BMJ Open

Failure to address potential bias in non-randomised controlled clinical trials may cause lack of evidence on patient-reported outcomes: a method study

Frank Peinemann,1 Alexander Michael Labeit,2 Christian Thielacher,3 Michael Pinkawa4


ABSTRACT

Objectives: We conducted a workup of a previously published systematic review and aimed to analyse why most of the identified non-randomised controlled clinical trials with patient-reported outcomes did not match a set of basic quality criteria.

Setting: There were no limits on the level of care and the geographical location.

Participants: The review evaluated permanent interstitial low-dose rate brachytherapy in patients with localised prostate cancer and compared that intervention with alternative procedures such as external beam radiotherapy, radical prostatectomy and no primary therapy.

Primary outcome measure: Fulfilment of basic inclusion criteria according to a Participants, Interventions, Comparisons, Outcomes (PICO) framework and accomplishment of requirements to contain superimposed risk of bias.

Results: We found that 21 of 50 excluded non-randomised controlled trials did not meet the PICO inclusion criteria. The remaining 29 studies showed a lack in the quality of reporting. The resulting flaws included attrition bias due to loss of follow-up, lack of reporting baseline data, potential confounding due to unadjusted data and lack of statistical comparison between groups.

Conclusions: With respect to the reporting of patient-reported outcomes, active efforts are required to improve the quality of reporting in non-randomised controlled trials concerning permanent interstitial low-dose rate brachytherapy in patients with localised prostate cancer.

Strengths and limitations of this study

- We conducted a comprehensive literature search and strictly adhered to the projected methodology.
- We identified a lack of quality in non-randomised controlled clinical trials reporting patient-reported outcomes, analysed the cause and suggested possible improvements in designing studies in the future.
- The analysis is confined to a single disease and a specific treatment and conclusions drawn from its results may not be generalisable to other diseases and treatments.
- The limits for the inclusion of studies are arbitrarily set.

INTRODUCTION

The present paper reports a workup of a previously published systematic review.1 It may be regarded as a methodological supplement adding information on a subset of excluded studies. We have compared permanent interstitial low-dose rate brachytherapy, with radical prostatectomy, external beam radiotherapy and ‘no primary therapy’ in patients with localised prostate cancer categorised as T1 to T2. We used the term ‘no primary therapy’ to accommodate different types of observation including active surveillance, watchful waiting and observing without a distinctive management. As a result, we included one randomised controlled trial (RCT) and 30 non-randomised controlled clinical trials (CCT). The primary outcome was overall survival. The secondary outcomes were clinically defined disease-free survival, biochemical recurrence-free survival, physician-reported severe adverse events and patient-reported outcomes (PROs) such as function and bother scores as well as generic and disease-related health-related quality of life. We concluded that the current evidence is insufficient to allow a definitive conclusion about overall survival. Radical prostatectomy and external beam radiotherapy can severely affect the structural integrity of neighbouring organs and their functions and can also cause considerable long-term impairment of health-related quality of life. With a view of

expecting similar survival but a tremendous difference of adverse events between treatment alternatives, valid data on health-related quality of life could tip the balance. At the least, we assume that shared decision-making and consideration of patients’ preferences in searching for the best individual treatment would rely on information on the health-related quality-of-life data. Of the 30 included non-randomised studies, 13 studies reported PROs, that is, only the patients provided the information. During the study selection process, we experienced that we excluded another 50 non-randomised PRO studies. We found it a pity that we could not use the data. We had the impression that a considerable number of studies were excluded because of a lack in the quality of reporting. Therefore, we wanted to summarise the reasons for excluding those PRO studies and make the authors of PRO studies aware of some basic requirements for reporting of comparative PRO data to achieve higher acceptance in the scientific community. The importance of reporting PRO has been addressed by the Consolidated Standards of Reporting Trials (CONSORT) group which recently published a PRO extension to their acclaimed previous statement. It may be wise to build a PRO extension to the STrengthening the Reporting of Observational studies in Epidemiology (STROBE) statement that addresses specific issues of observational studies.

