*International Journal of Technology Assessment in Health Care*

**cambridge.org/thc**

## Assessment

The guideline for the submission of HEEs to the Italian Medicines Agency was drafted during the revision of this manuscript and published the 27th of May 2020. It can be downloaded from https://www.aifa.gov.it/web/guest/linea-guida-capitolo-9. All the authors of this article were actively involved in the drafting and publication of the guideline.

**Cite this article:** Carletto A, Zanuzzi M, Sammarco A, Russo P (2020). Quality of health economic evaluations submitted to the Italian Medicines Agency: current state and future actions. *International Journal of Technology Assessment in Health Care* **36**, 560–568. https://doi.org/10.1017/S0266462320000641

Received: 16 December 2019
Revised: 4 August 2020
Accepted: 11 August 2020
First published online: 10 September 2020

**Key words:**
Methodological quality; Quality appraisal; Economic evaluation; Cost-effectiveness analysis; Pharmaceuticals

**Author for correspondence:**
Angelica Carletto,
E-mail: a.carletto@aifa.gov.it

© The Author(s), 2020. Published by Cambridge University Press

CAMBRIDGE UNIVERSITY PRESS

# Quality of health economic evaluations submitted to the Italian Medicines Agency: current state and future actions

Angelica Carletto, Matteo Zanuzzi, Annalisa Sammarco and Pierluigi Russo

Italian Medicines Agency (AIFA), Rome, Italy

**Objectives.** The purpose of this study was to evaluate the current state of health economic evaluations (HEEs) submitted by pharmaceutical companies to the Italian Medicines Agency (AIFA) as part of their pricing and reimbursement (P&R) dossiers, and to explore potential future actions in order to enhance their quality.

**Methods.** All company dossiers submitted from October 2016 to December 2018 were reviewed to select those containing pharmacoeconomic studies. The general characteristics of HEEs were described and their quality assessed based on a checklist adapted from Philips et al. (Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technol Assess. 2004;8: 1–158).

**Results.** Of the 299 dossiers submitted to AIFA, 105 included one or more pharmacoeconomic studies, of which fifty-three were cost-effectiveness analyses. Overall, the compliance of the HEEs with the quality checklist was highly variable: some studies reached high methodological standards whereas others had serious flaws (mean 59.22 percent, range 19.35–90.32 percent). The main weaknesses were the unjustified exclusion of relevant alternatives, poor description and justification of model data and assumptions, and insufficient exploration of uncertainty and study validity. Non-homogeneity across studies was found in study perspectives, discount rates, methods for costing, estimating quality-adjusted life-years and conducting sensitivity analyses.

**Conclusions.** Based on the results of this study, the recommended actions for increasing the quality of HEEs within reimbursement submissions in Italy are twofold: first, to set methodological standards for conducting and reporting HEEs; second, to strengthen the internal assessment process, also through the acquisition of companies' models and re-evaluation of results. These actions will hopefully provide greater contribution to the evidence-based P&R decision making.

Within the economic domain of health technology assessment (HTA), health economic evaluations (HEEs) provide a framework for comparing the relative benefit and cost of different treatments over a long-term time horizon and from a national healthcare system and societal perspective. Competent authorities for pricing and reimbursement (P&R) of medicines and other HTA bodies around the world usually require HEEs from the marketing authorization holder (MAH) to inform their decisions/recommendations. Most of them have issued official guidelines for setting methodological standards and ensuring consistency, relevance, and transparency of submitted studies (1). The Canadian State of Ontario and Australia were pioneers in this field, followed by many other European countries (2–4). In principle, HEEs, possibly integrated with budget impact analyses (BIA), are deemed useful tools for pursuing allocative efficiency in healthcare systems, which is one of the value-pillars of the emerging value(s)-based health care approach in the European public debate (5). Currently, in few countries the results from HEEs are the main driver of decisions (e.g., England and Australia); whereas, more generally, they have a less explicit role within a multi-criteria decision-analysis approach (e.g., Germany, France, and Italy) (6–8).

In Italy, the submission of HEEs within P&R applications is not mandatory, though officially encouraged by national regulations, especially for innovative products and medicines for orphan diseases (CIPE Resolution No. 3 of 1 February 2001) (9). The Italian Medicines Agency (AIFA) is the national authority for both the pricing and the reimbursement of pharmaceuticals. For decision making, it relies on the advice of two expert committees (i.e., the Technical Scientific Committee—CTS and the Pricing and Reimbursement Committee—CPR), which are in turn supported by the AIFA internal staff and the HTA Secretariat. As such, the decisions about the reimbursement criteria and the relative reimbursement price of medicines are strictly related to each other and simultaneously taken according to a stepwise approach: first, the CTS decides on the eligibility for reimbursement coverage; then, the price of reimbursable medicines is negotiated between the CPR and the MAH. However, in case an agreement on price is not reached, the medicine will not get reimbursement by the Italian

