PRACTICAL APPLICATION

Challenges in Modelling the Cost Effectiveness of Various Interventions for Cardiovascular Disease

Laura T. Burgers · William K. Redekop ·
Johan L. Severens

Abstract

Objectives Decision analytic modelling is essential in performing cost-effectiveness analyses (CEAs) of interventions in cardiovascular disease (CVD). However, modelling inherently poses challenges that need to be dealt with since models always represent a simplification of reality. The aim of this study was to identify and explore the challenges in modelling CVD interventions.

Methods A document analysis was performed of 40 model-based CEAs of CVD interventions published in high-impact journals. We analysed the systematically selected papers to identify challenges per type of intervention (test, non-drug, drug, disease management programme, and public health intervention), and a questionnaire was sent to the corresponding authors to obtain a more thorough overview. Ideas for possible solutions for the challenges were based on the papers, responses, modelling guidelines, and other sources.

Results The systematic literature search identified 1,720 potentially relevant articles. Forty authors were identified after screening the most recent 294 papers. Besides the

challenge of lack of data, the challenges encountered in the review suggest that it was difficult to obtain a sufficiently valid and accurate cost-effectiveness estimate, mainly due to lack of data or extrapolating from intermediate outcomes. Despite the low response rate of the questionnaire, it confirmed our results.

Conclusions This combination of a review and a survey showed examples of CVD modelling challenges found in studies published in high-impact journals. Modelling guidelines do not provide sufficient guidance in resolving all challenges. Some of the reported challenges are specific to the type of intervention and disease, while some are independent of intervention and disease.

L. T. Burgers · W. K. Redekop · J. L. Severens
Institute of Health Policy and Management, Erasmus
University Rotterdam, Rotterdam, The Netherlands
e-mail: burgers@bmg.eur.nl

L. T. Burgers · W. K. Redekop · J. L. Severens (✉)
Institute for Medical Technology Assessment,
Erasmus University Rotterdam, Burgemeester Oudlaan 50,
PO Box 1738, 3000 DR Rotterdam, The Netherlands
e-mail: severens@bmg.eur.nl

Key Points for Decision Makers
In the field of cardiovascular disease (CVD), modelling methods in economic evaluations come with challenges. Some of the reported challenges are specific to the type of intervention, but most challenges are present in all types of CVD interventions.
Modelling guidelines do not provide sufficient assistance in resolving all challenges in economic evaluations in CVD. In addition, many challenges require unique solutions. Other sources are proposed but further research is still needed since some challenges are still unsolved.
The main challenges that are present in modelling the cost effectiveness of cardiovascular interventions are lack of data and extrapolation of intermediate or surrogate outcomes.

1 Introduction

Decision analytic modelling when performing economic evaluations of interventions in cardiovascular disease (CVD) is challenging. For example, modelling is necessary if extrapolation of short- or intermediate-term results to long-term outcomes is required and numerous strategies need to be evaluated without direct evidence. Thus, modelling inherently poses challenges that need to be dealt with since models always represent a simplification of reality. The presence of challenges could be dependent on the type of intervention or the phase of disease in which the intervention would be used.

There are several ways to deal with the challenges in obtaining an accurate, precise and valid estimate of the cost effectiveness. The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and the Society for Medical Decision Making (SMDM) have recently published a series of recommendations for best practice in performing cost-effectiveness analyses (CEAs) based on a model [1–7]. These recommendations suggest some practical solutions to present challenges in modelling. However, they are not specific for any type of disease or intervention and therefore this review aims to identify and analyse challenges (e.g. multiple indications) in modelling the cost effectiveness of CVD interventions that currently exist in the field. Furthermore, we present ways to address the challenges based on current economic modelling guidelines and the opinions of experts from the field.

2 Methods

In order to identify current challenges in the field of CVD, a document analysis was performed of model-based CEAs of CVD interventions that were recently published (since January 2009) in disease-specific, health economical and general medical journals. In addition, a questionnaire was sent to the corresponding authors of the selected papers to obtain a more thorough overview of current CVD modelling challenges.

2.1 Selection

To select systematically relevant papers, we used a search string that contained both cost-effectiveness terms, based on the validated National Health Service Economic Evaluation Database (NHSEED) cost-effectiveness filter [8], and a disease-specific Medical Subject Heading (MeSH) term ('cardiovascular disease'). The search was performed on 8 May 2013. We assumed that the papers published in journals with a relatively high impact factor are more

susceptible to complicated challenges. Therefore, we limited the search results to 12 relatively high-impact factor journals in three categories: cardiovascular medicine, general medicine and health economics/health technology assessment (HTA). To select these 12 journals we sorted all possible journals per category on impact factor, based on Journal Citation Reports® [9], and included the four highest ranked journals that also published sufficient cardiovascular CEAs. The following journals were selected in the cardiovascular medicine category: *Circulation*, *European Heart Journal*, *Journal of the American College of Cardiology*, and *International Journal of Cardiology*; in the general medicine category: *The Lancet*, *New England Journal of Medicine*, *The Journal of the American Medical Association*, and *Annals of Internal Medicine*; and in the health economical category: *Value in Health*, *PharmacoEconomics*, *Health Technology Assessment*, and *Medical Decision Making*.

We sorted the search results (via Ovid MEDLINE®) of the 12 journals on entry date, and selected the most recent CEAs or methodological papers presenting results, both based on modelling methods and evaluating a CVD-related intervention, until we reached a convenience sample of 40 unique corresponding authors. Authors were selected if they met the following inclusion criteria: they should have evaluated a CVD intervention using modelling, and could only be included once as a corresponding author. The most recent publication of authors who met the inclusion criteria was included.