The first aim of this study was to assess whether the excluded studies met the basic inclusion criteria using the PICO framework. The second aim was to ensure whether the excluded studies met the requirements to contain high risk of bias.

MATERIALS AND METHODS

Study inclusion criteria

We defined the inclusion criteria according to the PICO framework that should include four essential constituents, that is, type of participants (P), intervention (I), comparator (C) and outcome (O). The four PICO items can be supplemented by timing (T) and setting (S), two other important features of a systematic review, to create the so-called PICOTS typology. A further extension embraces the study design (SD) to complete all major items of a search strategy (PICOTS-SD).

Population

Localised prostate cancer is defined by the categories T1 to T2 of the tumour-node-metastasis staging system if combined with the absence of regional lymph node metastasis and distant metastasis.

Intervention

Brachytherapy is short-distance radiotherapy placing radiation sources with different duration and rates of dose delivery in or near tumours. Permanent interstitial low-dose rate brachytherapy means implanting of low-energy radioactive sources emitting radiation, which are contained in titanium pellets of the size of rice grains called seeds.

Comparator

Initial and present publication

The European Association of Urology suggested three different treatment concepts for localised prostate cancer in addition to permanent interstitial low-dose rate brachytherapy: radical prostatectomy, external beam radiotherapy and different types of observation including active surveillance, watchful waiting and observing without distinctive management.

Outcome

Initial publication

Overall survival, cancer-specific survival, disease-free survival, biochemical recurrence-free survival, severe adverse events and PROs. PROs comprised function and bother scores as well as generic and disease-related health-related quality of life.

Present publication

Fulfilment of basic inclusion criteria according to a PICO framework by the excluded CCT. Accomplishment of requirements to contain superimposed risk of bias in addition to the high risk of bias caused by the lack of randomisation framework by the excluded CCT.

Timing

Initial and present publication

We did not set limits on the length of the observation period.

Setting

Initial and present publication

We did not set limits on the setting such as type of country, year of recruitment or level of healthcare.

Study design

Initial publication

We included RCT and CCT evaluating permanent interstitial low-dose rate brachytherapy as monotherapy in patients with localised prostate cancer. The proportion of relevant patients was required to be at least 80% of the study population and the response rate of questionnaires was expected to be at least 70%. For CCT to be included, comparable baseline characteristics between treatment groups or adjustment for imbalances of these data were required. Limits on year of publication or language were not applied.

Present publication

We included specifically the CCT that were excluded in the initial publication.
Search strategy
The search strategy was reported previously.¹

Study selection
In the present study, we selected only those 50 non-randomised studies on PRO that were excluded from the evaluation in the initial publication. In the study selection process, two reviewers independently judged whether a study was included or excluded. Differences were resolved by discussion without the need for a third opinion.

Data collection and analysis
The reasons for exclusion were extracted independently by two reviewers. We sought for the following data: the inclusion criteria using the PICO framework, the proportion of response of participants to questionnaires, which was required to be at least 70%, the reporting of separate baseline characteristics for each treatment group, the reporting of comparable baseline characteristics or adjustment for imbalances of these data such as the use of a Cox proportional hazard model and the reporting of statistics comparing treatment groups. Sufficient comparability was defined as a difference between baseline values that were not statistically significant. If a statistical test was not reported, we assumed two comparable values if the greater of the two values was less than 10% above the smaller one. We also required that authors reported effect measures and statistics testing the difference between treatment groups, for example, p values or effect measures including 95% CIs. Reporting of within group comparisons or before-and-after analyses was not deemed sufficient for inclusion. We did not apply a principal summary measure as we aimed to synthesise the information in a qualitative way.

