

Handling Missing Data in Within-Trial Cost-Effectiveness Analysis: A Review with Future Recommendations

Andrea Gabrio¹ · Alexina J. Mason² · Gianluca Baio¹

Published online: 9 February 2017

© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract Cost-effectiveness analyses (CEAs) alongside randomised controlled trials (RCTs) are increasingly designed to collect resource use and preference-based health status data for the purpose of healthcare technology assessment. However, because of the way these measures are collected, they are prone to missing data, which can ultimately affect the decision of whether an intervention is good value for money. We examine how missing cost and effect outcome data are handled in RCT-based CEAs, complementing a previous review (covering 2003–2009, 88 articles) with a new systematic review (2009–2015, 81 articles) focussing on two different perspectives. First, we provide guidelines on how the information about missingness and related methods should be presented to improve the reporting and handling of missing data. We propose to address this issue by means of a quality evaluation scheme, providing a structured approach that can be used to guide the collection of information, elicitation of the assumptions, choice of methods and considerations of possible limitations of the given missingness problem. Second, we review the description of the missing data, the statistical methods used to deal with them and the quality of the judgement underpinning the choice of these methods. Our review shows that missing data in within-RCT CEAs are still often inadequately handled and the overall

level of information provided to support the chosen methods is rarely satisfactory.

Key Points for Decision Makers

This is the first systematic review proposing a quality assessment of CEA studies based on how missingness in costs and effects is handled and reported (quality evaluation scheme, QES).

The recommendations, used in building the QES, guide the choice of the missing data method based on the description of the data and assumed missing data mechanism; this in turn should be informed by the available evidence and assessed using sensitivity analysis.

The review shows a currently inadequate handling of missingness for both outcomes in the literature.

This may in turn lead to bias or overconfidence in CEA results that decision makers should take into account in their evaluation of the quality of a study.

✉ Andrea Gabrio
andrea.gabrio.15@ucl.ac.uk

¹ Department of Statistical Science, University College London, 1–19 Torrington Place, London WC1E 7HB, UK

² Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, 15–17 Tavistock Place, London WC1H 9SH, UK

1 Introduction

A well-known issue in cost-effectiveness analysis (CEA), especially within a randomised controlled trial (RCT) setting, is the presence of large proportions of missing data in either or both outcome variables, i.e. the cost and the clinical effectiveness or utility measures. Removing the unobserved cases (a method usually referred to as “complete case analysis”, CCA) generally leads to a loss in

efficiency and possible serious biases in the parameter estimates [1–4]. Consequently, the final conclusions of the study may be strongly influenced by the way in which missingness is handled, thus possibly reversing the decision about the cost-effectiveness of a new treatment compared with the standard option [5, 6].

While the problem of missing data is widely discussed in the general statistical literature, it has been less extensively addressed in the health economics one. Notable exceptions include Graves et al. [7], Briggs et al. [8], Oostenbrink et al. [9], Oostenbrink et al. [10], Burton et al. [11] and Lambert et al. [12], mainly focussing on the cost measures; Richardson and Manca [13], Wood et al. [14], Groenwold et al. [15], Powney et al. [16], Simons et al. [17] and Rombach et al. [18], with reference to health outcome measures; and Manca and Palmer [5], Harkanen et al. [19], Diaz-Ordaz et al. [20] and Faria et al. [21], who consider both outcomes.

Interestingly, recent reviews on the methods applied in within-trial CEAs [22–24] have concluded that CCA has historically represented the standard approach in health economics. As a result, we should be naturally sceptical about the conclusions achieved by CEAs performed in a context where missingness is not addressed in a “principled” way. We make reference to the concept of a principled approach to handle missingness in Sect. 2. Within this framework, subsequent inferences are valid under the stated assumptions, which in turn can be varied to test their impact on the decision-making.

The objective of this article is twofold: first, we collect guidelines from the literature about the way in which missingness should be analysed and reported in within-trial CEAs and embed them within a structured scheme. Second, we review the methods used to handle missingness in the studies between 2003–2015 by updating and extending the work of Noble et al. [22]. This is done with a view to assessing whether the methods have evolved over time and whether the quality level of missing data analyses is adequate. The paper is structured as follows: Sect. 2 illustrates Rubin’s classification of missing data mechanisms [1]. In Sect. 3, we provide a brief summary of the most popular missingness methods. In Sect. 4, we provide a structured approach, which we name quality evaluation scheme, which summarises the most important recommendations collected from literature about the way information and assumptions on missing data should be provided in CEA studies. Based on these guidelines, we also suggest a systematic approach for grading the studies given the overall level of information about missingness provided. Section 5 presents the methodology used to select the review’s articles and the main results derived from the different descriptive analyses performed. In Sect. 6, we apply the quality evaluation scheme to the reviewed articles in the

period 2009–2015 and show how they perform in terms of missing data handling. Finally, Sect. 7 summarises our findings and recommendations.

2 Missing Data Mechanisms

When analysing partially observed data, it is essential to investigate the possible reasons behind the missingness. This formally translates into an assumed missing data mechanism that is linked to the data generating process, as a key concept to address missingness in a “principled” way. We specifically refer to “principled” methods for missing data as those based on a well-defined statistical model for the complete data, and explicit assumptions about the missing value mechanism.

We consider a sample of $i = 1, \dots, n$ individuals; for each, the relevant outcome is indicated as y_i , which is unobserved for some individuals. Typically, trial data also include a set of J covariates $\mathbf{x}_i = (x_{1i}, \dots, x_{Ji})$, e.g. sex, age or co-morbidities. While in general these may be partially or fully observed, in this section we consider only the latter case. In addition, we define a missingness indicator m_i taking value 1 if the i th subject is associated with missing outcome and 0 otherwise.

This setting can be modelled using two sub-models, or “modules”. The first module is the missing data mechanism, denoted as model of missingness (MoM). It describes a probability distribution for m_i , as a function of some unobserved parameters π_i and δ , defining the probability of missingness in the outcome variable y_i . The second module is the data generating process of the outcome variable, denoted as model of analysis (MoA). This contains the main parameters of interest (e.g. the population average costs and benefits) and describes a probability model for the outcome y_i . As a general example, we can think of a simple regression model where $y_i \sim N(\mu_i, \sigma)$, and $\mu_i = \beta_0 + \beta_1 x_i$. In this case, the parameters of the MoA are the regression coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1)$ representing respectively the intercept and the slope, and the individual standard deviation σ .

The most accepted classification of missingness mechanisms is given by Rubin [1] and is based on three classes, according to how the missingness probability in the MoM is modelled. A simple graphical representation of the three classes is provided in Fig. 1. Variables and parameters are denoted by nodes of different shapes and colours according to their nature. Parameters $(\beta_0, \beta_1, \sigma, \delta)$ are represented through grey circles. “Logical” quantities such as μ_i and π_i , which are defined as a function of other parameters, are denoted by a double circle notation. Fully observed variables (m_i) are represented with a white circle, while partially observed variables (y_i) are denoted by a darker grey

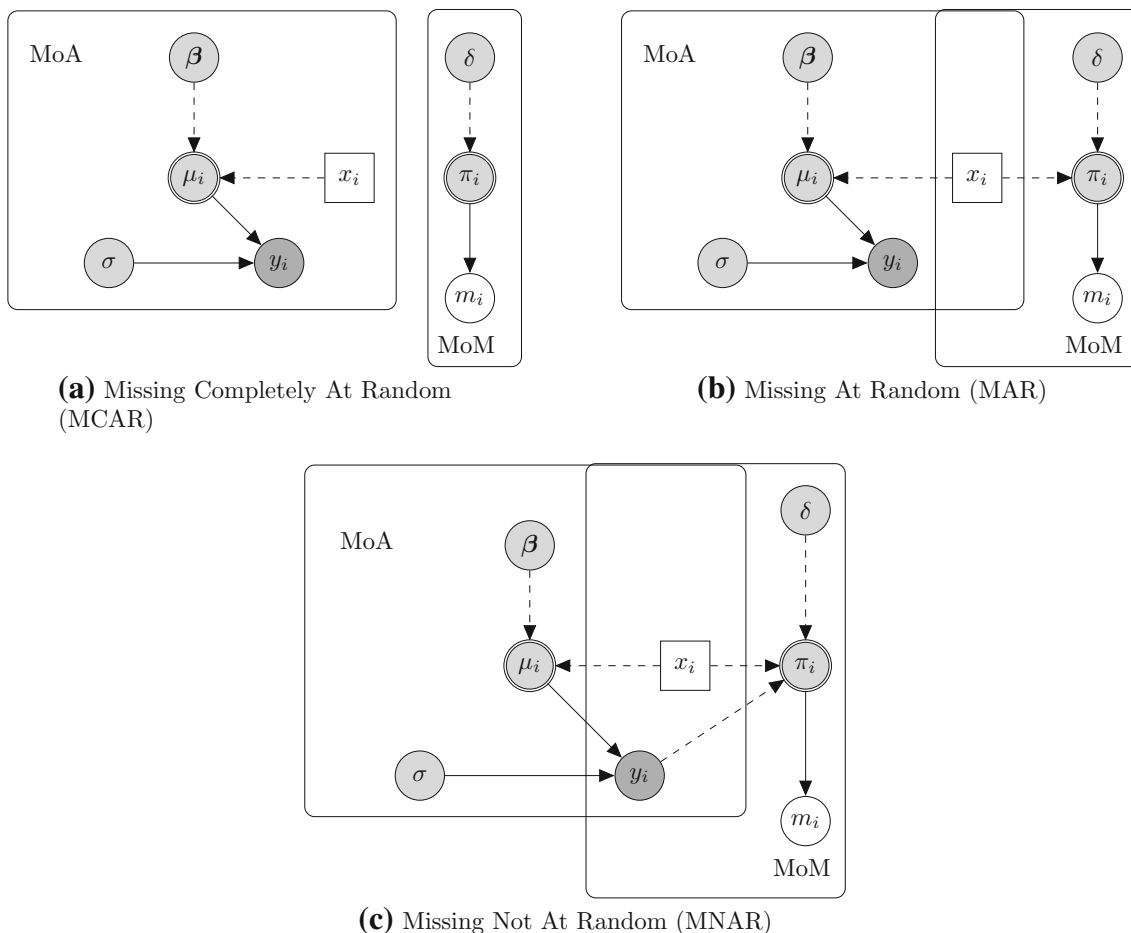


Fig. 1 Graphical representation of Rubin’s missing data mechanism classes, namely MCAR (a), MAR (b) and MNAR (c). Variables and parameters are represented through nodes of different shapes and colours. Parameters are indicated by *grey circles* with logical parameters defined by *double circles*, while predictor variables are

assumed fixed and drawn as *white squares*. Fully observed variables are denoted by *white circles*, partially observed variables by *darker grey circles*. Nodes are related to each other through *dashed* and *solid arrows* which respectively represent logical functions and stochastic dependence. *MoA* model of analysis, *MoM* model of missingness

circle. Finally, we show covariates (x_i) as white squares to indicate that they are fully observed and not modelled. Rounded rectangles are used to show the extent of the two modules in terms of variables/parameters included. Arrows show the relationships between the nodes, with dashed and solid lines indicating logical functions and stochastic dependence, respectively.

Figure 1a illustrates the class of “missing completely at random” (MCAR), in which the probability of missingness is fully independent of any other partially or fully observed variable. Consequently, in Fig. 1a, MoA and MoM are not connected and π_i does not depend on any quantity in the MoA. This amounts to assuming that there is no systematic difference between partially and fully observed individuals in terms of the outcome y_i . In other words, in this case, we would be assuming that observed cases are a representative sample of the full sample.

Figure 1b shows a case of “missing at random” (MAR), in which the missingness probability may depend on a fully

observed variable. As a result, MoA and MoM are connected by means of the predictor variable affecting both the mechanisms generating y_i and m_i . Because of this relationship, the partially observed cases are systematically different from the fully observed cases; crucially, however, the difference is fully captured by x_i .

Figure 1c provides an example of “missing not at random” (MNAR). This is characterised by dependence of the probability of missingness on both the partially and fully observed variables. Thus, in Fig. 1c, π_i depends on both the fully observed predictor x_i and the partially observed outcome y_i . This means that the difference between fully and partially observed cases still depends on the missing values, even after taking x_i into account. Therefore, it is necessary to make more structured assumptions about this relationship that go beyond the information contained in the data.

While intuitively helpful, this framework may be too simplistic in some cases. Since the scope of this section is to provide a broad overview for Rubin’s classification, we

assumed the simplest case where missingness is present in a single response variable only, which may not hold in real applications. This is particularly likely in the context of CEA, in which we are concerned with a multivariate outcome, made of suitable measures of clinical benefits and costs, i.e. $y_i = (e_i, c_i)$. Missingness can occur for either or both the relevant outcomes, and this can lead to as many missingness mechanisms as the number of partially observed quantities (covariate missingness must also be considered). Additional complexity is given by whether data are obtained in a cross-sectional or longitudinal setting, static or time-varying covariates and more importantly the possible correlation between variables and missingness mechanisms and between the mechanisms themselves.

3 Approaches to Handle Missing Data

There are many different statistical approaches to deal with missingness, each relying on different assumptions. It is important to carefully select the approach in line with the setting-specific assumptions we assume to hold. For the sake of simplicity here we only broadly categorise these techniques. More in-depth and complete presentation and analysis can be found for example in Molenberghs et al. [4] and Schafer and Graham [25].

