What do national pharmacoeconomic guidelines recommend regarding the statistical analysis of trial-based economic evaluations?

Johanna M van Dongen, Mohamed El Alili, Anita N Varga, Alejandra E Guevara Morel, Angela Jornada Ben, Mojdeh Khorrami, Maurits W van Tulder & Judith E Bosmans


To link to this article: https://doi.org/10.1080/14737167.2020.1694410
What do national pharmacoeconomic guidelines recommend regarding the statistical analysis of trial-based economic evaluations?

Johanna M van Dongen, Mohamed El Alili, Anita N Varga, Alejandra E Guevara Morel, Angela Jornada Ben, Mojdeh Khorrami, Maurits W van Tulder, and Judith E Bosmans

ABSTRACT

Introduction: The statistical quality of many trial-based economic evaluations is poor. When conducting trial-based economic evaluations, researchers often turn to national pharmacoeconomic guidelines for guidance. Therefore, this study reviewed which recommendations are currently given by national pharmacoeconomic guidelines on the statistical analysis of trial-based economic evaluations.

Areas covered: 40 national pharmacoeconomic guidelines were identified. Data were extracted on the guidelines' recommendations on how to deal with baseline imbalances, skewed costs, correlated costs and effects, clustering of data, longitudinal data, and missing data in trial-based economic evaluations. Four guidelines (10%) were found to include recommendations on how to deal with baseline imbalances, five (13%) on how to deal with skewed costs, and seven (18%) on how to deal with missing data. Recommendations were very general in nature and recommendations on dealing with correlated costs and effects, clustering of data, and longitudinal data were lacking.

Expert opinion: Current national pharmacoeconomic guidelines provide little to no guidance on how to deal with the statistical challenges to trial-based economic evaluations. Since the use of suboptimal statistical methods may lead to biased results, and, therefore, possibly to a waste of scarce resources, national agencies are advised to include more statistical guidance in their pharmacoeconomic guidelines.

1. Introduction

As resources are scarce, decision-makers need to decide which health-care interventions to reimburse with the available resources while at the same time maximizing health benefits [1]. Economic evaluations seek to inform decision-makers by assessing whether the additional health effects of an intervention justify its additional costs as compared to an alternative intervention [1]. Many grant organizations demand that economic evaluations are conducted alongside clinical trials and the results of such so-called trial-based economic evaluations are increasingly being used to inform health-care decision-making [2–4]. In various countries, cost-effectiveness is even established as a formal decision criterion for the pricing and/or reimbursement of pharmaceuticals and other health-care technologies [5–9].

As a consequence, national agencies, such as the English National Institute for Health and Care Excellence (NICE) and the Dutch Health Care Institute, developed guidelines containing principles and recommendations for the design, execution, and reporting of economic evaluations [5–8]. When conducting trial-based economic evaluations, researchers frequently turn to such pharmacoeconomic guidelines for guidance. Also, in an increasing number of countries, researchers and health technology manufacturers are obliged to follow pharmacoeconomic guidelines when preparing an economic evaluation as part of the application process for reimbursement of a new drug and/or health-care technology [5–8]. Research suggests that the introduction of national pharmacoeconomic guidelines has led to an improved reporting quality of trial-based economic evaluations in some areas (e.g. obstetrics) [10–13], whereas this effect was not observed in others (e.g. gynecology) [11,14,15].

Despite the existence of many national pharmacoeconomic guidelines, the statistical quality of trial-based economic evaluations is typically poor [11,16–19]. Baseline imbalances, the skewed nature of cost data, the correlation between costs and effects, the clustering of data, and the longitudinal nature of cost and/or data are often unaccounted for in trial-based economic evaluations [11,16–19]. On top of that, missing data are frequently handled using ‘naïve’ imputation methods, such as mean imputation and last observation carried forward [11,16–19]. This is worrisome, because the use of suboptimal statistical methods in trial-based economic evaluations reduces their credibility and usefulness to decision-makers [20], and may even lead to biased health-care decisions and, consequently, a possible waste of already scarce resources.