2.2 Document Analysis

Using the 40 publications, we extracted CVD modelling challenges explicitly mentioned in the Methods and Discussion sections of the papers, and determined the frequency of these challenges over all papers. Before data extraction, a list with challenges was created to identify possible challenges that were present in the studies. These challenges were based on our own experience and that of five other researchers with sufficient experience (3–15 years) in performing model-based CEAs. A challenge was added to the list when it was described in the paper but not included in the existing list. For each paper, two reviewers (LB and WR or JS) identified the challenges described in the Methods and Discussion sections. To be complete, we included the identified challenges of both reviewers after we carefully considered both sets of results. Challenges were initially analysed by type of intervention, although it is likely that challenges are not specific for one type of intervention and therefore present in several interventions. Interventions were categorized into tests (e.g. screening, diagnostic), non-drug interventions (e.g. surgically or non-diagnostic devices), drug interventions,

disease management programmes (DMPs) and public health interventions.

2.3 Questionnaire

To supplement our literature review, we sent a questionnaire (see electronic supplementary material) by e-mail to the corresponding authors of the 40 papers to identify challenges that were not described in the papers and to estimate the importance of the challenges. The same five modelling experts tested a pilot version of the questionnaire and indicated new challenges that were not previously identified. The first part of the questionnaire focussed on the solved and unsolved challenges that authors faced while creating and using the model used to perform their analyses. Authors were asked to provide data requirement and modelling challenges, and were also asked to provide their source of inspiration if they managed to solve the specific challenge. The second part of the questionnaire focussed on the challenges that the respondent may have experienced during any CEA modelling of a CVD intervention. The list used for the document analysis was also the basis for the questionnaire. Respondents were asked to indicate how often the challenge had occurred in model-based CVD intervention CEAs conducted by the respondent and how much impact it could have had on the incremental cost-effectiveness ratio (ICER).

2.4 Analysis

The challenges brought forward by the corresponding authors and identified in the literature review were analysed. We then examined ways to address them based on the current modelling guidelines [1–7], other literature, and the solutions provided by the authors.

3 Results

3.1 Selection

The systematic literature search identified 1,720 potentially relevant publications. In order to reach the target of 40
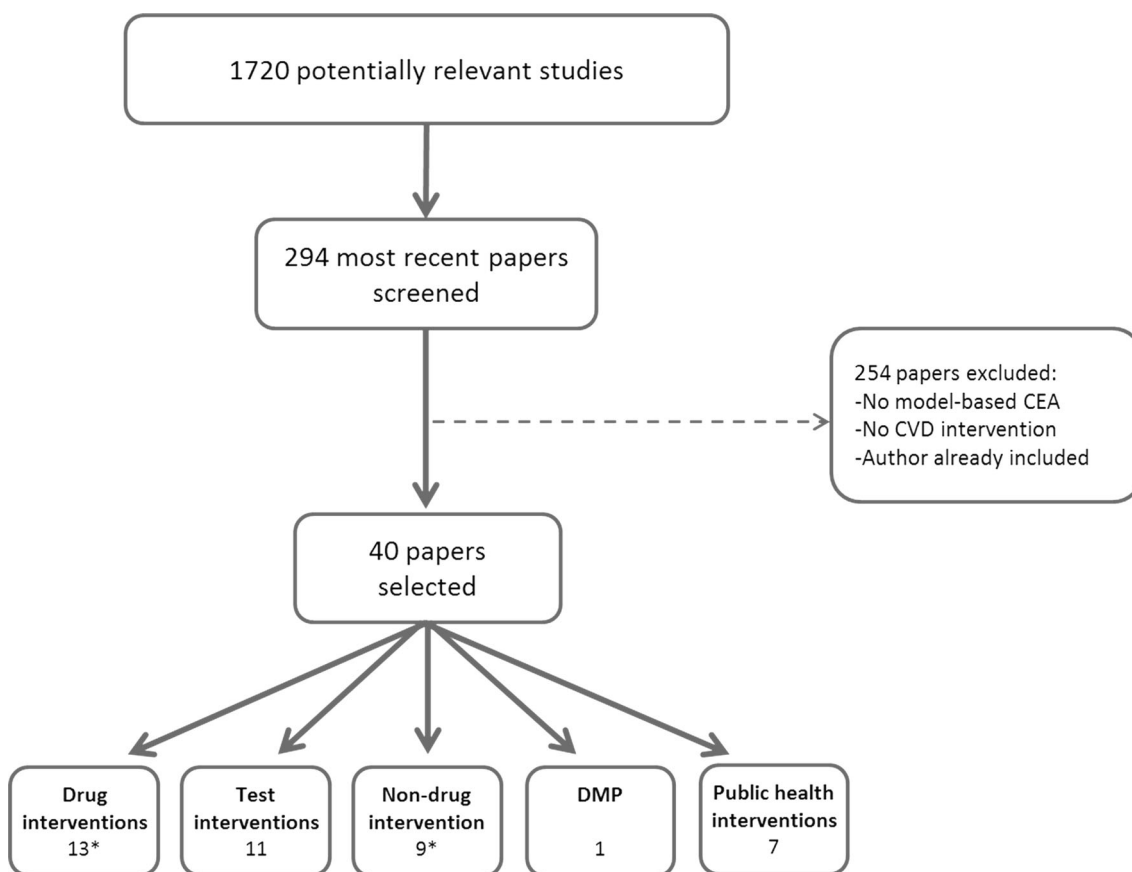

Fig. 1 Flowchart of selected papers. One study investigated two types of interventions (drug and non-drug interventions) [56]. Modelling; cost effectiveness; cardiovascular disease; challenges; review; survey; stents; pharmaceuticals; diagnostic test; public health. *DMP* disease management programme