3.1 Complete Case Analysis

This is a popular method in within-trial CEA studies, despite its limitations due to the strong assumption that only the fully observed cases are needed in order to correctly make inference. The critical disadvantage is that missing cases are simply discarded, thus reducing efficiency and possibly biasing the parameter estimates.

3.2 Single Imputation

Single imputation (SI) methods replace the missing data with a single predicted value, such as the unconditional or conditional mean or the last value observed for a given case. This category includes last value carried forward [26], linear extrapolation [27], mean and conditional imputation [28]. Although sometimes valid, these methods are never recommended, as they typically require stronger assumptions than MCAR and always fail to take account of the uncertainty underlying the imputation process; i.e. they do not recognise that the imputed values are estimated rather than known.

3.3 Multiple Imputation

A more sophisticated method is multiple imputation [MI, 1]. The underlying idea is to fill in the missing data with

plausible simulated values, drawn from the conditional predictive distribution of the missing given the observed values. Thus, the set of imputations can properly represent the information about the missing values that is contained in the observed data for the chosen model. This is repeated K times, leading to K imputed datasets that can be analysed via complete-data methods. The individual estimates are then combined into a single quantity, e.g. using Rubin's rules [1]; this captures the variability within and between imputations. However, the critical aspect is that valid inferences depend on the correct specification of the imputation model in terms of variable selection, distributions and correlations.

3.4 Sensitivity Analysis

Sensitivity analysis (SA) is a technique used to determine how different input values in a model will impact the output, under a given set of assumptions. Missing data will invariably increase the underlying uncertainty, because they induce a lower level of information in the data. SA represents an extremely valuable tool to deal with the uncertainty induced by missingness. Indeed, when applied specifically to missing data, SA corresponds to exploring as many plausible missing data assumptions as possible and then assessing how consistent results are across the different scenarios. In particular, it is generally recommended to set MAR as the reference assumption and then explore different MNAR departures from MAR, to assess the robustness of the results to different plausible alternative missingness mechanisms. The purpose of such analysis is to account more fully for the uncertainty about the missingness. Usually SA is implemented through more advanced methods that are able to explicitly model a MNAR mechanism such as selection or pattern mixture models [4, 30].

3.5 Others

Examples of other less commonly adopted methods that we found in the reviewed articles were: random draw, linear mixed-effects model [31], expectation-maximisation algorithm [32], input-case analysis, assumed zeros and two-part regression.

3.6 Alternatives

There is a long list of possible alternatives that could be explored. Among the most important we refer to inverse probability weighting [33], data augmentation [34], likelihood-based methods [35] and doubly robust [36] and full Bayesian methods [30, 37].

4 Recommendations for Missing Data Analyses

The main contribution of this paper to the literature for missing data analysis recommendations is provided by the quality evaluation scheme, as a structured reporting and analysis system which aims to embed all the most important guidelines for authors in terms of missing data handling in CEA. These are listed in Table 1 and summarise the guidelines provided by previous review articles [14, 21–23]. The contribution of the scheme consists in providing a unique, general and structured framework that authors can use to more easily check that all the relevant information about missing data is used to inform the assumptions relating to the missing data in their analysis.

4.1 Quality Evaluation Scheme

In order to judge whether missing data in CEAs have been adequately handled, a full description of the missingness problem, details of the methods used to address it and a discussion on the uncertainty in the conclusions resulting from the missingness are required. With this in mind, we have assembled guidelines on how information relating to the missing data should be reported (Table 1). We define three broad categories (description, method, limitations). For each, information that we consider vital for

transparency is listed under “key considerations”, while other details that could usefully be provided as supplementary material are suggested under “optimal considerations”.

Using the list of key considerations in Table 1, one first determines whether no (all key considerations absent), partial (one or more key considerations absent) or full (all key considerations present) information has been provided for each component. From this, one computes a numerical score to summarise the overall information provided on missingness, weighting the components in a ratio of 3:2:1. This weighting scheme has been chosen based on some reasonable assumptions in terms of how the different analysis components could impact the results in terms of missing data handling. Specifically, we think the “limitations” component should not be weighted as much as the other two because of its limited impact on the conclusions. In the same way, we suggest that the “description” should be weighted more than the “method” component, as it generally drives the choice for the initial assumptions about the missingness. We are aware that other weighting schemes could be chosen, but in light of the given considerations, we think the proposed weights are a reasonable choice that we recommend. In addition, the robustness of the scheme to a different weight choice, where each component has exactly the same weight, has been tested to assess the sensitivity of the score assignment. While the

Table 1 List of information content for each of the three components that one would like to observe in the studies to achieve full analysis reporting of missing data. The contents are divided into two subgroups: key and optimal considerations. The former are the statements to be considered as mandatory for transparency when conducting an economic evaluation in the presence of missing data.

The latter are additional considerations that further extend the analysis reporting of the missing data through supplementary materials. The lack of even one single key consideration is considered to indicate partial analysis reporting, while null analysis reporting is related to the absence of all key considerations

Description	Method	Limitations
<p>Key considerations</p> <ol style="list-style-type: none"> 1. Report the number of individuals with missing data for each variable in the reported analysis by treatment group 2. Describe the missing data patterns for all variables included in the economic analysis (is missingness on one variable associated with missingness on another variable? Is there a longitudinal aspect to the data?) 3. Discuss plausible reasons why values are missing (e.g. death) 	<p>Key considerations</p> <ol style="list-style-type: none"> 1. Identify a plausible missingness assumption for the specific patterns and setting analysed 2. State the method and software used in the base-case analysis 3. For more general methods, provide details about their implementation^a 4. Perform a plausible robustness analysis; provide and discuss the results 	<p>Key considerations</p> <ol style="list-style-type: none"> 1. Acknowledge and quantify the impact of the missing data on the results 2. State possible weaknesses and issues with respect to the method and assumptions
<p>Optimal considerations</p> <ol style="list-style-type: none"> 1. Provide supplementary material about the preliminary analysis on missingness (e.g. descriptive plots and tables) 	<p>Optimal considerations</p> <ol style="list-style-type: none"> 1. Provide supplementary material about the method implementation in the base-case and robustness analysis (e.g. software implementation code) 	

^a For example, in multiple imputation, state the imputation model specification and variables included, the number of imputations, post imputation checks

results do not show any significant changes in the classification of the articles, when using the alternative scheme, it is more difficult to discriminate across the articles due to a more limited score range.

Finally, one converts the scores into grades A–E. The relevance of the scores, and of the derived grades, in terms of decision analysis is mainly associated with a qualitative assessment of the articles. Studies that are graded in the top categories should be associated with a higher degree of confidence in their results, whereas more caution should be taken in the consideration of results coming from studies that are graded in the bottom categories. When qualitatively assessing the articles, the different grading assigned to each of them could be an indication of a lack in the robustness of the conclusions provided due to missingness uncertainty.

Figure 2 shows the process and weights used. Although the importance between the different components is subjective, we believe that the chosen structure represents a reasonable and relatively straightforward assessment scheme.

The resulting scores can be interpreted as follows:

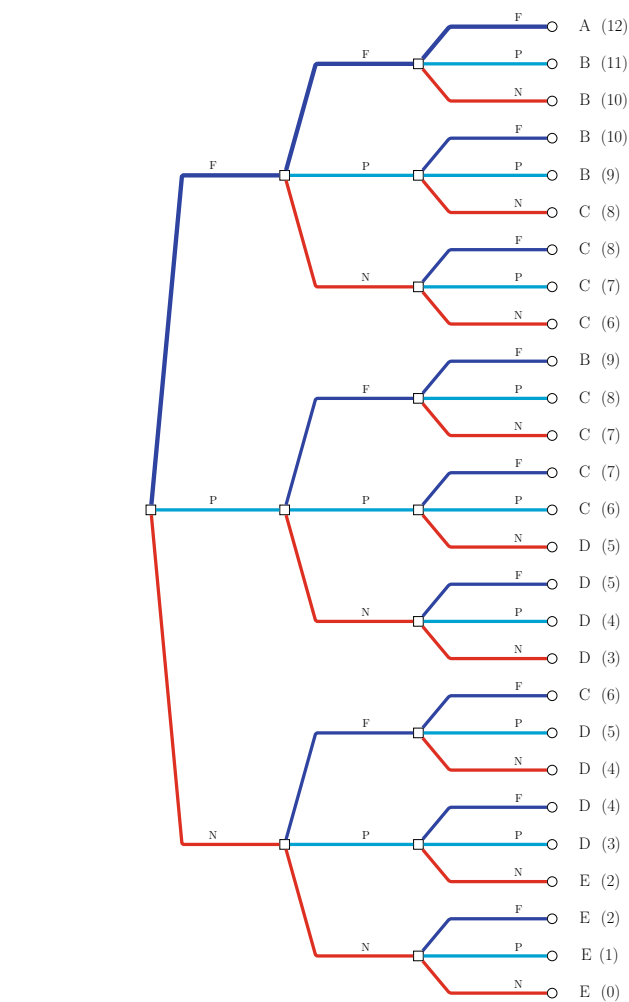
A (12) The highest quality judgement, identified by the upper thicker blue path in Fig. 2, including only those studies that simultaneously provide all the key considerations for all the components. It is the benchmark for a comprehensive explanation and justification of the adopted missing data method.

B (9–11) Includes studies providing full details for either the description or the method and at least partial information for the other components. Studies with no information about the limitations are only included in this category if full detail is provided for both the other components.

C (6–8) Studies for which information about missingness is not well spread across the components. All key considerations are provided either for the description or the method, but with only partial or no content in the other components. Alternatively, we can have partial content for description and method, and partial or full content for limitations.

D (3–5) Indicates a greater lack of relevant information about missingness. Despite possibly including key considerations on any of the components, the information provided will at most be partial for the description, in which case it will be combined with a total lack of content on either the method or the limitations.

E (0–2) The worst scenario, where the overall information about the missing data is considered to be totally unsatisfactory. No description is given, and we can observe at most only some of the key considerations for the method.



CONTENT \ SCORE	Description	Method	Limitations
Full (F)	6	4	2
Partial (P)	3	2	1
Null (N)	0	0	0

Fig. 2 Diagram representation for the quality score categories. The table at the bottom shows how scores have been weighted according to the information provided on each component. In accordance with the table, different branches of the diagram correspond to different components. From left to right, the initial branches are related to the description, branches in the middle to the method, and final branches to the limitations. The branches' colour represents the different way the information provided in each analysis component is evaluated: red no information (N), light blue partial information (P), blue full information (F). Final scores (0–12) with associated ordered categories (E–A) show the overall level of information provided for each combination of component and content evaluation

5 Literature Review

5.1 Methods

Noble et al. [22] (henceforth NHT) reviewed the methods used to handle missing cost measures in 88 articles published during the period 2003–2009. We extend their

review, to include missing effects. Further, we use NHT's strategy to identify papers in the subsequent period, 1 April 2009 to 31 December 2015. Articles were considered eligible for the review only if they were cost-effectiveness analyses within RCTs, used individual patient-level data and mentioned missing data in the text. We relied on the search engines of three online full-text journal repositories: (1) Science-Direct.com, (2) bmj.com, and (3) The Database of Abstracts of Reviews of Effects (DARE) and NHS Economic Evaluation Database (NHS EED). The key words used in the search strategy were (cost effectiveness OR economic evaluation) AND missing data AND trial AND (randomised OR randomized). The on-line databases identified 1129 articles, most of which were duplicates. After abstract review, 128 articles were considered, of which 81 fulfilled the eligibility criteria.

We present and compare the articles reviewed for the two periods by type of analysis performed. First, we look at the base-case methods implemented, i.e. those used in the main analysis embedding the assumptions about missing data. Second, we consider any alternative methods discussed; when present, these assess the robustness of the results obtained in the main analysis against departures from the initial assumptions on missingness.

5.2 Base-Case Analysis

As shown in Fig. 3a, NHT found that CCA was the most popular base-case method, used in 31% of the papers; 23% were unclear about the technique adopted. Single imputation methods were well represented, with mean imputation and conditional imputation used in 10 and 9% of the articles, respectively. MI was found in 9% of the articles. Our analysis of the methods for missing effectiveness measures shows a similar pattern in Fig. 3c. CCA was used in 27% of the cases and with a sizeable proportion of papers unclear about the technique adopted (24%). Single imputation methods are here dominated by last value carried forward (LVCF) (10%), while a slightly higher proportion uses MI (15%).

In 2009–2015, MI replaces CCA as the most frequently used base-case method in both costs and effects, at 33 and 34%, respectively (Fig. 2b, d). However, CCA is still the method of choice in many papers (15% for costs and 21% for effects). The proportion of papers that are unclear about the chosen method is similar over the two time periods for costs, but halves in the later period for effects. This is plausible since clinical effectiveness measures and the estimate of treatment effect have been the main focus of the analyses, whereas costs are less frequently included as a primary outcome in the study research questions. This will translate into a more careful and reasoned examination of

the missingness problem in the effect compared with the cost analysis.

5.3 Robustness Analysis

With the term “robustness analysis” we refer to a different concept with respect to the one introduced in Sect. 3.4. In particular, we make this distinction so we can specifically refer to robustness analysis as any alternative analysis, compared with the base-case analysis, based on different missing data methods. By contrast, when applied to missing data, SA can be thought of as structurally varying the assumptions about the missingness model, whose plausibility can be assessed in light of known information.