CONTACT Johanna M van Dongen j.m.van.dongen@vu.nl Department of Health Sciences, Faculty of Science, Amsterdam Movement Sciences research institute, Vrije Universiteit Amsterdam, the Netherlands

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.
A possible means to improve the statistical quality of trial-based economic evaluations is to include clearly formulated recommendations on the statistical analysis of such studies in national pharmacoeconomic guidelines. However, the extent to which pharmacoeconomic guidelines currently provide guidance on what statistical methods to use in trial-based economic evaluations is unknown. Therefore, this study aimed to review existing national pharmacoeconomic guidelines to assess what recommendations are made regarding the statistical analysis of trial-based economic evaluations.

2. Methods

A scoping review of existing national pharmacoeconomic guidelines was conducted to assess what these guidelines currently recommend regarding the statistical analyses of trial-based economic evaluations.

To identify existing guidelines, ISPOR’s ‘Pharmacoeconomic Guidelines Around the World’ website was searched up until February 20th, 2019 (https://tools.ispor.org/peguidelines/). This website aims to provide an overview of all available national pharmacoeconomic guidelines. In line with previous research, this website was considered to be the most comprehensive source of information because it is updated regularly with the help of health economic experts from about 60 countries [5,6,8,9]. Additionally, guidelines were identified by screening references of relevant review articles [5,6,8,9] and those of the included guidelines. Special efforts were made to identify the most recent versions of the included guidelines by searching websites of HTA agencies. If more than one guideline was available per country, the most recent guideline was included. Guidelines available in English, Dutch, Portuguese, German, Spanish, Persian, and Russian were included.

Data were extracted by one reviewer using a structured data extraction form, after which all of the extracted data were checked by a second reviewer (JvD, ME, AV, AGM, AJB, and MK participated in the data extraction process). Possible discrepancies between reviewers were discussed during a consensus-meeting. If only one reviewer was able to read and understand the guideline in question, a meeting was planned with a second reviewer. During these meetings, both reviewers went through the guideline together, with the second reviewer asking questions probing whether all required data were extracted and whether all extracted data were correct. Data were extracted on general guideline characteristics, such as language, organization, publication year, and type of guideline. In accordance with the ISPOR website, three types of guidelines were distinguished; (1) guidelines containing recommendations for economic evaluations that are not officially required to be followed for reimbursement (further referred to as ‘Recommendations’); (2) guidelines containing recommendations for economic evaluations that are required to be followed for reimbursement (further referred to as ‘Reimbursement guidelines’); and (3) guidelines concerning drug submission requirements with an economic evaluation part/section that are required to be followed for reimbursement (further referred to as ‘Drug submission guidelines’). For extracting the guidelines’ statistical recommendations, six issues were identified that typically complicate the statistical analysis of trial-based economic evaluations, including (1) baseline imbalances; (2) skewed costs; (3) correlated costs and effects; (4) clustering of data; (5) longitudinal data; and (6) missing data (Figure 1). These statistical challenges were based on previous research [1,2,16,17,19–36] as well as our own experiences with analyzing trial-based economic evaluation data. A more detailed description of the statistical challenges is provided in Box 1.

All of the results were narratively synthesized and presented in a Table.