National Health Service (NHS). In the Italian system the reimbursement decisions are primarily driven by the therapeutic value of a medicine (also compared to existing alternatives), whereas cost-effectiveness and budget impact estimates may have a role in the price negotiations, together with other factors (10). Despite the cost-effectiveness criterion was first introduced in 1997 (CIPE Resolution No. 5 of 30 January 1997) (11), and confirmed in the CIPE Resolution No. 3 of 1 February 2001 (9), as one of the criteria to be used for determining the prices of reimbursed medicines, it is only recently that HEEs have been formally integrated within the HTA activities performed at the central level by AIFA. In October 2016, the HEE Office was established to review the pharmacoeconomic studies submitted by manufacturers within the HTA process and provide pricing recommendations to the CPR based on cost-effectiveness results and budget impact estimates (12). Nevertheless, in Italy the cost-effectiveness criterion remains non-binding and rather vague to make P&R decisions; in fact, a cost-per-quality-adjusted life-year (QALY) threshold has never been explicitly defined, even due to the poor acceptance of the QALY as a unique measure of value. In 2018, AIFA was invited by the Ministry of Health to issue a position paper on these aspects, but it has not been following up on the request (13). Moreover, differently from many other countries, AIFA has never defined methodological requirements for the submission of HEEs. A proposal for guidelines for the economic evaluation of health interventions was issued in 2009 by the Italian Association of Health Economics (AIES) (14), but it has not officially been transposed into national guidelines. Therefore, in the Italian context, pharmaceutical companies face no explicit incentives to submit HEEs and anyhow, a high degree of discretion is granted in the choice of the study type and other methodological aspects. This could reasonably lead to scepticism toward the results of HEEs, given also the increasing use of electronic models for running pharmacoeconomic analyses, which implies many assumptions and other discretional choices (15;16). However, when designed, analyzed, and interpreted appropriately, HEEs could be important sources of information for decision makers.

In a previous study by Russo (17), the quality of cost-effectiveness analyses (CEAs) submitted to AIFA was considered extremely heterogeneous. Poor transparency and clarity in reporting were the main issues raised by the author. To assess the current state of pharmacoeconomic submissions to AIFA, we reviewed the general characteristics and quality of HEEs through the application of a checklist adapted from Philips et al. (18). Future actions to enhance the quality of HEEs submitted for P&R decisions of medicines in Italy will be proposed.

## Methods

For the purpose of this study, we reviewed all P&R dossiers submitted to AIFA by pharmaceutical companies from October 2016 (i.e., since the establishment of the HEE Office at AIFA) to December 2018. Dossiers were selected if related to (i) new medicinal products (never marketed before), (ii) orphan medicines, and (ii) new therapeutic indications. Each dossier might contain more than one pharmacoeconomic analysis for different therapeutic indications or subgroup populations. The general characteristics and quality of cost-effectiveness studies (if any) were further investigated. Our study focused only on HEEs, whereas the assessment of BIAs was excluded because it was outside the scope of this study.

### General Description of HEEs Submitted to AIFA

In order to collect data systematically, a data extraction sheet identifying all relevant items which needed to be extracted from each study was developed. The list of items was selected from the ISPOR Consolidated Health Economic Evaluation Reporting Standards (CHEERS) Checklist, which includes items to be taken into account when reporting HEEs (19). The investigated variables were grouped into five categories: (a) *general study characteristics*: study population, setting, study perspective, type of comparator, time horizon, and discount rate; (b) *health benefits*: types of health outcomes and source of clinical and utility data; (c) *costs*: types of costs, source of resource, and unit cost data; (d) *model structure* (if any); and (e) *sensitivity analysis*. Any previous publication of the study in peer-reviewed journals was also highlighted.

### Quality of HEEs Submitted to AIFA

The methodological quality of HEEs submitted to AIFA for P&R decisions was systematically assessed following a predefined checklist. Currently, there is no universally accepted instrument, but a number of checklists have been developed over time to guide the critical appraisal of model-based HEEs and the quality of their reporting (20–25). Some authors have also proposed a scoring system, where the final score is indicative of a study's overall quality. However, the use of such system is currently not recommended because no valid and reliable scoring approach has been found (23).

In this study, we applied a customized checklist adapted from Philips et al. (18), which is routinely used by the AIFA HEE Office. The Philips' checklist was selected because it is specifically designed and widely suggested for the assessment of modeling studies (23;26;27). However, it contains a relatively high number of items ($n = 61$) and its full application would be cumbersome in routine practice, given the large number of HEEs and the time constraints of the P&R procedures (28). Thus, a shorter version was created with the aim of making the assessment process more efficient, that is, balancing timeliness and comprehensiveness. Studies not adopting a model approach were excluded because a different checklist should be considered in these cases and results would not be comparable.

The customized checklist is composed of three dimensions, eighteen topics, and thirty-one items. For each item, four mutually exclusive responses were allowed: "yes" if the study complied with the criterion; "no" if the study substantially diverged from the criterion; "unclear" if the dossier provided insufficient information; "NA (not applicable)" if the criterion was not relevant in a particular instance. Because the application of this checklist may imply value judgments of the reviewer, each item was appraised by two independent reviewers (AC and MZ) and disagreements were resolved by consensus or through a third reviewer (AS or PR), where necessary. We calculated the overall compliance rate of each study by dividing the number of "yes" responses by the total number of items on our checklist. Moreover, the compliance rate of the individual items across studies was calculated by dividing the number of studies with "yes" responses by the total number of studies. In this study, we did not mention medicine names and other details that might reveal a specific product and only aggregated results were showed.

https://doi.org/10.1017/S0266462320000641 Published online by Cambridge University Press

**Table 1.** Number and type of pharmacoeconomic analyses in the P&R dossiers submitted by pharmaceutical companies to AIFA

| | Orphan medicines | | New medicinal products | | New therapeutic indications | | Total | |
|---|---|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % | *n* | % |
| All dossiers | 42 | 14.0 | 65 | 21.7 | 192 | 64.2 | 299 | 100 |
| Dossiers with pharmacoeconomic analyses of which: | 21 | 50.0 | 44 | 67.7 | 40 | 20.8 | 105 | 35.1 |
| (a) BIA | 12 | 22.2 | 25 | 46.3 | 17 | 31.5 | 54 | 51.4 |
| (b) CEA | 3 | 21.4 | 4 | 28.6 | 7 | 50.0 | 14 | 13.3 |
| (c) CEA and BIA | 6 | 16.2 | 15 | 40.5 | 16 | 43.2 | 37 | 35.2 |
| Total dossiers with BIAs (a + c) | 18 | 19.8 | 40 | 44.0 | 33 | 36.3 | 91 | 86.7 |
| Total dossiers with CEAs (b + c) | 9 | 17.6 | 19 | 37.3 | 23 | 45.1 | 51 | 32.7 |