However, in practice, even robustness analyses are rarely performed in CEAs. This poses an important question related to the reliability of the findings, as they may be affected by the specific assumptions about the missing data. From both review periods it seems that a robustness analysis is infrequently used and typically involves only one alternative scenario. This is not likely to be an optimal choice as the main objective of this analysis is to explore as many plausible alternative missing data assumptions as possible.

NHT found that 75% (66/88) of the articles did not include any robustness analysis, with the remaining papers typically performing an analysis by comparing CCA and MI (based on missing costs only). Similar findings apply to missing effects, with about 76% (67/88) of the studies lacking any alternative missing data method. Similarly, in the 2009–2015 review, we observe no robustness analysis in the majority of the articles for both costs (75% or 61/81) and effects (70%, 51/81).

Figure 4 provides a pictorial overview of the alternative methods used for cost and effect data. For costs, most articles describe no alternative analysis. In the earlier period, the choice of alternative missingness methods seems well spread across CCA, MI and the use of more than one method, with a slightly more frequent adoption of MI. By contrast, in the later period, more cases use CCA as a robustness method in combination with MI as the base-case method.

Figure 4c, d describes the effects, with most of the articles not reporting any robustness analysis and with a significant decrease in MI analyses used for robustness, opposed to an increase in CCA, between the two periods. There is a similar pattern to the cost graphs towards CCA used as a robustness method in combination with MI as the base-case method.

One of the possible reasons that lead authors to use CCA as a robustness method is its simplicity of implementation. Although in many cases its underlying assumptions may not be credible, authors may use it to make comparisons

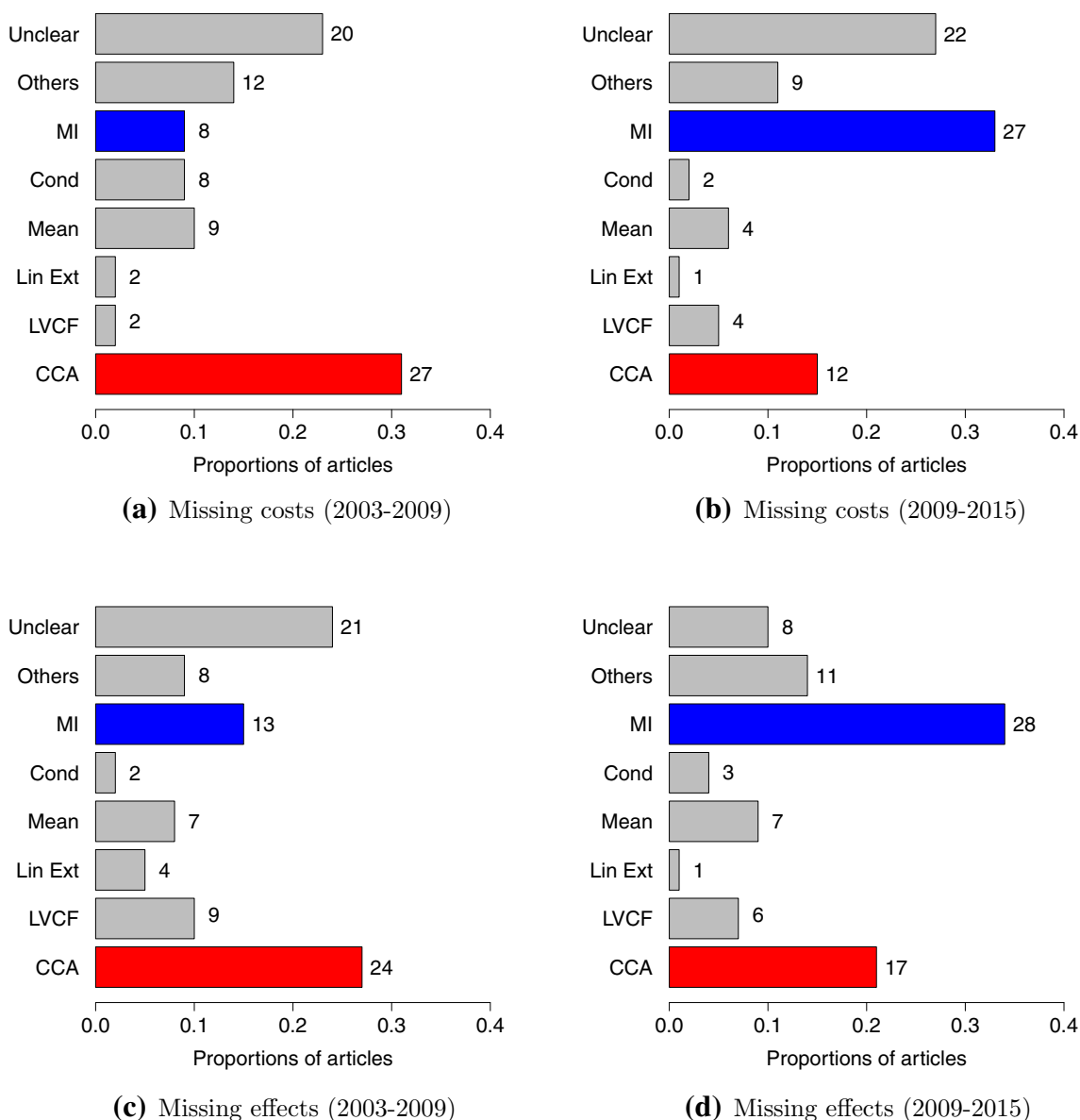


Fig. 3 Review of base-case methods used to handle missing cost and effect data between 2003–2009 and 2009–2015. Legend: complete case analysis (CCA), last value carried forward (LVCF), linear extrapolation (Lin Ext), mean imputation (Mean), conditional imputation (Cond), multiple imputation (MI), any other method present in fewer than four articles (Others), unspecified method (Unclear). The category “Unclear” includes those articles for which it was not

with the base-case method as it is a quick and easy-to-perform alternative option.

6 Grading the Articles

Comparing the information provided in the articles in our review against the list given in Sect. 4.1 allows us to qualitatively assess the quality of the reporting of how missingness has been handled in the reviewed CEA studies.

possible, based on the text, to understand the methodology used to deal with the missingness, while the category “Others” consists of the following methods: random draw, linear mixed-effects model, expectation-maximisation algorithm, input-case analysis, assumed zeros, two-part regression. The numbers to the right of the bars in the graphs are the number of papers including the corresponding method in the base-case analysis

To gain a fuller understanding of the current state of play, we also classify the articles from the perspective of the strength of the assumptions about the missingness mechanism. This is related to the choice of method, since each is underpinned by some specific missing data assumption. We can view the quality judgement and strength of assumptions as two dimensions providing a general mapping of how the missingness problem is handled. This applies to both the level of knowledge about the implications of a given missingness assumption on the

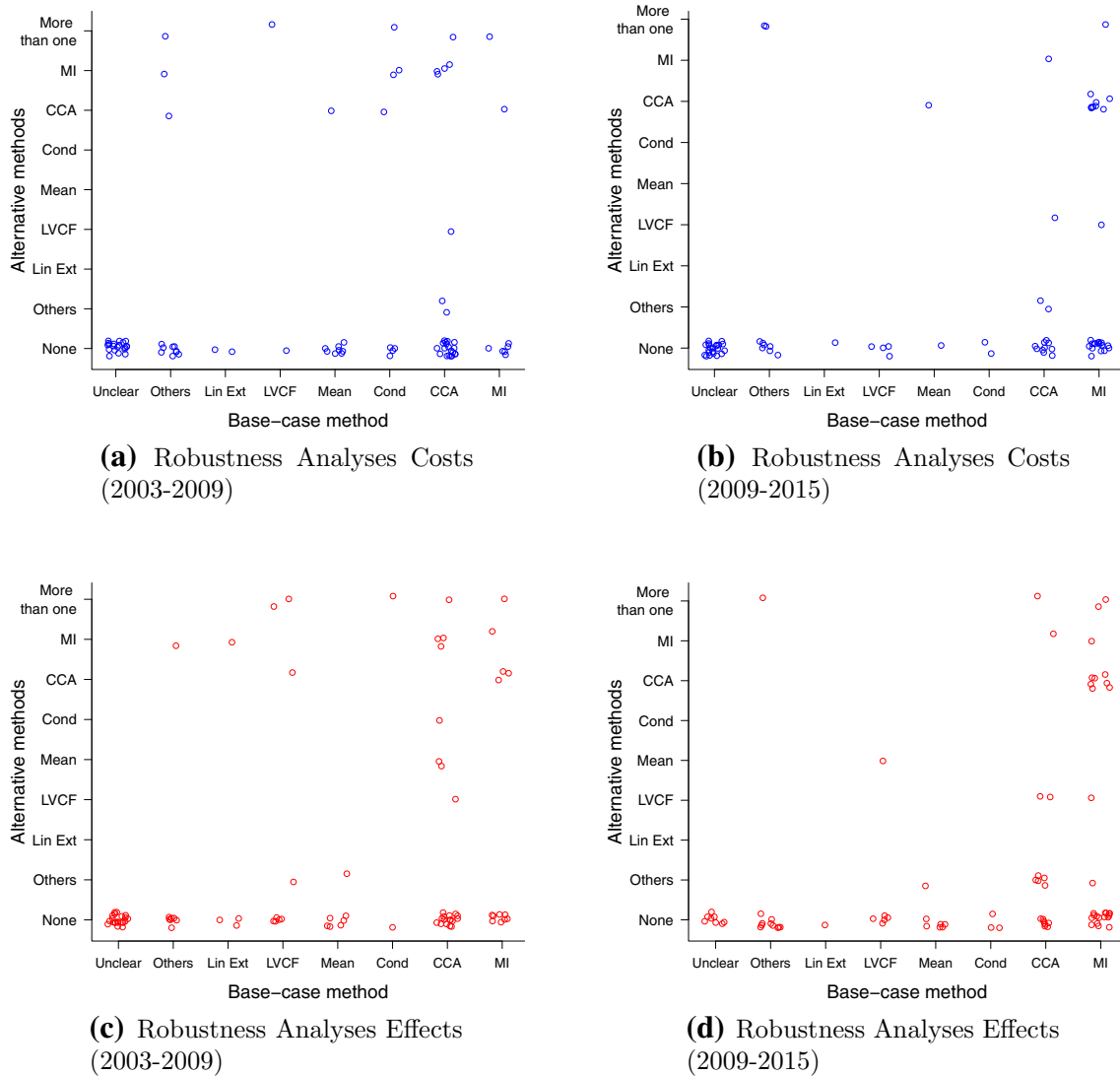


Fig. 4 Comparison of methods used in the base-case analysis (x axis) and those used as alternatives in a robustness analysis (y axis) for the articles between 2003–2009 and 2009–2015 for missing costs and effects. Legend: unspecified methods (Unclear), other methods (Others), linear extrapolation (Lin Ext), last value carried forward (LVCF), mean imputation (Mean), conditional imputation (Cond), complete case analysis (CCA), multiple imputation (MI). The

category “Unclear” includes those articles for which it was not possible, based on the text, to understand the methodology used to deal with the missingness, while the category “Others” consists of the following methods: random draw, linear mixed-effects model, expectation-maximisation algorithm, input-case analysis, assumed zeros, two-part regression

results and how these are translated into the chosen method. Details of our evaluation of both aspects are provided next, starting with the strength of assumptions.

We group the methods into five categories, broadly ordered according to the strength of the associated missingness assumptions. These are: single imputation (SI); complete case analysis; multiple imputation; and unknown (UNK), a residual group in which we classify studies that do not explicitly mention the method used. We associate this class with the strongest level of assumptions, since the lack of any method description may implicitly suggest (over-)confidence in a small effect of missingness on the results. By

contrast, we define sensitivity analysis (SA) as the least restrictive approach, which can assess the robustness of the results to different alternative missing data assumptions.

