CrossMark

SYSTEMATIC REVIEW

Unremarked or Unperformed? Systematic Review on Reporting of Validation Efforts of Health Economic Decision Models in Seasonal Influenza and Early Breast Cancer

Pieter T. de Boer[1] · Geert W. J. Frederix[2] · Talitha L. Feenstra[3,4] · Pepijn Vemer[1,3]

Abstract

Background Transparent reporting of validation efforts of health economic models give stakeholders better insight into the credibility of model outcomes. In this study we reviewed recently published studies on seasonal influenza and early breast cancer in order to gain insight into the reporting of model validation efforts in the overall health economic literature.

Methods A literature search was performed in Pubmed and Embase to retrieve health economic modelling studies published between 2008 and 2014. Reporting on model validation was evaluated by checking for the word validation, and by using AdViSHE (Assessment of the Validation Status of Health Economic decision models), a tool containing a structured list of relevant items for validation. Additionally, we contacted corresponding authors to ask whether more validation efforts were performed other than those reported in the manuscripts.

Results A total of 53 studies on seasonal influenza and 41 studies on early breast cancer were included in our review. The word validation was used in 16 studies (30 %) on seasonal influenza and 23 studies (56 %) on early breast cancer; however, in a minority of studies, this referred to a model validation technique. Fifty-seven percent of seasonal influenza studies and 71 % of early breast cancer studies reported one or more validation techniques. Cross-validation of study outcomes was found most often. A limited number of studies reported on model validation efforts, although good examples were identified. Author comments indicated that more validation techniques were performed than those reported in the manuscripts.

Conclusions Although validation is deemed important by many researchers, this is not reflected in the reporting habits of health economic modelling studies. Systematic reporting of validation efforts would be desirable to further enhance decision makers' confidence in health economic models and their outcomes.

✉ Pepijn Vemer
p.vemer@rug.nl; healthecon@pepijnvemer.nl

[1] Department of Pharmacy, PharmacoTherapy, -Epidemiology and -Economics (PTEE), University of Groningen, Groningen, The Netherlands

[2] Pharmacoepidemiology and Clinical Pharmacology, University of Utrecht, Utrecht, The Netherlands

[3] Department of Epidemiology, University Medical Center Groningen, University of Groningen, PO Box 30.001, 9700 RB Groningen, The Netherlands

[4] Centre for Nutrition, Prevention and Health Services Research, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

Key Points for Decision Makers

All stakeholders have a vested interest in a high validation status of health economic models since they play an important role in the economic evaluation of therapeutic interventions. Transparent reporting of validation efforts and their outcomes will allow stakeholders to make their own judgment of a model's validation status.

Only a limited number of studies reported on validation efforts, although good examples were identified. To further increase transparency, more explicit and structured attention to the reporting of validation efforts by authors and journals seems worthwhile.

1 Introduction

Health economic decision analytic models play an important role in the economic evaluation of therapeutic interventions [1]. Since policy decisions are influenced by the results of such models, all stakeholders have a vested interest in a high validation status of these models. Transparent reporting of validation efforts and their outcomes will give the stakeholder better insight into the model's credibility (is the model scientifically sound?), salience (is the model applicable within the context?) and legitimacy (are all stakeholder concerns, values and views included properly?) [2, 3]. Proper information regarding these aspects allows stakeholders to make their own judgement of the models' validation status.

Several systematic reviews of health economic evaluations in different disease areas indicated that little was reported on model validation [4–10]; however, most of these reviews were not focused on the general quality of modelling aspects and contained little details on model validation performances. Only one review, focusing on interventions on cardiovascular diseases, provided a clear overview on which part of the included studies reported on model validation tests distinguishing model validation techniques according to the International Society for Pharmacoeconomics and Outcomes Research–Society for Medical Decision Making (ISPOR–SMDM) guidelines [9]. However, modelling evaluation processes might vary between different disease areas, therefore more studies assessing model validation efforts are needed.

In this study we aimed to systematically review the reporting of validation efforts of recently published health economic decision models, explicitly distinguishing

between different validation techniques. For this purpose, we chose two example diseases, namely seasonal influenza (SI) and early breast cancer (EBC). These two diseases are well-defined and by choosing both a communicable disease, which is often modelled using dynamic models [11], and a non-communicable disease, which is often modelled using static models, we expected to cover a wide range of model types [1]. This should provide a good overview of the current standard in the reporting of validation efforts in the health economic literature.

Since validation is an integral part of the modelling process (see, for example, Fig. 2 in Sargent [12]), low reporting of validation efforts does not have to mean that they were not performed. For instance, impromptu checking of bits of computer code while coding may not always be reported. In order to gain insight into the discrepancy between the performance and reporting of validation efforts, we also reached out to the corresponding author of each of the included papers in this review for comments.

2 Methods

2.1 Search Strategy and Study Selection

We searched the PubMed and Embase databases to identify studies focusing on the health economic evaluations of SI and EBC. The full search strings for both diseases can be found in Appendix 1 and contained free-text searching terms as well as exploded (Medical Subject Heading [MeSH]) terms. For both disease areas, the health economic evaluations had to meet the following criteria: (1) published in peer reviewed journals from January 2008 to December 2014; (2) presented results of costs as well as health effects; and (3) used a computer simulation model to generate these results. We also screened reference lists of selected articles. Review papers, meta-analyses, letters and non-full-text such as abstracts were excluded, and we restricted our selection to the English language. For SI, studies focusing only on pandemic influenza were excluded, as well as studies analyzing interventions against multiple infectious diseases without showing separated results for SI. For EBC, we excluded studies on metastatic breast cancer, breast cancer screening and diagnostic systems to stage breast cancer.

2.2 Study Characteristics

General characteristics of the studies extracted included year of publication, country, income level of the country according to the classification of the World Bank [13], funding source, type of intervention, type of evaluation and model type. For SI, studies that incorporated disease

transmission dynamics (i.e. from carrier/infected to a susceptible individual), or used discrete event simulation were categorised as dynamic models. In all other cases, the model was categorised as static.

2.3 Reporting of Model Validation Efforts

In this study, we defined validation as the act of evaluating whether a model is a proper and sufficient representation of the system it is intended to represent in view of an application, where 'proper' was defined as "the model is in accordance with what is known about the system" and 'sufficient' was defined as "the results can serve as a solid basis for decision making" [14]. We first searched the publication's text and appendices for the word validation and its conjugate forms (valid*, verif*). When present, we reported the context in which the word was used. Then, the reporting of model validation efforts were systematically assessed, using the outline presented in the validation-assessment tool AdViSHE (Assessment of the Validation Status of Health Economic decision models) [15]. This tool was designed to provide model users with structured information regarding the validation status of health economic decision models, and therefore enables systematic extraction of the reporting of model validation efforts. An added advantage of this tool is that it explicitly presents

clear definitions of validation techniques since there is little, if any, consensus on terminology in the validation literature [16]. An abbreviated form of the AdViSHE tool is shown in Table 1 and includes five validation categories, i.e. validation of the conceptual model (A, the theories and assumptions underlying the model concepts, and the model's structure and causal relationships), input data (B, available input data and data transformations), computerised model (C, implemented software program, including code, mathematical calculations and implementation of the conceptual model), and operational model (D, behaviour of the model outcomes). Remaining validation techniques, such as, for instance, double programming, are assigned to category E. Assessment of studies on validation reporting was performed by two of the authors (PdB and PV for influenza, and PdB and GF for EBC) separately. After comparing results, differences between the two authors were resolved in a consensus meeting. Examples of each validation technique found were collected and presented.