Assessment of risk of bias and quality of reporting
Two reviewers independently assessed the quality of reporting of CCT according to the criteria specified in the previous paragraph. We did not specifically assess the risk of bias because we decided to exclude all papers with regard to a lack of reporting essential data.

RESULTS
Of a total of 462 full-text articles assessed for eligibility in the previously published systematic review, 31 studies were included and 431 studies were excluded. Among the 431 excluded articles, we identified 50 non-randomised studies that were reporting on PRO (figure 1). We evaluated the reasons for exclusion of those 50 studies and documented the results in table 1. In 42% (21 of 50) studies, the essential PICO framework was simply not met. In the majority of 58% (29 of 50) studies, the predefined requirement to apply measures to contain high risk of bias was not met. Of these 29 studies, 19 reported a proportion of patients responding to questionnaires of less than 70% or did not address this item. Baseline characteristics were not presented for treatment groups in three studies. In another six studies, baseline characteristics were not comparable between treatment groups or there was no confounder control in the analysis adjusting for important different factors such as mean age. The statistical comparison between treatment groups was deemed not appropriate in one study.

DISCUSSION
Main results
In summary, we found that roughly 4 of 10 excluded PRO studies did not meet the essential inclusion criteria using the PICO framework. This result is consistent with the problem of information retrieval aiming at a high recall and ending up with a low precision. The papers were obviously not relevant to the research question and we did not further examine the quality of reporting. We also found that roughly 6 of 10 excluded PRO studies met the PICO framework but did not provide the predefined requirements to care sufficiently enough for a low response of patients to questionnaires, for reporting baseline characteristics between treatment groups, for adjusting differences in those baseline characteristics between treatment groups and to use appropriate statistics to compare the outcome between treatment groups.

Quality of reporting of PROs
We identified a lack in the quality of reporting in many excluded CCT and wish to stress the importance of considering a series of requirements while conducting a study on PRO. Other authors have reported recently that, concerning disease-specific mortality or disease-free survival, the available studies did not show significant differences between treatment groups.¹³ ¹⁴ In the view of unknown or small differences in survival measures, the results of PRO studies could have a noticeable impact on medical decision-making.¹⁵ ¹⁶ None of the 50 excluded studies reported a non-responder analysis, although it is known that non-responders may have different attitudes than responders. Etter and Perneger¹⁷ concluded that low response rates may be associated with overestimating an effect and that the strength and direction of a non-response bias may depend on the mechanism of non-response. Therefore, results may be confounded if the proportion of included data not available for analysis such as data from non-responders or due to loss to follow-up is considerable. We believe that a value of 30% or more can be denoted as considerable. Lowering this threshold, for example, to 20%, would have resulted in less included studies. However, others suggested that 20% or more loss would be sufficient for a high risk of bias threatening the validity of results.¹⁸ Concerning questionnaires, we recommend taking measures that are known to improve response rates.¹⁹ ²⁰ Edwards²¹ conducted a systematic review to identify effective strategies to increase the response to postal and electronic questionnaires. The authors found several strategies to increase the response, for example,
prenotification, follow-up contact, shorter questionnaires, mentioning an obligation to respond, university sponsorship, non-monetary incentives, a statement that others had responded, an offer of survey results, giving a deadline. We did not use a strict algorithm to differentiate between comparable and not comparable baseline values between treatment groups. A statistically significant difference was judged as not comparable. Non-significant differences were also regarded as not comparable if the difference was at least 10% of the lower of two values. Using this approach we tried to reduce subjective decisions. We are not aware of published strict algorithms in this matter.