3. Results

3.1. Guidelines

In total, 40 national pharmacoeconomic guidelines were identified that met our inclusion criteria (Figure 2). Of them, the majority was published in Europe (n = 22; 55%), followed by Asia (n = 6; 15%), North America and South America (both; n = 4; 10%), and Africa and Oceania (both; n = 2; 5%). Four guidelines were published prior to 2005 (10%), four guidelines were published between 2005 and 2009 (10%), 16 guidelines were published between 2010 and 2014 (40%), and 16 guidelines were published between 2015 and 2019 (40%). Many of the guidelines published after 2010 are updates of previous guidelines, rather than completely new guidelines. Most guidelines were published in English (n = 31; 78%) and by a public organization (n = 38; 95%), such as the English ‘NICE’, the Dutch ‘Health Care Institute’, the French ‘National Authority for Health’, and the Mexican ‘General Health Council’. Nine guidelines (23%) were categorized as ‘Recommendations’, meaning that they are not officially required to be followed for reimbursement; 22 guidelines (55%) were categorized as ‘Reimbursement guidelines’, meaning that they are required to be followed for reimbursement; and nine guidelines (23%) were categorized as being ‘Drug submission guidelines’, meaning that they are required to be followed as part of a drug submission (Table 1).
Box 1. Statistical challenges to trial-based economic evaluations.

1) Baseline imbalances
Analysts of trial-based economic evaluations frequently assume that the random allocation of participants ensures that mean baseline clinical, cost, and prognostic variables are well-balanced across study groups. Nonetheless, there will inevitably be some between-group differences in baseline values and/or prognostic factors [22]. If such baseline imbalances are unaccounted for in trial-based economic evaluations, their results are likely to be biased [21,22]. Several methods have been proposed to account for baseline imbalances, including mean difference adjustment, regression-based adjustment, and matching methods, such as Genetic Matching [21–24,37]. It should be noted that if such methods are used, they are more frequently used to correct for baseline differences in health effects than in costs [11]. Reasons for this include the relatively high research burden associated with the measurement and valuation of baseline costs and the fact that the period over which baseline costs are measured typically differs from the complete duration of follow-up, whereas it is unknown how this difference in follow-up duration is best accounted for.

2) Skewed costs
Cost data are typically right-skewed, as there are usually relatively few patients with (very) high costs and it is impossible to incur negative costs. Thus, the assumption of normality of standard parametric tests, such as t-tests and linear regression analyses, is violated [20,26]. Unfortunately, log-transformations and standard non-parametric tests (e.g., Mann-Whitney U) are also inappropriate for trial-based economic evaluations, because both methods fail to provide an estimate of the mean difference in costs. Estimates of mean differences in costs are typically required by decision-makers to allow for estimations of the total budget impact of a new intervention (i.e. total budget impact = mean difference in costs * size of the target population) [1,2,17,25]. Non-parametric bootstrapping and generalized linear models with a log or identity link and a gamma distribution are proposed as appropriate means to account for the skewness of costs, while simultaneously providing an estimate of the mean difference in costs [1,2,17,20,25,26].

3) Correlated costs and effects
Costs and effects are typically correlated. To illustrate, patients who died during the follow-up of a trial may cost more than those who survived, because they were sicker to begin with [27–29]. As a consequence, it is inappropriate to assume that differences in costs and effects are independent and their possible correlation should thus be accounted for [27–29]. This can be done using non-parametric bootstrapping, which can account for the correlation between costs and effects by resampling them in pairs [18]. Another proposed method is seemingly unrelated regression (SUR), in which two separate regression models are specified (e.g., one for costs/one for effects). SUR accounts for the correlation between costs and effects through correlated error terms.

4) Clustering of data
Trial-based economic evaluations are often conducted alongside cluster-randomized controlled trials, in which participants are randomized on the cluster-level, rather than on the patient-level. Clusters could refer to groups of patients within different hospitals, practices or even countries. Participants from one cluster tend to be more similar than patients of other clusters. As standard statistical methods assume independent observations, they likely underestimate the level of uncertainty and are therefore not appropriate for analyzing clustered data. Several methods have been proposed for analyzing clustered trial-based economic evaluations, such as two-stage non-parametric bootstrapping and multilevel models [30–34].