2.4 Comments from Authors

The corresponding authors of the included studies were contacted by email in August 2015, followed by a reminder in November 2015. In this email, we explained the aim of our study, provided details of the corresponding author's

Table 1 Validation aspects included in the AdViSHE validation status assessment tool [15]

Validation categories	Subcategory	Questions
(A) Conceptual model	A1 Face validity	Have experts been asked to judge the appropriateness of the conceptual model?
	A2 Cross validity	Has this model been compared with other conceptual models found in the literature or clinical textbooks?
(B) Input data	B1 Face validity	Have experts been asked to judge the appropriateness of the input data?
	B2 Model fit	When input parameters are based on regression models, have statistical tests been performed?
(C) Computerised model	C1 External review	Has the computerised model been examined by modelling experts?
	C2 Extreme value testing	Has the model been run for specific, extreme sets of parameter values in order to detect any coding errors?
	C3 Testing of traces	Have patients been tracked through the model to determine whether its logic is correct?
	C4 Unit testing	Have individual submodules of the computerised model been tested?
(D) Operational model	D1 Face validity	Have experts been asked to judge the appropriateness of the model outcomes?
	D2 Cross validity	Have the model outcomes been compared with the outcomes of other models that address similar problems?
	D3 Alternative input	Have the model outcomes been compared with the outcomes obtained when using alternative input data?
	D4 Empirical data	Have the model outcomes been compared with empirical data?
		(A) Comparison against the data sources on which the model is based (dependent validation)
		(B) Comparison against a data source that was not used to build the model (independent validation)
(E) Other techniques	–	Have any other validation techniques been performed?

paper that was included in our review, and enquired whether authors had performed validation efforts other than those reported in their manuscript. To provide help on validation techniques that could have been performed, we attached the AdViSHE tool to this email. Any answers from the authors were reported.

3 Results

3.1 Study Selection

The searches resulted in 53 SI studies [17–70] and 41 EBC studies [71–111] that were eligible for inclusion in our review. More details on the study selection process are shown in Fig. 1a, b.

3.2 Study Characteristics

General characteristics of the included studies are shown in Table 2. For both disease areas, most studies were performed in countries in North America, followed by countries in Europe and Asia. Studies were predominantly performed for high-income countries, although we also found studies for middle-income countries such as China, Taiwan and Argentina (SI), and China, Brazil, Colombia

and Iran (EBC). We found no studies for low-income countries.

All SI studies analyzed pharmaceutical interventions, i.e. influenza vaccines or antiviral drugs. One study additionally assessed non-pharmaceutical mitigation strategies, including ventilation, face masks, hand washing and ultraviolet irradiation [21]. For EBC, antineoplastic drugs were predominantly studied, although we also included four studies on radiation or surgery treatments [72, 84, 100, 111].

SI was analyzed using static models in 43 (81 %) of the studies, mostly a decision tree model (36, 68 %) or a state transition model with Markov properties ['Markov model'] (6, 11 %). One study used a multicohort model, in which cohorts of different ages were simultaneously followed over their lifetimes [64]. A total of eight studies (15 %) used a dynamic transmission model, with all compartmental models using an SIR structure [17, 21, 25, 28, 29, 50, 51, 54]. Such models divide the population between susceptible (S), infected (I) and recovered (R), and include a (often age-stratified) mixing pattern between different groups. One study did not elaborate on model structure [57], and another study called the model 'spreadsheet based', with no further information on model type or structure [70]. Thirty-seven EBC studies (90 %) used a Markov model [71–83, 85–95, 97–102, 104–107, 109–
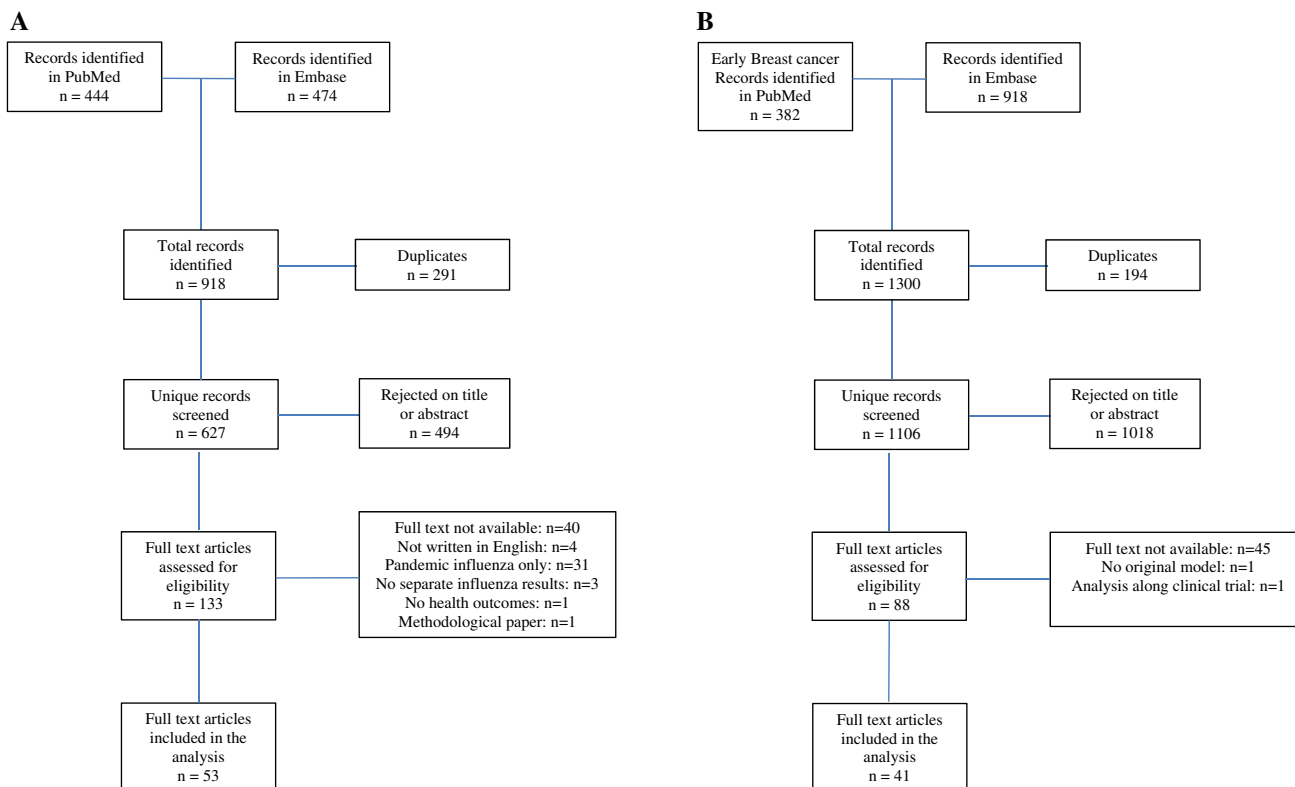

Fig. 1 a Seasonal influenza literature search. **b** Early breast cancer literature search

Table 2 General statistics of the studies included in the review

Study characteristic	Seasonal influenza		Early breast cancer	
	N	% of total	N	% of total
Total studies included	53	100	41	100
Region				
Europe	8	15	14	34
North America	34	64	16	39
Asia	6	11	7	17
South America	1	2	3	7
Australia	4	8	1	2
Income level				
High	50	94	35	85
Middle	3	6	6	15
Low	0	0	0	0
Funding				
Public health sources	28	53	17	41
Industry	15	28	17	41
No external funding	2	4	3	7
Not stated	7	13	4	10
Type of intervention[a]				
Pharmaceutical-related				
Vaccine	45	85	–	–
Drug	8	15	37	90
Non-pharmaceutical related				
Personal protection	1	2	–	–
Radiation	–	–	3	7
Surgical intervention	–	–	1	2
Type of evaluation				
Cost-benefit	6	11	0	0
Cost-effectiveness	6	11	4	10
Cost-utility	41	77	37	90
Model type				
Dynamic				
Compartmental model	8	15	0	0
Semi-Markov	0	0	1	2
Static				
Decision tree	36	68	1	2
Multicohort	1	2	0	0
Markov	6	11	37	90
Not stated	2	4	2	5

[a] Multiple interventions might be studied in one study

111], one study used a decision tree [96], and one study used a semi-Markov model [84]. Two studies did not provide information on model type, but either defined the model as a 'decision analytic model' [108] or stated that 'the model took a state-transition approach' [103].

