A, Petticrew M, et al. Testing methodological guidance on the conduct of narrative synthesis in systematic reviews. *Evaluation*. 2009;15:49-73.
40. Briggs A, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. *Health Econ*. 1998;7:723-740.
41. Nixon J, Khan KS, Kleijnen J. Summarising economic evaluations in systematic reviews: A new approach. *BMJ*. 2001;322:1596-1598.
42. Shemilt I, Mugford M, Byford S, et al. Meta-analysis of resource use and cost data. In: Higgins JP, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions - Version 5.1.0*. <http://handbook-5-1.cochrane.org/> <http://handbook-5-1.cochrane.org/> (accessed September 28, 2018).
43. EUnetHTA. Common questions - Health technology assessment (HTA). <http://5026.makemeweb.net/faq/Category%20-%20#t287n73> (accessed November 28, 2016).
44. HTA glossary. Health technology assessment (HTA). <http://htaglossary.net/health+technology+assessment+%28HTA%29> (accessed November 28, 2016).
45. World Health Organization. Medical devices - Health technology assessment. http://www.who.int/medical_devices/assessment/en/ (accessed November 28, 2016).
46. Institute for Clinical and Economic Review. Methodology: ICER integrated evidence rating. <http://icer-review.org/wp-content/uploads/2016/01/ICER-Rating-System-Apr-2013-Update-FINAL.pdf>. published 2013 (accessed January 19, 2017).
47. Evers S, Goossens M, de Vet H, et al. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *Int J Technol Assess Health Care*. 2005;21:240-245.
48. Critical Appraisal Skills Programme. Economic Evaluations Checklist. http://media.wix.com/ugd/dded87_3b2bd5743feb4b1aaac6ebdd68771d3f.pdf. published 2013 (accessed January 19, 2017).