5) Longitudinal data
In clinical trials, health effects are typically analyzed using longitudinal methods. This is done because the outcome of interest is often repeatedly measured for each participant and observations of one participant are not independent of each other. As a consequence, it is necessary to apply longitudinal methods to account for the fact that the repeated observations of each participant are correlated [38]. In trial-based economic evaluations, the longitudinal nature of data is typically ignored, most often by only looking at the difference in health effects at follow-up, instead of looking at the difference in health effects over time [11]. This is likely because it is not straightforward to implement longitudinal methods in trial-based economic evaluations. Furthermore, the application of longitudinal methods to cost data results in an estimate of the mean cost difference per time period (i.e. the period between two measurements), whereas an estimate of the mean total cost difference during the entire follow-up period is required. Recently, some researchers started using longitudinal methods in trial-based economic evaluations by summing up the mean cost differences per time period to obtain an estimate of the total cost difference during follow-up [39]. It is unknown, however, whether this method is appropriate, particularly when the duration of time periods is variable.

6) Missing data
Missing data are common in clinical trials [1,2,25]. This is a particular concern for trial-based economic evaluations, because total costs are typically the sum of numerous cost components, measured over different follow-up periods. Thus, if one resource use item is missing, total costs will also be missing [1,2,25]. Historically, missing data were handled by deleting participants with missing values (i.e. complete-case analysis) [1,25]. This approach, however, reduces a study’s power and precision, and only produces unbiased estimates if missing values are not dependent on any observed or unobserved variable (i.e., missing completely at random) [1,2,25]. Possible means to account for missing data in trial-based economic evaluations include “naive” imputation methods, such as mean imputation and last observation carried forward, and more advanced methods, such as multiple imputation and maximum likelihood estimation [16,19,34,33]. Of them, “naive” imputation methods are currently discouraged, amongst others because they do not account for the uncertainty related to filling in missing values [16,19,34–36].

3.2. Recommendations on how to deal with the statistical challenges to trial-based economic evaluations

3.2.1. Baseline imbalances
Four guidelines (10%) included recommendations on dealing with baseline imbalances (i.e., the Scottish, English, Swedish and Dutch guideline; see Table 1). Of them, the Scottish guideline recommends to report baseline differences between groups, the English guideline recommends to adjust treatment effects for their baseline values, and the Swedish and Dutch guidelines recommend the use of regression techniques to adjust for baseline differences between groups. All other guidelines did not provide any recommendations on how to deal with baseline imbalances in trial-based economic evaluations.

3.2.2. Skewed costs
Five guidelines (13%) included recommendations on dealing with skewed costs (i.e., the Danish, Catalanian, Swedish, MERCOSUR, and Dutch guideline; see Table 1). All of these guidelines recommend the use of non-parametric bootstrapping for estimating uncertainty, and the Dutch guideline added that the number of sampling draws should demonstrably lead to a stable result. All other guidelines did not provide any recommendations on how to deal with the skewed nature of cost data in trial-based economic evaluations.

3.2.3. Correlated costs and effects
None of the guidelines included recommendations on how to deal with correlated costs and effects in trial-based economic evaluations.

3.2.4. Clustering of data
None of the guidelines included recommendations on how to deal with the clustering of data in trial-based economic evaluations.
3.2.5. **Longitudinal data**
None of the guidelines included recommendations on how to deal with longitudinal data in trial-based economic evaluations.

3.2.6. **Missing data**
Seven guidelines (18%) included recommendations on dealing with missing data (i.e., the Australian, Belgian, New Zealand, Polish, Scottish, Swedish, and Dutch guideline; see Table 1). The Australian, Belgian, Swedish, and Dutch guidelines recommend to explore and describe the possible missing data mechanism(s). The Swedish guideline added that missing data should never be assumed to occur completely at random. The Australian, Belgian, Polish, Scottish, and Dutch guidelines recommend to describe the applied imputation method. Furthermore, the New Zealand guideline discourages the use of last observation carried forward, and the Dutch guideline recommends to assess the robustness of the applied imputation method by performing sensitivity analyses with different imputation techniques. All
Table 1. National pharmacoeconomic guidelines’ recommendations for dealing with the statistical challenges to trial-based economic evaluations.

