Author’s response to reviews

Title: The Six-minute walk test is an excellent predictor of functional ambulation after total knee arthroplasty.

Authors:

Victoria W Ko (vwmko@yahoo.com)
Justine M Naylor (justine.naylor@sswhs.nsw.gov.au)
Ian A Harris (jaharris1@gmail.com)
Jack Crosbie (J.Crosbie@uws.edu.au)
Anthony ET Yeo (anthony.yeo@uni.sydney.edu.au)

Version: 4 Date: 8 March 2013

Author’s response to reviews: see over
8th March, 2013.

Prof Gregoire Millet
Editor
BMC Musculoskeletal Disorder.

Re: MS: 1296994407790849
Research article
The Six-minute walk test is an excellent predictor of functional ambulation after total knee arthroplasty.

Dear Prof Millet,

Thank you for your interest in this manuscript and the opportunity to address the questions and concerns raised by the Consultant Referees. The name of the ethics committee has been inserted as per Editorial requirement. Each referee question is addressed below (referee questions are in italics). We have also made some minor changes in the manuscript for clarity (denoted in bold in the manuscript).

Referee 1:
Major compulsory revisions:

1. For sake of clarity, the authors should report only the ‘overarching aim’ in the Abstract (p. 3, at the end of the first paragraph “Background”). Please change to (or similar): “The overarching aim of this study was to compare the performance of total knee arthroplasty (TKA) recipients in an extended walk test to healthy, aged-matched controls in order to determine the utility of this extended walk test as a research tool to evaluate longer-term functional mobility in TKA recipients.”.

We have adjusted the Abstract (Background) to accommodate the reviewer’s suggestion.

Abstract

Background: The Six-minute walk (6MW) and Timed-Up-and-Go (TUG) are short walk tests commonly used to evaluate functional recovery after total knee arthroplasty...
(TKA). However, little is known about walking capacity of TKA recipients over extended periods typical of everyday living and whether these short walk tests actually predict longer, more functional distances. Further, short walk tests only correlate moderately with patient-reported outcomes. The overarching aim of this study was to compare the performance of TKA recipients in an extended walk test to healthy age-matched controls and to determine the utility of this extended walk test as a research tool to evaluate longer term functional mobility in TKA recipients.

2. The authors should strengthen the point concerning the greater functional validity of the ‘longer walk test’ quoting some references that investigated this issue in other similar populations (p. 5, ll. 6-11). This may improve the construction of the rationale of the study. For instance, Butland et al. (British Medical Journal, 1982) have previously compared 2-, 6-, 12-min walking tests in respiratory disease and they concluded that: “The time chosen to assess exercise tolerance by walking tests is not critical. Shorter times are easier for both patient and investigator and are as reproducible but discriminate slightly less well and have less of a training role. The six-minute walk may represent a sensible compromise.”. Are there other similar studies concerning this type of comparison in the literature? Please check and modify this paragraph.

Thank you for this suggestion. We have inserted the following paragraph describing 2 studies which have investigated the use of a longer duration walk test in other patient populations in the Introduction section.

“The use of extended walk tests has been investigated in other populations. In a study of the two-, six-and 12-minute walk tests in patients following stroke, the 12-minute walk test was observed to be the most responsive to change [14]. Further, despite high correlation between the three tests, the two-minute assessment was reported to overestimated 6 and 12 minute walking distances. The authors suggested the overestimation may be due to fatigue during the longer tests. In another study investigating patients with stable chronic airflow obstruction, a 12-minute walk test was also found to be highly correlated to six-minute and two-minute walk tests[15]. Variance of the 12-minute walk test was greater than the shorter walk tests, suggesting that the longer test was perhaps more discriminating. However, as the shorter walk tests were easier for both the patient and researcher, the authors recommended the six-minute walk as a fair compromise. Compared to other patient populations, TKA recipients may be affected by different factors which limit walking capacity, such as residual knee pain and reduced muscle strength. As such, it cannot be assumed that the results of these studies can be generalized to the TKA population.”
3. The authors should at least write in the Introduction or in the aim of the study that the functional duration of the longer walk test would have been determined during a previous, pilot survey of patients awaiting knee and hip replacement.