Table 1 Characteristics of the studies included in the literature review

	Frequency Intervention strategy[a]										
	Test		Drug		Non-drug		DMP		Public health intervention		
	N	%	N	%	N	%	N	%	N	%	Total
Generic journals	1	17	1	17	0	0	0	0	4	67	6
The Lancet	1	50							1	50	2
New England Journal of Medicine									1	100	1
Annals of Internal Medicine			1	33					2	67	3
Journal of the American Medical Association											0
Cardiovascular disease journals	5	33	4	27	4	27	0	0	2	13	15
European Heart Journal					1	100					1
Circulation	3	43	2	29	1	14			1	14	7
Journal of the American College of Cardiology	2	50			1	25			1	25	4
International Journal of Cardiology			2	67	1	33					3
Economic evaluation journals	5	25	8	40	5	25	1	5	1	5	20
Health Technology Assessment	1	33			1	33			1	33	3
Medical Decision Making			1	100							1
Value in Health	3	21	6	43	4	29	1	7			14
PharmacoEconomics	1	50	1	50							2
Total	11	27	13	32	9	22	1	2	7	17	41

DMP disease management programme

[a] One study investigated two types of interventions (drug and non-drug interventions) [56]

relevant papers with unique corresponding authors, we read the title, abstract and full paper of 294 publications in case the title and abstract were non-conclusive (Fig. 1). Table 1 provides an overview of the publications written by these authors. Most (49 %) of the publications involved CEAs of tests and drugs; analyses of non-drug interventions, DMPs and public health interventions were less common. Health economics journals accounted for approximately half of the publications included in this study.

3.2 Challenges

Table 2 presents the presence and frequency of each challenge in each type of intervention and provides ways to address them. Furthermore, papers that have presented a solution for a specific challenge are also identified in Table 2.

3.2.1 Data Requirement Challenges

Challenges such as lack of data (e.g. effectiveness, costs, adverse events or parameter distributions) and difficulties in evidence synthesis are usually addressed by performing sensitivity analyses which show the impact on the outcomes. Univariate or multivariate sensitivity analyses, scenario analyses or probabilistic sensitivity analyses are often used in addressing data requirement challenges.

Furthermore, it is important to recognize the problem of publication bias when a meta-analysis is performed to estimate the effectiveness of an intervention. Trial registries (e.g. ClinicalTrials.gov) can be searched to identify clinical trials that have not published their results, to reduce the risk of bias from selective publication. Furthermore, funnel plots can be used to identify if publication bias exists [10]. When publication bias is an issue then it could be useful to adjust for this in the meta-analysis [11].

3.2.2 Structural Uncertainty

Besides parameter uncertainty, our document analysis showed that modelling studies often encountered structural uncertainties. This means that it was difficult to include or consider (1) all relevant comparators; (2) all relevant disease states or events; and (3) a sufficient time horizon to capture all relevant differences in costs and consequences. Often, not all relevant comparators are included in the model due to data requirements, as could be seen in the study performed by Magnuson et al. [12], which estimated the cost effectiveness of drug-eluting stents compared with coronary artery bypass based on the Future Revascularization Evaluation in Patients with Diabetes Mellitus: Optimal Management of Multivessel Disease (FREEDOM) trial. The trial did not include the second generation of drug-eluting stents, which meant that the authors were

Table 2 Frequencies of challenges per type of interventions

	Test intervention (N = 11)		Non-drug intervention (N = 9)		Drug intervention (N = 13)		DMP (N = 1)		Public health intervention (N = 7)		Total (N = 41[b])		Methodological paper	Example paper
	N	%	N	%	N	%	N	%	N	%	N	%	References	References
Data requirement challenges														
1. Treatment effectiveness	8	73	8	89	10	77	1	100	5	71	32	80	[2, 11, 57]	
2. Prevalence	2	18	NA	NA	NA	NA	NA	NA	3	43	5	31	[2, 57]	
3. Accuracy data	5	45	NA	NA	NA	NA	NA	NA	4	57	9	56	[2, 42, 57]	
4. Compliance[a]	3	27	1	11	5	38	1	100	1	14	11	37	[2, 43, 57]	[44]
5. Quality of life	7	64	8	89	7	54	0	0	3	43	24	71	[2, 57]	
6. Resource use	6	55	7	78	7	54	1	100	3	43	24	60	[2, 57]	
7. Unit costs	6	55	5	56	9	69	0	0	4	57	23	56	[2, 57]	
8. Indirect costs	1	9	1	11	0	0	0	0	0	0	2	5	[58]	
9. Missing values	1	9	4	44	0	0	0	0	0	0	5	13	[59]	
10. Parameter distributions	4	36	4	44	3	23	1	100	2	29	14	35	[2, 57, 60]	
11. Adverse events	5	45	6	67	6	46	0	0	1	14	18	45		
12. Subpopulation data	4	36	1	11	2	15	0	0	2	29	8	20	[23]	
13. Evidence synthesis	1	9	3	33	5	38	0	0	1	14	10	25	[1, 3, 6, 13]	[61]
Modelling challenges														
1. Structure	3	27	4	44	3	23	0	0	4	57	13	32	[21, 22]	
1a. Comparators	3	27	3	33	0	0	0	0	0	0	6	15	[2, 13]	[20, 62]
1b. Disease pathway	2	18	2	22	0	0	0	0	2	29	6	15	[1, 63]	[64]
1c. Time horizon	0	0	1	11	1	8	0	0	0	0	2	5	[2, 3, 65]	[66]
2. Heterogeneity	2	18	0	0	2	15	0	0	0	0	4	10	[23, 57]	
3. History	0	0	0	0	0	0	0	0	0	0	0	0	[2]	
4. Extrapolating short/intermediate results	7	64	7	78	7	54	1	100	4	57	26	65	[2]	[24]
5. Competing risks	2	18	0	0	1	8	0	0	0	0	3	8	[27]	[28]
6. Multiple testing	3	27	0	0	NA	0	NA	0	2	29	5	31	[39, 40]	
7. Multiple interventions effects	0	0	1	11	0	0	0	0	2	29	3	8		[54]
8. Learning curve	2	18	0	0	NA	NA	NA	NA	0	0	2	8	[34]	[35]
9. Wait time (e.g. capacity constraints)	0	0	2	22	NA	NA	NA	NA	0	0	2	8	[67]	[36]
10. Multiple indications	0	0	0	0	0	0	0	0	1	14	1	3	[33]	
11. Lead time	0	0	NA	NA	0	0	0	0	0	0	0	0	[49]	
12. Reusability	NA	NA	0	0	NA	NA	NA	NA	NA	NA	0	0	[48]	
13. Process utilities	NA	NA	0	0	NA	NA	NA	NA	NA	NA	0	0	[48]	
14. Scenario analyses	0	0	0	0	0	0	0	0	0	0	0	0		