<table>
<thead>
<tr>
<th>Country</th>
<th>Guideline characteristics</th>
<th>Baseline imbalances</th>
<th>Skewed costs</th>
<th>Correlated costs and effects</th>
<th>Clustering of data</th>
<th>Longitudinal data</th>
<th>Missing data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia [40]</td>
<td>Year of publication: 2016 Language: English Institution: Department of Health Type of guideline: Drug submission guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Explore and describe possible missing data mechanisms. Describe applied imputation method and its assumptions.</td>
</tr>
<tr>
<td>Austria [41]</td>
<td>Year of publication: 2006 Language: English Institution: Institute for Pharmaeconomic Research Type of guideline: Recommendations</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Brazil [42]</td>
<td>Year of publication: 2014 Language: Portuguese Institution: Ministry of Health Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Baltic [Latvia, Lithuania, Estonia] [43]</td>
<td>Year of publication: 2002 Language: English Institution: Medicines’ Pricing and Reimbursement Agency – Latvia; Health Insurance Fund – Estonia; Ministry of Health – Lithuania Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Belgium [82]</td>
<td>Year of publication: 2012 Language: English Institution: Belgian Health Care Knowledge Center Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Explore and describe possible missing data mechanisms. Describe applied imputation method and its assumptions.</td>
</tr>
<tr>
<td>Canada [44]</td>
<td>Year of publication: 2017 Language: English Institution: Canadian Agency for Drugs and Technologies in Health Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Chile [45]</td>
<td>Year of publication: 2013 Language: Spanish Institution: Ministry of Health Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Colombia [46]</td>
<td>Year of publication: 2014 Language: Spanish Institution: Institute of Health Technology Assessment Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Croatia [47]</td>
<td>Year of publication: 2011 Language: English Institution: Agency for Quality and Accreditation in Health Care and Social Welfare Type of guideline: Recommendations</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cuba [48]</td>
<td>Year of publication: 2003 Language: Spanish Institution: Ministry of Health Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Denmark [49]</td>
<td>Year of publication: 2007 Language: English Institution: Danish Center for Health Technology Assessment Type of guideline: Recommendations</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Use bootstrapping for assessing uncertainty.</td>
</tr>
<tr>
<td>Egypt [50]</td>
<td>Year of publication: 2013 Language: English Institution: Ministry of Health Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>England and Wales [83]</td>
<td>Year of publication: 2013 Language: English Institution: National Institute of Health and Clinical Excellence Type of guideline: Drug submission guideline</td>
<td>-</td>
<td>-</td>
<td>Treatment effects should be adjusted for baseline.</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Finland [51]</td>
<td>Year of publication: 2017 Language: English Institution: Ministry of Social Affairs and Health Type of guideline: Drug submission guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>France [78]</td>
<td>Year of publication: 2012 Language: English Institution French National Authority for Health Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Country</td>
<td>Guideline characteristics</td>
<td>Baseline imbalances</td>
<td>Skewed costs</td>
<td>Correlated costs and effects</td>
<td>Clustering of data</td>
<td>Longitudinal data</td>
<td>Missing data</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>--------------</td>
<td>------------------------------</td>
<td>--------------------</td>
<td>-------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Germany [52]</td>
<td>Year of publication: 2009 Language: English Institution: German National Institute for Quality and Efficiency in Health Care Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hungary [53]</td>
<td>Year of publication: 2017 Language: English Institution: Ministry of Human Resources Type of guideline: Recommendations</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Iran [54]</td>
<td>Year of publication: 2017 Language: Persian Institution: Iran FDA, Medicine selecting committee secretariate Type of guideline: Drug submission guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ireland [55]</td>