**High risk of bias inherent in non-RCTs**

With a view to include only one RCT, the initial publication was based almost exclusively on CCT. However, the lack of randomisation poses a very large challenge to the authors who are advised to deal with essential problems such as selection bias and confounding. Otherwise, the findings may not be valid and of limited usefulness and the many efforts may be in vain. We wish to stress that the non-randomised design is associated with a high risk of bias because known and unknown characteristics may be distributed unequally between groups. 22 Certain study characteristics, such as prospective design, concurrent control group, adjustment of results with respect to different baseline values and confounder control, can limit additional bias. For example, Ioannidis et al reported that discrepancies between RCT and CCT were less common when only CCT with a prospective design were considered. The Cochrane Collaboration offers a guide for inclusion of non-randomised studies and it has developed a tool for assessing the risk of bias in RCT.
<table>
<thead>
<tr>
<th>Non-randomised studies</th>
<th>Inclusion criteria</th>
<th>Requirements to contain high risk of bias</th>
<th>Statistical comparison between groups</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Response ≥70%</td>
<td>Baseline each group</td>
<td>Baseline comparable/or adjusted</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Table 1 Reasons for excluding PRO articles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inclusion criteria</td>
<td>Requirements to contain high risk of bias</td>
<td>Statistical comparison between groups</td>
<td>Comments</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI CO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacon et al.</td>
<td>Yes Yes No</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Ball et al.</td>
<td>Yes Yes No</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Befort et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Low response</td>
</tr>
<tr>
<td>Bergman et al.</td>
<td>Yes Yes Yes Yes Yes</td>
<td>Yes Yes No</td>
<td>No</td>
<td>No appropriate test</td>
</tr>
<tr>
<td>Bergman et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Low response</td>
</tr>
<tr>
<td>Brandeis et al.</td>
<td>Yes No – – – – –</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Brown et al.</td>
<td>Yes No – – – – –</td>
<td>–</td>
<td>–</td>
<td>EBRT 29% LDR-BT+EBRT</td>
</tr>
<tr>
<td>Burnett et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Response not reported</td>
</tr>
<tr>
<td>Chaikin et al.</td>
<td>No – – – – – – –</td>
<td>–</td>
<td>–</td>
<td>Staging not reported</td>
</tr>
<tr>
<td>Chen et al.</td>
<td>Yes Yes Yes Yes Yes</td>
<td>Yes Yes No</td>
<td>–</td>
<td>No confounder control</td>
</tr>
<tr>
<td>Choo et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Baseline not reported</td>
</tr>
<tr>
<td>Clark et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Low response</td>
</tr>
<tr>
<td>Downs et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>Yes No</td>
<td>–</td>
<td>No confounder control</td>
</tr>
<tr>
<td>Eton et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Low response</td>
</tr>
<tr>
<td>Frank et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Low response</td>
</tr>
<tr>
<td>Fulmer et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Response not reported</td>
</tr>
<tr>
<td>Gore et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Low response</td>
</tr>
<tr>
<td>Guedea et al.</td>
<td>Yes Yes Yes Yes Yes</td>
<td>Yes No</td>
<td>–</td>
<td>No confounder control</td>
</tr>
<tr>
<td>Hashine et al.</td>
<td>Yes Yes Yes Yes Yes</td>
<td>Yes No</td>
<td>–</td>
<td>No confounder control</td>
</tr>
<tr>
<td>Hashine et al.</td>
<td>Yes Yes Yes Yes Yes</td>
<td>Yes No</td>
<td>–</td>
<td>No confounder control</td>
</tr>
<tr>
<td>Hervouet et al.</td>