DMP disease management programme, *NA* not applicable

The proportions were calculated based on the number of studies that could have been exposed to the challenges

[a] Compliance in studies that evaluate test interventions applies to the drug treatment, which is part of the strategy

[b] One study investigated two types of interventions (drug and non-drug interventions)

unable to evaluate stents that were available during the trial. When direct evidence between comparators is lacking then mixed treatment comparison or network meta-analysis could be a solution. Recently, seven tutorial papers were published on evidence synthesis methods for decision making, including network meta-analysis [13–19]. Stettler et al. [20] compared the safety and effectiveness of bare-metal stents and drug-eluting stents by means of a network

meta-analysis. Bojke et al. [21] discussed ways (model averaging, model selection and parameterizing structural uncertainty) to address structural uncertainty. Frederix et al. [22] have explored the influence of model structures in breast cancer treatment on the estimated cost effectiveness of an intervention.

3.2.3 Patient Heterogeneity

The difficulty of incorporating patient heterogeneity is reported in some of the papers. Sufficient incorporation of heterogeneity in a model requires a great deal of data that is often not available. Recently, a review by Grutters et al. [23] provided a comprehensive overview of the current knowledge regarding patient heterogeneity within economic evaluations of healthcare programmes, and provided guidance for researchers to address heterogeneity.

3.2.4 Extrapolation of Short- or Intermediate-Term Results

Modelling guidelines recommend that models should include long-term or final outcomes [2]. One common problem in modelling is that the length of follow-up of a clinical study used in the model is shorter than the time horizon of the model. Another problem is the fact that only intermediate outcomes (e.g. sensitivity and specificity of coronary angiography or surrogate outcomes (e.g. the effect of statins on low-density lipoprotein [LDL] cholesterol) are presented. Methods to extrapolate intermediate and surrogate outcomes are: (1) using population-level data (e.g. national mortality statistics); (2) long-term epidemiological (observational) studies or registries that reflect the natural history of disease; (3) extrapolating survival curves; and (4) assuming different scenarios for extrapolation (based on, for example, expert opinion). A common approach in CVD is to use trial-based results (short-term) and extrapolate them by using literature or assuming different scenarios for extrapolation. Furthermore, final outcomes (lifetime costs or survival) of previously published models that focus on a later stage in disease progression can be used. For example, Lieu et al. [24] used published results from a modelling study to estimate the cost effectiveness of primary angioplasty. In addition, a CEA evaluating new-generation computed tomography (CT) scanners for the diagnosis of coronary artery disease (CAD) combined five existing models to extrapolate test outcomes [25]. However, combining existing models introduces additional uncertainty since these are often designed for different populations/interventions [26].

3.2.5 Competing Risks

Some studies (8 %) recognized the challenge of competing risks (events that preclude or alter the likelihood of another event occurring [4]. The paper of Putter et al. [27] reviewed statistical methods for the analysis of competing risks and how to model them. Wolbers et al. [28] has considered three models to account for competing risks in risk prediction models for coronary heart disease.

3.2.6 Multiple Intervention Effects

Some treatment strategies consist of multiple interventions and some single interventions have an effect on more than one clinical outcome. Estimating the effectiveness of such interventions could be a challenge. It is more likely that CEAs evaluating public health interventions or DMPs have more difficulty in estimating the effectiveness since they often exist of multiple interventions. This challenge can also arise when estimating the cost effectiveness of lifestyle interventions or drug interventions such as the 'poly pill' that combines several pills (e.g. statins, aspirin, blood pressure lowering drugs, folic acid) [29]. It is important to recognize the problem of interrelating outcomes and overestimation of the 'real' treatment effect. Interrelating outcomes are present if an intervention (e.g. cardiovascular DMP) has an effect on multiple outcomes (e.g. blood pressure and smoking) which interact in some way to improve health. When it comes to primary CVD prevention, the Framingham risk score [30] or the Systematic COronary Risk Evaluation (SCORE) risk function [31] are often used to overcome the problem of interrelating outcomes. Both risk functions estimate the risk of developing a (non-) fatal event in the coming 10 years based on several risk factors (e.g. smoking, cholesterol or age). While none of the included studies consisted of multiple interventions, some of them recognized that their single intervention could have an effect on more than one clinical outcome. When treatment effectiveness of such a strategy is lacking, it is possible to use a synergy factor. Van Kempen et al. [32] multiplied the individual relative risks of the single interventions (aspirin and statins) and multiplied this with a synergy factor. This factor can be varied through sensitivity analyses, and incorporates the interaction between drugs (synergy or dyssynergy).

3.2.7 Learning Curve

Obtaining effectiveness evidence of new tests and other non-drug interventions (e.g. endovascular aneurysm repair vs. open aneurysm repair) is often difficult due to the presence of a learning curve [33]. It is important to include the consequences of learning effects for new (invasive) procedures. Learning effects could influence operating time of a procedure, diagnostic accuracy, and the frequency of adverse events. Ramsay et al. [34] presented statistical techniques to incorporate learning effects of tests and

procedures but concluded that new statistical techniques should be developed. The impact of incorporating a learning curve on the cost effectiveness of strategies will vary depending on various factors. Woods et al. [35] investigated the impact of the learning curve of heart transplantation on the operative and postoperative hospital costs.