## Results

Overall, 299 P&R dossiers were submitted to AIFA for coverage and pricing decisions from October 2016 to December 2018. Among them, only 105 (35.1 percent) included one or more pharmacoeconomic studies, referring to 21 orphan medicines, 44 new medicinal products, and 40 new therapeutic indications. A higher frequency of pharmacoeconomic analyses was observed in P&R dossiers related to new medicinal products (67.7 percent) and orphan medicines (50.0 percent), whereas only a few dossiers included a pharmacoeconomic analysis for new therapeutic indications (20.8 percent). Table 1 shows that dossiers with CEAs, including cost-utility analyses, were less frequently submitted than those with BIAs (32.7 percent, $n = 51/105$ and 86.7 percent, $n = 91/105$, respectively). Moreover, when CEAs were conducted, in the majority of cases, they were accompanied by a BIA (72.5 percent, $n = 37/51$) for all types of P&R applications.

### General Description and Quality Assessment of HEEs Submitted to AIFA

Overall, the P&R dossiers with CEAs ($n = 51$) included fifty-three model-based HEEs (five non-model-based analyses were excluded from our sample). The general characteristics of HEEs are listed in Table 2 and compliance with quality items is summarized in Table 3.

A summary description of all reviewed studies ($n = 53$) is provided below, based on the five categories mentioned in the "Methods" section. Moreover, the main methodological issues emerged from the application of the quality checklist were further discussed.

#### General Study Characteristics

*Target population and subgroups*. Overall, the target population was generally well described and in line with the authorized indication requested for reimbursement ($n = 48/53$, 90.6 percent). Among these, three studies explored specific subgroups in addition. A narrower population compared to the reimbursement request was rarely used, often without a clear justification.

*Perspective*. All but one of the fifty-three reviewed studies was conducted from the perspective of the Italian NHS. In few cases the societal perspective was presented in addition, whereas the regional perspectives were never explored. The perspective was always clearly stated. However, in two studies it was found that the model inputs were inconsistent with the declared perspective (i.e., indirect costs and NHS perspective).

*Comparators*. The alternative options under evaluation were usually well defined. Studies against active comparators were observed in twenty-seven studies and more than one comparator was included in five studies. In the remaining cases (all but one with before-after design) the medicinal product was compared to placebo, standard of care, or best supportive care. The quality checklist highlighted that nearly 36 percent of studies did not include all the most relevant comparators used in clinical practice and very few of them provided justifications for their exclusion.

*Time horizon*. A long-term time horizon was generally adopted in accordance with international methodological guidelines (thirty-nine lifetime horizon, nine greater than 10 yr), with a minority of studies presenting results over a period of less than 10 years ($n = 3$) and two studies with an unknown time horizon.

*Discount rate*. The same discount rate was always used for both future costs and health outcomes ($n = 48$), ranging from 1.5 to 3.5 percent; in the remaining five studies it was not applied or not reported, despite each study having a time horizon of more than 1 year.

#### Health Benefits

*Choice of health outcomes*. All studies expressed the effect measure as QALYs gained or life-years (LYs) gained and forty-one of them reported both measures. The number of avoided events was used in two cases as a further measure of effectiveness in the models.

*Source of efficacy and utility data*. The prevalent source of efficacy data was the main pivotal trial ($n = 40$); in twelve cases, indirect comparison studies (i.e., network meta-analysis) were also carried out to populate the model, especially in the absence of head-to-head clinical trial evidence. Description of methods and limitations of indirect comparisons were often poorly reported. Methods and assumptions used to extrapolate short-term results to final outcomes were well documented and justified in only 45.3 percent of the studies. The transition probabilities across health states were often missing or not clearly described in the dossiers, and only for eight studies they were judged appropriately by the assessors. Health-related quality of life data were generally retrieved from the pivotal trial or from a literature review, whereas observational studies, expert opinion, or assumptions were rarely used. The sources of utility weights were judged to be well

https://doi.org/10.1017/S0266462320000641 Published online by Cambridge University Press