Within the study selection process, multiple studies had the same first author. For SI, one author conducted ten studies [32–42], two authors conducted three studies [49–51, 67–69] and three authors conducted two studies [23, 24, 55, 56, 60, 61], while for EBC, one author conducted three studies [93–95] and three authors conducted two studies [80, 81, 103, 104, 109, 110].

3.3 Model Validation

3.3.1 Free-Text Search

For SI, the word 'validation' or its conjugates was found in 16 studies (30 %). The context in which 'validation' was used diverged widely. Three studies did not use validation in a model validation context [27, 52, 53]; two studies mentioned that the evidence level of some input data was low and not validated [47, 57]; one study stated that picking a starting date for the simulation between two influenza seasons would be useful "to demonstrate model validity", but did not specify how this was the case [48]; two studies used the word 'validation' in a context that might be linked to a validation technique, but did not provide information on which parts of the model were validated, by whom, or which techniques were used [44, 64]; and three studies stated that a previously validated model was used, without stating whether the model would be valid for the new purpose [56, 60, 61]. Consulting the prior publications these studies were based on did not provide further clarification on validation efforts performed. In seven studies (13 %), the word 'validation' was used in such a way that we were able to link this directly to a validation technique [17, 20, 25, 30, 48, 62, 64]; these are discussed in the next paragraph.

For EBC, 23 studies (56 %) reported the word 'validation' or its conjugate forms, also in various contexts. In four studies we found 'validation' was not related to validation techniques of the health economic model [73, 78, 101, 106]. A fifth study debated "[t]he validity of the assumption of differences in effectiveness with letrozole and anastrozole" [91], while a sixth study mentioned that "[n]o relevant cost and/or utility data were identified against which the model's outputs could be validated" [88]. In two studies by the same first author it was stated that that "[t]he validity of the model is presented using cost effectiveness-acceptance curves" [93, 95]. A total of 16 (39 %) of the included EBC studies used the word 'validation' in a context linked to a model validation technique, which will be addressed below [71, 72, 74, 76, 77, 80, 82, 84, 85, 87, 88, 92, 100, 104, 109, 110].

3.3.2 Validation Techniques

We identified 30 studies (57 %) on SI that reported one or more validation techniques, and 28 (68 %) EBC studies. Two or more validation techniques were found in five
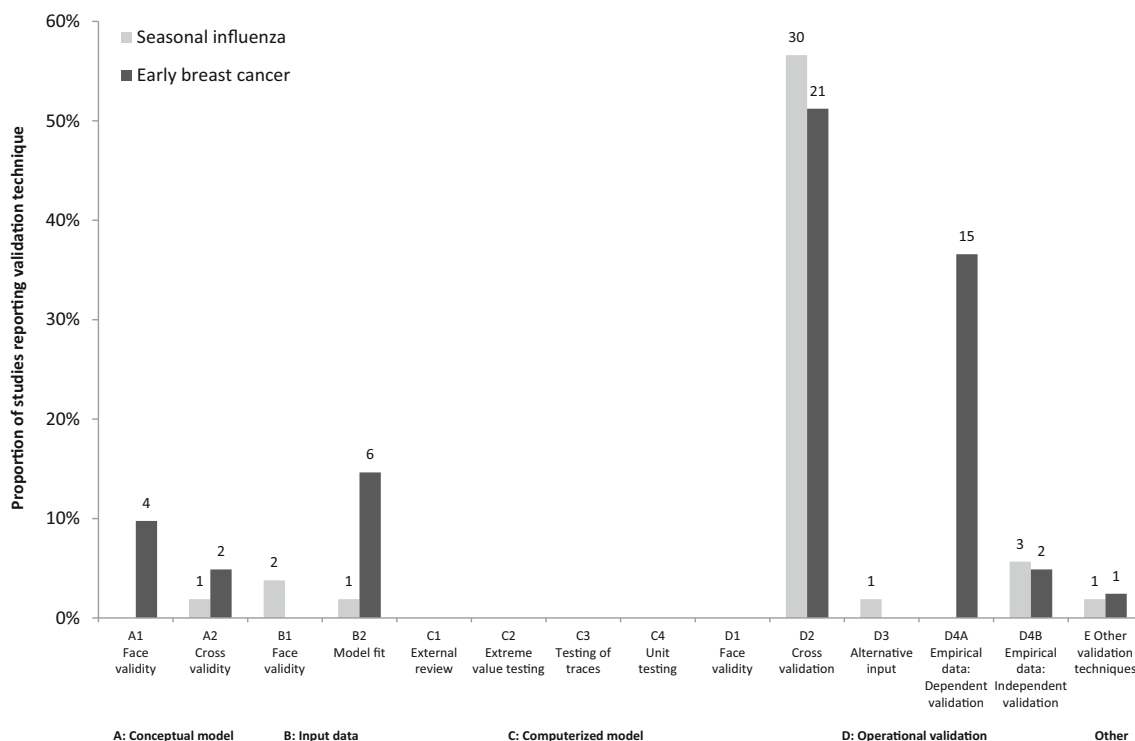
Fig. 2 Model validation performances of 53 studies on seasonal influenza, as reported in publications, using the classification of the AdViSHE tool. *Numbers* above the *bars* represent absolute numbers

of studies. *AdViSHE* Assessment of the Validation Status of Health Economic decision models

studies (9 %) of SI and 15 studies (37 %) on EBC. Figure 2 shows the model validation performance stratified by (sub)category.

Five studies reported on validation of the conceptual model (category A). Four EBC studies reported on face validity of (part of) the conceptual model [74, 85, 88, 92] (A1). For example, Au et al. [74] reported that "validation of [treatment] strategies were achieved by consensus of a Canadian panel of breast cancer oncologists", including the identities of the concerned oncologists. Hall et al. [85] reported that "[t]he structure of the model was developed by consensus between clinical experts, health economists and medical statisticians", and reported the background as well as selection procedure. This study also performed cross-validation of the conceptual model (A2) by performing a systematic review to identify all previously published models of the same intervention and subsequently comparing the conceptual model. We also found cross-validation of the conceptual model in one other study on EBC [88] and one study on SI [47].

Reporting on validation of input data (category B) was found in nine studies. Face validity (B1) was described in two SI studies [20, 62]. For example, Tarride et al. [62] mentioned that "144 Canadian physicians were surveyed to validate [complication rate] estimates for children aged 2-5 years old". Testing of the model fit (B2) was addressed

in one study on SI [30] and six studies on EBC [77, 81, 92, 94, 98, 107]. Jit et al. [30] performed multiple linear regressions to estimate the proportion of hospitalisations that was caused by influenza and provided details on the goodness of fit (R^2) of the model. They also indicated that different regression models to estimate the proportion of healthcare attendances related to influenza gave similar outcomes, which provided internal validation of this input into the health economic model. Purmonen et al. [98] used different parametric survival models (Weibull, exponential, log-logistic) to optimally fit the trace of the curves; however, how they decided on optimal fit was not reported.

Validation of the computerised model (category C) was not found in any of the reviewed articles.

A total of 56 studies reported on validation of the operational model (category D). Cross-validation of the results (D2) was the validation technique found most often in SI (57 %) as well as EBC (51 %). For example, Chit et al. [22] performed an extensive comparison of the number of influenza cases and various other clinical outcomes with data published in a model from the Centers for Disease Control and Prevention. Next to cross-validation, validation of model outcomes to empirical data (D4) was often found, namely in 20 studies. Three studies on SI reported on independent validation (D4B) by validating the predicted incidence of a specified influenza-related clinical

event against data from a national surveillance system or a national registration agency [17, 25, 48]. Seventeen EBC studies performed dependent validation (D4A) by comparing the results with data of the clinical trial the study was based on. Two different EBC studies performed validation against independent data sources, namely against the 'Adjuvant! Online' prediction tool [84, 100], an online tool used in the US to help oncologists estimate the risks of mortality and side effects, given different clinical and treatment scenarios [112]. Face validity testing of the results (D1) was not reported in either of the disease areas, and validation by using alternative input data (D3) was found in one study [47]. This study used all parameters of a similar study analyzing the same intervention in the same country to compare the results.