3.2.8 Waiting Time

In real life it is likely that tests and procedures are not performed immediately after each other, sometimes because of capacity issues. However, CEAs usually do not incorporate capacity constraints and the time delay between tests and procedures. Two papers identified waiting time as a challenge [36, 37], and one of these examined this issue in detail [36]. Neglecting waiting time may lead to an overestimate of the effectiveness and cost effectiveness of an intervention since a delay in assessment may prolong suffering or increase the risk of cardiovascular events (e.g. myocardial infarction or cardiac arrest). If waiting time is deemed important, a disutility for the quality-of-life loss due to postponed treatment can be added. When the delay in treatment is due to capacity constraints, then modelling guidelines suggest using an agent-based simulation model or a discrete-event simulation (DES) to incorporate competition for resources [2], as performed by Jahn et al. [36] who evaluated the cost effectiveness of drug-eluting stents versus bare-metal stents. A tree that precedes a Markov model can also be used to include waiting time from capacity constraints by modelling a proportion of the cohort that suffers from these constraints.

3.2.9 Multiple Indications

Performing an economic evaluation of an intervention that can be used for several indications is also considered a challenge (3 %). Usually, interventions are evaluated for a specific indication; however, many tests and drugs can be used for several indications. Drugs are divisible and can be evaluated for each indication separately and thus are not a challenge per se. However, for example, the cost effectiveness of a CT scanner in diagnosing CAD can be evaluated but this intervention can also be used in various ways, such as brain CT scans. The weighted average of its use in multiple applications can be used to estimate the overall value of both costs and effects of the intervention [33] in order to decide whether or not to purchase the scanner. In order to estimate a weighted average of its use for all applicants we need to know the relative frequency of each application and have a sufficient understanding of the alternative strategy, including the health and economic consequences of correct and incorrect diagnoses. Furthermore, all effects would preferably be expressed using the same unit of health gain (i.e. life-years or quality-adjusted life-years).

3.2.10 Diagnostic Performance

Several authors (56 %) have indicated that it is a challenge to obtain values for the diagnostic accuracy (sensitivity and specificity) of a test. These input parameters to a model usually have an important impact on the cost effectiveness of the test since they are key in extrapolating an initial disease status assessment, either being correct or incorrect (false positive or false negative diagnosis). It is even more of a challenge when tests are performed in combination or sequence and previous test results need to be incorporated in the model, as recognized by Denchev et al. [38]. It is a challenge to estimate the sensitivity and specificity of each individual test and of the whole strategy. It is very common to assume that tests are independent; however, this does not allow for already-known test results leading to misinterpretation of the test results (posterior probability) since the interpretation of a test depends on the prior probability (known test results and prevalence of outcome) and the accuracy of the test. Hunink [39] explored the influence of assuming independence for multiple test strategies. Weintraub et al. [40] examined the application of Bayes theorem in non-invasive diagnosis of CAD and showed that it is not always appropriate to assume independency. This challenge is mainly due to lack of data since it is almost impossible to perform an observational diagnostic evaluation study in order to derive reliable estimates of diagnostic, therapeutic and health status outcomes for multiple test strategies [41]. Furthermore, estimates of the sensitivity and specificity of a test (e.g. CT angiography) could also be invalid since they are often derived by comparing the result with a 'gold standard' (e.g. invasive coronary angiography) which may not necessarily be 100 % accurate [25]. When data is not available, the best strategy is to vary the accuracy estimates in sensitivity analyses to estimate the impact on cost effectiveness. However, the sensitivity and specificity of a test might be linked, so that improvements in one parameter may be achieved at the expense of reductions in the other [42]. Berry et al. [42] incorporated the link in a decision tree evaluating magnetic resonance angiography. In addition, the ISPOR-SMDM modelling guidelines recommend incorporating test results that are prognostic in the states or as tracker variables in state-transition models [3].

3.2.11 Compliance and Persistence

Compliance and persistence of interventions (e.g. drug intervention or lifestyle interventions) is usually higher in clinical trials than in daily practice due to close monitoring.

Non-compliance can have an impact on medicine acquisition costs and subsequent overall healthcare resource utilization and costs [43]. However, non-compliance may not always result in clinically meaningful differences between efficacy and effectiveness due to long duration of action in relation to its dosing interval (e.g. statins) [43]. Sokol et al. [44] investigated the impact of medication adherence on hospitalization risk and healthcare costs for four conditions: hypertension, diabetes, chronic heart failure and hypercholesterolemia. Drummond et al. [45] proposed some suggestions for trials to become more generalizable to a real-world situation. Guidelines on compliance measurements are provided by Peterson et al. [46]. While compliance could also be based on what is seen during observational studies, these must be adjusted for confounding through multivariate regression techniques or propensity scoring [3]. If no 'real' estimates of compliance are available, it could be useful to perform several scenario analyses with different assumptions to estimate the impact of compliance on cost effectiveness [47]. Ideally, these scenarios should be based on expert opinion, partly to determine how any observed short-term compliance rates could be extrapolated. Modelling guidelines recommend considering dynamic characteristics such as compliance in the states or as tracker variables [6]. Hughes et al. [43] also recognized the challenge of incorporating compliance into models, and provided some techniques to implement compliance in decision models, DES models and Markov models.

3.2.12 Other Challenges

Other challenges not identified in the document analysis but included in our survey are: (1) reusability; (2) lead time bias; (3) incorporating history; (4) process utilities; and (5) defining appropriate scenario analyses. Some interventions such as telemonitoring devices for heart failure can be reused in several patients and this element of reusability might be incorporated in economic evaluations [48]. Lead-time bias could be present in test interventions in the CVD field [49] and should be considered in CEAs. Incorporating history in a model may be a challenge and could be solved by increasing the number of states in state-transition models or modelling cost effectiveness with a DES model [2]. Process utilities (ease, comfort of use or the unpleasantness of a device) could be a potential challenge, although this was not identified in any of the studies included in the document analysis.