Figure 5 gives a graphical representation of both aspects for the articles reviewed between 2009–2015 in terms of the assumptions and justifications (quality scores) on missingness. In both graphs, more studies lie in the lower than in the upper part, indicating that fewer studies can be classified as high quality in terms of the considerations about missingness. This is highlighted by a greater concentration of points at the bottom of the figures (grade E). As we move along the vertical axis, this tends to reduce up

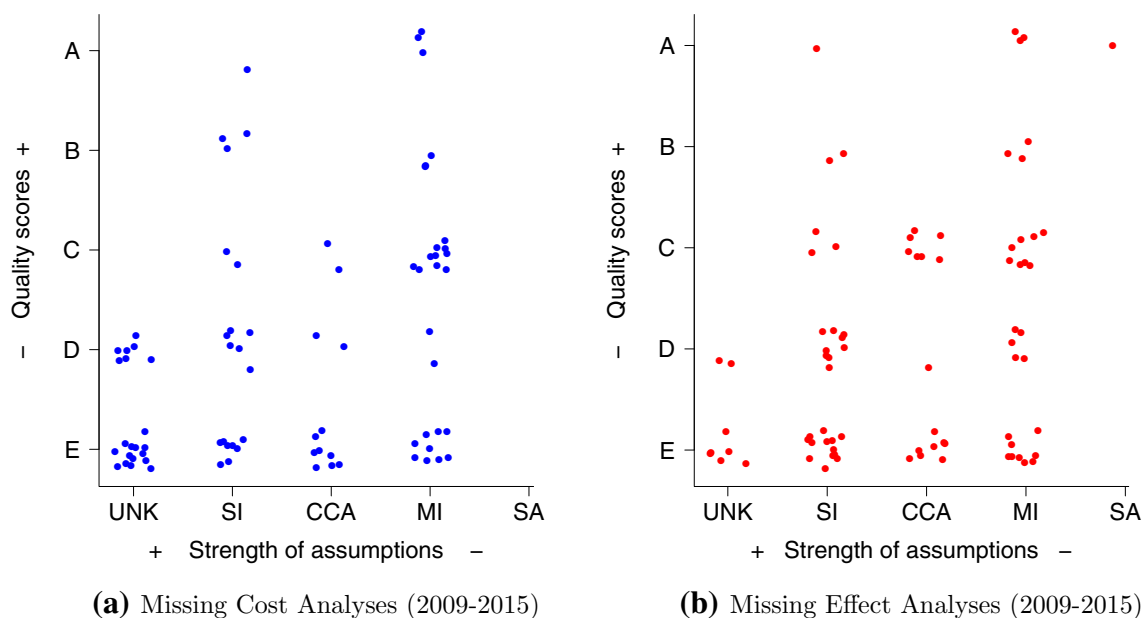


Fig. 5 Joint assessment, in the reviewed articles between 2009–2015, for missing costs and effects, of two components. The *x*-axis is the missingness method assumptions: unknown (UNK), single imputation (SI), complete case analysis (CCA), multiple imputation (MI) and

sensitivity analysis (SA). The *y*-axis is the ordered classification for the quality judgement (scores) to support these assumptions: *E, D, C, B, A*

to the top level (grade A), where only very few cases are shown. Of particular interest is the (almost) total absence of articles that performed a sensitivity analysis (SA), clearly indicating very slow uptake of this technique.

A shift along the vertical axis in the graphs indicates an increase in the level of understanding about the implications for the results of different choices of the missing data assumptions. Therefore, we can argue that an upward movement in the plot will always improve the justification of a specific assumption. However, to be able to follow this path, we may have to rely on more sophisticated methods that can match the given missingness assumption; e.g. if we think our data are MNAR, then CCA assumptions are less likely to hold. The aim of an optimal analysis should be to select a method that can be fully justified by matching the description of the missing data problem to the assumptions underpinning the chosen method, i.e. map onto the upper section of the graphs.

As a concrete example about the importance of exploring different missingness assumptions in terms of the impact they may have on the CEA, we consider one of the reviewed studies that has been graded as “A” by our scheme [187]. The authors provide an assessment of the probability of accepting a given treatment against a comparator (CEAC) for different willingness to pay thresholds and missingness methods. For all the thresholds considered, substantial uncertainty is reflected by significant variations in the CEAC values according to the different missing data methods used. Specifically, the incorporation of external information leads

to missingness assumptions that significantly affect the uncertainty in the results, producing a decrease in the value of the CEAC, at the target willingness to pay threshold, from 71% in the base-case (CCA) to 53% in one alternative scenario (MI). This example should encourage authors to recognise the importance that a comprehensive examination of missingness via SA may have on the uncertainty around CEA conclusions.

7 Discussion

The objective of this paper is to critically appraise the issue of missing data analysis in within-trial CEAs. In addition, we aim at providing a set of recommendations to guide future studies towards a more principled handling and reporting of missingness. It is important that assumptions about missing data are clearly stated and justified. Sensitivity analysis is also important, in order to explore the impact of plausible alternative missing data assumptions on the results of the CEA. Often, a variety of techniques and analyses are used but not reported because of space limits; on-line appendices and supplementary material could be used to report these alternatives.

7.1 Descriptive Review

Figure 3 highlights a shift in the most popular base-case missingness method from CCA to MI, between the two

periods of the review. The reasons behind this change may be related to some drawbacks of CCA and the relatively recent wide development of software to perform MI.

First, even under a strong missing data assumption such as MCAR, CCA remains inefficient because it ignores the predictive information contained in the partially observed cases. Non-negligible rates of missingness on a few variables of interest may cause large portions of the sample to be discarded. Second, CCA may cause serious biases in the parameter estimates. Indeed, the condition for validity of CCA does not fit neatly into Rubin's classes [38] in the important cases when: missing data affect the covariates; or the partially observed outcome has a longitudinal nature.

Arguably, a very important factor in the increasing popularity of MI is the recent availability of specific computer routines or packages (e.g. STATA or R; see [2]). This probably led to some abuse of the method as noted by Molenberghs et al. [4]. On the one hand, MI generally allows the inclusion of a larger number of variables/predictors in the imputation model than used in the analysis model, which potentially makes the assumption of MAR more plausible and thus the overall analysis less likely to be biased. On the other, the performance of MI depends on the correct specification of the imputation model (i.e. complexity in the analysis model is reflected in the imputation model) and care is required in its construction. Although essential, these details can be overlooked and are not often included in the reporting of the analysis, undermining its reliability.

From the comparison of the base-case methods used for the costs and effects between 2009 and 2015 (Fig. 3), we observe a marked reduction in the number of methods not clearly described for the effects, compared with those for the costs. A possible reason for this is that, while clinical effectiveness measures are often collected through self-reported questionnaires, which are naturally prone to missingness, cost measures rely more on clinical patient files, which may ensure a higher completeness rate. It was not possible to confirm this interpretation in the reviewed studies due to the high proportions of articles not clearly reporting the missing rates in both 2003–2009 and 2009–2015 periods, for effect (≈ 45 and $\approx 38\%$) and cost data (≈ 50 and $\approx 62\%$). In addition, clinical outcomes are almost invariably the main objective of RCTs, and as such they are usually subject to more advanced and standardised analyses. Arguably, costs are often considered as an add-on to the standard trial: for instance, sample size calculations are almost always performed with the effectiveness measure as the only outcome of interest. Consequently, missing data methods are less frequently well thought through for the analysis of the costs. However, this situation is likely to change as cost data from different perspectives (e.g.

caregivers, patients, society, etc.) are being increasingly used in trials, leading to the more frequent adoption of self-report cost data, which may start to exhibit similar missingness characteristics to effect data.

Our review identified only a few articles using more than one alternative method (Fig. 4). This situation indicates a gap in the literature associated with an under-implementation of sensitivity analyses, which may significantly affect the whole decision-making process outcome, under the perspective of a body who is responsible for providing recommendations about the implementation of alternative interventions for health care matters.

Limiting the assessment of missingness assumptions to a single case is unlikely to provide a reliable picture of the underlying mechanism. This, in turn, may have a significant impact on the CEA and mislead its conclusions, suggesting the implementation of non-cost-effective treatments, whose definition will depend on the assumption the modeller thinks is more likely to hold in the given setting. Sensitivity analysis represents an important tool to properly account for more structured uncertainty related to the missing data, and its implementation may provide a more realistic picture of the impact that the assumptions have on the final conclusions.

7.2 Quality Assessment

Despite the fact that our quality evaluation scheme has not been tested/validated by independent assessors, we believe Sect. 4.1 provides some reasonable justification for its application to the reviewed articles. Generally speaking, most papers in our review achieved an unsatisfactory quality score under our classification (Fig. 5). Indeed, our benchmark area on the top-right corner of the graphs is barely reached by less than 7% of the articles, for both cost and effect data. The opportunity of reaching such a target might be precluded by the choice of the method adopted, which may not be able to support less restrictive assumptions about missingness, even when this would be desirable. As a result, when simple methods cannot be fully justified, it is necessary to replace them with more flexible ones that can relax assumptions and incorporate more alternatives. In settings such as those involving MNAR, sensitivity analysis might represent the only possible approach to account for the uncertainty due to the missingness in a principled way. However, due to the lack of studies either performing a SA or providing high quality scores on the assumptions, we argue that missingness is not adequately addressed in most studies. This could have the serious consequence of imposing too restrictive assumptions about missingness and affect the outcome of the decision-making process.

We suggest that authors may use Table 1 as a convenient tool to check that all relevant information on missing data is taken into account in determining the assumptions in the analysis. All the key considerations should be fully satisfied, if possible. These criteria summarise previously published missing data recommendations from various settings, drawing them together within a general, simple and easy-to-read checklist table.

The classification of the studies into ordered categories (Fig. 2) according to the information provided on missing data (Table 1) is potentially a valuable tool for meta-analysis. The proposed quality evaluation scheme could be used by analysts to assign scores and grade individual studies based on their overall quality level in terms of missingness handling. These grades could then be taken into account in assigning different weights to the individual studies within the meta-analysis framework, in order to reflect a different degree of confidence in their results.

7.3 Conclusions

Given the common problem of missing effect and cost data in within-trial CEAs, many study conclusions could be based on imprecise economic evidence. This is a potentially serious issue for bodies such as the National Institute for Health and Care Excellence (NICE) who use these evaluations in their decision making, thus possibly leading to incorrect policy decisions about the cost-effectiveness of new treatment options.

Our review shows, over time, a significant change from more to less restrictive methods in terms of the assumptions on the missingness mechanism. This is an encouraging movement towards a more suitable and careful missing data analysis. Nevertheless, improvements are still needed, as only a small number of articles provide transparent information or perform a sensitivity analysis.

Our guidelines could represent a valuable tool to improve missing data handling. By carefully thinking about each component in the analysis, we are forced to explicitly consider all the assumptions we make about missingness and assess the impact of their variation on final conclusions. The main advantage is a more comparable formalisation of the uncertainty as well as a better indication of possible issues in assessing the cost-effectiveness of new treatments.

Acknowledgements We would like to acknowledge the journal reviewers and editor for their important contributions that led to the final version of the article.

Author contributions PhD student Andrea Gabrio gathered the reviewed articles, prepared the literature review and the manuscript, and contributed to the construction of the recommendations. Dr. Gianluca Baio and Dr. Alexina J. Mason assisted with the manuscript

preparation and contributed to the construction of the recommendations and preparation of the review. All authors contributed to and approved the final version of the manuscript.

Compliance with Ethical Standards

Funding PhD student Andrea Gabrio is partially funded by a research grant sponsored by The Foundation BLANCEFLOR Boncompagni Ludovisi, née Bildt. Dr. Gianluca Baio is partially funded by a research grant sponsored by Mapi.

Conflict of interest PhD student Andrea Gabrio, Dr. Gianluca Baio and Dr. Alexina J. Mason have no conflicts of interest to declare that are directly relevant to the content of this article.

Data availability statement The datasets generated during and/or analysed as part of the current study are available on the UCL Statistics for Health Economics evaluation group website [<http://www.ucl.ac.uk/statistics/research/statistics-health-economics/current-projects/ag>].

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.
2. Schafer JL. Analysis of incomplete multivariate data. New York: Chapman and Hall; 1997.
3. Little RJA, D'Agostino R, Dickersin K, Emerson SS, Farrar JT, Frangakis C, Hogan JW, Molenberghs G, Murphy SA, Neaton JD, Rotnitzky A, Scharfstein D, Shih W, Siegel JP, Stern H. The prevention and treatment of missing data in clinical trials. Panel on handling missing data in clinical trials. Committee on National Statistics, Division of Behavioral and Social Sciences and Education; 2010.
4. Molenberghs G, Fitzmaurice G, Kenward MG, Tsiatis A, Verbeke G. Handbook of missing data methodology. New York: Chapman and Hall; 2015.
5. Manca P, Palmer S. Handling missing values in cost effectiveness analyses that use data from cluster randomized trials. *Appl Health Econ Health Policy*. 2006;4:65–75.
6. Marshall A, Billingham LJ, Bryan S. Can we afford to ignore missing data in cost-effectiveness analyses? *Eur J Health Econ*. 2009;10:1–3.
7. Graves N, Walker D, Raine R, Hutchings A, Roberts JA. Cost data for individual patients included in clinical studies: no amount of statistical analysis can compensate for inadequate costing method. *Health Econ*. 2002;11:735–739
8. Briggs A, Clark T, Wolstenholme J, Clarke P. Missing. Presumed at random: cost-analysis of incomplete data. *Health Econ*. 2003;12:377–392
9. Oostenbrink JB, Al MJ, Rutten-van Molken PMH. Methods to analyse cost data of patients who withdraw in a clinical trial setting. *PharmacoEconomics*. 2003;21:1003–12.