<td>No – – – – – – –</td>
<td>–</td>
<td>–</td>
<td>≥20% T3–T4 in control groups</td>
</tr>
<tr>
<td>Hollenbeck et al.</td>
<td>Yes No – – – – –</td>
<td>–</td>
<td>–</td>
<td>LDR-BT+EBRT</td>
</tr>
<tr>
<td>Jo et al.</td>
<td>Yes No – – – – –</td>
<td>–</td>
<td>–</td>
<td>High-dose rate brachytherapy</td>
</tr>
<tr>
<td>Johnstone et al.</td>
<td>Yes No – – – – –</td>
<td>–</td>
<td>–</td>
<td>EBRT</td>
</tr>
<tr>
<td>Joly et al.</td>
<td>Yes No – – – – –</td>
<td>–</td>
<td>–</td>
<td>LDR-BT+EBRT</td>
</tr>
<tr>
<td>Kakehi et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Baseline not reported</td>
</tr>
<tr>
<td>Lev et al.</td>
<td>Yes No – – – – –</td>
<td>–</td>
<td>–</td>
<td>LDR-BT+EBRT</td>
</tr>
<tr>
<td>Lilleby et al.</td>
<td>Yes No – – – – –</td>
<td>–</td>
<td>–</td>
<td>EBRT</td>
</tr>
<tr>
<td>Litwin et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Low response</td>
</tr>
<tr>
<td>Litwin et al.</td>
<td>Yes No – – – – –</td>
<td>–</td>
<td>–</td>
<td>25% LDR-BT+EBRT</td>
</tr>
<tr>
<td>Mehta et al.</td>
<td>Yes Yes Yes No –</td>
<td>–</td>
<td>–</td>
<td>“Fear of cancer”*</td>
</tr>
<tr>
<td>Miller et al.</td>
<td>Yes No – – – – –</td>
<td>–</td>
<td>–</td>
<td>44% LDR-BT+EBRT</td>
</tr>
<tr>
<td>Miller et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Baseline not reported</td>
</tr>
<tr>
<td>Monahan et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Low response</td>
</tr>
<tr>
<td>Namiki et al.</td>
<td>Yes Yes Yes No –</td>
<td>–</td>
<td>–</td>
<td>Low response</td>
</tr>
<tr>
<td>Namiki et al.</td>
<td>Yes Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>Low response</td>
</tr>
<tr>
<td>Non-randomised studies</td>
<td>Inclusion criteria</td>
<td>Requirements to contain high risk of bias</td>
<td>Comments</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------</td>
<td>------------------------------------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P I C O</td>
<td>Response ≥70%</td>
<td>Baseline each group</td>
<td>Baseline comparable/or adjusted</td>
</tr>
<tr>
<td>Ohashi et al&lt;sup&gt;65&lt;/sup&gt;</td>
<td>Yes Yes Yes Yes</td>
<td>No</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Pinkawa et al&lt;sup&gt;66&lt;/sup&gt;</td>
<td>Yes Yes No –</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Roach et al&lt;sup&gt;67&lt;/sup&gt;</td>
<td>Yes No – –</td>
<td>No</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Sanda et al&lt;sup&gt;68&lt;/sup&gt;</td>
<td>Yes Yes Yes Yes</td>
<td>No</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Schover et al&lt;sup&gt;69&lt;/sup&gt;</td>
<td>Yes Yes Yes Yes</td>
<td>No</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Soderdahl et al&lt;sup&gt;70&lt;/sup&gt;</td>
<td>Yes Yes Yes Yes</td>
<td>No</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Speight et al&lt;sup&gt;71&lt;/sup&gt;</td>
<td>Yes Yes Yes Yes</td>
<td>No</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Stone et al&lt;sup&gt;72&lt;/sup&gt;</td>
<td>Yes Yes No –</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Trojan et al&lt;sup&gt;73&lt;/sup&gt;</td>
<td>Yes Yes Yes Yes</td>
<td>No</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Tward et al&lt;sup&gt;74&lt;/sup&gt;</td>
<td>Yes Yes Yes No</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Valicenti et al&lt;sup&gt;75&lt;/sup&gt;</td>
<td>Yes Yes Yes Yes</td>
<td>No</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Van de Poll-</td>
<td>Yes No – –</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Franse et al&lt;sup&gt;76&lt;/sup&gt;</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Wyler et al&lt;sup&gt;77&lt;/sup&gt;</td>
<td>Yes Yes Yes Yes</td>
<td>Yes</td>
<td>No</td>
<td>–</td>
</tr>
<tr>
<td>Zagar et al&lt;sup&gt;78&lt;/sup&gt;</td>
<td>Yes No – –</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>NO’ counts</strong></td>
<td>2 13 4 2</td>
<td>19 3</td>
<td>6 1</td>
<td></td>
</tr>
</tbody>
</table>