**Table 2.** General description of model-based HEEs submitted to AIFA ($n = 53$)

| | *n* | % |
|---|---|---|
| (a) *General study characteristics* | | |
| Target population and subgroups | | |
| Population in line with the reimbursement request | 45 | 85 |
| Population in line with the reimbursement request, plus subgroups | 3 | 6 |
| Narrower population compared with the reimbursement request | 5 | 9 |
| Study perspective | | |
| National Health Service | 45 | 85 |
| Society | 1 | 2 |
| National Health Service and society | 7 | 13 |
| Comparators | | |
| Active comparator | 27 | 51 |
| Placebo/best supportive care/standard of care | 21 | 40 |
| More than one comparator | 5 | 9 |
| Time horizon | | |
| Lifetime | 39 | 74 |
| >10 yr | 9 | 17 |
| <10 yr | 3 | 6 |
| Not specified | 2 | 4 |
| Discount rate (costs and benefits) | | |
| 1.50% | 1 | 2 |
| 3.00% | 32 | 60 |
| 3.50% | 15 | 28 |
| Not used | 2 | 4 |
| Not specified | 3 | 6 |
| (b) *Health benefits* | | |
| Choice of health outcomes[a] | | |
| QALY and LY | 41 | 77 |
| QALY | 10 | 19 |
| LY | 2 | 4 |
| Avoided event | 2 | 4 |
| Sources of effectiveness data[a] | | |
| Direct comparison studies—clinical trials | 40 | 75 |
| Direct comparison studies—meta-analyses | 1 | 2 |
| Indirect comparison studies (i.e., network meta-analysis) | 12 | 23 |
| Observational studies | 2 | 4 |
| Other (e.g., expert panel, assumptions) | 10 | 19 |
| Sources of utility data[a] | | |
| Clinical trial | 29 | 55 |
| Literature | 25 | 47 |
| Other (e.g., expert panel, assumptions) | 4 | 8 |
| Sources not specified | 6 | 11 |
| No utility data | 2 | 4 |
| (c) *Costs* | | |
| Sources of resource use data[a] | | |
| Clinical trial | 10 | 19 |
| Medical records | 4 | 8 |
| Literature | 33 | 62 |
| Other (e.g., expert panel, assumptions) | 32 | 60 |
| Not specified | 14 | 26 |
| Types of costs[a] | | |
| Pharmaceuticals | 53 | 100 |
| Other direct healthcare costs | 53 | 100 |
| Direct non-healthcare costs | 4 | 8 |
| Indirect costs | 7 | 13 |
| (d) *Model structure* | | |
| Decision tree | 1 | 2 |
| Markov model | 23 | 43 |
| Partitioned survival model | 20 | 38 |
| Other | 6 | 11 |
| Not specified | 3 | 6 |
| (e) *Sensitivity analysis* | | |
| One-way | 27 | 51 |
| Multivariate | 0 | 0 |
| Probabilistic | 36 | 68 |
| Scenario | 9 | 17 |
| No sensitivity analysis | 9 | 17 |
| Publication on peer-reviewed journals | | |
| Yes | 3 | 6 |
| No | 50 | 94 |

[a]More than one option could be selected for each study.
% = Proportions were calculated by dividing the number of relevant submissions on the individual items by the total number of submissions.

referenced in over two third of the studies (67.9 percent). However, no studies in our sample made explicit use of Italian-specific preference weights, such as the published Italian value set of the EQ-5D health states questionnaire (29;30).

## Costs

*Sources of resource use and cost data*. The typical approach to costing was to retrieve resource use data from multiple sources in the literature and then assign national tariffs to each item. Generally, disease-specific costs were attached to model health states, regardless of the type of intervention, whereas therapy-specific costs (acquisition and administration costs) were differently assigned to each model arm. No trial-based economic evaluations were found in our sample. Expert panels or assumptions made by the authors were commonly adopted to generate

https://doi.org/10.1017/S0266462320000641 Published online by Cambridge University Press

https://doi.org/10.1017/S0266462320000641 Published online by Cambridge University Press