10. Oostenbrink JB, Al MJ. The analysis of incomplete cost data due to dropout. *Health Econ.* 2005;14:763–76.
11. Burton A, Billingham LJ, Bryan S. Cost-effectiveness in clinical trials: using multiple imputation to deal with incomplete cost data. *Clin Trials.* 2007;4:154–61.
12. Lambert PC, Billingham LJ, Cooper NJ, Sutton AJ, Abrams KR. Estimating the cost-effectiveness of an intervention in a clinical trial when partial cost information is available: a Bayesian approach. *Health Econ.* 2008;17:67–81.
13. Richardson G, Manca A. Calculation of quality adjusted life years in the published literature: a review of methodology and transparency. *Health Econ.* 2004;13:1203–10.
14. Wood AM, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clin Trials.* 2004;1:368–376.
15. Groenwold RHH, Rogier A, Donders T, Roes KCB, Harrell FE, Moons KGM. Dealing with missing outcome data in randomized trials and observational studies. *Am J Epidemiol.* 2012;175:210–217.
16. Powney M, Williamson P, Kirkham J, Kolarnunnage-Dona R. Multiple imputation to deal with missing EQ-5D-3L data: should we impute individual domains or the actual index? *Trials.* 2014;15.
17. Simons CL, Arias OR, Yu LM, Simon J. Multiple imputation to deal with missing EQ-5D-3L data: Should we impute individual domains or the actual index? *Qual Life Res.* 2015;24:805–15.
18. Rombach I, Rivero-Arias O, Gray AM, Jenkinson C, Burke O. The current practice of handling and reporting missing outcome data in eight widely used PROMs in RCT publications: a review of the current literature. *Qual Life Res.* 2016.
19. Harkanen T, Maljanen T, Lindfors O, Virtala E, Knekt P. Confounding and missing data in cost-effectiveness analysis: comparing different methods. *Health Econ Rev.* 2013;3.
20. Diaz-Ordaz K, Kenward MG, Grieve R. Handling missing values in cost effectiveness analyses that use data from cluster randomized trials. *J R Stat Soc.* 2014;177:457–474.
21. Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *Pharmacoeconomics.* 2014;32:1157–70.
22. Noble SM, Hollingworth W, Tilling K. Missing data in trial-based cost-effectiveness analysis: the current state of play. *Health Econ.* 2012;21:187–200.
23. Diaz-Ordaz K, Kenward MG, Cohen A, Coleman CL, Eldridge S. Are missing data adequately handled in cluster randomised trials? a systematic review and guidelines. *Clin Trials.* 2014;11:590–600.
24. Hughes D, Charles J, Dawoud D, Edwards RT, Holmes E, Jones C, Parham P, Plumpton C, Ridyard C, Lloyd-Williams H, Wood E, Yeo ST. Conducting economic evaluations alongside randomised trials: current methodological issues and novel approaches. *Pharmacoeconomics.* 2016;34:447–61.
25. Schafer JL, Graham JW. Missing data: Our view of the state of the art. *Psychol Methods.* 2002;7:147–77.
26. Shao J, Zhong B. Last observation carry-forward and last observation analysis. *Stat Med.* 2003;22:2429–41.
27. Twisk J, de Vente W. Attrition in longitudinal studies: How to deal with missing data. *J R Stat Soc.* 2002;55:329–37.
28. Buck SF. A method of estimation for missing values in multivariate data suitable for use with an electronic computer. *J R Stat Soc.* 1960;22:302–306.
29. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res* 1999;8:3–15.
30. Daniels MJ, Hogan JW. Missing data in longitudinal studies: strategies for Bayesian modeling and sensitivity analysis. New York: Chapman and Hall; 2008.
31. Schafer JL, Yucel RM. Computational strategies for multivariate linear mixed-effects models with missing values. *J Comput Graph Stat.* 2002;11:437–57.
32. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc.* 1977;39:1–38.
33. Robins JM, Rotnitzky A. Recovery of information and adjustment for dependent censoring using surrogate markers. *AIDS epidemiology: methodological issues*, Boston: Birkhauser; 1992.
34. Tanner MA, Wong WH. The calculation of posterior distributions by data augmentation. *J Am Stat Assoc.* 1987;82:528–50.
35. McLachlan GL, Krishnan T. *The EM Algorithm and Extensions*. Hoboken: Wiley; 2008.
36. Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. *Biometrics.* 2005;61:962–73.
37. Mason A, Richardson S, Plewis I, Best N. Strategy for modelling nonrandom missing data mechanisms in observational studies using Bayesian methods. *J Off Stat.* 2012;28:279–302.
38. White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. *Stat Med.* 2010;29:2920–31.
39. Van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *J Stat Soft.* 2011;45:1–67.
40. Baio G. *Bayesian methods in health economics*. London: Chapman and Hall/CRC, University College London; 2013.

Review 2003–2009

41. Araya R, Flynn T, Rojas G, Fritsch R, Simon G. Cost-effectiveness of a primary care treatment program for depression in low-income women in Santiago, Chile. *Am J Psychiatry.* 2006;163:1379–87.
42. Athobari J, Asselbergs F, Boersma C, De Vries R, Hillege H, Van Gilst W, Gansevoort R, De Jong P, de Jong-van den Berg L, Postma M. Cost-effectiveness of screening for albuminuria with subsequent foscipril treatment to prevent cardiovascular events: a pharmacoeconomic analysis linked to the prevention of renal and vascular endstage disease (PREVEND) study and the prevention of renal and vascular endstage disease intervention trial (PREVEND IT). *Clin Ther.* 2006;28:432–44.
43. Barrett B, Byford S, Crawford M, Patton R, Drummond C, Henry J, Touquet R. Cost-effectiveness of screening and referral to an alcohol health worker in alcohol misusing patients attending an accident and emergency department: a decision-making approach. *Drug Alcohol Depend.* 2006;81:47–54.
44. Bos I, Hoving J, van Tulder M, Molken M, Ader H, de Vet H, Koes B, Vondeling H, Bouter L, Mullner M. Cost effectiveness of physiotherapy, manual therapy, and general practitioner care for neck pain: economic evaluation alongside a randomised controlled trial commentary: Bootstrapping simplifies appreciation of statistical inferences. *BMJ.* 2003;326:911.
45. Brouwers E, Bruijne M, Terluin B, Verhaak P. Cost-effectiveness of an activating intervention by social workers for patients with minor mental disorders on sick leave: a randomized controlled trial. *Eur J Public Health.* 2007;17:214–20.
46. Burton A, Billingham L, Bryan S. Cost-effectiveness in clinical trials: using multiple imputation to deal with incomplete cost data. *Clin Trials.* 2007;4:154–61.
47. Byford S, Barrett B, Roberts C, Clark A, Edwards V, Smethurst N, Gowers S. Economic evaluation of a randomised controlled trial for anorexia nervosa in adolescents. *Br J Psychiatry.* 2007;191:436–40.
48. Byford S, Barrett B, Roberts C, Wilkinson P, Dubicka B, Kelvin R, White L, Ford C, Breen S, Goodyer I. Cost-effectiveness of

- selective serotonin reuptake inhibitors and routine specialist care with and without cognitive behavioural therapy in adolescents with major depression. *Br J Psychiatry*. 2007;191:521–7.
49. Coast J, Noble A, Horrocks S, Asim O, Peters T, Salisbury C. Economic evaluation of a general practitioner with special interests led dermatology service in primary care. *BMJ*. 2005;331:1444–9.
 50. Coupe V, Veenhof C, van Tulder M, Dekker J, Bjlmsa J, Van den Ende C. The cost effectiveness of behavioural graded activity in patients with osteoarthritis of hip and/or knee. *Ann Rheum Dis*. 2007;66:215–21.
 51. Delaney B, Qume M, Moayyedi P, Logan R, Ford A, Elliott C, McNulty C, Wilson S, Hobbs F. *Helicobacter pylori* test and treat versus proton pump inhibitor in initial management of dyspepsia in primary care: multicentre randomised controlled trial (MRC-CUBE trial). *BMJ*. 2008;336:651–4.
 52. Dennis M, Godley S, Diamond G, Tims F, Babor T, Donaldson J, Liddle H, Titus J, Kaminer Y, Webb C, Hamilton N, Funk R. The cannabis youth treatment (CYT) study: main findings from two randomized trials. *J Subst Abuse Treat*. 2004;27:197–213.
 53. Djkgraaf M, van der Zanden B, de Borgie C, Blanken P, van Ree J, van den Brink W. Cost utility analysis of co-prescribed heroin compared with methadone maintenance treatment in heroin addicts in two randomised trials. *BMJ*. 2005;330:1297.
 54. Dornelas E, Magnavita J, Beazoglou T, Fischer E, Oncken C, Lando H, Greene J, Barbagallo J, Stepnowski R, Gregonis E. Efficacy and cost-effectiveness of a clinic-based counseling intervention tested in an ethnically diverse sample of pregnant smokers. *Patient Educ Couns*. 2006;64:342–9.
 55. Drummond M, Becker D, Hux M, Chancellor J, Duprat-Lomon I, Kubin R, Sagnier P. An economic evaluation of sequential iv/po moxifloxacin therapy compared to iv/po co-amoxiclav with or without clarithromycin in the treatment of community-acquired pneumonia. *Chest*. 2003;124:526–35.
 56. Edwards R, Ceilleachair A, Bywater T, Hughes D, Hutchings J. Parenting programme for parents of children at risk of developing conduct disorder: cost effectiveness analysis. *BMJ*. 2007;334:682.
 57. Emmons K, Puleo E, Park E, Gritz E, Butterfield R, Weeks J, Mertens A, Li F. Peer-delivered smoking counseling for childhood cancer survivors increases rate of cessation: the partnership for health study. *J Clin Oncol*. 2005;23:6516–23.
 58. Fals-Stewart W, Klostermann K, Yates B, O'Farrell T, Birchler G. Brief relationship therapy for alcoholism: a randomized clinical trial examining clinical efficacy and cost-effectiveness. *Psychol Addict Behav*. 2005;19:363–71.
 59. Fals-Stewart W, Lam W. Brief behavioral couples therapy for drug abuse: a randomized clinical trial examining clinical efficacy and cost-effectiveness. *Fam Syst Health*. 2008;26:377–92.
 60. Furze G, Dumville J, Miles J, Irvine K, Thompson D, Lewin R. "Prehabilitation" prior to CABG surgery improves physical functioning and depression. *Int J Cardiol*. 2009;132:51–8.
 61. Gilbert F, Grant A, Gillan M, Vale L, Campbell M, Scott N, Knight D, Wardlaw D. Low back pain: influence of early MR imaging or CT on treatment and outcome—multicenter randomized trial. *Radiology*. 2004;231:343–51.
 62. Goodacre S, Nicholl J, Dixon S, Cross E, Angelini K, Arnold J, Revill S, Locker T, Capewell S, Quinney D, Campbell S, Morris F. Randomised controlled trial and economic evaluation of a chest pain observation unit compared with routine care. *BMJ*. 2004;328:254.
 63. Graff M, Adang E, Vernooij-Dassen M, Dekker J, Jonsson L, Thijssen M, Hoefnagels W, Rikkert M. Community occupational therapy for older patients with dementia and their care givers: cost effectiveness study. *BMJ*. 2008;336:134–8.
 64. Group, A. C. Long-term donepezil treatment in 565 patients with Alzheimer's disease (AD2000): randomised double-blind trial. *Lancet*. 2004;363:2105–15.
 65. Haddock G, Barrowclough C, Tarrier N, Moring J, O'Brien R, Schofield N, Quinn J, Palmer S, Davies L, Lowens I, McGovern J, Lewis S. Cognitive-behavioural therapy and motivational intervention for schizophrenia and substance misuse: 18-month outcomes of a randomised controlled trial. *Br J Psychiatry*. 2003;183:418–26.
 66. Hartman M, van Ede A, Severens J, Laan R, van de Putte L, van der Wilt G. Economic evaluation of folate supplementation during methotrexate treatment in rheumatoid arthritis. *J Rheumatol*. 2004;31:902–8.
 67. Hollinghurst S, Redmond N, Costelloe C, Montgomery A, Fletcher M, Peters T, Hay A. Paracetamol plus ibuprofen for the treatment of fever in children (PITCH): economic evaluation of a randomised controlled trial. *BMJ*. 2008;337:1490.
 68. Hollinghurst S, Sharp D, Ballard K, Barnett J, Beattie A, Evans M, Lewith G, Middleton K, Oxford F, Webley F, Little P. Randomised controlled trial of Alexander technique lessons, exercise, and massage (ATEAM) for chronic and recurrent back pain: economic evaluation. *BMJ*. 2008;337:2656.
 69. Hollis J, McAfee T, Fellows J, Zbikowski S, Stark M, Riedlinger K. The effectiveness and cost effectiveness of telephone counselling and the nicotine patch in a state tobacco quitline. *Tob Control*. 2007;16:53–9.
 70. Hurskainen R, Teperi J, Rissanen P, Aalto A, Grenman S, Kivela A, Kujansuu E, Vuorma S, Yliskoski M, Paavonen J. Clinical outcomes and costs with the levonorgestrel-releasing intrauterine system or hysterectomy for treatment of menorrhagia: randomized trial 5-year followup. *J Am Med Assoc*. 2004;291:1456–63.
 71. Jones K, Colson P, Holter M, Lin S, Valencia E, Susser E, Wyatt R. Cost-effectiveness of critical time intervention to reduce homelessness among persons with mental illness. *Psychiatr Serv*. 2003;54:884–90.
 72. Katon W, Schoenbaum M, Fan M, Callahan C, Williams J, Hunkeler E, Harpole L, Zhou X, Langston C, Unutzer J. Cost-effectiveness of improving primary care treatment of late-life depression. *Arch Gen Psychiatry*. 2006;62:1313–20.
 73. Katon W, Unutzer J, Fan M, Williams J, Schoenbaum M, Lin E, Hunkeler E. Cost-effectiveness and net benefit of enhanced treatment of depression for older adults with diabetes and depression. *Diabetes Care*. 2006;29:265–70.
 74. Kattan M, Stearns S, Crain E, Stout J, Gergen P, Evans III R, Visness C, Gruchalla R, Morgan W, O'Connor G, Mastin J, Mitchell H. Cost-effectiveness of a home-based environmental intervention for inner-city children with asthma. *J Allergy Clin Immunol*. 2005;116:1058–63.
 75. Kendrick T, Peveler R, Longworth L, Baldwin D, Moore M, Chatwin J, Thornett A, Goddard J, Campbell M, Smith H, Buxton M, Thompson C. Cost-effectiveness and cost-utility of tricyclic antidepressants, selective serotonin reuptake inhibitors and lofepramine: randomised controlled trial. *Br J Psychiatry*. 2006;188:337–45.
 76. Kendrick T, Simons L, Mynors-Wallis L, Gray A, Lathlean J, Pickering R, Harris S, Rivero-Arias O, Gerard K, Thompson C. Cost-effectiveness of referral for generic care or problem-solving treatment from community mental health nurses, compared with usual general practitioner care for common mental disorders: randomised controlled trial. *Br J Psychiatry*. 2006;189:50–9.
 77. Kilonzo M, Vale L, Cook J, Milne A, Stephen A, Avenell A. A cost-utility analysis of multivitamin and multimineral supplements in men and women aged 65 years and over. *Clin Nutr*. 2007;26:364–70.

