PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Current Practices in Patient-Reported Outcome (PRO) Data Collection in Clinical Trials: A Cross-Sectional Survey of UK trial staff and management.
AUTHORS	Kyte, Derek; Ives, Jonathan; Draper, Heather; Calvert, Melanie

VERSION 1 - REVIEW

REVIEWER	José Armando Mangione Beneficência Portuguesa Hospital São Paulo/Brazil
REVIEW RETURNED	30-Apr-2016

GENERAL COMMENTS	Very interesting study trying to check the quality of PROMs data collection in an adequate and complete questionary with 767 participants responded the online survey. PROMs is fundamental for the reliability of the study. Unfurtunately the results os this study showed inconsistencies in the way PROMs are administered which may reduce the quality of the trial. Congratulations to the authors for the quality of the work.

REVIEWER	Michael Leu University of Washington United States
REVIEW RETURNED	24-May-2016

GENERAL COMMENTS	In this well-written manuscript, authors performed a cross-sectional survey of staff/investigators of multiple UK patient-centered outcome reporting teams. Survey development informed by a pilot qualitative study. Questions for the most part seemed relevant. Background and references seemed passable - I would wonder if there was anything from the US PCORI organization that could be cited. Methods: Seemed appropriate, although in lines 176-181, the investigators had already noted that the qualitative pilot study was used to develop the research questions. Since the second study was so much bigger it seems really a matter of convenience to have started from the pilot framework instead of de novo; I would like to better understand how much of the data from the pilot vs how much data from the full study contributed to the framework. Results: Lines 197-200: It would seem that you know the denominators for each of the 4 subcategories and the numerator, so you could calculate the response rate for the 4 individual groups. This would

	be more informative than a generic 43%. While not a fatal flaw, it does seem that a limitation of the study setup/methodology to not be able to better state how many study teams were represented - and from the write up, I can't really tell if some study teams were super overrepresented vs other study teams that might not have answered at all. Table 1: It was confusing to say "participant characteristics" as I was thinking of patients, but this is really surveyed study team characteristics. Table 2: While I appreciate the exhaustive reporting of every response, perhaps highlights could be selected with the full results available in the supplemental materials. I think Table 3 can be omitted, or reported with the supplemental table which showed the full analysis. Perhaps the authors could be a little more ambitious in using these findings to define a checklist of items which should be used to 1) design and/or 2) evaluate studies with patient-centered outcomes. The supplemental materials are exhaustive - I did briefly peruse them and they look acceptable.

REVIEWER	Daniel Strech and Holger Langhof Hannover Medical School, Germany I participated in a workshop organised by one of the co-authors (JI) and contributed to a publication (with JI as first author but not yet submitted) that aims to present workshop results.
REVIEW RETURNED	04-Jun-2016

GENERAL COMMENTS	The objective of the presented paper is to assess the practice of current PROM data administration in the UK. Three major questions are addressed: 1) the extent of inconsistencies in PROM administration as experienced by the survey respondents, 2) survey respondents perceptions on PRO-specific trial protocol content and training, and 3) future needs for PRO-specific trial protocol content and training. A cross-sectional survey was conducted, targeting UK-based research nurses, data managers/coordinators, trial managers and chief/principal investigators. The results are presented separately for each of the four professional groups. 767 respondents finished the survey, with research nurses as the main professional group. The main findings are that inconsistencies in PROM administration are widely perceived by respondents. In addition, the PRO-specific trial protocol content and training are perceived to be adequate, but more incorporation of PRO-specific content in SOPs is desired. The authors conclude that the found inconsistencies may reduce the quality of trial data and should be addressed by further enhancing PROM guidances. The rationale for the survey is clearly described and comprehensible. Though certainly an international issue, it was a reasonable decision to limit the survey to UK based trial professionals, due to the innovative research question and resource considerations. The conduct of the survey seems methodologically sound. Better knowledge on the response rate and on demographic

