Author's response to reviews

**Title:** Enhancing the Quality of International Orthopedic Medical Mission Trips Using The Blue Distinction Criteria for Knee and Hip Replacement Centers

**Authors:**

Kyle E Dempsey (kyle_dempsey@hms.harvard.edu)
Roya Ghazinouri (rghazinouri@mghihp.edu)
Desiree Diez (desiree.diez@gmail.com)
Luis Alcantara (lalcantara@hgps.org.do)
Carolyn Beagan (cbegan@partners.org)
Barbara Aggouras (baggouras@partners.org)
Monica Hoagland (monica.hoagland@childrens.harvard.edu)
Thomas S Thornhill (tthornhill@partners.org)
Jeffrey N Katz (jnkatz@partners.org)

**Version:** 3  **Date:** 27 June 2013

**Author's response to reviews:** see over
June 3, 2013

Editorial Team
BMC Musculoskeletal Disorders

Dear BMC Musculoskeletal Disorder Editorial Team:

My coauthors and I have addressed the well-considered critiques from the reviewers and we are pleased to resubmit our manuscript entitled “Enhancing the Quality of International Orthopedic Medical Mission Trips Using The Blue Distinction Criteria for Knee and Hip Replacement Centers” to BMC Musculoskeletal Disorder. The reviewers’ suggestions and criticisms were very helpful and have improved the manuscript.

We respond point by point to the reviewer’s comments below.

Thank you for your consideration.

Reviewer 1 (Flavia Cicuttini):

This is an interesting study. The main issue is that lack of information about the study methodology. This makes it difficult to assess the methodological quality of the paper.

Major compulsory revisions.

1. What were the study procedures?
2. How was the data collected?

Response: We appreciate the suggestion to add more details to the study’s methods section. We have clarified how the data was collected, who collected the data, and who the data sources were. In essence, the study’s primary author compiled the Blue Cross/Blue Shield criteria and then met with two of Op-Walk Boston’s program directors. The director’s responded to each criteria with “meet”, “do not meet”, or “do not know”. The study’s primary author then compiled this data and scored it accordingly. We have included a more detailed description of our data collection methods in the methods section under the subheading “Data Collection and Analysis” (page 7):

The study’s main author compiled the Blue Cross/Shield’s criteria and independently reviewed each criterion with two of Op-Walk Boston’s program directors during in-person interviews. The directors independently responded to each criterion with “meet”, “do not meet”, or “unsure”. For criteria that produced conflicting results, the criteria were sent to an independent person familiar with the specific criteria in question and this third party’s response served as a tie-breaker.
Similarly, for all criteria that lacked responses from both directors, the study’s main author redirected the question to an independent team member who had the most knowledge of that criterion.

3. What were the data sources for each of the study categories?

Response: The majority of data (~90%) regarding criteria that were met or unmet was collected from the program’s directors. There were a small number of questions that the directors gave conflicting answers to or that they were unable to answer, so Op-Walk Boston organizers with more knowledge of that specific criteria were questioned with the same binary format to see if the program met or did not meet the criteria. We altered the methods section under the “Data Collection and Analysis” subsection to further elaborate on the data sources (Page 7):

The study’s main author compiled the Blue Cross/Shield’s criteria and independently reviewed each criterion with two of Op-Walk Boston’s program directors during in-person interviews. The directors independently responded to each criterion with “meet”, “do not meet”, or “unsure”. For criteria that produced conflicting results, the criteria were sent to an independent person familiar with the specific criteria in question and this third party’s response served as a tie-breaker. Similarly, for all criteria that lacked responses from both directors, the study’s main author redirected the question to an independent team member who had the most knowledge of that criterion.

4. Who collected the data? Could they have been conflicted or biased in how they collected to interpreted responses. How did the investigators ensure that bias was minimized?

Response: The study’s primary author collected the data. The Blue Cross/Blue Shield criteria are objective criteria that leave little room for interpretation. Furthermore, the primary author asked for binary responses from the program’s directors, minimizing data collection interpretation bias. To further clarify the binary nature of the responses gathered, we included the following sentence in the methods section under the data collection and analysis subheading (page 7):

The directors independently responded to each criterion with “meet”, “do not meet”, or “unsure”.

5. How much variation might one expect around the measures examined?

Response: The Blue Cross/Blue Shield criteria explicitly define each measured parameter, so there is very little room for interpretation. Variation should therefore be minimal in the measured criteria. We have added a sentence to our newly-added “limitations” subsection of the Discussion section to explicitly state that respondent variability is one of our study’s limitations (page 15):
Furthermore, the data for this study was collected from a small number of respondents who provided key information to evaluate if the program meets or does not meet the criteria; it is therefore possible that some data variability would exist if more people were surveyed but this variability should be limited by the objective and binary nature of the Blue Cross/Shield’s criteria.

6. Given this is a scientific paper, what are potential limitations of the study?

This should be added to the discussion.