78. Kuyken W, Byford S, Taylor R, Watkins E, Holden E, White K, Barrett B, Byng R, Evans A, Mullan E, Teasdale J. Mindfulness-based cognitive therapy to prevent relapse in recurrent depression. *J Consult Clin Psychol*. 2008;76:966–78.
79. Lam D, McCrone P, Wright K, Kerr N. Cost-effectiveness of relapse-prevention cognitive therapy for bipolar disorder: 30-month study. *Br J Psychiatry*. 2005;186:500–6.
80. Lewis M, James M, Stokes E, Hill J, Sim J, Hay E, Dziedzic K. An economic evaluation of three physiotherapy treatments for non-specific neck disorders alongside a randomized trial. *Rheumatology*. 2007;46:1701–8.
81. Manca A, Dumville J, Toregerson D, Klaber Moffett J, Mooney M, Jackson D, Eaton S. Randomized trial of two physiotherapy interventions for primary care back and neck pain patients: cost effectiveness analysis. *Rheumatology*. 2007;46:1495–501.
82. Manca A, Sculpher M, Ward K, Hilton P. A cost-utility analysis of tension-free vaginal tape versus colposuspension for primary urodynamic stress incontinence. *Int J Obstet Gynaecol*. 2003;110:255–62.
83. Mandelblatt J, Cullen J, Lawrence W, Stanton A, Yi B, Kwan L, Ganz P. Economic evaluation alongside a clinical trial of psycho-educational interventions to improve adjustment to survivorship among patients with breast cancer. *J Clin Oncol*. 2008;26:1684–90.
84. Maniadakis N, Dafni U, Fragoulakis V, Grimani I, Galani E, Fragkoulidi A, Fountzilias G. Economic evaluation of taxane-based first-line chemotherapy in the treatment of patients with metastatic breast cancer in Greece: an analysis alongside a multicenter, randomized phase III clinical trial. *Ann Oncol*. 2009;20:278–85.
85. Marson A, Al-Kharusi A, Alwaidh M, Appleton R, Baker G, Chadwick D, Cramp C, Cockerell O, Cooper P, Doughty J, Eaton B, Gamble C, Goulding P, Howell S, Hughes A, Jackson M, Jacoby A, Kellett M, Lawson G, Leach J, Licolaides P, Roberts R, Shackley P, Shen J, Smith D, Smith P, Smith C, Vanoli A, Williamson P. The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: an unblinded randomised controlled trial. *Lancet*. 2003;363:1016–26.
86. Marson A, Al-Kharusi A, Alwaidh M, Appleton R, Baker G, Chadwick D, Cramp C, Cockerell O, Cooper P, Doughty J, Eaton B, Gamble C, Goulding P, Howell S, Hughes A, Jackson M, Jacoby A, Kellett M, Lawson G, Leach J, Licolaides P, Roberts R, Shackley P, Shen J, Smith D, Smith P, Smith C, Vanoli A, Williamson P. The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: an unblinded randomised controlled trial. *Lancet*. 2003;363:1000–15.
87. Mary Davies L, Anne Fargher E, Tricker K, Dawnes P, Scott D, Symmons D. Is shared care with annual hospital review better value for money than predominantly hospital-based care in patients with established stable rheumatoid arthritis? *Ann Rheum Dis*. 2007;66:658–63.
88. McCrone P, Knapp M, Proudfoot J, Ryden C, Cavanagh K, Shapiro D, Ilson S, Gray J, Goldberg D, Mann A, Marks I, Everitt B, Tylee A. Cost-effectiveness of computerised cognitive-behavioural therapy for anxiety and depression in primary care: randomised controlled trial. *Br J Psychiatry*. 2004;185:55–62.
89. McKenna C, Bojke L, Manca A, Adebajo A, Dickson J, Helliwell P, Morton V, Russell I, Torgerson D, Watson J (2009) Shoulder acute pain in primary health care: is retraining GPs effective? The SAPPHERE randomized trial: a cost-effectiveness analysis. *Rheumatology*.
90. Melis R, Adang E, Teerenstra S, van Eijken M, Wimo A, Achterberg T, Lisdonk E, Rikkert M. Multidimensional geriatric assessment: back to the future cost-effectiveness of a multidisciplinary intervention model for community-dwelling frail older people. *J Gerontol Ser A Biol Med Sci*. 2008;63:275–82.
91. Munro J, Nicholl J, Brazier J, Davey R, Cochrane T. Cost effectiveness of a community based exercise programme in over 65 year olds: cluster randomised trial. *J Epidemiol Community Health*. 2004;58:1004–10.
92. Najafzadeh M, Marra C, Sadatsafavi M, Aaron S, Sullivan S, Vandemheen K, Jones P, Fitzgerald J. Cost effectiveness of therapy with combinations of long acting bronchodilators and inhaled steroids for treatment of COPD. *Thorax*. 2008;63:962–7.
93. Nathoe H, van Dijk D, Jansen E, Suyker W, Diephuis J, van Boven WJ, de la Riviere A, Borst C, Kalkman C, Grobbee D, Buskens E, de Jaegere PPT. A comparison of on-pump and off-pump coronary bypass surgery in low-risk patients. *N Engl J Med*. 2003;348:394–402.
94. Noyes K, Dick A, Holloway R. Pramipexole v. levodopa as initial treatment for Parkinson's disease: a randomized clinical-economic trial. *Med Decis Mak*. 2004;24:472–85.
95. Olmstead T, Sindelar J, Petry N. Cost-effectiveness of prize-based incentives for stimulant abusers in outpatient psychosocial treatment programs. *Drug Alcohol Depend*. 2007;87:175–82.
96. Olsson A, Casciano R, Stern L, Svangren P. A pharmacoeconomic evaluation of aggressive cholesterol lowering in Sweden. *Int J Cardiol*. 2004;96:51–7.
97. Oosternbrink J, Rutten-van Molken M, Al M, Van Noord J, Vincken W. One-year Cost-effectiveness of tiotropium versus ipratropium to treat chronic obstructive pulmonary disease. *Eur Respir J*. 2004;23:241–9.
98. O'Reilly J, Lowson K, Young J, Forster A, Green J, Small N. A cost effectiveness analysis within a randomised controlled trial of post-acute care of older people in a community hospital. *BMJ*. 2006;333:228.
99. Patel A, Knapp M, Evans A, Perez I, Kalra L. Training care givers of stroke patients: economic evaluation. *BMJ*. 2004;328:1102.
100. Petrou S, Bischof M, Bennett C, Elbourne D, Field D, McNally H. Cost-effectiveness of neonatal extracorporeal membrane oxygenation based on 7-year results from the United Kingdom collaborative ECMO trial. *Pediatrics*. 2006;117:1640–9.
101. Prinssen M, Buskens E, de Jong S, Buth J, Mackaay A, Sambeek M, Blankensteijn J. Cost-effectiveness of conventional and endovascular repair of abdominal aortic aneurysms: results of a randomized trial. *J Vasc Surg*. 2007;46:883–90.
102. Raftery J, Yao G, Murchie P, Campbell N, Ritchie L. Cost effectiveness of nurse led secondary prevention clinics for coronary heart disease in primary care: follow up of a randomised controlled trial. *BMJ*. 2005;330:707.
103. Ratcliffe J, Thomas K, MacPherson H, Brazier J. A randomised controlled trial of acupuncture care for persistent low back pain: cost effectiveness analysis. *BMJ*. 2006;333:626.
104. Reed S, Radeva J, Glendenning G, Saad F, Schulman K. Cost-effectiveness of zoledronic acid for the prevention of skeletal complications in patients with prostate cancer. *J Urol*. 2004;171:1537–42.
105. Revicki D, Siddique J, Frank L, Chung J, Green B, Krupnick J, Prasad M, Miranda J. Cost-effectiveness of evidence-based pharmacotherapy or cognitive behavior therapy compared with community referral for major depression in predominantly low-income minority women. *Arch Gen Psychiatry*. 2005;62:868–75.
106. Richardson G, Bloor K, Williamns J, Russell I, Durai D, Cheung W, Farrin A, Coulton S. Cost effectiveness of nurse delivered endoscopy: findings from randomised multi-institution nurse endoscopy trial (MINuET). *BMJ*. 2009;338:270.

107. Richardson G, Kennedy A, Reeves D, Bower P, Lee V, Middleton E, Gardner C, Gately C, Rogers A. Cost effectiveness of the expert patients programme (EPP) for patients with chronic conditions. *J Epidemiol Community Health*. 2008;62:361–7.
108. Richardson G, Sculpher M, Kennedy A, Nelson E, Reeves D, Roberts C, Robinson A, Rogers A, Thompson D. Is self-care a cost-effective use of resources? evidence from a randomized trial in inflammatory bowel disease. *J Health Serv Res Policy*. 2006;11:225–30.
109. Rocca H, Kaiser C, Bernheim A, Zellweger M, Jeger R, Buser P, Osswald S, Pfisterer M. Cost-effectiveness of drug-eluting stents in patients at high or low risk of major cardiac events in the Basel Stent KostenEffektivitäts Trial (BASKET): an 18-month analysis. *Lancet*. 2003;370:1552–9.
110. Roijen L, Van Straten A, Al M, Rutten F, Donker M. Cost-utility of brief psychological treatment for depression and anxiety. *Br J Psychiatry*. 2006;188:323–9.
111. Rosenheck R, KasproW W, Frisman L, Liu-Mares W. Cost-effectiveness of supported housing for homeless persons with mental illness. *Arch Gen Psychiatry*. 2003;60:940–51.
112. Schweikert B, Jacobi E, Seitz R, Cziške R, Ehlert A, Knab J, Leidl R. Effectiveness and Cost-effectiveness of adding a cognitive behavioral treatment to the rehabilitation of chronic low back pain. *J Rheumatol*. 2006;33:2519–26.
113. Scott J, Palmer S, Paykel E, Teasdale J, Hayhurst H. Use of cognitive therapy for relapse prevention in chronic depression: cost-effectiveness study. *Br J Psychiatry*. 2003;182:221–7.
114. Seivewright H, Green J, Salkovskis P, Barrett B, Nur U, Tyrer P. Cognitive-behavioural therapy for health anxiety in a genitourinary medicine clinic: randomised controlled trial. *Br J Psychiatry*. 2008;193:332–7.
115. Severens J, Prins J, van der Wilt G, van der Meer J, Bleijenberg G. Cost-effectiveness of cognitive behaviour therapy for patients with chronic fatigue syndrome. *Q J Med*. 2004;97:153–61.
116. Sevick M, Napolitano M, Papandonatos G, Gordon A, Reiser L, Marcus B. Cost-effectiveness of alternative approaches for motivating activity in sedentary adults: results of Project STRIDE. *Prev Med*. 2007;45:54–61.
117. Simon J, Gray A, Clarke P, Wade A, Neil A, Farmer A. Cost effectiveness of self monitoring of blood glucose in patients with non-insulin treated type 2 diabetes: economic evaluation of data from the DIGEM trial. *BMJ*. 2008;336:1177–80.
118. Smeets R, Severens J, Beelen S, Vlaeyen J, Knottnerus J. More is not always better: Cost-effectiveness analysis of combined, single behavioral and single physical rehabilitation programs for chronic low back pain. *Eur J Pain*. 2009;13:71–81.
119. Smit F, Willemse G, Koopmanschap M, Onrust S, Cuijpers P, Beekman A. Cost-effectiveness of preventing depression in primary care patients: randomised trial. *Br J Psychiatry*. 2006;188:330–6.
120. Sullivan S, Buxton M, Andersson L, Lamm C, Liljas B, Chen Y, Pauwels R, Weiss K. Cost-effectiveness analysis of early intervention with budesonide in mild persistent asthma. *Allergy Clin Immunol*. 2003;112:1229–36.
121. Taimela S, Justen S, Aronen P, Sintonen H, Laara E, Malmivaara A, Tiekso J, Aro T. An occupational health intervention programme for workers at high risk for sickness absence. Cost effectiveness analysis based on a randomised controlled trial. *Occup Environ Med*. 2008;65:242–8.
122. Teng J, Mayo N, Latimer E, Hanley J, Wood-Dauphinee S, Cote R, Scott S. Costs and caregiver consequences of early supported discharge for stroke patients. *Stroke*. 2003;34:528–36.
123. Witt C, Jena S, Selim D, Brinkhaus B, Reinhold T, Wruck K, Liecker B, Linde K, Wegscheider K, Willich S. Pragmatic randomized trial evaluating the clinical and economic effectiveness of acupuncture for chronic low back pain. *Am J Epidemiol*. 2006;164:487–96.
124. Witt C, Reinhold T, Brinkhaus B, Roll S, Jena S, Willich S. Acupuncture in patients with dysmenorrhea: a randomized study on clinical effectiveness and cost-effectiveness in usual care. *Am J Obstet Gynecol*. 2008;198:166.
125. Wolfs C, Dirksen C, Kessels A, Severens J, Verhey F. Economic evaluation of an integrated diagnostic approach for psychogeriatric patients: results of a randomized controlled trial. *Arch Gen Psychiatry*. 2009;66:313–23.
126. Wonderling D, Vickers A, Grieve R, McCarney R. Cost effectiveness analysis of a randomised trial of acupuncture for chronic headache in primary care. *BMJ*. 2004;328:747.
127. Wu E, Birnbaum H, Mareva M, Le T, Robinson R, Rosen A, Gelwicks S. Cost-effectiveness of duloxetine versus routine treatment for U.S. patients with diabetic peripheral neuropathic pain. *J Pain*. 2006;7:399–407.
128. Zwanziger J, Hall W, Dick A, Zhao H, Mushlin A, Hahn R, Wang H, Andrews M, Mooney C, Wang H, Moss A. The cost effectiveness of implantable cardioverter-defibrillators: results from the Multicenter Automatic Defibrillator Implantation Trial (MADIT)-II. *J Am Coll Cardiol*. 2006;47:2310–8.