3.3 Questionnaire

In total, six (15 %) corresponding authors completed and returned the questionnaire after one reminder had been sent to all corresponding authors and an additional reminder

sent to the authors who were initially willing to complete the questionnaire. Of the six papers they had authored, two focussed on drugs, two on tests, one on a population-based strategy, and one on a device. These authors had previously conducted an average of four CEAs of CVD interventions using an economic model. The challenges found in the document analysis were confirmed by the respondents; in particular, lack of data and extrapolating short-term results. However, the questionnaire also led to the identification of challenges that were not described in the paper. For example, one respondent indicated that extrapolation of initial surrogate outcomes to later clinical events was a challenge, even though this issue was not prominently documented in the paper written by that author. This was very likely because it was a methodological paper that focussed on the impact of utilities on the incremental effects [50].

4 Discussion

Model-based CEAs of CVD interventions are always accompanied by challenges in modelling methods and data requirements. This review identified and analysed the challenges that currently exist in the CVD area. Furthermore, some ways to address the identified challenges based on the literature and on expert opinion are mentioned.

Lack of effectiveness data and quality-of-life data, determining a model structure, and extrapolating short- or intermediate-term results, are very frequently reported or implied challenges in the document analysis. However, frequency is not necessarily an indicator of importance (i.e. degree of impact on the ICER). Less-frequently reported challenges are difficulties in incorporating patient heterogeneity and including waiting time for an intervention in the model. The document analysis also showed that more complex interventions are associated with more and more complex challenges. Public health interventions, DMPs and tests are interventions becoming more difficult to evaluate since they combine several interventions (e.g. companion diagnostic) instead of one single therapy. CVD interventions have become more complex (including DMP, targeted treatments and devices) over the years, meaning that more parameters and relationships between parameters have to be included in the analysis. Consequently, the complexity of models has also increased over the years [51]; for example, a simple decision tree is now used much less often than before. Despite the increased complexity of models, authors of the included studies did not report any challenges that were impossible to be solved. One solution that is often used to overcome challenges is the use of sensitivity analysis. The influence of structural uncertainty and patient heterogeneity on outcomes is less often

assessed in model-based economic evaluations than uncertainty regarding parameters and methodology.

4.1 Limitations

Since we limited our review by including only 40 papers in a limited number of journals, the papers do not represent all model-based CEAs of CVD interventions. However, we do not expect the identified challenges identified using the document analysis to differ substantially from those reported in other papers since we have used the most recent papers and those that were published in the most relevant and highest impact journals.

The identification of challenges from papers is a subjective process which may result in inconsistencies when estimating their frequencies. However, we tried to eliminate this subjectivity by having two reviewers score all studies. Furthermore, the identification of challenges was dependent on whether authors reported all of the challenges they encountered in their study. Consequently, some of the challenges actually encountered might not have been identified by the reviewers and this might have led to an underestimation of the frequencies. To identify challenges that were not reported in the articles we sent the questionnaire to the corresponding authors. However, the response rate was 15 % and therefore the usefulness of the questionnaire results may be limited. For the responders, the questionnaire did confirm the results of our own document analysis.

4.2 Recommendations

This review identified challenges that were present in recently published model-based CEAs in the field of CVD. However, most of these challenges and the ways to address them could also be applied to interventions in other disease areas. Challenges are often the result of data that is not available, particularly relating to CEAs of test and non-drug interventions. In our opinion, there are two main reasons for this. First, the current regulatory framework in the US and Europe for tests and non-drug interventions is less stringent than for pharmaceuticals, i.e. the European Medicines Agency does not require a randomized study design for market approval, while the US FDA requires only a single randomized controlled trial demonstrating safety and effectiveness for high-risk tests or non-drug interventions [48]. We recommend the regulations concerning pharmaceuticals should also be applied to these types of interventions since they are also subject to the same budget constraints and should therefore meet the same requirements for appraisal [48, 52]. Second, tests and non-drug interventions are generally also associated with clinical research-limiting factors such as the impossibility of double blinding. We recommend that primary studies on test and non-drug interventions should pursue rigid research methods, as in drug efficacy studies.

We found that the validity of economic models concerning the challenges and assumptions are often not described in papers. However, if the validity of models is described, this generally concerns face validity and technical validity (debugging) only, instead of disclosing how challenges are addressed. We recommend authors report their findings according to the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement [53], making it easier to investigate the validity of the model. While modellers are generally very resourceful when it comes to overcoming challenges, one could question whether those challenges are adequately addressed since there are often many ways to do so. For instance, Van Kempen et al. [54] showed that the use of different methods to model the treatment effectiveness of statins (through lipid-level modification, fixed risk reduction of CVD events or risk reduction of CVD events proportional to individual change in LDL cholesterol) led to different results. Consequently, they also addressed the importance of carefully considering the assumptions underlying a simulation model and performing extensive model validation. As in the case of breast cancer modelling [55], we recommend standardization of and better guidance for disease-specific modelling in economic evaluations

5 Conclusions

Modelling is unavoidable when performing comprehensive economic evaluations and always comes with challenges. This study provides examples of CVD modelling challenges encountered during studies published in high-impact journals. Some of the reported challenges are specific for CVD, but most challenges are present in all types of diseases. Modelling guidelines do not provide sufficient assistance in resolving all challenges but it is probably unrealistic to expect this. Besides identifying where more research is needed, this review provides some directions for researchers about how to deal with modelling challenges when performing CEAs in the area of CVD.