We assume this comment is based on the fact that the pilot survey was conducted using a different patient cohort and this is not clear in the aims. For clarity and brevity, we have removed ‘the establishment of a pertinent time frame’ from the Aims and simply referred to the establishment of a relevant time parameter using a pilot survey as a preliminary stage of the study in the Methods section.

“Prior to undertaking the definitive study, a pilot survey of patients awaiting knee and hip replacement was performed to determine a time parameter for an extended walk test which may be functionally relevant for this patient group.”

4. The working hypothesis is lacking and the authors should add it at the end of the Introduction.

A statement about our main hypotheses has been added to the Introduction section.

“The main hypotheses to be addressed are that: TKA recipients one year post surgery will perform significantly worse than age-matched healthy controls in an extended walk test; the shorter walk tests will not predict performance in the longer walk test, and; the longer walk test will be a stronger predictor of self-reported function and physical activity than the shorter walk tests.”

5. The structure of the Methods seems to be inappropriate. The authors should use this structure to avoid also the repetitions in the text: 1) Preliminary assessments for defining the functional duration of the longer walk test; 2) Participants with description of the anthropometric measurements (p. 8, last paragraph ll. 1-2); 3) Experimental design (rapid and brief timing description of the experimental design); 4) Tests (30MW, 6MW, TUG, WOMAC and physical activity questionnaire: test descriptions and main outcomes for each test); 5) Statistical analysis.

The structure of the Methods has been re-organised under the headings as recommended by the reviewer (see Manuscript).

6. The description of 6MW and TUG tests should be improved and more detailed in the manuscript in the suggested new structure of the Methods (Paragraph “Tests” here above). Which was the length of the lap for 6MW? How many repetitions did each patient perform in 6MW session? Etc.
Greater descriptions of the TUG and 6MW tests have been added to the Methods section.

“**TUG.** The procedures for the TUG test was based on published protocols [22]. A standard height chair of 45cm with arm rests was placed on an outdoor level footpath, and a line 3 metres from the chair was drawn. The participant was instructed to stand up from the chair, walk to the line, turn around, walk back to the chair and sit down as quickly and safely as possible. The assessor commenced timing as the participant leaned forward to stand up, and ceased when the participant’s hips made contact with the seat to sit down. A minimum of two tests was performed for each participant with the best time used in the analysis. High repeatability of the TUG has been established in patients awaiting TKA (ICC 0.75) [23].

**6MW test.** The procedure for the 6MW test was based on published guidelines[24]. A 25 metre section of the outdoor footpath was demarcated for this test. The participant was instructed to walk as far as possible for six minutes, up and down the demarcated footpath, pivoting to turn at the end of each lap. Timing commenced as the participant stepped over the start line. Standardized encouragement was given to the patient after each minute. The participant was instructed to stop at six minutes, or prior if they were unable to complete six minutes, and to maintain their position whilst the assessor measured the final partial lap with a trundle wheel. The use of a walking aid and standing rests were permitted. One test was performed for each participant as we did not want fatigue from completing two 6MWs to undermine performance in the longer test. Standardized encouragement was given at every minute. High repeatability of the 6MW test has been established in patients awaiting TKA (ICC 0.94) [23].”

7. The paragraph “Testing/protocol/procedure” should be incorporated in the paragraph on the 6MW test.

The paragraph “Testing protocol / procedures” has been incorporated in the new Methods structure.

8. The authors well described and reported the criteria of the selection of the sample size.

Comment only – no response required.
9. Did the authors assess Rating Perceived Effort (RPE) with Borg scale during and/or after 30MW and 6MW tests?

Rating of Perceived Effort (RPE) was not recorded for the 6MW or 30MW. It is not clear whether this index is useful after arthroplasty as per the following studies.


10. How many stops did the individuals make during 30MW and 6MW tests? Is there any significant difference between the two tests and/or the 2 groups?