**Table 3.** Evaluation of the HEEs submitted to the AIFA based on a checklist adapted from Philips et al. (18)

| | | | Yes | | UC | | NA | | *N* | |
|---|---|---|---|---|---|---|---|---|---|---|
| Dimensions | Topics | Items | *n* | % | *n* | % | *n* | % | *n* | % |
| Structure | Statement of decision problem/objective | Is there a clear statement of the decision problem? | 53 | 100.0 | 0 | 0.0 | 0 | 0.0 | 0 | 0.0 |
| | | Is the objective of the model specified and consistent with the stated decision problem? | 49 | 92.5 | 0 | 0.0 | 0 | 0.0 | 4 | 7.5 |
| | Statement of scope/ perspective | Is the perspective of the model stated clearly? | 53 | 100.0 | 0 | 0.0 | 0 | 0.0 | 0 | 0.0 |
| | | Are the model inputs consistent with the stated perspective? | 51 | 96.2 | 0 | 0.0 | 0 | 0.0 | 2 | 3.8 |
| | Rationale for structure | Is the structure of the model consistent with a coherent theory of the health condition under evaluation? | 49 | 92.5 | 2 | 3.8 | 1 | 1.9 | 1 | 1.9 |
| | | Are the sources of the data used to develop the structure of the model specified? | 19 | 35.8 | 0 | 0.0 | 0 | 0.0 | 34 | 64.2 |
| | Structural assumptions | Are the structural assumptions reasonable given the overall objective, perspective and scope of the model? | 48 | 90.6 | 4 | 7.5 | 1 | 1.9 | 0 | 0.0 |
| | Strategies/comparators | Is there a clear definition of the options under evaluation? | 51 | 96.2 | 0 | 0.0 | 0 | 0.0 | 2 | 3.8 |
| | | Have all feasible and practical options been evaluated? | 34 | 64.2 | 0 | 0.0 | 0 | 0.0 | 19 | 35.8 |
| | | Is there justification for the exclusion of feasible options? | 2 | 3.8 | 0 | 0.0 | 34 | 64.2 | 17 | 32.1 |
| | Model type | Is the chosen model type appropriate given the decision problem and specified casual relationships within the model? | 48 | 90.6 | 4 | 7.5 | 0 | 0.0 | 1 | 1.9 |
| | Time horizon | Is the time horizon of the model sufficient to reflect all important differences between the options? | 51 | 96.2 | 2 | 3.8 | 0 | 0.0 | 0 | 0.0 |
| | Disease states/pathways | Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions? | 49 | 92.5 | 2 | 3.8 | 1 | 1.9 | 1 | 1.9 |
| | Cycle length | Is the cycle length defined and justified in terms of the natural history of disease? | 31 | 58.5 | 10 | 18.9 | 1 | 1.9 | 11 | 20.8 |
| Data | Data identification | Are the data identification methods transparent and appropriate given the objectives of the model? | 9 | 17.0 | 0 | 0.0 | 0 | 0.0 | 44 | 83.0 |
| | | Where choices have been made between data sources, are these justified appropriately? | 10 | 18.9 | 0 | 0.0 | 3 | 5.7 | 40 | 75.5 |
| | | Where expert opinion has been used are the methods described and justified? | 6 | 11.3 | 2 | 3.8 | 27 | 50.9 | 18 | 34.0 |
| | Baseline data | Is the choice of baseline data described and justified? | 48 | 90.6 | 0 | 0.0 | 0 | 0.0 | 5 | 9.4 |
| | | Are transition probabilities calculated appropriately? | 8 | 15.1 | 23 | 43.4 | 21 | 39.6 | 1 | 1.9 |
| | | Has a half-cycle correction been applied to both costs and outcomes? | 7 | 13.2 | 44 | 83.0 | 0 | 0.0 | 2 | 3.8 |
| | | If not, has the omission been justified? | 0 | 0.0 | 0 | 0.0 | 7 | 13.2 | 46 | 86.8 |
| | Treatment effects | Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified? | 24 | 45.3 | 1 | 1.9 | 0 | 0.0 | 28 | 52.8 |
| | Costs | Are the costs incorporated into the model justified? | 47 | 88.7 | 0 | 0.0 | 0 | 0.0 | 6 | 11.3 |
| | | Has the source for all costs been described? | 39 | 73.6 | 0 | 0.0 | 0 | 0.0 | 14 | 26.4 |
| | | Have discount rates been described and justified given the target decision maker? | 44 | 83.0 | 0 | 0.0 | 0 | 0.0 | 9 | 17.0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Quality of life weights (utilities) | Are the utilities incorporated into the model appropriate? | 38 | 71.7 | 10 | 18.9 | 2 | 3.8 | 3 | 5.7 |
| | Is the source of utility weights well referenced? | 36 | 67.9 | 9 | 17.0 | 2 | 3.8 | 6 | 11.3 |
| Data incorporation | If data have been incorporated as distributions, has the choice of distributions for each parameter been described and justified? | 17 | 32.1 | 0 | 0.0 | 17 | 32.1 | 19 | 35.8 |
| Heterogeneity | Has heterogeneity been dealt with by running the model separately for different subgroups? | 7 | 13.2 | 0 | 0.0 | 0 | 0.0% | 46 | 86.8 |
| Parameter | Are the methods of assessment of parameter uncertainty appropriate? If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified? | 31 | 58.5 | 0 | 0.0 | 0 | 0.0% | 22 | 41.5 |
| Consistency | External consistency | Have the results been compared with those of previous models and any differences in results explained? | 1 | 1.9 | 0 | 0.0 | 0 | 0.0% | 52 | 98.1 |

resource use data. On the contrary, data from clinical trials had a limited use in our sample and were generally confined to the calculation of costs of treatment, adverse events, and hospitalizations. In fourteen out of fifty-three studies the source of resource use data was missing for all or for certain items.

*Types of costs*. All studies included medicine acquisition and other direct healthcare costs, whereas only a few studies considered direct non-healthcare costs ($n = 4$) and indirect costs ($n = 7$), either in the base-case analysis or in the sensitivity analysis.

### Model Structure

Modeling approaches consisted of Markov models ($n = 23$), partition survival models ($n = 20$), decision trees ($n = 1$), and other decision models (i.e., hybrid models; $n = 6$); in three cases the type of model was not specified. Even though most studies thoroughly reported on the type of model and description of health states, only a small number gave reasons for the choice of the model structure and data used to develop it ($n = 19$). Information on the cycle length of health-state transition models was missing or unjustified in over half of the studies (58.5 percent). Half-cycle corrections were explicitly applied only in seven studies.

### Sensitivity Analyses

Sensitivity analyses were performed for forty-four of the fifty-three reviewed studies. Most were conducted in a probabilistic way ($n = 36$) and results presented using scatter plots on the cost-effectiveness plane or a cost-effectiveness acceptability curve. A one-way sensitivity analysis was performed in twenty-seven studies, either as a unique analysis or in addition to a probabilistic sensitivity analysis. However, ranges of point estimates and distributions assigned to each parameter in the model were frequently not reported or not justified. The exploration of heterogeneity in sensitivity analyses was properly performed in seven studies.

*Publication on peer-reviewed journals*. Ninety-four percent of submitted studies was not published at the time of evaluation and decision making, hence these studies were not subjected to the scrutiny of peer-reviewers.

Overall, the quality of reviewed studies was highly variable, with a compliance rate ranging between 19.35 and 90.32 percent (mean 59.22 percent). Similarly, the mean compliance rate of each item was on average 58.43 percent (range 0–100 percent). Out of the thirty-one items on our checklist, sixteen showed compliance rates above 67 percent, five had a compliance rate >33 and <67 percent, and the remaining ten items revealed critical flaws with a compliance rate of less than 33 percent. The main weaknesses emerged from the quality assessment were the following: (i) unjustified exclusion of some relevant alternatives; (ii) lack of transparency or justification of data sources and assumptions; (iii) poor reporting of transition probabilities; (iv) half-cycle correction not used or not reported; (v) distribution of parameters often omitted or not justified; (vi) heterogeneity issues not adequately addressed; and (vii) study validity not checked or not reported.