<td>Year of publication: 2018 Language: English Institution: Health Information and Quality Authority Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Italy [56]</td>
<td>Year of publication: 2001 Language: English Institution: Italian Group of Pharmacoeconomic Studies Type of guideline: Recommendations</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Israel [57]</td>
<td>Year of publication: 2010 Language: English Institution: Ministry of Health Type of guideline: Drug submission guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Japan [58]</td>
<td>Year of publication: 2016 Language: English Institution: Central Social Insurance Medical Council Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Malaysia [59]</td>
<td>Year of publication: 2012 Language: English Institution: Ministry of Health Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mexico [60]</td>
<td>Year of publication: 2015 Language: Spanish Institution: General Health Council Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MERCOSUR (Argentina, Brazil, Paraguay, Uruguay) [61]</td>
<td>Year of publication: 2015 Language: Spanish Institution: Common Market of the Southern Cone (MERCOSUR) Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>New Zealand [62]</td>
<td>Year of publication: 2015 Language: English Institution: Pharmaceutical Management Agency Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Norway [63]</td>
<td>Year of publication: 2012 Language: English Institution: The Norwegian Medicines Agency Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Poland [64]</td>
<td>Year of publication: 2016 Language: English Institution: The Agency for Health Technology Assessment and Tariff System Type of guideline: Drug submission guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Portugal [65]</td>
<td>Year of publication: 1998 Language: English Institution: National Authority of Medicines and Health Products Type of guideline: Reimbursement guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Russia [66]</td>
<td>Year of publication: 2016 Language: Russian Institution: Ministry of Health Type of guideline: Recommendations</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Scotland [67]</td>
<td>Year of publication: 2018 Language: English Institution: Scottish Medicines Consortium Type of guideline: Drug submission guideline</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>South Africa [68]</td>
<td>Year of publication: 2013 Language: English Institution: National Department of Health Type of guideline: Recommendations</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Spain [69]</td>
<td>Year of publication: 2010 Language: English Institution: Ministry of Health and Social Policy Type of guideline: Recommendations</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Country</th>
<th>Year of publication</th>
<th>Language</th>
<th>Institution</th>
<th>Type of guideline</th>
<th>Guideline characteristics</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>2016/2017</td>
<td>English</td>
<td>The Dental and Pharmaceutical Benefits Agency</td>
<td>Reimbursement guideline</td>
<td>Regression techniques and propensity scores may be used to correct for differences between groups/confounders.</td>
<td>Use bootstrapping for assessing uncertainty.</td>
</tr>
<tr>
<td>Switzerland</td>
<td>2011</td>
<td>German</td>
<td>Swiss Federal Office of Public Health</td>
<td>Reimbursement guideline</td>
<td>-</td>
<td>Explore and describe possible missing data mechanisms. Missing data should never be assumed to occur at random.</td>
</tr>
<tr>
<td>Taiwan</td>
<td>2006</td>
<td>English</td>
<td>Taiwan Society for Pharmacoeconomics and Outcomes Research</td>
<td>Reimbursement guideline</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Thailand</td>
<td>2014</td>
<td>English</td>
<td>Various</td>
<td>Drug submission guideline</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>2016</td>
<td>English</td>
<td>Dutch Health Care Institute</td>
<td>Reimbursement guideline</td>
<td>Regression techniques may be used to correct for differences between subgroups.</td>
<td>Use bootstrapping, with a sufficient number of sampling draws for assessing uncertainty.</td>
</tr>
<tr>
<td>United States</td>
<td>2016</td>
<td>English</td>
<td>Academy of Managed Care Pharmacy</td>
<td>Recommendations</td>
<td>-</td>
<td>Explore and describe possible missing data mechanisms. Describe applied imputation method and its assumptions. Use different methods to deal with missing data to assess their relative robustness.</td>
</tr>
</tbody>
</table>
other guidelines did not provide recommendations on how to deal with missing data in trial-based economic evaluations.