Response: We appreciate the suggestion to elaborate on the study’s potential limitations. We have added a new subsection to the Discussion section (page 15):

**Limitations**

This study’s data were potentially subject to observer bias, as data were collected by an investigator rather than a research assistant blinded to the study’s hypotheses and objectives. Anticipating this bias, we used data elements from the Blue Cross/Shield’s excellence criteria that were objective, binary, and subject to little interpretation. Furthermore, the data for this study was collected from a small number of respondents who provided key information to evaluate if the program meets or does not meet the criteria; it is therefore possible that responses would have been more heterogeneous if more people were surveyed but this variability should be limited by the objective and binary nature of the Blue Cross/Shield’s criteria.

Reviewer 2 (Megan Bohensky):

This well-written paper highlighted the evaluation of a joint replacement mission service operating in the Dominican Republic. With a growing burden of arthritis in developing countries, TJRs are likely to become more common. Processes for ensuring the quality of these services would be very valuable. This paper has applied a US standard (the Blue Cross/Shield Blue Distinction criteria) and discusses how the program performed against each of the criteria. Overall, I think this topic is very interesting. The authors describe methods by which they evaluated the barriers to adopting the criteria, but I did not think they were clearly presented in the results (apart from brief descriptions in the Tables). If other developing nations are seeking to use these criteria or develop their own, I think a more in depth discussion of some of the barriers and common emerging themes would be quite useful.

Discretionary Revisions

Methods, Setting
1. Just by way of background, it would be helpful to know if there are other centers performing TJR in DR (i.e. are TJRs only performed when visiting clinicians come to DR?) If not, what proportion are performed when there are visiting clinicians?

Response: We agree that providing more information regarding currently existing joint replacement programs in the DR would be helpful in providing readers with more context. Unfortunately, data regarding joint replacement programs and the outcomes of these programs is not readily available. We have included an additional sentence located in the methods section under the “setting” sub-heading to provide as much information as we know (page 6):

Other private hospitals throughout the country also provide hip and knee replacements to patients who are able to pay for the procedures’ high costs, but information regarding the number of joint replacements and the outcomes of these other joint replacement programs is not available.

Major essential revisions

Abstract, Results

2. The results describe the criterion categories which were met and not met. It would be useful to describe some of the actual criteria where the program didn’t meet the requirements. I realise space does not permit the listing of all of these, but some examples might be useful for readers to understand what an “informational” criteria is, for instance.

Response: The reviewer suggests that we more explicitly identify examples of criteria that were not met. While such examples are interspersed throughout the text, we appreciate the suggestion to more specifically mention these examples. In response, we indicate that the criteria are specifically mentioned in Tables 1 and 2, and we have also provided examples below that are pasted verbatim from the manuscript. For example, in the “General Criteria” category (page 9):

The program did not meet one of the required criteria because the host hospital is not accredited by a CMS-deemed national accreditation organization. The two unmet informational criteria relate to using a Surgical Care Improvement Project (SCIP) database to produce procedure specific performance reports and to tracking FDA-recalled prostheses and contacting patients with these prostheses. The hospital does not have a SCIP database and it also has difficulty tracking patients, which makes it difficult to contact patients whose prosthesis are recalled.

From the “Structure” category (page 9):

The only criterion that the program did not fully meet involved reporting to surgical quality improvement registries and databases. The program also lost three points because it lacks three out of eight required multi-disciplinary team members: psychiatrists and psychologists, pain management specialists, and dedicated case
managers. Despite lacking psychiatrists and psychologists and formally trained pain management specialists, the program compensates by having doctors and nurses work directly with patients to address their mental health needs and by having well-trained anesthesiologists who commonly provide pain management services.

Similar entries exist in the “Process” and “Outcomes and Volume” categories.

To further clarify what each category means (required vs informational), we have included an explanation in the last two lines of the third paragraph from the “Data Collection and Analysis” subsection of the Methods section (page 8):

Points marked as “informational” represent criteria that are not formally scored in the Blue Cross/Shield’s criteria but are still important when evaluating joint replacement programs. Points marked as “required” are deemed essential for achieving the Blue Cross/Shield’s distinction.

Methods, Analysis

3. Who scored the program criteria? Was the reliability of this scoring assessed in any way? If the scoring was not assessed for reliability, this should be mentioned in a limitations section.

Response: The first reviewer raised a similar question. We repeat our response here: The study’s primary author used data collected through interviewers to score the criteria. The reliability of the data was verified by the fact that two independent people (program directors) in separate interviews would have to respond similarly to each scoring criteria. We have clarified the scoring process in the methods section, under the “Data Collection and Analysis” subheading (page 7):

The study’s main author compiled the Blue Cross/Shield’s criteria and independently reviewed each criterion with two of Op-Walk Boston’s program directors during in-person interviews. The directors independently responded to each criterion with “meet”, “do not meet”, or “unsure”. For criteria that produced conflicting results, the criteria were sent to an independent person familiar with the specific criteria in question and this third party’s response served as a tie-breaker. Similarly, for all criteria that lacked responses from both directors, the study’s main author redirected the question to an independent team member who had the most knowledge of that criterion.