## Discussion and Conclusions

In the current status, where the submission of HEEs within medicines P&R applications is voluntary and no methodological standards have been established by AIFA, the availability of health economics evidence for informing price negotiations in Italy is fairly limited. As a general observation, the number of P&R

https://doi.org/10.1017/S0266462320000641 Published online by Cambridge University Press

dossiers with CEAs has not increased over time compared with the frequency reported by Russo (17) (about 35 percent in both studies). Moreover, full economic evaluations were less used by pharmaceutical companies to support their applications compared to BIAs. These elements suggest a low level of acceptance and use of HEEs by both private and public parties involved in the price negotiation process in Italy.

The critical assessment of all fifty-three HEEs submitted to AIFA, between October 2016 and December 2018, revealed that the quality level is widely variable. On one hand, numerous HEEs met high methodological standards, both in the rigor of the analysis and the quality of reporting. To a certain extent, this finding corresponds to our expectations, given the great effort dedicated by the international community in establishing good-practice modeling guidelines, as well as the industry's widespread practice of developing very accurate global cost-effectiveness models subsequently adapted to local contexts (23;31;32). On the other hand, none of the reviewed studies performed impeccably with respect to all checklist items and some critical issues were identified. The most frequent methodological flaws were the unjustified exclusion of relevant alternatives, the insufficient description and justification of model inputs and assumptions, and the poor exploration of uncertainty and study validity. Moreover, non-homogeneity across studies was found in the choice of the study perspectives, discount rates, methods for costing, estimating QALYs, and conducting sensitivity analyses. In many cases, some relevant information was unclear/not available in the reporting of HEEs and this did not allow a proper critical assessment of models throughout the checklist.

The limitations of this study should be acknowledged. First, the appraisal of quality inevitably involves subjectivity. The use of a checklist was helpful to identify the main critical issues, but it consisted of general questions whose interpretation often relied on a value judgment by the reviewer (e.g., when methods could be considered "justified"). Another limitation is that the models were assessed against a customized checklist consisting of thirty-one selected items from the Philips's checklist. The methodological approach adopted reflects the concern of maintaining a trade-off between scientific and operational needs in routine practice at AIFA. In addition, the calculation of the compliance rates of each study against the checklist items implicitly involved that each item weighted equally, even though flaws observed in some crucial items might affect the overall quality of a study more than others. For these reasons, the results in terms of compliance rates should be used with caution and always interpreted along with a qualitative descriptive analysis of all criticalities.

Overall, high variability of quality in HEEs submitted for reimbursement decisions was also detected by means of a checklist in other studies (33;34). Unfortunately, a direct comparison with our study was not possible given the methodological differences, mainly in the type of checklist adopted. A limited comparison with the results obtained by Ramsberg et al. (33), regarding the quality of reimbursement submissions in Sweden, showed a similar quality score of approximately 60 percent (range 24–83 percent). Moreover, common shortcomings were related to the choice of the relevant comparators, inconsistencies between types of costs and study perspective, insufficient exploration of uncertainty and study validity. Another study published by Yim et al. (34) about HEEs submitted for reimbursement decisions in South Korea reported a broadly higher quality level to our study, having a compliance rate of 70.9 percent (range 35.0–100 percent) according to their specific checklist, which however does not overlap much with ours.

Despite all the aforementioned limitations, relevant conclusions could be drawn from our study. First, the presence of variation in methodological approaches across studies and poor reporting of relevant information could be overcome by the publication of AIFA guidelines. Second, the checklist allowed the identification of items that should be carefully addressed by future guidelines and better fulfilled by the applicants in order to increase the quality of HEEs. In general, the majority of reimbursement authorities and HTA agencies have produced HEE guidelines to clarify their own position with regard to aspects of methodology (e.g., preferred perspective, discount rate, methods for valuing health outcomes, etc.), which often differ from each other due to different national contexts and cultural values (35–37). However, it is worth noting that the experiences gained from other countries revealed that having such guidelines, although it was helpful in setting a minimum of standards, may not be sufficient to guarantee higher-quality evidence. Wherever they were delivered, a variety of quality issues and weak compliance with the established requirements was found in the HEEs submitted to several reimbursement or HTA bodies in Canada (British Columbia), Australia, the Netherlands, Belgium, and France (38–42). For these reasons, the critical assessment of HEEs by internal reviewers, which goes beyond the simple application of a checklist, remains a crucial step in many countries (e.g., Australia and England) (39;43), with different levels of accuracy based on a context-specific trade-off between scientific rigor and available resources (44). Whenever possible, companies' models should be requested and analyzed using different assumptions or inputs, because industry-sponsored studies are likely to overestimate cost-effectiveness (44;45).

In conclusion, our study underscored that the quality of pharmacoeconomic studies submitted to AIFA within the reimbursement dossiers has still room for improvement. A much greater effort would be needed by both parties involved in the P&R process in Italy: AIFA should clarify its position with regard to the economic evidence required for decision making, whereas industry should strive to provide more accurate analyses and increase the transparency of their models. Considering the results of this study and the knowledge learned from other contexts abroad, we believe that the issue of AIFA guidelines for reimbursement submissions, together with the strengthening of the internal assessment process, could contribute to enhance the quality of manufacturers' HEEs, as well as the reliability of their results. Overall, both actions will hopefully represent a significant step toward a greater use of HEEs for evidence-based decision making on reimbursement and prices of medicines in Italy.