4. Expert opinion

National pharmacoeconomic guidelines were found to provide little to no guidance on how to deal with the statistical challenges to trial-based economic evaluations. That is, the majority of guidelines did not provide recommendations on how to deal with baseline imbalances, skewed costs, correlated costs and effects, the clustering of data, the longitudinal nature of data, and missing data in trial-based economic evaluations; and where recommendations were provided, they were very general in nature. This finding is consistent with previous research showing that, even though national pharmacoeconomic guidelines view transferability of, and dealing with patient heterogeneity in, economic evaluations as being important, very few of them proposed specific methods for dealing with these issues [5, 77].

Various reasons may explain the lack of guidance by national pharmacoeconomic guidelines on how to deal with the statistical challenges to trial-based economic evaluations. First, the majority of guidelines was found to focus predominantly on model-based economic evaluations, and some even see model-based economic evaluations as the ‘gold standard’. The latter is evidenced by the French guideline that states that ‘modelling is the preferred approach in health economic evaluation’ [78]. As a consequence, many guidelines only included detailed recommendations on how to construct a decision analytic model, but lacked guidance on how to appropriately analyze trial-based economic evaluations. Second, many of the identified guidelines seem to be modeled after, or included, reporting guidelines, such as the CHEC-list [79], CHEERS statement [80], or a self-developed one. The Belgian guideline, for example, included a checklist for assessing the quality of trial-based economic evaluations that mainly included reporting items, such as whether there was a comprehensive description of alternatives [81, 82]. Even though reporting guidelines are important to ensure uniformity in the reporting of (trial-based) economic evaluations, they do not provide guidance on how to appropriately analyze such studies. Third, some guidelines were published over a decade ago, while in the meantime, methods for analyzing trial-based economic evaluations have improved considerably [26]. Fourth, some guidelines only included very general statistical recommendations, instead of recommending specific methods for dealing with one or more of the statistical challenges to trial-based economic evaluations. The English guideline, for example, stated that “Data should be analyzed in a way that is methodologically sound and, in particular, minimizes any bias”, but did not indicate what those methodologically sound methods are. Guideline developers may have opted for such a strategy due to a lack of overall consensus on how to deal with statistical issues, such as skewed data and missing data, or in an effort to allow future analysts the freedom to choose different methods. It is unlikely, however, that such general recommendations will improve the quality of trial-based economic evaluations. In an effort to deal with this issue, the English guideline commissioned a series of Technical Support Documents with the aim of providing further information on how to implement the approaches described in the guideline. However, even though a broad range of Technical Support Documents is currently available, they are mostly focussed on the measurement and valuation of health and the conduct of modeling studies, and lack specific recommendations on how to statistically analyze trial-based economic evaluations [83].

A possible limitation of this study is that we had to exclude five national pharmacoeconomic guidelines due to language restrictions (i.e., the Chinese, South Korean, Slovakian, Slovenian, and Czech guidelines). Nonetheless, with 40 guidelines, the number of included guidelines is relatively high compared to similar studies [5, 77, 84] and we do not expect that the exclusion of some guidelines affected our overall conclusion. Also, as in previous reviews, there may have been some level of subjectivity in our assessment of what parts of the guidelines could be considered guidance and what not [77]. We tried to tackle this limitation by using a structured data extraction form (available upon request) and by having all extracted data checked by a second reviewer. Another limitation might be that overall consensus does not exist regarding the most important statistical challenges to trial-based economic evaluations. Based on the literature as well as our own experiences with analyzing trial-based economic evaluation data, we selected the handling of baseline imbalances, skewed data, correlated costs and effects, clustering of data, longitudinal data, and missing data. Some researchers may be of the opinion that issues, such as sample size calculations and patient heterogeneity should have also been included. We decided to exclude sample size calculations because they are typically performed prior to the start of a clinical trial instead of during the analysis phase. Moreover, if studies are powered to detect relevant cost differences they will almost certainly be overpowered for the clinical outcome, while it may be unethical to continue recruiting patients into a trial beyond the point at which clinical superiority has been determined beyond reasonable doubt [85]. Furthermore, the recommendations of national pharmacoeconomic guidelines regarding patient heterogeneity have already been assessed in a previous review, which found similar results [77].