To explicitly address the fact that our data may be variable due to the small number of people surveyed, we included the following sentences in a new subsection (Limitations) of the discussion section (page 15):

Furthermore, the data for this study was collected from a small number of respondents who provided key information to evaluate if the program meets or does
not meet the criteria; it is therefore possible that some data variability would exist if more people were surveyed.

4. The last sentence of this paragraph states, “We also asked team members if there are barriers…” Were formal interviews or surveys conducted? Was a wide range of team members targeted to obtain a representative view?

Response: We interviewed one clinical team leader from each respective clinical team. To clarify this point, we altered the last two sentences of paragraph 2 under the “Data Collection and Analysis” of the Methods section. It now reads (page 8):

For all criteria that are not replicated, we interviewed a leader from each clinical and administrative team to see if they compensate for failing to meet the criterion by introducing an alternative strategy or process to enhance quality. We also interviewed these key informants to see if there are barriers that prevent them from adopting certain criteria.

We recognize that interviewing only one team member may introduce bias, which we addressed in the Limitations section of the discussion section (page 15):

Furthermore, the data for this study was collected from a small number of respondents who provided key information to evaluate if the program meets or does not meet the criteria; it is therefore possible that some data variability would exist if more people were surveyed.

Results section

5. Can you please present some of the information that was collected on the barriers in the results section? Are there common themes that have emerged within these, such as resourcing issues or medical cultural? Are there plans to address any of these?

Response: We are pleased to elaborate on common themes that prevented Op-Walk from meeting the criteria. In the results section, under the subheading “Overall Evaluation”, we have included a new paragraph that elaborates on the common barriers that prevented Op-Walk from meeting many criteria (page 12):

Common barriers were identified that prevented some criteria from being met. First, many of the criteria require registration and participation in organizations that only exist in the United States or other developed countries. Second, many criteria were not met because they require frequent patient contact and PCP follow up. Third, some criteria require outcomes and patient tracking through online databases that are not currently available in many hospitals in the D.R.

It is not possible to meet some of these criteria because they do not apply to hospitals outside of the U.S. This fact is noted in the last paragraph of the discussion section:
Since some aspects of the Blue Distinction criteria do not apply to programs outside the US, future work should focus on designing criteria that are more applicable to international contexts.

Discussion section, paragraph 3

6. The authors mention that initiating a patient navigation program and evaluating patients’ discharge needs are two areas where the program lost points and these areas have been highlighted for improvement. I would like to see more of a discussion of the barriers that were identified in addressing these issues—e.g. is there a lack of resourcing to be able to provide many of these services? Should the Blue Distinction criteria then be modified for health services in developing country settings or should these issues ultimately be addressed?

Response: Identifying common obstacles to achieving Blue Distinction criteria is an important aspect of this paper—we agree that more discussion should focus on how the existing criteria can be altered for wider applicability in international settings. We have therefore altered the sixth paragraph of the discussion (page 14) to now read:

Since some aspects of the Blue Distinction criteria require organizations to report to US-based quality improvement organizations, future work should alter the existing criteria so that these organizations can earn points for reporting to equivalent international quality improvement organizations or provide waivers for organizations that operate in countries without equivalent quality improvement organizations. Furthermore, some Blue Distinction criteria require investment in patient navigators or expensive health care infrastructure, which is difficult because cost is a common barrier for most international medical missions. The criteria should therefore be redesigned so that they can be implemented with a level of investment more congruent with the resource capacity in the country being evaluated.

7. Could the authors please discuss whether there are criteria which will be addressed as a result of this project? For example, using shared decision making could be implemented at a relatively low cost, but it seems that there are cultural barriers. Can these be changed within this context? Or where cost is a barrier, is it possible to attract resourcing to address the issues?

Response: We appreciate the reviewer's observation that these analyses illuminate potential areas that the Op-Walk program can address to improve care. The example of shared decision making is apt and we plan to discuss with Op-Walk’s medical, nursing, and rehabilitation colleagues in the Dominican Republic ways of amplifying the patient’s voice in decision making, in ways that are culturally acceptable. The reviewer also points out correctly that financial resources can overcome some of the barriers. We have discussed our findings carefully with administrative and clinical leaders in the
Dominican Republic so that the observed deficiencies can be considered for funding in subsequent budgeting processes. To make this information clear to the readers of this study, we have added three new sentences to the third paragraph of the discussion section (page 13):

We plan to discuss with Op-Walk’s medical, nursing and rehabilitation colleagues in the Dominican Republic culturally acceptable ways of amplifying the patient's voice in decision making. Some of the barriers we identify could be overcome with greater funding. We have discussed our findings with administrative and clinical leaders in the Dominican Republic so that the observed deficiencies can be considered for funding in subsequent budgeting processes.

Table 1

8. The criterion entitled “If facility does not report to Leapfrog, facility participates in other initiatives that encourage the sharing of best practices, incorporates data feedback for objective analysis, and promotes collaborative improvement” scored 0 points but it says that the criterion has been met. If this is an optional/unscored criteria, this should be specified for clarity.

Response: Thank you for the very close reading of our tables! This should be labeled as an optional criterion. We have made the following change to table 1:

“0/0” to “Optional”