## Disclaimer

The views expressed in this work are personal and should not be understood or quoted as being made on behalf of or reflecting the position of the Italian Medicines Agency or of one of their committees or working parties. All authors are employed by the Italian Medicines Agency (Health Economic Evaluation Office) and bound by the obligation of professional secrecy according to the AIFA's Code of Conduct. All content of P&R dossiers submitted to AIFA by pharmaceutical companies is regarded as confidential information. Therefore, medicine names and other details that might reveal a specific product were not mentioned in the article and only aggregated results were showed.

https://doi.org/10.1017/S0266462320000641 Published online by Cambridge University Press

**Acknowledgments.** The authors thank Marta Toma for providing language editing and proofreading.

**Financial Support.** This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

**Conflict of Interest.** The authors have nothing to disclose.

## References

1. **International Society for Pharmacoeconomics and Outcomes Research (ISPOR) [Internet]. c2019** [cited 2019 Jun 27]. Available from: https://tools.ispor.org/peguidelines/
2. **Ontario, Ministry of Health**. Guidelines for preparation of economic analysis to be included in submission to drug programs branch for listing in the Ontario Benefit Formulary/Comparative Drug Index, Toronto: Ontario Ministry of Health; 1991.
3. **Department of Health, Housing and Community Services**. Guidelines for the pharmaceutical industry on preparation of submissions to the Pharmaceutical Benefits Advisory Committee, Canberra: Commonwealth of Australia; 1992.
4. **European Network for Health Technology Assessment (EUnetHTA)** [Internet] Methods for health economic evaluations—A guideline based on current practices in Europe. Methodological Guideline: EUnetHTA; 2015 [cited 2020 April 27] Available from: https://www.eunethta.eu/wp-content/uploads/2018/03/Methods_for_health_economic_evaluations.pdf
5. **Expert Panel on effective ways of investing in Health (EXPH)** [Internet] Defining value in "value-based healthcare". Luxembourg: Publications Office of the European Union, 2019 [cited 2019 Jun 27]. Available from: https://ec.europa.eu/health/expert_panel/sites/expertpanel/files/docsdir/024_defining-value-vbhc_en.pdf
6. **Ngo P**. The influence of cost-effectiveness evaluations on reimbursement in Australia: A retrospective study of decisions made by the pharmaceutical benefits advisory committee. *Pharm Med*. 2014;**28**:187–93.
7. **Dakin H, Devlin N, Feng Y, Rice N, O'Neill P, Parkin D**. The influence of cost-effectiveness and other factors on NICE decisions. *Health Econ*. 2015;**24**:1256–71.
8. **Jommi C, Armeni P, Costa F, Bertolani A, Otto M**. Implementation of value-based pricing for medicines. *Clin Ther*. 2020;**42**:15–24.
9. **Comitato Interministeriale per la Programmazione Economica (CIPE)**. [Internet] Delibera CIPE No. 3 of 1 February 2001. Individuazione dei criteri per la contrattazione del prezzo dei farmaci. (GU n. 73 del 28 marzo 2001). c2001 [cited 2020 Apr 30]. Available from: http://ricerca-delibere.programmazioneeconomica.gov.it/3-01-febbraio-2001/
10. **Villa F, Tutone M, Altamura G, Antignani S, Cangini A, Fortino I et al.**. Determinants of price negotiations for new drugs. The experience of the Italian medicines agency. *Health Policy*. 2019; **123**: 595–600.
11. **Comitato Interministeriale per la Programmazione Economica (CIPE)**. [Internet] Delibera CIPE No. 5 of 30 January 1997. Individuazione dei criteri per la contrattazione del prezzo dei farmaci innovativi. c1997 [cited 2020 Apr 30]. Available from: http://ricerca-delibere.programmazioneeconomica.gov.it/5-30-gennaio-1997/
12. **Agenzia Italiana del Farmaco (AIFA)** [Internet] Nuovo regolamento di organizzazione, del funzionamento e dell'ordinamento del personale e della nuova dotazione organica. (16A04575) (GU Serie Generale n. 140 del 17-06-2016) [cited 2020 April 30] Available from: https://www.aifa.gov.it/documents/20142/629739/Regolamento_AIFA_2016_3.pdf/18550ff8-1e5e-0261-98a9-1eb404100003
13. **Ministero della Salute** [Internet] Documento in materia di Governance Farmaceutica [cited 2019 Dec 05]. Available from: http://www.salute.gov.it/imgs/C_17_notizie_3567_listaFile_itemName_0_file.pdf
14. **Italian Health Economics Association (Associazione Italiana di Economia Sanitaria—AIES)**. Proposta di linee guida per la valutazione economica degli interventi sanitari in Italia. [Italian guidelines proposal on how to conduct economic evaluation studies of health programs] (only available in Italian). *Pharmacoeconomics*. 2009;**11**:83–93.
15. **Cornago D, Li Bassi L, De Compadri P, Garattini L**. Pharmacoeconomic studies in Italy: A critical review of the literature. *Eur J Health Econ*. 2007;**8**:89–95.
16. **Hoffmann C, von der Schulenburg JMG**. The influence of economic evaluation studies on decision making: A European survey. *Health Policy*. 2000;**52**:179–92.
17. **Russo P**. La valutazione farmacoeconomica nel contesto regolatorio italiano. Analisi quali-quantitativa dei dossier di richiesta del prezzo e della rimborsabilità. *Pharmacoeconomics*. 2008;**10**:59.
18. **Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R et al.** Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess*. 2004;**8**:iii-iv, ix-xi, 1–158.
19. **Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D et al.**; ISPOR health economic evaluation publication guidelines-CHEERS good reporting practices task force. Consolidated health economic evaluation reporting standards (CHEERS)—explanation and elaboration: A report of the ISPOR health economic evaluation publication guidelines good reporting practices task force. *Value Health*. 2013;**16**:231–50.
20. **Sculpher M, Fenwick E, Claxton K**. Assessing quality in decision analytic cost-effectiveness models: A suggested framework and example of application. *Pharmacoeconomics*. 2000;**17**:461–77.
21. **Soto J**. Health economic evaluations using decision analytic modeling. Principles and practices—utilization of a checklist to their development and appraisal. *Int J Technol Assess Health Care* Winter. 2002;**18**:94–111.
22. **Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C et al.**; ISPOR task force on good research practices—modeling studies. 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.
23. **Walker DG, Wilson RF, Sharma R, Bridges J, Niessen L, Bass EB et al. Best Practices for Conducting Economic Evaluations in Health Care: A Systematic Review of Quality Assessment Tools [Internet].** Rockville (MD): Agency for Healthcare Research and Quality (US); 2012 Oct. Report No.: 12(13)-EHC132-EF.
24. **Caro J, Eddy DM, Kan H, Kaltz C, Patel B, Eldessouki R et al.**; ISPOR-AMCP-NPC modeling CER task forces. Questionnaire to assess relevance and credibility of modeling studies for informing health care decision making: An ISPOR-AMCP-NPC good practice task force report. *Value Health*. 2014;**17**:174–82. Erratum in: *Value Health*. 2016;19(1):121.
25. **Adarkwah CC, van Gils PF, Hiligsmann M, Evers SM**. Risk of bias in model-based economic evaluations: The ECOBIAS checklist. *Expert Rev Pharmacoecon Outcomes Res*. 2016;**16**:513–23.
26. **Shemilt I, Mugford M, Byford S, Drummond M, Eisenstein E, Knapp M et al.** Incorporating economics evidence (Chap. 15). In: Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions: Cochrane book series. Wiley; 2008. p. 449.
27. **National Institute for Clinical Excellence (NICE)**. [Internet] Developing NICE guidelines: the manual; 2014; [cited 2020 Apr 29]. Available from: http://www.nice.org.uk/article/PMG20/chapter/7-Incorporating-economic-evaluation.
28. **Wijnen B, Van Mastrigt G, Redekop W, Majoie H, De Kinderen R, Evers S**. How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: Data extraction, risk of bias, and transferability (part 3/3). *Expert Rev Pharmacoecon Outcomes Res*. 2016;**16**:723–32.
29. **Scalone L, Cortesi PA, Ciampichini R, Belisari A, D'Angiolella LS, Cesana G et al.** Italian population-based values of EQ-5D health states. *Value Health*. 2013;**16**:814–22.
30. **Scalone L, Cortesi PA, Ciampichini R, Cesana G, Mantovani LG**. Health related quality of life norm data of the Italian general population: Results using the EQ-5D-3L and EQ-5D-5L instruments. *Epidemiol, Biostatist Public Health*. 2015;**12**:e11457 1–15.
31. **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. *Med Decis Making*. 2012;**32**:667–77.