PICO not met: 21 High risk of bias: 29

*: not appropriate.
†Pinkawa et al<sup>66</sup>; Stone et al<sup>72</sup>: neoadjuvant hormonal therapy.
§Tward et al<sup>74</sup>: non-disease-related mortality differs greatly.
©: comparison of interest is radical prostatectomy, external beam radiotherapy, or no primary therapy; EBRT, external beam radiotherapy; I, intervention of interest is low-dose rate brachytherapy as monotherapy; LDR-BT, permanent interstitial low-dose rate brachytherapy; O, outcome of interest is function, bother, or generic health-related quality of life; P, patients with localised prostate cancer; PRO, patient-reported outcomes.
and CCT. Guidelines for reporting observational studies have been published to improve their quality. Cox regression analysis, propensity-score-based analysis and instrumental variable analysis are methods that have been used for correction of confounding bias in non-randomised studies. Different values of various outcome measures between groups may be simply caused by different baseline data in lieu of absent significant treatment effects. We accepted any type of method adjusting or stratifying for one or more known differences in baseline characteristics. Nevertheless, it should be kept in mind that methods of adjustment do not guarantee removal of bias and that residual confounding may remain high. Concerning the non-randomised design, we strongly recommend the use of methods for adjusting the results for confounders to aim for a less-biased estimation of the treatment effect and the adoption of guidelines for the reporting of observational studies.

**Strengths and limitations**

The strengths of the present study are a comprehensive literature search, strict adherence to the projected methodology, the identification of a lack of quality in PRO studies and addressing the specific problems of PRO studies. We should consider some limitations: the study is confined to a single disease, so conclusions drawn from its results may not be generalisable to other diseases. The arbitrary limits set for inclusion of studies are responsible for the extent of excluded studies. These limits may be questioned by other investigators. During the re-evaluation of study quality, we found that one study fulfilled all criteria, although, this study was excluded in previous reports. The minimum follow-up of 70% for inclusion was set arbitrarily and others might find this threshold too low. We did not endorse the recently published reporting of PRO in randomised trials, an extension of the CONSORT statement. All included studies in the present review are non-randomised. We think that the lack of randomisation is the prevailing issue. We did not endorse the CONSORT PRO extension for another reason. The included studies were published many years before this extension was published. There might be a need to develop an extension of the STROBE statement with the aim of improving the reporting of PRO in non-randomised studies. This extension could emphasise the specific challenges of reporting PRO with respect to lack of randomisation.

**CONCLUSIONS**

We found that a considerable number of non-randomised controlled reporting PROs were excluded from a systematic review because of a lack of predefined reporting requirements. The assumed overall risk of bias was regarded as too high to consider the data of these studies for inclusion in the systematic review. With respect to the reporting of PROs, active efforts are required to improve the quality of reporting in non-randomised controlled trials and to increase the number of randomised controlled trials.

**Author affiliations**


Failure to address potential bias in non-randomised controlled clinical trials may cause lack of evidence on patient-reported outcomes: a method study

Frank Peinemann, Alexander Michael Labeit, Christian Thielscher and Michael Pinkawa

BMJ Open 2014 4:
doi: 10.1136/bmjopen-2013-004720

Updated information and services can be found at:
http://bmjopen.bmj.com/content/4/6/e004720

These include:

References
This article cites 67 articles, 8 of which you can access for free at:
http://bmjopen.bmj.com/content/4/6/e004720#BIBL

Open Access
This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/3.0/

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

Evidence based practice (698)
Patient-centred medicine (448)
Urology (70)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/