	data and opinions from non-respondents would be desirable but the authors appropriately justify why they had to accept their limitations. Beside its relevant content the paper is a very good example of implementation/quality assurance research in the field of research oversight/governance. While implementation research is more and more established to understand and improve how health care is delivered etc. we still lack studies that looks at how the clinical research procedures can be strengthened. Furthermore, the method-mix applied by this study nicely demonstrates how better qualitative understanding of the problem (e.g. what sources of biased PROM administration exist?) helps to develop meaningful quantitative surveys. Some minor comments: • “Moreover, neither the UK CRC-CTUs nor the NIHR CLRNs held data regarding the number of staff involved in trials with a primary or secondary PRO, so there was no way to determine a denominator. However, we estimate the total number of researchers receiving our survey invite, which includes those individuals ineligible for the study, was approximately 1,800; our sample represented 43% of this number.” (page 11). It would be interesting to give a brief explanation how the number of 1,800 researchers was estimated. • Table 2: In the footnote the authors declare that the “Columns may not add up to n due to missing values”, which is of course appropriate. The column titles indicate that the numbers presented show the number of respondents and, in brackets, the percentage rate of the respondents. In this case, however, most of the percentage rates given do not refer to the total number of respondents. For instance, the question “I read the questions out to the participants” (table row 3) presents a total number of 194 respondents, which would mean a percentage rate of 34,64% of all N=560 responding nurses, instead of the given rate of 36,9%. This is of particular interest, as another footnotes indicates that “Participants were able to select multiple categories”. It is unclear to the reader, what the denominator for the calculations is. • “Some research nurses in our survey questioned the usefulness of PRO-specific trial guidance, appearing to rely on their experience and judgment instead. Also, our regression model indicated that staff with greater experience (≥10 years) tended to report dealing more appropriately with missing PRO data, but trial protocol content or training were not significant predictors. It is possible, however, that the protocol content and training received by the respondents did not contain adequate information on the management of missing data. A review of PRO protocol content reported that under half of the n=75 included protocols detailed plans to minimize levels of avoidable missing PRO data.” (page 31). Here, the authors refer to the results of the performed logistic regression. The choice of independent variables (as described on page 9) seems reasonable. However, the finding that the only factor statistically significant associated with the appropriate management of missing PRO data is the professional experience in clinical trials should be discussed a little bit more. Having more than 10 years of professional experience as research nurse implies having conducted a greater amount of trials. Against the background of the other findings (e.g. “In general, all groups supported the inclusion of the majority of proposed PRO guidance within trial training and/or a SOP, but there was less support for including items within the protocol”, page 26) this finding would underline the need for a more general guidance for accurate PROM administration, independent from particular trials, but maybe embedded in SOPs. • Although briefly addressed in the background and discussion

	section, the importance of PROMs for regulatory as well as reimbursement decisions should be discussed against the background of the survey findings. A sound management and administration of PROMs should be guaranteed throughout the conduct of the trial, otherwise the possible data contaminations and biases (as outlined in the paper) might lead to misinformation of policy, regulatory, and reimbursement decisions.

VERSION 1 – AUTHOR RESPONSE

Reviewer 2

Comment: “Methods: Seemed appropriate, although in lines 176-181, the investigators had already noted that the qualitative pilot study was used to develop the research questions. Since the second study was so much bigger it seems really a matter of convenience to have started from the pilot framework instead of de novo; I would like to better understand how much of the data from the pilot vs how much data from the full study contributed to the framework.”

Author Response: as stated on pg 3, development of the survey questions was informed by our previous qualitative study. The pilot survey was aimed at testing the feasibility of collection methods and clarity of the questions; the pilot did not alter the framework or the survey questions. This has now been made clearer on pg 9.

Comment: “Lines 197-200: It would seem that you know the denominators for each of the 4 subcategories and the numerator, so you could calculate the response rate for the 4 individual groups. This would be more informative than a generic 43%.”

Author Response: As this was an anonymised survey (for reasons explained on page 9) and neither the UK CRC-CTUs nor the NIHR CLRNs held data regarding the number of staff involved in trials with a primary or secondary PRO, we did not know the denominators for each of the 4 subcategories. In discussion with the CRC-CTU’s and CLRN’s, we estimate the total number of researchers receiving our survey invite, including those individuals ineligible for the study, was approximately 1,800; our sample represented 43% of this number.

Comment: While not a fatal flaw, it does seem that a limitation of the study setup/methodology to not be able to better state how many study teams were represented - and from the write up, I can't really tell if some study teams were super overrepresented vs other study teams that might not have answered at all.

Author Response: thank you for this comment. We have now addressed this in the limitations section on page 33/34.

Comment: Table 1: It was confusing to say "participant characteristics" as I was thinking of patients, but this is really surveyed study team characteristics.

Author Response: We have changed ‘participant characteristics’ to ‘respondent characteristics’ throughout.

Comment: Table 2: While I appreciate the exhaustive reporting of every response, perhaps highlights could be selected with the full results available in the supplemental materials.

Author response: Table 2 provides the main survey results and as such we would prefer to retain Table 2 within the manuscript.

Comment: I think Table 3 can be omitted, or reported with the supplemental table which showed the full analysis.