Review 2009–2015

129. Aasa M, Henriksson M, Dellborg M, Grip L, Herlitz J, Levin L, Svensson L, Janzon M. Cost and health outcome of primary percutaneous coronary intervention versus thrombolysis in acute ST-segment elevation myocardial infarction—results of the Swedish Early Decision Reperfusion Study (SWEDES) trial. *Am Heart J*. 2010;160:322–8.
130. Andersson E, Ljotsson B, Hedman E, Mattson S, Enander J, Andersson G, Kaldor V, Lindfors N, Ruck C. Cost-effectiveness of an internet-based booster program for patients with obsessivecompulsive disorder: Results from a randomized controlled trial. *J Obsessive Compuls Relat Disord*. 2015;4:14–9.
131. Asha S, Chan A, Walter E, Kelly P, Morton R, Ajami A, Wilson R, Honneyman D. Impact from point-of-care devices on emergency department patient processing times compared with central laboratory testing of blood samples: a randomised controlled trial and cost-effectiveness analysis. *Emerg Med J*. 2014;31:714–9.
132. Barton G, Fairall L, Bachmann M, Uebel K, Timmerman V, Lombard C, Zwarenstein M. Cost-effectiveness of nurse-led versus doctor-led antiretroviral treatment in South Africa: pragmatic cluster randomised trial. *Trop Med Int Health*. 2013;18:769–77.
133. Barton G, Hodjekins J, Mugford M, Jones P, Croudace T, Fowler D. Cognitive behaviour therapy for improving social recovery in psychosis: Cost-effectiveness analysis. *Schizophr Res*. 2009;112:158–63.
134. Barton G, Sach T, Jenkinson C, Doherty M, Avery A, Muir K. Lifestyle interventions for knee pain in overweight and obese adults aged over 45: economic evaluation of randomised controlled trial. *BMJ*. 2009;339:2273.
135. Berkhof F, Hesselink A, Vaessen D, Uil S, Kerstjens H, van der Berg J. The effect of an outpatient care on-demand-system on health status and costs in patients with COPD. A randomized trial. *Respir Med*. 2014;108:1163–70.
136. Bijen C, Vermeulen K, Mourits M, Arts H, ter Brugge H, van der Sijde R, Wijma J, Bongers M, van der Zee A, de Bock G. Cost effectiveness of laparoscopy versus laparotomy in early stage endometrial cancer: A randomised trial. *Gynecol Oncol*. 2011;121:76–82.

137. Campbell A, Nunes E, Miele G, Metthews A, Polsky D, Ghitza U, Turrigiano E, Bailey G, VanVeldhuisen P, Chapdelaine R, Froias A, Stitzer M, Carroll K, Winhusen T, Clingerman S, Perez L, McClure E, Goldman B, Crowell A. Design and methodological considerations of an effectiveness trial of a computer-assisted intervention: An example from the NIDA clinical trials network. *Contemp Clin Trials*. 2012;33:386–95.
138. Carr A, Cooper C, Campbell M, Rees J, Moser J, Beard D, Fitzpatrick R, Gray A, Dawson J, Murphy J, Bruhn H, Cooper D, Ramsay C. Clinical effectiveness and cost-effectiveness of open and arthroscopic rotator cuff repair [the UK Rotator Cuff Surgery (UKUFF) randomised trial]. *Health Technol Assess*. 2015;19.
139. CLOTS Trials Collaboration. Effect of intermittent pneumatic compression on disability, living circumstances, quality of life, and hospital costs after stroke: secondary analyses from CLOTS 3, a randomised trial. *Lancet Neurol*. 2014;13:1186–92.
140. Costa M, Achten J, Parsons N, Edlin R, Foguet P, Prakash U, Griffin D. Total hip arthroplasty versus resurfacing arthroplasty in the treatment of patients with arthritis of the hip joint: single centre, parallel group, assessor blinded, randomised controlled trial. *BMJ*. 2012;344.
141. Cuthbertson B, Rattray J, Campbell M, Gager M, Roughton S, Smith A, Hull A, Breeman S, Norrie J, Jenkinson D, Hernandez R, Johnston M, Wilson E, Waldmann C. The practical study of nurse led, intensive care follow-up programmes for improving long term outcomes from critical illness: a pragmatic randomised controlled trial. *BMJ*. 2009;339.
142. Dakin H, Wordsworth S, Rogers C, Abangma G, Raftery J, Harding S, Lotery A, Downes S, Chakravarthy U, Reeves, B. Cost-effectiveness of ranibizumab and bevacizumab for age-related macular degeneration: 2-year findings from the IVAN randomised trial. *BMJ*. 2014;4.
143. D'Amico F, Rehill A, Knapp M, Aguirre E, Donovan H, Hoare Z, Hoe J, Russell I, Spector A, Streater A, Whitaker C, Woods R, Orrell M. Maintenance cognitive stimulation therapy: An economic evaluation within a randomized controlled trial. *JAMDA*. 2015;16:63–70.
144. De Beurs D, Bosmans J, de Groot M, de Keijser J, van Duijn E, de Winter R, Kerkhof A. Training mental health professionals in suicide practice guideline adherence: cost-effectiveness analysis alongside a randomized controlled trial. *J Affect Disord*. 2015;186:203–10.
145. Domino M, Foster E, Vitiello B, Kratochvil C, Burns B, Silva S, Reinecke M, March J. Relative cost-effectiveness of treatments for adolescent depression: 36-week results from the TADS randomized trial. *J Am Acad Child Adolesc Psychiatry*. 2009;48.
146. Duarte A, Walker J, Walker S, Richardson G, Hansen C, Martin P, Murray G, Sculpher M, Sharpe M. Cost-effectiveness of integrated collaborative care for comorbid major depression in patients with cancer. *J Psychosom Res*. 2015;79:465–70.
147. Felker G, Ahmad T, Anstrom K, Adams K, Cooper L, Ezekowitz J, Fiuzat M, Houston-Miller N, Januzzi J, Leifer E, Mark D, Desvigne-Nickens P, Paynter G, Pina I, Whellan D, O'Connor C. Rationale and design of the GUIDE-IT study guiding evidence based therapy using biomarker intensified treatment in heart failure. *JACC: Heart Fail*. 2014;2:457–65.
148. Forster A, Dickerson J, Young J, Patel A, Kalra A, Nixon J, Smithard D, Knapp M, Holloway I, Anwar S, Farrin A. A structured training programme for caregivers of inpatients after stroke (TRACS): a cluster randomised controlled trial and cost-effectiveness analysis. *Lancet*. 2013;382:2069–76.
149. Fuller N, Colagiuri S, Schofield D, Olson A, Shrestha R, Holzappel C, Wolfenstetter S, Holle R, Ahern A, Hauner H, Jebb S, Caterson I. A within-trial cost-effectiveness analysis of primary care referral to a commercial provider for weight loss treatment, relative to standard care—an international randomised controlled trial. *Int J Obes*. 2013;37:828–34.
150. Gillespie P, O'Shea E, Casey D, Murphy K, Devane D, Cooney A, Mee L, Kirwan C, McCarthy B, Newell J. The cost-effectiveness of a structured education pulmonary rehabilitation programme for chronic obstructive pulmonary disease in primary care: the PRINCE cluster randomised trial. *BMJ*. 2013;3.
151. Gillett M, Dallosso H, Dixon S, Brennan A, Carey M, Campbell M, Heller S, Khunti K, Skinner T, Davies M. Delivering the diabetes education and self management for ongoing and newly diagnosed (DESMOND) programme for people with newly diagnosed type 2 diabetes: cost effectiveness analysis. *BMJ*. 2010;341.
152. Godley S, Garner B, Passetti L, Funk R, Dennis M, Godley M. Adolescent outpatient treatment and continuing care: Main findings from a randomized clinical trial. *Drug Alcohol Depend*. 2010;110:44–54.
153. Green J, Wood A, Kerfoot M, Trainor G, Roberts C, Rothwell J, Woodham A, Ayodeji E, Barrett B, Byford S, Harrington R. Group therapy for adolescents with repeated self harm: randomised controlled trial with economic evaluation. *BMJ*. 2011;342.
154. Group, P. M. C. Long-term effectiveness of dopamine agonists and monoamine oxidase B inhibitors compared with levodopa as initial treatment for Parkinson's disease (PD MED): a large, open-label, pragmatic randomised trial. *Lancet*. 2014;384:1196–205.
155. Group T. Options for managing low grade cervical abnormalities detected at screening: cost effectiveness study. *BMJ*. 2009;339.
156. Hedman E, El Alaoui S, Lindefors N, Andersson E, Ruck C, Ghaderi A, Kaldo V, Lekander M, Andersson G, Ljotsson B. Clinical effectiveness and cost-effectiveness of internet- vs. group-based cognitive behavior therapy for social anxiety disorder: 4-year follow-up of a randomized trial. *Behav Res Ther*. 2014;59:20–9.
157. Heliövaara-Peippo S, Hurskainen R, Teperi J, Aalto A, Grenman S, Halmesmaki K, Jokela M, Kivela A, Tomas E, Tuppurainen M, Paavonen J. Quality of life and costs of levonorgestrel-releasing intrauterine system or hysterectomy in the treatment of menorrhagia: a 10-year randomized controlled trial. *Am J Obstet Gynecol*. 2013;535.
158. Henderson C, Knapp M, Fernandez J, Beecham J, Hirani S, Beynon M, Cartwright M, Rixon L, Doll H, Bower P, Steventon A, Rogers A, Fitzpatrick R, Barlow J, Bardsley M, Newman S. Cost effectiveness of telehealth for patients with long term conditions (whole systems demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial. *BMJ*. 2013;346.
159. Henderson C, Knapp M, Fernandez J, Beecham J, Hirani S, Beynon M, Cartwright M, Rixon L, Doll H, Bower P, Steventon A, Rogers A, Fitzpatrick R, Barlow J, Bardsley M, Newman S. Cost-effectiveness of telecare for people with social care needs: the whole systems demonstrator cluster randomised trial. *Age Ageing*. 2014;1–7.
160. Higginson I, McCrone P, Burman R, Silber E, Edmonds P. Is short-term palliative care cost-effective in multiple sclerosis? a randomized phase ii trial. *J Pain Symptom Manag*. 2009;38:816–26.
161. Honkoop P, Loijmans R, Termeer E, Snoeck-Stroband J, van den Hout W, Bakker M, Assendelft W, ter Riet G, Sterk P, Schermer T, Sont J. Symptom- and fraction of exhaled nitric oxide-driven strategies for asthma control: A cluster-randomized trial in primary care. *J Allergy Clin Immunol*. 2014;135:683–8.
162. Jones L, FitzGerald G, Leurent B, Round J, Eades J, Davis S, Gishen F, Holman A, Hopkins K, Tookman A. Rehabilitation in