Finally, two studies reported validation techniques that were not categorised in the AdViSHE tool (Category E) [64, 81]. Van Bellinghen et al. [64] performed double programming by programming one cohort in another software package, and Delea et al. [81] conducted an extensive comparison of the input data compared with the input data of other models, which might be regarded as cross-validation of the input data.

3.4 Comments from Authors

We reached out twice to the corresponding authors to ask whether, in practice, more validation efforts were performed than those reported in the manuscript. We were able to reach 77/94 corresponding authors, and of these we received a total of ten responses. Three responding authors informed us that they were not able to answer the enquiry due to various reasons, such as no access to project files anymore, study was conducted a long time ago, or current workload was too high. Comments from the remaining seven responding authors are shown below.

A first author response of an SI study included that, additionally to what had been reported in the manuscript, the complete model was double programmed (category E) using a different software package, and results were compared. When modifications were completed, the affected modules were checked for face validity by another programmer (C1) and run against test data to ensure consistency (D3). A second author response of an SI study reported that face validity of the conceptual model (A1), input data (B1) and model outcomes (D1) were assessed internally and externally by two different panels of independent researchers. Moreover, tests such as likelihood ratio testing, Akaike information criterion (AIC), Bayesian information criterion (BIC) and goodness of fit (B2) were performed to select the optimum regression model to determine the attributable fraction of influenza within surveillance data of 'influenza-like illness'. The authors

indicated that all input data have been varied outside their ranges to detect coding errors (C2), and mentioned that this was not reported in the manuscript since this was considered a natural part of model development. A third author response of a SI study indicated that face validity of the conceptual model (A1), input data (B1) and outcomes (D1) were tested within the team containing clinical experts of different fields. Moreover, the entire model was double programmed (E). Additionally, the authors indicated that validation of the computerised model (C) was not reported as this was considered standard procedure.

Concerning EBC, a first author response mentioned that the majority of the validation techniques that are found on the AdViSHE tool were performed. The model was constructed to the standards of the National Institute for Health and Care Excellence (NICE) guideline on technical appraisal, which, according to the author, in turn implies a certain standard of validation. A second author response indicated that not all validation performances were reported due to a very restrictive word limit of the journal and the audience of the journal being mainly clinical. However, the structure of the model was reviewed by clinical and health economic experts within and outside the team (A1). In addition, a quality control on input data (B1) and model programming (C1) was performed by health economists not involved in the original model design. Additionally, they attempted to perform cross-validity of the model outcomes (D2) but were unable to do so due to a lack of suitable comparison studies. A third author responded that face validity of model structure (A1), input parameters (B1) and code checking (C1) were performed, and that links between different submodules were also tested (C4). A third author indicated that no other additional validation efforts than those described in the manuscript were performed.

4 Discussion

In this study, we assessed the reporting of model validation efforts in the disease areas of SI and EBC within the period 2008 to 2014. Overall, reporting of model validation efforts was found to be limited. Reviewing the papers systematically using the AdViSHE tool, demonstrated that 57 % of the studies on SI and 71 % of the ECB models performed at least one validation technique; however, only 9 and 37 % of studies on SI and EBC, respectively, performed two or more validation techniques. A limited number of author's responses to our enquiry on model validation efforts performed, indicated that, in practice, considerably more validation techniques might be used than those reported in the manuscripts, provided these few responders are representative of the majority who did not reply to our request for additional information.

The most performed validation technique was cross-validation of the model outcomes. A first explanation for this might be that many general guidelines for writing scientific papers state that the discussion section should include a comparison of the study outcomes with the existing literature (e.g. Hall [113]). Moreover, Eddy et al. [114] specifically name cross-validation as one of the five types of validation. Few reports were identified regarding validation of the conceptual model, and no reports regarding validation of the computerised model. As indicated in two author responses, validation techniques such as code checking and extreme value testing might be regarded as implicit in the model development process and were therefore not reported [12]. This might also partly explain why face validity of the input data and results was not often reported. Moreover, the peer-review process before publication might be regarded by some authors as a way of testing face validity. Another reason why little is reported on validation of the conceptual model might be that many studies of SI used a basic decision-tree model; however, even in case of such simple models, validation remains important. Conceptual model validation in that case might possibly be even more important since the choice of such a simple structure should be justified. A final explanation might be that the word count or the (clinical) audience of the journal might restrict authors on reporting of validation efforts.

In addition to simply describing the conduct of validation, it may be useful for model users to report what was done with the outcomes of the validation techniques. For instance, did the authors make any changes to (parts of) the model when faced with the validation outcomes? Such outcomes may emphasize the importance of validation. Unfortunately, none of the studies included in this review reported this aspect of model validation.

The main difference between SI and EBC was found in validation of the model outcomes by using empirical data. Dependent validation of the model outcomes was found in several studies of EBC but not in studies on SI. This may be due to the nature of SI, which, as a communicable disease, requires complex transmission dynamics and should therefore be studied on a population level rather than on a cohort level. For such dynamic SI models, understanding disease transmission dynamics is complex and the level of indirect protection caused by herd immunity is dependent on vaccine uptake levels. This complicates direct validation of model outcomes using randomised clinical trials of influenza vaccines, further enhanced by the variation of influenza activity by season or nation, and that the vaccine might not match the prevalent circulating strain. Independent validation of model outcomes to incidence data of national healthcare registries might therefore be more suitable compared with randomised clinical trial data, although the quality of influenza monitoring systems should then be taken into account. Such monitoring systems might not always be available in the studied countries.

We feel that simply mentioning that the model was previously validated does not guarantee a high validation status as new validation efforts are necessary when a model is used in a different setting or with different data (a 'new application'). On the other hand, indicating that some particular input data was not validated due to a lack of suitable data sources was found to be useful for the reader as they can distinguish which parts of the model include high uncertainty [1]. Reporting that the model is validated according to the guidelines of ISPOR–SMDM Task Force on Good Research Practices–Modelling Studies [115], as was reported by Van Bellingen et al. [64], or to the standards of the NICE, as was communicated through author comments, is insufficient. Although such guidelines give guidance on how model validation should be performed [115], these guidelines are general in nature and it is not clear which parts of validation have been performed, how and by whom. Therefore, simply following these guidelines does not guarantee that a model has a high enough validation status for its purpose, nor that model users can assess the validation status themselves.

Although a probabilistic sensitivity analysis is the most important technique to demonstrate the uncertainty around the model outcomes, using cost effectiveness acceptability curves to demonstrate validity of the results [93, 95] does not evaluate how accurately the model simulates what occurs in reality. The model that was applied by the author presenting with ten papers in our review was very similar in nine of these studies; however, no cross-validation of the conceptual model, or additional testing for variations, was described in any of these papers. For instance, the model type and structure were similar in most of the studies but were adapted because of a different target group, perspective or vaccination strategy. Explanations on these deviations or validation of their outcomes were lacking.

A positive example of reporting of model validation in the eyes of the authors was the study of Campbell et al. [77]. In this study, the underlying probability of a first recurrent breast cancer was based on a regression-based survival model that was externally validated against two online prediction tools: Nottingham Prognostic Index and Adjuvant! Online. A separate paper was devoted to the estimation and validation of this model [116]. Moreover, the web appendix contained an extensive report on external validation of the model used to estimate health-related quality-of-life during and after receiving chemotherapy.

Our finding that reporting of validation activities is limited was confirmed by other studies. Carrasco et al. [10] assessed the validation of dynamic transmission models

evaluating the epidemiology of pandemic influenza, and found that 16 % of the compartmental models and 22 % of the agent-based models reported on validation. As validation of model outcomes might be more difficult in cases where pandemic influenza is studied, reporting validation efforts of conceptual models or model inputs might be relatively more valuable. A study by Haji Ali Afzali et al. [9] reviewed the validation performances of 81 studies on therapeutic interventions for cardiovascular diseases, and found that 73 % of the papers reported some form of model performance. The most executed form of validity was cross-validation (55 %), similar to this study. Reporting of face validity (7 %), internal validity (12 %) and external validity (16 %) was low. Although the review process was carried out using a different checklist and was therefore not completely comparable, these results at least support our findings that cross-validation is the most reported validation procedure, and reporting of other validation efforts is rare. Moreover, it demonstrates that limited validation documentation is not restricted to the disease areas of SI and EBC.