Author Response: This table has been reported as supplemental material as suggested.

Comment: Perhaps the authors could be a little more ambitious in using these findings to define a checklist of items which should be used to 1) design and/or 2) evaluate studies with patient-centered outcomes.

Author Response: The study team are currently developing the aforementioned checklist – we have now included this information on page 32.

Reviewer 3

Comment: “Moreover, neither the UK CRC-CTUs nor the NIHR CLRNs held data regarding the number of staff involved in trials with a primary or secondary PRO, so there was no way to determine a denominator. However, we estimate the total number of researchers receiving our survey invite, which includes those individuals ineligible for the study, was approximately 1,800; our sample represented 43% of this number.” (page 11). It would be interesting to give a brief explanation how the number of 1,800 researchers was estimated.

Author Response: We have now made this clearer on page 12

Comment: Table 2: In the footnote the authors declare that the “Columns may not add up to n due to missing values”, which is of course appropriate. The column titles indicate that the numbers presented show the number of respondents and, in brackets, the percentage rate of the respondents. In this case, however, most of the percentage rates given do not refer to the total number of respondents. For instance, the question “I read the questions out to the participants” (table row 3) presents a total number of 194 respondents, which would mean a percentage rate of 34,64% of all N=560 responding nurses, instead of the given rate of 36,9%. This is of particular interest, as another footnotes indicates that “Participants were able to select multiple categories”. It is unclear to the reader, what the denominator for the calculations is.

Author Response: thank you for this comment. We have now included the raw data file as an online supplementary appendix which gives transparent information regarding all numerators/denominators. It is not possible to add this level of detail to table 2 without significantly reducing its clarity and interpretability.

Comment: “Some research nurses in our survey questioned the usefulness of PRO-specific trial guidance, appearing to rely on their experience and judgment instead. Also, our regression model indicated that staff with greater experience (≥10 years) tended to report dealing more appropriately with missing PRO data, but trial protocol content or training were not significant predictors. It is possible, however, that the protocol content and training received by the respondents did not contain adequate information on the management of missing data. A review of PRO protocol content reported that under half of the n=75 included protocols detailed plans to minimize levels of avoidable missing PRO data.” (page 31). Here, the authors refer to the results of the performed logistic regression. The choice of independent variables (as described on page 9) seems reasonable. However, the finding that the only factor statistically significant associated with the appropriate management of missing PRO data is the professional experience in clinical trials should be discussed a little bit more. Having more than 10 years of professional experience as research nurse implies having conducted a greater amount of trials. Against the background of the other findings (e.g. “In general, all groups supported

the inclusion of the majority of proposed PRO guidance within trial training and/or a SOP, but there was less support for including items within the protocol", page 26) this finding would underline the need for a more general guidance for accurate PROM administration, independent from particular trials, but maybe embedded in SOPs.

Author response: thank you for this comment, we have now addressed this on page 32.

Comment: Although briefly addressed in the background and discussion section, the importance of PROMs for regulatory as well as reimbursement decisions should be discussed against the background of the survey findings. A sound management and administration of PROMs should be guaranteed throughout the conduct of the trial, otherwise the possible data contaminations and biases (as outlined in the paper) might lead to misinformation of policy, regulatory, and reimbursement decisions.

Author Response: we have now addressed this on page 30.

VERSION 2 – REVIEW

REVIEWER	Daniel Strech Hannover Medical School, Germany I know one co-author (JI) from workshops and training sessions where we both were involved but we did not cooperate in research projects.
REVIEW RETURNED	08-Aug-2016

GENERAL COMMENTS	The authors generally agreed with the proposed comments and addressed them in the revised manuscript. For instance, it was unclear, how the number of 1,800 researchers in the UK receiving the survey invites was estimated. The authors now address this question, but the additional information given on this question in the revised manuscript is still scarce (see page 12: " However, in discussion with the CRC-CTU coordinating centre and with CLRN contacts, we estimate the total number of researchers receiving our survey invite, including those individuals ineligible for the study, was approximately 1,800; our sample represented 43% of this number"). The authors further declare, that no valid data (i.e. a register) exists. However, this is only a minor matter and does neither relate to the overall rationale of the study nor does any influence on the interpretation of results seem likely. Other reviewer comments are answered similarly. The authors addressed all comments and questions. Sometimes more detailed explanations would have been desirable, but again, in the end they would not add substantially to the overall rationale of the study. The study is of significant scientific interest and value, the methodology and analysis are sound, and the results are presented and discussed comprehensively.