- advanced, progressive, recurrent cancer: A randomized controlled trial. *J Pain Symptom Manag.* 2013;46:315–25.
163. Kilonzo M, Sambrook A, Cook J, Campbell M, Cooper K. A cost-utility analysis of microwave endometrial ablation versus thermal balloon endometrial ablation. *Value Health.* 2010;13:528–34.
 164. Knapp M, King D, Romeo R, Schehl B, Barber J, Griffin M, Rapaport P, Livingston D, Mummery C, Walker Z, Hoe J, Sampson E, Cooper C, Livingston G. Cost effectiveness of a manual based coping strategy programme in promoting the mental health of family carers of people with dementia (the START (STrAtegies for RelaTives) study): a pragmatic randomised controlled trial. *BMJ.* 2013;347.
 165. Kolu P, Raitanen J, Rissanen P, Luoto R. Cost-effectiveness of lifestyle counselling as primary prevention of gestational diabetes mellitus: findings from a cluster-randomised trial. *PLoS One.* 2013;8.
 166. Krist M, van Beijsterveldt A, Backx F, de Wit G. Preventive exercises reduced injury-related costs among adult male amateur soccer players: a cluster-randomised trial. *J Physiother.* 2013;59:15–23.
 167. Kuyken W, Hayes R, Barrett B, Byng R, Dalgleish T, Kessler D, Lewis G, Watkins E, Brejcha C, Cardy J, Causley A, Cowderoy S, Evans A, Gradinger F, Kaur S, Lanham P, Morant N, Richards J, Shah P, Sutton H, Vicary R, Weaver A, Wilks J, Williams M, Taylor R, Byford S. Effectiveness and cost-effectiveness of mindfulness-based cognitive therapy compared with maintenance antidepressant treatment in the prevention of depressive relapse or recurrence (PREVENT): a randomised controlled trial. *Lancet.* 2015;386:63–73.
 168. Ladapo J, Elliott M, Bogart L, Kanouse D, Vestal K, Klein D, Ratner J, Schuster M. Cost of talking parents, healthy teens: A worksite-based intervention to promote parent-adolescent sexual health communication. *J Adolesc Health.* 2013;53:595–601.
 169. Lall R, Hamilton P, Young D, Hulme C, Hall P, Shah S, MacKenzie I, Tunnicliffe W, Rowan K, Cuthbertson B, McCabe C, Lamb S. A randomised controlled trial and cost-effectiveness analysis of high-frequency oscillatory ventilation against conventional artificial ventilation for adults with acute respiratory distress syndrome. The OSCAR (OSCillation in ARDS) study. *Health Technol Assess.* 2015;19.
 170. Lamb S, Gates S, Williams M, Williamson E, Mt-Isa S, Withers E, Castelnovo E, Smith J, Ashby D, Cooke M, Petrou S, Underwood M. Emergency department treatments and physiotherapy for acute whiplash: a pragmatic, two-step, randomised controlled trial. *Lancet.* 2013;381:546–56.
 171. Lamb S, Hansen Z, Lall R, Castelnovo E, Withers E, Nichols V, Potter R, Underwood M. Group cognitive behavioural treatment for low-back pain in primary care: a randomised controlled trial and cost-effectiveness analysis. *Lancet.* 2010;375:916–23.
 172. Lambeek L, Bosmans J, Van Royen B, Van Tulder M, Van Mechelen W, Anema J. Effect of integrated care for sick listed patients with chronic low back pain: economic evaluation alongside a randomised controlled trial. *BMJ.* 2010;341.
 173. Luik A, Merkel M, Hoeren D, Riexinger T, Kieser M, Schmitt C. A randomized controlled noninferiority trial comparing isolation of the pulmonary veins with the cryoballoon catheter versus open irrigated radiofrequency ablation in patients with paroxysmal atrial fibrillation. *Am Heart J.* 2010;159:555–60.
 174. Lynch F, Dickerson J, Saldana L, Fisher P. Incremental net benefit of early intervention for preschool-aged children with emotional and behavioral problems in foster care. *Child Youth Serv Rev.* 2014;36:213–9.
 175. Mak S, Lee M, Cheung J, Choi K, Chung T, Wong T, Lam K, Lee D. Pressurised irrigation versus swabbing method in cleansing wounds healed by secondary intention: A randomised controlled trial with cost-effectiveness analysis. *Int J Nurs Stud.* 2015;52:88–101.
 176. Maljanen T, Kenkt P, Lindfors O, Virtala E, Tillman P, Harkanen T. The cost-effectiveness of short-term and long-term psychotherapy in the treatment of depressive and anxiety disorders during a 5-year follow-up. *J Affect Disord.* 2015;190:254–63.
 177. Manca A, Asseburg C, Vergel Y, Seymour M, Meade A, Stephens R, Parmar M, Sculpher M. The cost-effectiveness of different chemotherapy strategies for patients with poor prognosis advanced colorectal cancer (MRC FOCUS). *Value Health.* 2012;15:22–31.
 178. McCollister K, Yang X, McKay J. Cost-effectiveness analysis of a continuing care intervention for cocaine-dependent adults. *Drug Alcohol Depend.* 2015;158:38–44.
 179. Meuldijk D, Carlier I, van Vliet I, van der Akker-Marle M, Zitman F. A randomized controlled trial of the efficacy and cost-effectiveness of a brief intensified cognitive behavioral therapy and/or pharmacotherapy for mood and anxiety disorders: Design and methods. *Contemp Clin Trials.* 2012;33:983–92.
 180. Ninot G, Moullec G, Picot M, Jaussent A, Hayot M, Desplan M, Brun J, Mercier J, Prefaut C. Cost-saving effect of supervised exercise associated to COPD self-management education program. *Respir Med.* 2011;105:377–85.
 181. Noben C, Smit F, Nieuwenhuijsen K, Ketelaar S, Gartner F, Boon B, Sluiter J, Evers S. Comparative cost-effectiveness of two interventions to promote work functioning by targeting mental health complaints among nurses: Pragmatic cluster randomised trial. *Int J Nurs Stud.* 2014;51:1321–31.
 182. Nyman M, Gustafsson M, Langius-Eklof A, Johansson J, Norlin R, Hagberg L. Intermittent versus indwelling urinary catheterisation in hip surgery patients: A randomised controlled trial with cost-effectiveness analysis. *Int J Nurs Stud.* 2013;50:1589–98.
 183. Parry G, Cooper C, Moore J, Yadegarfar G, Campbell M, Esmonde L, Morice A, Hutchcroft B. Cognitive behavioural intervention for adults with anxiety complications of asthma: Prospective randomised trial. *Respir Med.* 2012;106:802–10.
 184. Peek G, Mugford M, Tiruvoipati R, Wilson A, Allen E, Thalanany M, Hibbert C, Truesdale A, Clemens F, Cooper N, Firmin R, Elbourne D. Efficacy and economic assessment of conventional ventilatory support versus extracorporeal membrane oxygenation for severe adult respiratory failure (CESAR): a multicentre randomised controlled trial. *Lancet.* 2009;374:1351–63.
 185. Pennington M, Grieve R, Sekhon J, Gregg P, Black N, van der Meulen J. Cemented, cementless, and hybrid prostheses for total hip replacement: cost effectiveness analysis. *BMJ.* 2013;346.
 186. Petrou S, Dakin H, Abangma G, Benges S, Williamson I. Cost-utility analysis of topical intranasal steroids for otitis media with effusion based on evidence from the gnometrial. *Value Health.* 2010;13:543–51.
 187. Pickard R, Starr K, MacLennan G, Kilonzo M, Lam T, Thomas R, Burr J, Norrie J, McPherson G, McDonald A, Shearer K, Gillies K, Anson K, Boachie C, N'Dow J, Burgess N, Clark T, Cameron S, McClinton S. Use of drug therapy in the management of symptomatic ureteric stones in hospitalised adults: a multicentre, placebo-controlled, randomised controlled trial and cost-effectiveness analysis of a calcium channel blocker (nifedipine) and an alpha-blocker (tamsulosin) (the SUSPEND trial). *Health Technol Assess.* 2015;19.
 188. Prestmo A, Hagen G, Sletvold O, Helbostad J, Thingstad P, Taraldsen K, Lydersen S, Halsteinli V, Saltnes T, Lamb S, Johnsen L, Saltvedt I. Comprehensive geriatric care for patients with hip fractures: a prospective, randomised, controlled trial. *Lancet.* 2015;385:1623–33.
 189. Regier D, Petrou S, Henderson J, Eddama O, Patel N, Strohm B, Brocklehurst P, Edwards AD, Azzopardi D. Cost-effectiveness

- of therapeutic hypothermia to treat neonatal encephalopathy. *Value Health*. 2010;13:695–702.
190. Ryan D, Price D, Musgrave S, Malhotra S, Lee A, Ayansina D, Sheikh A, Tarassenko L, Pagliari C, Pinnock H. Clinical and cost effectiveness of mobile phone supported self monitoring of asthma: multicentre randomised controlled trial. *BMJ*. 2012;344.
 191. Sabes-Figuera R, McCrone P, Hurley M, King M, Donaldson A, Risdale L. Cost-effectiveness of zoledronic acid for the prevention of skeletal complications in patients with prostate cancer. *BMC Health Serv Res*. 2012;12:264.
 192. Scott D, Ibrahim F, Farewell V, O’Keeffe A, Walker D, Kelly C, Birrell F, Chakravarty K, Maddison P, Heslin M, Patel A, Kingsley G. Tumour necrosis factor inhibitors versus combination intensive therapy with conventional disease modifying anti-rheumatic drugs in established rheumatoid arthritis: TACIT non-inferiority randomised controlled trial. *BMJ*. 2015;350.
 193. Smit F, Willemse G, Meulenbeek P, Koopmanschap M, van Balkom A, Spinhoven P, Cuijpers P. Preventing panic disorder: cost-effectiveness analysis alongside a pragmatic randomised trial. *Cost Eff Resour Allocation*. 2009;7:8.
 194. Stoddart A, Hanley J, Wild S, Pagliari C, Peterson M, Lewis S, Sheikh A, Krishan A, Padfield P, McKinsty B. Telemonitoring-based service redesign for the management of uncontrolled hypertension (HITS): cost and cost-effectiveness analysis of a randomised controlled trial. *BMJ*. 2013;3.
 195. Taylor A, Thompson T, Greves C, Taylor R, Green C, Warren F, Kandiyali R, Aveyard P, Ayres R, Byng R, Campbell J, Ussher H, Michie S, West R. A pilot randomised trial to assess the methods and procedures for evaluating the clinical effectiveness and cost-effectiveness of exercise assisted reduction then stop (EARS) among disadvantaged smokers. *Health Technol Assess*. 2014;18.
 196. Thompson S, Ashton H, Gao L, Scott R. Screening men for abdominal aortic aneurysm: 10 year mortality and cost effectiveness results from the randomised multicentre aneurysm screening study. *BMJ*. 2009;338.
 197. Turner D, Little P, Raftery J, Turner S, Smith H, Rumsby K, Mullee M. Cost effectiveness of management strategies for urinary tract infections: results from randomised controlled trial. *BMJ*. 2010;340.
 198. Tyrer P, Cooper S, Salkovkis Tyrer H, Crawford M, Byford S, Dupont S, Finnis S, Green J, McLaren E, Murphy D, Reid S, Smith G, Wang D, Warwick H, Petkova H, Barrett B. Clinical and cost-effectiveness of cognitive behaviour therapy for health anxiety in medical patients: a multicentre randomised controlled trial. *Lancet*. 2014;383:219–25.
 199. Underwood M, Lamb S, Eldridge S, Sheehan B, Slowther A, Spencer A, Thorogood M, Atherton N, Bremner S, Devine A, Diaz-Ordaz K, Ellard D, Potter R, Spanjers K, Taylor S. Exercise for depression in care home residents: a randomised controlled trial with cost-effectiveness analysis (OPERA). *Health Technol Assess*. 2013;196:319–25.
 200. Van Rossem C, Spigt M, Smit E, Viechtbauer W, Mijnheer K, van Schayck C, Kotz D. Combining intensive practice nurse counselling or brief general practitioner advice with varenicline for smoking cessation in primary care: Study protocol of a pragmatic randomized controlled trial. *Contemp Clin Trials*. 2015;41:298–312.
 201. Van Wijk S, van Asselt A, Rickli H, Estlinbaum W, Erne P, Rickenbacher P, Vuillomenet A, Peter M, Pfisterer M, Brunner-La Rocca H. Cost-effectiveness of n-terminal pro-b-type natriuretic-guided therapy in elderly heart failure patients. *JACC*. 2013;1:64–71.
 202. Viksveen P, Relton C. Depression treated by homeopaths: a study protocol for a pragmatic cohort multiple randomised controlled trial. *Homeopathy*. 2014;103:147–52.
 203. Wagner T, Hattler B, Bishawi M, Baltz J, Collins J, Quin J, Grover F, Shroyer A. On-pump versus off-pump coronary artery bypass surgery: Cost-effectiveness analysis alongside a multisite trial. *Ann Thorac Surg*. 2013;96:770–7.
 204. Wake M, Baur L, Gerner B, Gibbons K, Gold L, Gunn J, Levickis P, McCallum Z, Naughton G, Sanci L, Ukoumunne O. Outcomes and costs of primary care surveillance and intervention for overweight or obese children: the LEAP 2 randomised controlled trial. *Prev Med*. 2009;339.
 205. Ward S, Wang K, Serlin R, Peterson S, Murray L. A randomized trial of a tailored barriers intervention for cancer information service (CIS) callers in pain. *Pain*. 2009;144:49–56.
 206. Yardley L, Barker F, Muller I, Turner D, Kirby S, Mullee M, Morris A, Little P. Clinical and cost effectiveness of booklet based vestibular rehabilitation for chronic dizziness in primary care: single blind, parallel group, pragmatic, randomised controlled trial. *BMJ*. 2012;344.
 207. Yoo S, Nyman J, Cheville A, Kroenke K. Cost effectiveness of telecare management for pain and depression in patients with cancer: results from a randomized trial. *Gen Hosp Psychiatry*. 2014;36:599–606.
 208. Zwerink M, Am Kerstjens H, van der Palen J, van der Valk P, Brusse-Keizer M, Zielhuis G, Effing T. (Cost-)effectiveness of self-treatment of exacerbations in patients with COPD: 2 years follow-up of a RCT. *Respirology*. 2015.