A strong point of this study is that we systematically assessed model validation performances. We looked at two disease areas, thereby covering a wider range of models than previous studies. Moreover, compared with previous studies analyzing reporting of validation efforts, we judged validation not only by technique but also by model aspect: conceptual model, input data, computerised model and the model outcomes. This made our findings more specific on which model aspects are generally validated and which are not. A final strong point is that we provided authors an opportunity to comment on whether more validation techniques were performed than those reported in the paper, which gave insight into the difference between performance and reporting of validation.

A limitation of our study was that for most studies we could only evaluate published validation efforts, rather than actual validation efforts undertaken. Thus, these models may have seemed less well-validated to the reader than they actually were. Although the responses of contacted authors of non-reported validation efforts were helpful, the response rate was low. Moreover, authors who performed more validation efforts might have been more aware of the importance of model validation and therefore more eager to respond to our enquiry. On the other hand, the poor response rate might also indicate that authors do not record which model validation tests were performed at the time the analysis was carried out. This was illustrated by three authors responses indicating that they were not able to provide additional information because the analysis was performed many years ago or because their current workload was too high. Next, we included ten papers on SI from the same first author, which might have had an effect on the

total model validation performance within SI. Finally, although our search algorithm was extensive, we still may have missed publications that were not included in PubMed or EmBase. However, the main focus of the current review was to give insight into present practice with regard to reporting of model validation, rather than a complete comprehensive overview of model-based publications in the fields of SI and EBC.

The main implication of our findings is that readers have no structured insight into the validation status of health economic models, which makes it difficult for them to evaluate the credibility of the model and its outcomes. In order to prevent making wrong decisions due to improper model validation status, readers might therefore be forced to perform validity checks themselves, which is highly inefficient. To date, we are not aware of any studies that have looked into the impact of the model's validation status on the correctness of the model outcomes; however, we are aware of a case in The Netherlands in which the validation status of the health economic model can have a decisive effect on the reimbursement status of a drug. In this case, a vaccine against human papillomavirus was rejected for reimbursement because of a lack of model transparency and non-face-valid model inputs and model outcomes [117, 118].

Based on our results, we have several recommendations. First, better attention should be given to validation efforts in scientific publications. A more systematic use of model reporting guidelines might be useful [119, 120], possibly aided by reporting tools specifically aimed at validation efforts, such as AdViSHE [15]. In order to circumvent space limitations, inclusion of a small summary on model validation techniques in the Methods and Result sections would be desirable, in combination with a full model validation report in online appendices. Moreover, validation is important for all published health economic models, even if the model was validated for an earlier purpose. In addition, the choice of validation techniques reported deserves more attention and should be less guided by general publication guidelines, which now seem to imply undue attention for cross-validation only. Finally, it will be interesting to see whether the reporting of validation efforts will improve in time. A similar publication in a few years' time will be very welcome.

5 Conclusions

Although validation is deemed important by many researchers, this is not reflected in the reporting habits of health economic modelling studies. A limited number of studies reported on model validation efforts, although good examples were identified. This lack of transparency might

reduce the credibility of study outcomes and hamper decision makers in interpreting and translating these study outcomes to policy decisions. Since authors have indicated that much more is undertaken than is reported, there is room for quick improvement of the reporting practices if stakeholders such as journals, editors and policy makers start explicitly requesting the results of the validation efforts. Therefore, systematic reporting of validation efforts would be desirable to further enhance decision makers' confidence in health economic models and their outcomes.

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;15:796–803.

2. Cash D, Clark WC, Alcock F, Dickson NM, Eckley N, Jäger J. Salience, credibility, legitimacy and boundaries: linking research, assessment and decision making. RWP02-046. Cambridge, MA: John F. Kennedy School of Government, Harvard University; 2002.

3. van Voorn GA, Vemer P, Hamerlijnck D, Ramos IC, Teunissen GJ, Al M, Feenstra TL. The Missing Stakeholder Group: why patients should be involved in health economic modelling. Appl Health Econ Health Policy. 2016;14:129–33.

4. Dams J, Bornschein B, Reese JP, Conrads-Frank A, Oertel WH, Siebert U, et al. Modelling the cost effectiveness of treatments for Parkinson's disease: a methodological review. Pharmacoeconomics. 2011;29:1025–49.

5. Bolin K. Economic evaluation of smoking-cessation therapies: a critical and systematic review of simulation models. Pharmacoeconomics. 2012;30:551–64.

6. Rochau U, Schwarzer R, Jahn B, Sroczynski G, Kluibenschaedl M, Wolf D, et al. Systematic assessment of decision-analytic models for chronic myeloid leukemia. Appl Health Econ Health Policy. 2014;12:103–15.

7. Leung HW, Chan AL, Leung MS, Lu CL. Systematic review and quality assessment of cost-effectiveness analysis of pharmaceutical therapies for advanced colorectal cancer. Ann Pharmacother. 2013;47:506–18.

8. Kim SY, Goldie SJ. Cost-effectiveness analyses of vaccination programmes: a focused review of modelling approaches. Pharmacoeconomics. 2008;26:191–215.

9. Afzali HHA, Gray J, Karnon J. Model performance evaluation (validation and calibration) in model-based studies of therapeutic interventions for cardiovascular diseases: a review and suggested reporting framework. Appl Health Econ Health Policy. 2013;11:85–93.

10. Carrasco LR, Jit M, Chen MI, Lee VJ, Milne GJ, Cook AR. Trends in parameterization, economics and host behaviour in influenza pandemic modelling: a review and reporting protocol. Emerg Themes Epidemiol. 2013;10:3.

11. Pitman R, Fisman D, Zaric GS, Postma M, Kretzschmar M, Edmunds J, ISPOR-SMDM Modeling Good Research Practices Task Force, et al. Dynamic transmission modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-5. Value Health. 2012;15:828–34.

12. Sargent RG. Validation and verification of simulation models. Proceedings of the 2004 Winter Simulation Conference. 2004. http://www.informs-sim.org/wsc04papers/004.pdf.

13. The World Bank. Country and lending groups. 2015. http://data.worldbank.org/about/country-and-lending-groups. Accessed 1 Oct 2015.

14. Vemer P, van Voom GA, Ramos IC, Krabbe PF, Al MJ, Feenstra TL. Improving model validation in health technology assessment: comments on guidelines of the ISPOR-SMDM Modeling Good Research Practices Task Force. Value Health. 2013;16:1106–7.

15. Vemer P, Ramos IC, van Voorn GA, Al MJ, Feenstra TL. AdViSHE: a validation-assessment tool of health-economic models for decision makers and model users. Pharmacoeconomics. 2016;34:349–61.

16. Chilcott J, Tappenden P, Rawdin A, Johnson M, Kaltenthaler E, Paisley S, et al. Avoiding and identifying errors in health technology assessment models: qualitative study and methodological review. Health Technol Assess. 2010;14:iii–iv, ix–xii, 1–107.

17. Baguelin M, Jit M, Miller E, Edmunds WJ. Health and economic impact of the seasonal influenza vaccination programme in England. Vaccine. 2012;30:3459–62.

18. Beigi RH, Wiringa AE, Bailey RR, Assi TM, Lee BY. Economic value of seasonal and pandemic influenza vaccination during pregnancy. Clin Infect Dis. 2009;49:1784–92.

19. Blommaert A, Bilcke J, Vandendijck Y, Hanquet G, Hens N, Beutels P. Cost-effectiveness of seasonal influenza vaccination in pregnant women, health care workers and persons with underlying illnesses in Belgium. Vaccine. 2014;32:6075–83.