References

1. Caro JJ, Briggs AH, Siebert U, Kuntz KM, ISPOR-SMDM Modeling Good Research Practices Task Force. Modeling good research practices—overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. Value Health. 2012;15(6):796–803.
2. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M, et al. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. Value Health. 2012;15(6):804–11.
3. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. Value Health. 2012;15(6):812–20.
4. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Moller J, et al. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. Value Health. 2012;15(6):821–7.
5. Pitman R, Fisman D, Zaric GS, Postma M, Kretzschmar M, Edmunds J, et al. Dynamic transmission modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-5. Value Health. 2012;15(6):828–34.
6. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD, et al. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. Value Health. 2012;15(6):835–42.
7. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB, et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. Value Health. 2012;15(6):843–50.
8. Centre for Reviews and Dissemination. NHS EED economics filter: Medline (Ovid) monthly search; 2013. http://www.crd.york.ac.uk/crdweb/AboutPage.asp.
9. ISI Web of Knowledge. Journal Citation Reports 2012; 2013. http://admin-apps.webofknowledge.com/JCR/JCR?SID=P2O5aGPGn3SVdnPEIKN&locale=en_US.
10. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011;22(343):d4002.
11. Sutton AJ, Song F, Gilbody SM, Abrams KR. Modelling publication bias in meta-analysis: a review. Stat Methods Med Res. 2000;9(5):421–45.
12. Magnuson EA, Farkouh ME, Fuster V, Wang K, Vilain K, Li H, et al. Cost-effectiveness of percutaneous coronary intervention with drug eluting stents versus bypass surgery for patients with diabetes mellitus and multivessel coronary artery disease: results from the FREEDOM trial. Circulation. 2013;127(7):820–31.
13. Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 1: introduction. Med Decis Making. 2013;33(5):597–606.
14. Ades AE, Caldwell DM, Reken S, Welton NJ, Sutton AJ, Dias S. Evidence synthesis for decision making 7: a reviewer's checklist. Med Decis Making. 2013;33(5):679–91.
15. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Med Decis Making. 2013;33(5):607–17.
16. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 6: embedding evidence synthesis in probabilistic cost-effectiveness analysis. Med Decis Making. 2013;33(5):671–8.
17. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 3: heterogeneity—subgroups, meta-regression, bias, and bias-adjustment. Med Decis Making. 2013;33(5):618–40.
18. Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 5: the baseline natural history model. Med Decis Making. 2013;33(5):657–70.
19. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. Med Decis Making. 2013;33(5):641–56.
20. Stettler C, Wandel S, Allemann S, Kastrati A, Morice MC, Schomig A, et al. Outcomes associated with drug-eluting and bare-metal stents: a collaborative network meta-analysis. Lancet. 2007;370(9591):937–48.
21. Bojke L, Claxton K, Palmer S, Sculpher M. Defining and characterising structural uncertainty in decision analytic models. CHE Research Paper 9; 2006. http://www.york.ac.uk/che/pdf/rp9.pdf
22. Frederix GW, Severens JL, Hovels AM, Raaijmakers JA, Schellens JH. Reviewing the cost-effectiveness of endocrine early breast cancer therapies: influence of differences in modeling methods on outcomes. Value Health. 2012;15(1):94–105.
23. Grutters JP, Sculpher M, Briggs AH, Severens JL, Candel MJ, Stahl JE, et al. Acknowledging patient heterogeneity in economic evaluation: a systematic literature review. Pharmacoeconomics. 2013;31(2):111–23.
24. Lieu TA, Gurley RJ, Lundstrom RJ, Ray GT, Fireman BH, Weinstein MC, et al. Projected cost-effectiveness of primary angioplasty for acute myocardial infarction. J Am Coll Cardiol. 1997;30(7):1741–50.
25. Westwood ME, Raatz HD, Misso K, Burgers L, Redekop K, Lhachimi SK, et al. Systematic review of the accuracy of dual-source cardiac CT for detection of arterial stenosis in difficult to image patient groups. Radiology. 2013;267(2):387–95.
26. Burgers LT, Redekop W, Westwood M, Lhachimi S, Severens JL, Armstrong N, et al. Using five existing models to comprehensively model the cost-effectiveness of a high definition ct scanner in a coronary artery disease population: a nice diagnostic guidance project. Value Health. 2011;14(7):A259.
27. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Stat Med. 2007;26(11):2389–430.
28. Wolbers M, Koller MT, Witteman JC, Steyerberg EW. Prognostic models with competing risks: methods and application to coronary risk prediction. Epidemiology. 2009;20(4):555–61.
29. van Gils PF, Over EA, Hamberg-van Reenen HH, de Wit GA, van den Berg M, Schuit AJ, et al. The polypill in the primary prevention of cardiovascular disease: cost-effectiveness in the Dutch population. BMJ Open. 2011;1(2):e000363.
30. Anderson KM, Odell PM, Wilson PWF, Kannel WB. Cardiovascular-disease risk profiles. Am Heart J. 1991;121(1):293–8.
31. Conroy RM, Pyorala K, Fitzgerald AP, Sans S, Menotti A, De Backer G, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. Eur Heart J. 2003;24(11):987–1003.
32. van Kempen BJ, Spronk S, Koller MT, Elias-Smale SE, Fleischmann KE, Ikram MA, et al. Comparative effectiveness and cost-effectiveness of computed tomography screening for coronary artery calcium in asymptomatic individuals. J Am Coll Cardiol. 2011;58(16):1690–701.
33. Drummond M, Griffin A, Tarricone R. Economic evaluation for devices and drugs—same or different? Value Health. 2009;12(4):402–4.
34. Ramsay CR, Grant AM, Wallace SA, Garthwaite PH, Monk AF, Russell IT. Statistical assessment of the learning curves of health technologies. Health Technol Assess. 2001;5(12):1–79.
35. Woods JR, Saywell RM Jr, Nyhuis AW, Jay SJ, Lohrman RG, Halbrook HG. The learning curve and the cost of heart transplantation. Health Serv Res. 1992;27(2):219–38.

36. Jahn B, Theurl E, Siebert U, Pfeiffer KP. Tutorial in medical decision modeling incorporating waiting lines and queues using discrete event simulation. Value Health. 2010;13(4):501–6.

37. Brown LC, Powell JT, Thompson SG, Epstein DM, Sculpher MJ, Greenhalgh RM. The UK EndoVascular Aneurysm Repair (EVAR) trials: randomised trials of EVAR versus standard therapy. Health Technol Assess. 2012;16(9):1–218.

38. Denchev P, Kaltman JR, Schoenbaum M, Vitiello B. Modeled economic evaluation of alternative strategies to reduce sudden cardiac death among children treated for attention deficit/hyperactivity disorder. Circulation. 2010;121(11):1329–37.