In order to improve the statistical quality of trial-based economic evaluations, we recommend to include more detailed guidance on how to statistically analyze such studies in national pharmacoeconomic guidelines. National guidelines may, for example, adopt and/or refer to the latest statistical recommendations of the ISPOR RCT-CEA Task Force and/or the Panel on Cost-Effectiveness in Health and Medicine [26, 84, 85]. The ISPOR RCT-CEA Task Force recommends to adjust for baseline imbalances using multivariable analyses, to deal with skewed costs using non-parametric bootstrapping, and to deal with missing data using multiple imputation. The task force also recommends to analyze multinational studies, in which participants are clustered within countries, using multilevel random effects models with shrinkage estimators. It remains unclear, however, whether this recommendation also applies to other kinds of clustered data (e.g. that of economic evaluations performed alongside...
a cluster-randomized controlled trial). If clinical endpoints are used in a trial-based economic evaluation, the task force recommends to use the statistical methods that are also applied in the clinical analysis, which are typically longitudinal in nature. However, the task force does not provide guidance on how to implement such longitudinal methods in trial-based economic evaluations. Recommendations on how to deal with correlated costs and effects are not included in the latest task force report [26]. The latest Panel on Cost-Effectiveness in Health and Medicine did not provide recommendations on how to deal with the statistical challenges to trial-based economic evaluations [86], but did discuss the issues of skewed costs and missing data into more detail in their complete report [87]. Also, if statistical recommendations are included, and/or referred to, in national pharmacoeconomic guidelines, it is important that these guidelines are updated frequently in order not to stifle further scientific development and/or restrict analysts to use outdated statistical methods.

Next to the inclusion of statistical recommendations in national pharmacoeconomic guidelines, other measures could also be taken to improve the statistical quality of trial-based economic evaluations. First, even though various methodological studies assessed the (relative) performance of certain statistical methods in trial-based economic evaluations (e.g. [27–34,80]), more research into this area is warranted. Particular focus should be given to statistical methods for dealing with correlated costs and effects, clustered data, and longitudinal data as well as strategies for combining statistical methods, such as non-parametric bootstrapping and multiple imputation and/or longitudinal methods and multiple imputations. Second, increased research efforts should be focused on reaching consensus about the ‘best’ methods for dealing with the statistical challenges to trial-based economic evaluations. For achieving this, it might be useful to perform a scoping review to provide researchers with an overview of the (relative) performance of available statistical methods as well as existing gaps in knowledge. The results of such a scoping review may in turn serve as a starting point for the development of a consensus-based statistical quality checklist for trial-based economic evaluations.

In conclusion, current national pharmacoeconomic guidelines provide little to no guidance on what statistical methods to use in trial-based economic evaluations. Since the use of suboptimal statistical methods may lead to biased results, and, therefore, possibly to a waste of scarce resources, national agencies are advised to include statistical guidance and/or recommendations in their pharmacoeconomic guidelines.

**Funding**

This paper was not funded.

**Declaration of interest**

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

**Reviewer Disclosures**

Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

**References**


47. Agency for Quality and Accreditation in Health Care, Department for development, research and health technology assessment: the Croatian guideline for health technology assessment process and reporting. Agency for Quality and Accreditation in Health Care, Zagreb (2011).


51. Pharmaceuticals Pricing Board, Ministry of Social Affairs and Health. Preparing a health economic evaluation to be attached to the application for reimbursement status and wholesale price for a medical product. 2017.


57. Ministry of Health Pharmaceutical Administration. Guidelines for the submission of a request to include a pharmaceutical product in the national list of health services. 2010.


76. Academy of Managed Care Pharmacy. The AMCP format for formulary submissions. Version 4.0.