20. Brydak L, Roiz J, Faivre P, Reygrobellet C. Implementing an influenza vaccination programme for adults aged ≥65 years in Poland: a cost-effectiveness analysis. Clin Drug Invest. 2012;32:73–85.

21. Chen S, Liao C. Cost-effectiveness of influenza control measures: a dynamic transmission model-based analysis. Epidemiol Infect. 2013;141:2581–94.

22. Chit A, Roiz J, Briquet B, Greenberg DP. Expected cost effectiveness of high-dose trivalent influenza vaccine in US seniors. Vaccine. 2015;33:734–41.

23. Clements KM, Chancellor J, Nichol K, DeLong K, Thompson D. Cost-effectiveness of a recommendation of universal mass vaccination for seasonal influenza in the United States. Value Health. 2011;14:800–11.

24. Clements KM, Meier G, McGarry LJ, Pruttivarasin N, Misurski DA. Cost-effectiveness analysis of universal influenza vaccination with quadrivalent inactivated vaccine in the United States. Hum Vaccin Immunother. 2014;10:1171–80.

25. Damm O, Eichner M, Rose MA, Knuf M, Wutzler P, Liese JG, et al. Public health impact and cost-effectiveness of intranasal live attenuated influenza vaccination of children in Germany. Eur J Health Econ. 2015;16:471–88.

26. Ding Y, Zangwill KM, Hay JW, Allred NJ, Yeh SH. Cost-benefit analysis of in-hospital influenza vaccination of postpartum women. Obstet Gynecol. 2012;119:306–14.

27. Dugas AF, Coleman S, Gaydos CA, Rothman RE, Frick KD. Cost-utility of rapid polymerase chain reaction-based influenza testing for high-risk emergency department patients. Ann Emerg Med. 2013;62:80–8.

28. Fisman DN, Tuite AR. Estimation of the health impact and cost-effectiveness of influenza vaccination with enhanced effectiveness in Canada. PLoS One. 2011;6:e27420.

29. Giglio N, Gentile A, Lees L, Micone P, Armoni J, Reygrobellet C, et al. Public health and economic benefits of new pediatric influenza vaccination programs in Argentina. Hum Vaccines Immunother. 2012;8:302–12.

30. Jit M, Cromer D, Baguelin M, Stowe J, Andrews N, Miller E. The cost-effectiveness of vaccinating pregnant women against seasonal influenza in England and Wales. Vaccine. 2010;29:115–22.

31. Lavelle TA, Uyeki TM, Prosser LA. Cost-effectiveness of oseltamivir treatment for children with uncomplicated seasonal influenza. J Pediatr. 2012;160(67–73):e6.

32. Lee BY, Bailey RR, Wiringa AE, Assi TM, Beigi RH. Antiviral medications for pregnant women for pandemic and seasonal influenza: an economic computer model. Obstet Gynecol. 2009;114:971–80.

33. Lee BY, Ercius AK, Smith KJ. A predictive model of the economic effects of an influenza vaccine adjuvant for the older adult (age 65 and over) population. Vaccine. 2009;27:2251–7.

34. Lee BY, Tai JHY, Bailey RR, Smith KJ. The timing of influenza vaccination for older adults (65 years and older). Vaccine. 2009;27:7110–5.

35. Lee BY, Tai JH, Bailey RR, Smith KJ, Nowalk AJ. Economics of influenza vaccine administration timing for children. Am J Manag Care. 2010;16(e75):e85.

36. Lee BY, Bailey RR, Wiringa AE, Afriyie A, Wateska AR, Smith KJ, et al. Economics of employer-sponsored workplace vaccination to prevent pandemic and seasonal influenza. Vaccine. 2010;28:5952–9.

37. Lee BY, Tai JH, Bailey RR, McGlone SM, Wiringa AE, Zimmer SM, et al. Economic model for emergency use authorization of intravenous peramivir. Am J Manag Care. 2011;17:e1–9.

38. Lee BY, Bacon KM, Donohue JM, Wiringa AE, Bailey RR, Zimmerman RK. From the patient perspective: the economic value of seasonal and H1N1 influenza vaccination. Vaccine. 2011;29:2149–58.

39. Lee BY, Stalter RM, Bacon KM, Tai JHY, Bailey RR, Zimmer SM, et al. Cost-effectiveness of adjuvanted versus nonadjuvanted influenza vaccine in adult hemodialysis patients. Am J Kidney Dis. 2011;57:724–32.

40. Lee BY, Bartsch SM, Willig AM. The economic value of a quadrivalent versus trivalent influenza vaccine. Vaccine. 2012;30:7443–6.

41. Lee BY, Bartsch SM, Willig AM. Corrigendum to the economic value of a quadrivalent versus trivalent influenza vaccine [Vaccine 30 (2012) 7443–7446]. Vaccine. 2013;31:2477–9.

42. Lee BY, Tai JHY, McGlone SM, Bailey RR, Wateska AR, Zimmer SM, et al. The potential economic value of a 'universal' (multi-year) influenza vaccine. Influenza Other Respir Viruses. 2012;6:167–75.

43. Lin HH, Hsu KL, Ko WW, Yang YC, Chang YW, Yu MC, Chen KT. Cost-effectiveness of influenza immunization in adult cancer patients in Taiwan. Clin Microbiol Infect. 2010;16:663–70.

44. Luce BR, Nichol KL, Belshe RB, Frick KD, Li SX, Boscoe A, et al. Cost-effectiveness of live attenuated influenza vaccine versus inactivated influenza vaccine among children aged 24–59 months in the United States. Vaccine. 2008;26:2841–8.

45. Mamma M, Spandidos DA. Economic evaluation of the vaccination program against seasonal and pandemic A/H1N1 influenza among customs officers in Greece. Health Policy. 2013;109:71–7.

46. Michaelidis CI, Zimmerman RK, Nowalk MP, Smith KJ. Estimating the cost-effectiveness of a national program to eliminate disparities in influenza vaccination rates among elderly minority groups. Vaccine. 2011;29:3525–30.

47. Mogasale V, Barendregt J. Cost-effectiveness of influenza vaccination of people aged 50-64 years in Australia: results are inconclusive. Aust N Z J Public Health. 2011;35:180–6.

48. Myers ER, Misurski DA, Swamy GK. Influence of timing of seasonal influenza vaccination on effectiveness and cost-effectiveness in pregnancy. Am J Obstet Gynecol. 2011;204:S128–40.

49. Newall AT, Scuffham PA, Kelly H, Harsley S, MacIntyre CR. The cost-effectiveness of a universal influenza vaccination program for adults aged 50–64 years in Australia. Vaccine. 2008;26:2142–53.

50. Newall AT, Dehollain JP, Creighton P, Beutels P, Wood JG. Understanding the cost-effectiveness of influenza vaccination in children: methodological choices and seasonal variability. Pharmacoeconomics. 2013;31:693–702.

51. Newall AT, Dehollain JP. The cost-effectiveness of influenza vaccination in elderly Australians: an exploratory analysis of the vaccine efficacy required. Vaccine. 2014;32:1323–5.

52. Nosyk B, Sharif B, Sun H, Cooper C, Anis AH. The cost-effectiveness and value of information of three influenza vaccination dosing strategies for individuals with human immunodeficiency virus. PLoS One. 2011;6:e27059.

53. Patterson BW, Khare RK, Courtney DM, Lee TA, Kyriacou DN. Cost-effectiveness of influenza vaccination of older adults in the ED setting. Am J Emerg Med. 2012;30:1072–9.

54. Pitman RJ, Nagy LD, Sculpher MJ. Cost-effectiveness of childhood influenza vaccination in England and Wales: results from a dynamic transmission model. Vaccine. 2013;31:927–42.

55. Prosser LA, O'Brien MA, Molinari N-M, Hohman KH, Nichol KL, Messonnier ML, et al. Non-traditional settings for influenza vaccination of adults: costs and cost effectiveness. Pharmacoeconomics. 2008;26:163–78.

56. Prosser LA, Meltzer MI, Fiore A, Epperson S, Bridges CB, Hinrichsen V, et al. Effects of adverse events on the projected

population benefits and cost-effectiveness of using live attenuated influenza vaccine in children aged 6 months to 4 years. Arch Pediatr Adolesc Med. 2011;165:112–8.