39. Hunink M. Multiple test results. In: Hunink M, Glasziou P, Siegel J, Weeks J, Pliskin J, Elstein A, et al., editors. Decision making in health and medicine. New York: Cambridge university press; 2010. p. 177.

40. Weintraub WS, Madeira SW Jr, Bodenheimer MM, Seelaus PA, Katz RI, Feldman MS, et al. Critical analysis of the application of Bayes' theorem to sequential testing in the noninvasive diagnosis of coronary artery disease. Am J Cardiol. 1984;54(1):43–9.

41. Dijkgraaf M. Evaluation of diagnostic device clinical outcomes. In: Ackerman S, Smith M, Ehreth J, Eldessouki R, Sullivan E, editors. Therapeutic and diagnostic device outcomes research. USA: International Society for Pharmacoeconomics and Outcomes Research; 2011. p. 139.

42. Berry E, Kelly S, Hutton J, Lindsay HS, Blaxill JM, Evans JA, et al. Intravascular ultrasound-guided interventions in coronary artery disease: a systematic literature review, with decision-analytic modelling, of outcomes and cost-effectiveness. Health Technol Assess. 2000;4(35):1–117.

43. Hughes D, Cowell W, Koncz T, Cramer J, International Society for Pharmacoeconomics & Outcomes Research Economics of Medication Compliance Working Group. Methods for integrating medication compliance and persistence in pharmacoeconomic evaluations. Value Health. 2007;10(6):498–509.

44. Sokol MC, McGuigan KA, Verbrugge RR, Epstein RS. Impact of medication adherence on hospitalization risk and healthcare cost. Med Care. 2005;43(6):521–30.

45. Drummond MF, Sculpher MJ, Torrance G, O'Brien BJ, Stoddart G. Economic evaluation using patient-level data. Methods for the economic evaluation of health care programmes. 3rd ed. Oxford: Oxford University Press; 2005. p. 247.

46. Peterson AM, Nau DP, Cramer JA, Benner J, Gwadry-Sridhar F, Nichol M. A checklist for medication compliance and persistence studies using retrospective databases. Value Health. 2007;10(1):3–12.

47. Heart Protection Study Collaborative, Mihaylova B, Briggs A, Armitage J, Parish S, Gray A, et al. Lifetime cost effectiveness of simvastatin in a range of risk groups and age groups derived from a randomised trial of 20,536 people. BMJ. 2006;333(7579):1145.

48. Kirisits A, Redekop WK. The economic evaluation of medical devices: challenges ahead. Appl Health Econ Health Policy. 2013;11(1):15–26.

49. Goldman L, Mudge GH Jr, Cook EF. The changing "natural history" of symptomatic coronary artery disease: basis versus bias. Am J Cardiol. 1983;51(3):449–54.

50. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. Value Health. 2010;13(5):509–18.

51. Weinstein MC. Recent developments in decision-analytic modelling for economic evaluation. Pharmacoeconomics. 2006;24(11):1043–53.

52. Taylor RS, Iglesias CP. Assessing the clinical and cost-effectiveness of medical devices and drugs: are they that different? Value Health. 2009;12(4):404–6.

53. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. Pharmacoeconomics. 2013;31(5):361–7.

54. van Kempen BJ, Ferket BS, Hofman A, Spronk S, Steyerberg E, Hunink MG. Do different methods of modeling statin treatment effectiveness influence the optimal decision? Med Decis Making. 2012;32(3):507–16.

55. Frederix GW, Severens JL, Hovels AM, Raaijmakers JA, Schellens JH. The cloudy crystal ball of cost-effectiveness studies. Value Health. 2013;16(6):1100–2.

56. Turner D, Raftery J, Cooper K, Fairbank E, Palmer S, Ward S, et al. The CHD challenge: comparing four cost-effectiveness models. Value Health. 2011;14(1):53–60.

57. Spiegelhalter DJ, Best NG. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. Stat Med. 2003;22(23):3687–709.

58. Koopmanschap MA, Rutten FF. A practical guide for calculating indirect costs of disease. Pharmacoeconomics. 1996;10(5):460–6.

59. Briggs A, Clark T, Wolstenholme J, Clarke P. Missing… presumed at random: cost-analysis of incomplete data. Health Econ. 2003;12(5):377–92.

60. Briggs AH. Handling uncertainty in cost-effectiveness models. Pharmacoeconomics. 2000;17(5):479–500.

61. Downs SM. Technical report: urinary tract infections in febrile infants and young children. The Urinary Tract Subcommittee of the American Academy of Pediatrics Committee on Quality Improvement. Pediatrics. 1999;103(4):e54.

62. Robinson M, Palmer S, Sculpher M, Philips Z, Ginnelly L, Bowens A, et al. Cost-effectiveness of alternative strategies for the initial medical management of non-ST elevation acute coronary syndrome: systematic review and decision-analytical modelling. Health Technol Assess. 2005;9(27):iii–iv, ix–xi, 1–158.

63. Reed SD, Anstrom KJ, Bakhai A, Briggs AH, Califf RM, Cohen DJ, et al. Conducting economic evaluations alongside multinational clinical trials: toward a research consensus. Am Heart J. 2005;149(3):434–43.

64. Jones L, Griffin S, Palmer S, Main C, Orton V, Sculpher M, et al. Clinical effectiveness and cost-effectiveness of clopidogrel and modified-release dipyridamole in the secondary prevention of occlusive vascular events: a systematic review and economic evaluation. Health Technol Assess. 2004;8(38):iii–iv, 1–196.

65. Briggs A, Claxton K, Sculpher M. Introduction. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006. p. 1.

66. Hill R, Bagust A, Bakhai A, Dickson R, Dundar Y, Haycox A. Coronary artery stents: a rapid systematic review and economic evaluation. Health Technol Assess. 2004;8(35):1–256.

67. Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment: an overview and guide. Pharmacoeconomics. 2008;26(2):131–48.