https://doi.org/10.1017/S0266462320000641 Published online by Cambridge University Press

32. **Mullins CD, Onwudiwe NC, de Araújo GTB, Chen W, Xuan J, Tichopád A et al.**. Guidance document: Global pharmacoeconomic model adaption strategies. *Value Health Reg Issues*. 2014;**5**:7–13.
33. **Ramsberg J, Odeberg S, Engström A, Lundin D**. Examining the quality of health economic analyses submitted to the pharmaceutical benefits board in Sweden. *Eur J Health Econo*. 2004;**49**:351–56.
34. **Yim EY, Lim SH, Oh MJ, Park HK, Gong JR, Park SE et al.**. (2012). Assessment of pharmacoeconomic evaluations submitted for reimbursement in Korea. *Value Health*. 2012;**15**:S104–S110.
35. **European Network for Health Technology Assessment (EUnetHTA) Joint Action**. Is there a European view on health economic evaluations? Results from a synopsis of methodological guidelines used in the EUnetHTA partner countries. *Pharmacoeconomics*. 2016;**34**:59–76.
36. **Hjelmgren J, Berggren F, Andersson F**. Health economic guidelines—similarities, differences and some implications. *Value Health*. 2001;**4**:225–50.
37. **Mauskopf J, Walter J, Birt J, Bowman L, Copley-Merriman C, Drummond M**. Differences among formulary submission guidelines: Implications for health technology assessment. *Int J Technol Assess Health Care*. 2011;**27**:261–70.
38. **Anis AH, Gagnon Y**. Using economic evaluations to make formulary coverage decisions. So much for guidelines. *Pharmacoeconomics*. 2000;**18**:55–62.
39. **Hill SR, Mitchell AS, Henry DA**. Problems with the interpretation of pharmacoeconomic analyses: A review of submissions to the Australian pharmaceutical benefits scheme. *JAMA* 2000;**283**:2116–21.
40. **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.
41. **Simoens S**. Assessment of methodological quality of economic evaluations in Belgian drug reimbursement applications. *PLoS ONE*. 2013;**8**:e85411 1–8.
42. **Toumi M, Motrunich A, Millier A, Rémuzat C, Chouaid C, Falissard B et al.** Analysis of health economics assessment reports for pharmaceuticals in France—Understanding the underlying philosophy of CEESP assessment. *J Mark Access Health Policy*. 2017;**5**:1344088.
43. **Barbieri M, Hawkins N, Sculpher M**. Who does the numbers? The role of third-party technology assessment to inform health systems' decision-making about the funding of health technologies. *Value Health*. 2009;**12**:193–201.
44. **Johannesen KM, Claxton K, Sculpher MJ, Wailoo AJ**. How to design the cost-effectiveness appraisal process of new healthcare technologies to maximise population health: A conceptual framework. *Health Econ*. 2018;**27**:e41–54.
45. **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.


https://doi.org/10.1017/S0266462320000641 Published online by Cambridge University Press