57. Schmier J, Li S, King JC Jr, Nichol K, Mahadevia PJ. Benefits and costs of immunizing children against influenza at school: an economic analysis based on a large-cluster controlled clinical trial. Health Aff (Millwood). 2008;27:w96–104.

58. Skedgel C, Langley JM, MacDonald NE, Scott J, McNeil S. An incremental economic evaluation of targeted and universal influenza vaccination in pregnant women. Can J Public Health. 2011;102:445–50.

59. Smolen LJ, Klein TM, Bly CA, Ryan KJ. Cost-effectiveness of live attenuated versus inactivated influenza vaccine among children. Am J Pharm Benefits. 2014;6:171–82.

60. Talbird SE, Brogan AJ, Winiarski AP, Sander B. Cost-effectiveness of treating influenzalike illness with oseltamivir in the United States. Am J Health Syst Pharm. 2009;66:469–80.

61. Talbird SE, Brogan AJ, Winiarski AP. Oseltamivir for influenza postexposure prophylaxis: economic evaluation for children aged 1-12 years in the US. Am J Prev Med. 2009;37:381–8.

62. Tarride JE, Burke N, Von Keyserlingk C, O'Reilly D, Xie F, Goeree R. Cost-effectiveness analysis of intranasal live attenuated vaccine (LAIV) versus injectable inactivated influenza vaccine (TIV) for Canadian children and adolescents. Clin Outcomes Res. 2012;4:287–98.

63. Teufel RJ II, Basco WT Jr, Simpson KN. Cost effectiveness of an inpatient influenza immunization assessment and delivery program for children with asthma. J Hosp Med. 2008;3:134–41.

64. Van Bellinghen LA, Meier G, Van Vlaenderen I. The potential cost-effectiveness of quadrivalent versus trivalent influenza vaccine in elderly people and clinical risk groups in the UK: a lifetime multi-cohort model. PLoS One. 2014;9:e98437.

65. Wailoo AJ, Sutton AJ, Cooper NJ, Turner DA, Abrams KR, Brennan A, et al. Cost-effectiveness and value of information analyses of neuraminidase inhibitors for the treatment of influenza. Value Health. 2008;11:160–71.

66. Werker GR, Sharif B, Sun H, Cooper C, Bansback N, Anis AH. Optimal timing of influenza vaccination in patients with human immunodeficiency virus: a Markov cohort model based on serial study participant hemoagglutination inhibition titers. Vaccine. 2014;32:677–84.

67. You JHS, Wong WCW, Ip M, Lee NLS, Ho SC. Cost-effectiveness analysis of influenza and pneumococcal vaccination for Hong Kong elderly in long-term care facilities. J Epidemiol Community Health. 2009;63:906–11.

68. You JHS, Chan ESK, Leung MYK, Ip M, Lee NLS. A cost-effectiveness analysis of "test" versus "treat" patients hospitalized with suspected influenza in Hong Kong. PLoS One. 2012;7:e33123.

69. You J, Ming WK, Chan P. Cost-effectiveness analysis of quadrivalent influenza vaccine versus trivalent influenza vaccine for elderly in Hong Kong. BMC Infect Dis. 2014;14:618.

70. Zhou L, Situ S, Feng Z, Atkins CY, Fung IC, Xu Z, et al. Cost-effectiveness of alternative strategies for annual influenza vaccination among children aged 6 months to 14 years in four provinces in China. PLoS ONE. 2014;9:e87590.

71. Aboutorabi A, Hadian M, Ghaderi H, Salehi M, Ghiasipour M. Cost-effectiveness analysis of trastuzumab in the adjuvant treatment for early breast cancer. Glob J Health Sci. 2014;7:98–106.

72. Alvarado MD, Mohan AJ, Esserman LJ, Park CC, Harrison BL, Howe RJ, et al. Cost-effectiveness analysis of intraoperative radiation therapy for early-stage breast cancer. Ann Surg Oncol. 2013;20:2873–80.

73. Attard CL, Pepper AN, Brown ST, Thompson MF, Thuresson PO, Yunger S, et al. Cost-effectiveness analysis of neoadjuvant pertuzumab and trastuzumab therapy for locally advanced, inflammatory, or early HER2-positive breast cancer in Canada. J Med Econ. 2015;18:173–88.

74. Au HJ, Golmohammadi K, Younis T, Verma S, Chia S, Fassbender K, et al. Cost-effectiveness analysis of adjuvant docetaxel, doxorubicin, and cyclophosphamide (TAC) for node-positive breast cancer: modeling the downstream effects. Breast Cancer Res Treat. 2009;114:579–87.

75. Braun S, Mittendorf T, Menschik T, Greiner W, von der Schulenburg JM. Cost effectiveness of exemestane versus tamoxifen in post-menopausal women with early breast cancer in Germany. Breast Care (Basel). 2009;4:389–96.

76. Buendia JA, Vallejos C, Pichon-Riviere A. An economic evaluation of trastuzumab as adjuvant treatment of early HER2-positive breast cancer patients in Colombia. Biomedica. 2013;33:411–7.

77. Campbell HE, Epstein D, Bloomfield D, Griffin S, Manca A, Yarnold J, et al. The cost-effectiveness of adjuvant chemotherapy for early breast cancer: a comparison of no chemotherapy and first, second, and third generation regimens for patients with differing prognoses. Eur J Cancer. 2011;47:2517–30.

78. Candon D, Healy J, Crown J. Modelling the cost-effectiveness of adjuvant lapatinib for early-stage breast cancer. Acta Oncol. 2014;53:201–8.

79. Chen W, Jiang Z, Shao Z, Sun Q, Shen K. An economic evaluation of adjuvant trastuzumab therapy in HER2-positive early breast cancer. Value Health. 2009;12(Suppl 3):S82–4.

80. Delea TE, El-Ouagari K, Karnon J, Sofrygin O. Cost-effectiveness of letrozole versus tamoxifen as initial adjuvant therapy in postmenopausal women with hormone-receptor positive early breast cancer from a Canadian perspective. Breast Cancer Res Treat. 2008;108:375–87.

81. Delea TE, Taneja C, Sofrygin O, Kaura S, Gnant M. Cost-effectiveness of zoledronic acid plus endocrine therapy in premenopausal women with hormone-responsive early breast cancer. Clin Breast Cancer. 2010;10:267–74.

82. Erman A, Nugent A, Amir E, Coyte PC. Cost-effectiveness analysis of extended adjuvant endocrine therapy in the treatment of post-menopausal women with hormone receptor positive breast cancer. Breast Cancer Res Treat. 2014;145:267–79.

83. Fonseca M, Araujo GT, Saad ED. Cost-effectiveness of anastrozole, in comparison with tamoxifen, in the adjuvant treatment of early breast cancer in Brazil. Rev Assoc Med Bras. 2009;55:410–5.

84. Gold HT, Hayes MK. Cost effectiveness of new breast cancer radiotherapy technologies in diverse populations. Breast Cancer Res Treat. 2012;136:221–9.

85. Hall PS, McCabe C, Stein RC, Cameron D. Economic evaluation of genomic test-directed chemotherapy for early-stage lymph node-positive breast cancer. J Natl Cancer Inst. 2012;104:56–66.

86. Hedden L, O'Reilly S, Lohrisch C, Chia S, Speers C, Kovacic L, et al. Assessing the real-world cost-effectiveness of adjuvant trastuzumab in HER-2/neu positive breast cancer. Oncologist. 2012;17:164–71.

87. Ito K, Elkin E, Blinder V, Keating N, Choudhry N. Cost-effectiveness of full coverage of aromatase inhibitors for Medicare beneficiaries with early breast cancer. Cancer. 2013;119:2494–502.

88. Karnon J, Delea T, Barghout V. Cost utility analysis of early adjuvant letrozole or anastrozole versus tamoxifen in post-menopausal women with early invasive breast cancer: the UK perspective. Eur J Health Econ. 2008;9:171–83.

89. Lee SG, Jee YG, Chung HC, Kim SB, Ro J, Im YH, et al. Cost-effectiveness analysis of adjuvant therapy for node positive breast cancer in Korea: docetaxel, doxorubicin and cyclophosphamide (TAC) versus fluorouracil, doxorubicin and cyclophosphamide (FAC). Breast Cancer Res Treat. 2009;114:589–95.

90. Lidgren M, Jonsson B, Rehnberg C, Willking N, Bergh J. Cost-effectiveness of HER2 testing and 1-year adjuvant trastuzumab therapy for early breast cancer. Ann Oncol. 2008;19:487–95.

91. Lipsitz M, Delea TE, Guo A. Cost effectiveness of letrozole versus anastrozole in postmenopausal women with HR + early-stage breast cancer. Curr Med Res Opin. 2010;26:2315–28.

92. Liubao P, Xiaomin W, Chongqing T, Karnon J, Gannong C, Jianhe L, et al. Cost-effectiveness analysis of adjuvant therapy for operable breast cancer from a Chinese perspective: doxorubicin plus cyclophosphamide versus docetaxel plus cyclophosphamide. Pharmacoeconomics. 2009;27:873–86.

93. Lux MP, Reichelt C, Wallwiener D, Kreienberg R, Jonat W, Gnant M, et al. Results of the Zometa cost-utility model for the german healthcare system based on the results of the ABCSG-12 study. Onkologie. 2010;33:360–8.

94. Lux MP, Wockel A, Benedict A, Buchholz S, Kreif N, Harbeck N, et al. Cost-effectiveness analysis of anastrozole versus tamoxifen in adjuvant therapy for early-stage breast cancer: a health-economic analysis based on the 100-month analysis of the ATAC trial and the German health system. Onkologie. 2010;33:155–66.

95. Lux MP, Reichelt C, Karnon J, Tanzer TD, Radosavac D, Fasching PA, et al. Cost-benefit analysis of endocrine therapy in the adjuvant setting for postmenopausal patients with hormone receptor-positive breast cancer, based on survival data and future prices for generic drugs in the context of the german health care system. Breast Care (Basel). 2011;6:381–9.

96. Martin-Jimenez M, Rodriguez-Lescure A, Ruiz-Borrego M, Segui-Palmer MA, Brosa-Riestra M. Cost-effectiveness analysis of docetaxel (Taxotere) vs. 5-fluorouracil in combined therapy in the initial phases of breast cancer. Clin Transl Oncol. 2009;11:41–7.

97. Neyt M, Huybrechts M, Hulstaert F, Vrijens F, Ramaekers D. Trastuzumab in early stage breast cancer: a cost-effectiveness analysis for Belgium. Health Policy. 2008;87:146–59.

98. Purmonen TT, Pankalainen E, Turunen JH, Asseburg C, Martikainen JA. Short-course adjuvant trastuzumab therapy in early stage breast cancer in Finland: cost-effectiveness and value of information analysis based on the 5-year follow-up results of the FinHer Trial. Acta Oncol. 2011;50:344–52.

99. Sasse AD, Sasse EC. Cost-effectiveness analysis of adjuvant anastrozol in post-menopausal women with breast cancer. Rev Assoc Med Bras. 2009;55:535–40.

100. Sher DJ, Wittenberg E, Suh WW, Taghian AG, Punglia RS. Partial-breast irradiation versus whole-breast irradiation for early-stage breast cancer: a cost-effectiveness analysis. Int J Radiat Oncol Biol Phys. 2009;74:440–6.

101. Shih V, Chan A, Xie F, Ko Y. Economic evaluation of anastrozole versus tamoxifen for early stage breast cancer in Singapore. Value Health Reg Issues. 2012;1:46–53.

102. Shiroiwa T, Fukuda T, Shimozuma K, Ohashi Y, Tsutani K. The model-based cost-effectiveness analysis of 1-year adjuvant trastuzumab treatment: based on 2-year follow-up HERA trial data. Breast Cancer Res Treat. 2008;109:559–66.

103. Skedgel C, Rayson D, Younis T. Is adjuvant trastuzumab a cost-effective therapy for HER-2/neu-positive T1bN0 breast cancer? Ann Oncol. 2013;24:1834–40.

104. Skedgel C, Rayson D, Younis T. The cost-utility of sequential adjuvant trastuzumab in women with Her2/Neu-positive breast cancer: an analysis based on updated results from the HERA Trial. Value Health. 2009;12:641–8.

105. Van Vlaenderen I, Canon JL, Cocquyt V, Jerusalem G, Machiels JP, Neven P, et al. Trastuzumab treatment of early stage breast cancer is cost-effective from the perspective of the Belgian health care authorities. Acta Clin Belg. 2009;64:100–12.

106. Webber-Foster R, Kvizhinadze G, Rivalland G, Blakely T. Cost-effectiveness analysis of docetaxel versus weekly paclitaxel in adjuvant treatment of regional breast cancer in New Zealand. Pharmacoeconomics. 2014;32:707–24.

107. Wolowacz SE, Cameron DA, Tate HC, Bagust A. Docetaxel in combination with doxorubicin and cyclophosphamide as adjuvant treatment for early node-positive breast cancer: a cost-effectiveness and cost-utility analysis. J Clin Oncol. 2008;26:925–33.

108. Yang JJ, Park SK, Cho LY, Han W, Park B, Kim H, et al. Cost-effectiveness analysis of 5 years of postoperative adjuvant tamoxifen therapy for korean women with breast cancer: retrospective cohort study of the Korean Breast Cancer Society database. Clin Ther. 2010;32:1122–38.

109. Younis T, Rayson D, Sellon M, Skedgel C. Adjuvant chemotherapy for breast cancer: a cost-utility analysis of FEC-D vs. FEC 100. Breast Cancer Res Treat. 2008;111:261–7.

110. Younis T, Rayson D, Skedgel C. The cost-utility of adjuvant chemotherapy using docetaxel and cyclophosphamide compared with doxorubicin and cyclophosphamide in breast cancer. Curr Oncol. 2011;18:e288–96.

111. Zendejas B, Moriarty JP, O'Byrne J, Degnim AC, Farley DR, Boughey JC. Cost-effectiveness of contralateral prophylactic mastectomy versus routine surveillance in patients with unilateral breast cancer. J Clin Oncol. 2011;29:2993–3000.

112. Adjuvant! Online. Decision making tools for health care professionals. 2011. https://www.adjuvantonline.com. Accessed 1 Sep 2015.

113. Hall GM. How to write a paper. 5th ed. Chichester, UK: Wiley; 2012. 10.1002/9781118488713.ch1

114. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB, ISPOR-SMDM Modeling Good Research Practices Task Force. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. Value Health. 2012;15:843–50.

115. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, ISPOR Task Force on Good Research Practices-Modeling Studies, et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices-Modeling Studies. Value Health. 2003;6:9–17.

116. Campbell HE, Gray AM, Harris AL, Briggs AH, Taylor MA. Estimation and external validation of a new prognostic model for predicting recurrence-free survival for early breast cancer patients in the UK. Br J Cancer. 2010;103:776–86.

117. Hoomans T, Severens JL, van der Roer N, Delwel GO. Methodological quality of economic evaluations of new pharmaceuticals in The Netherlands. Pharmacoeconomics. 2012;30:219–27.

118. Dutch Healthcare Institute. Herbeoordeling HPV-vaccin Gardasil [in Dutch]. 2009. https://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/publicaties/geneesmiddelbeoordelingen/2009/0903-hpv-vaccin-type-6-11-16-en-18-gardasil/0903-hpv-vaccin-type-6-11-16-en-18-gardasil/HPV-vaccin+(type+6,+11,+16+en+18)+(Gardasil).pdf. Accessed 21 Mar 2016.

119. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. CHEERS task force. consolidated health economic evaluation reporting standards (CHEERS) statement. Int J Technol Assess Health Care. 2013;29:117–22.

120. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics. 2006;24:355–71.