Description and comparison of the rests taken by participants in the 30MW have been added to the Results section.

“No participants stopped to rest during the 6MW test. Significantly more participants amongst the TKA cohort required rests compared to controls during the 30MW test (12/32 TKA; 0/43 Controls; p<0.001). The number of rests ranged from one to more than ten occasions.”

11. The most important concern is the similar average speed of the 2 walking tests in the 2 groups. It is difficult to explain that the average speed was the same in 30 min and in 6 min walking tests in TKA recipients and control counterparts. Which were the instructions given to the participants? The authors discuss at the p. 14 (first paragraph) that TKA recipients adopted the same pacing strategies to complete both tests and that “the factors limiting their mobility (as described above) also limit walking speed”. However, the walking speed in both tests is similar also in the control group. Therefore, the explanation used by the authors is not suitable to discuss this result in both groups. How may the authors explain this finding in control group? Please clarify and change in the text.

Thank you for this query. We have reflected more on this and have come up with, we believe, a more satisfactory explanation.

The healthy cohort had a greater average walking speed in the 30MW test (1.71m/s) than the 6MW test (1.61m/s, P < 0.001). It is likely that the greater speed was facilitated by the longer lap used in the 30MW test. That is, the frequency of deceleration and turning is far less in the 200m lap compared to the 25 m lap.
The TKA cohort manifested similar walk speeds between the two tests. On reflection, when we applied the logic that walk speed should have been facilitated by the longer lap, it was apparent that the walk speed of the longer test of the TKA cohort was relatively slower.

To investigate this further, we separated TKA recipients who rested from those who did not, and for each group, compared the 6MW and 30MW speeds. For TKA recipients who rested (n=12), 6MW speed was 0.90m/s compared to 30MW speed 0.82m/s (p=0.08). For these participants, it is likely that the momentum gained by a longer track was negated by the rest periods taken in the longer test. For TKA recipients who did not rest (n = 20), there was a trend for a faster speed in the 30 MW test, but the difference did not quite reach statistical significance (1.34m/s vs 1.38m/s, p = 0.05). For these participants, the similar walking speeds between the two tests suggest that whilst there may have been some facilitation in gait speed owing to the longer lap, this effect was possibly countered by fatigue, hence there was no measurable change in average gait speed between the two tests.

Hence, regardless of whether TKA recipients rested or not, we infer then, that the longer test was more fatiguing than the short test for the TKA cohort; average speed was protected by the facilitating effect of a longer lap. Consequently, distance covered in the 6MW predicted distances covered in the 30MW, thus, there were very strong associations between the two tests.

To clarify these points, we have inserted the following in the Results and Discussion sections.

“Control subjects had a significantly faster walking speed in the 30MW test than the 6MW test (t = 4.82, P < 0.001), while TKA recipients had very similar walking speeds (t = 0.22, P = 0.83). Subgroup analysis found that for TKA participants who rested (n= 12), 6MW speed was 0.90m/s compared to 30MW speed of 0.82m/s (P = 0.08). For TKA participants who did not rest (n = 20), the difference between 6MW speed and 30MW speed was minor and did not quite reach significance (1.34 m/s vs 1.38 m/s, P = 0.05).”

“The near-perfect correlation between the 6MW and 30MW tests and the high predictability of 30MW distance by 6MW distance were unexpected. This is because participants were instructed to walk as far as possible but safely, and as such we expected that some participants would not be able to maintain the same speed in both tests, presuming a comparatively poor fitness level in the cohort. We observed that the healthy cohort walked faster in the longer test. This may be due to less frequent deceleration and turning associated with the longer lap in the 30MW test. In contrast, the walking speeds of TKA recipients were identical in both tests. The gain in walk speed associated with a longer track was not seen in the TKA cohort, suggesting that the 30MW walk test was indeed more fatiguing for this cohort. For some participants, the overall walk speed was reduced as a result of the rests taken during the 30MW test and these people demonstrated the slowest speeds in the 6MW test. For those who did not rest, there was a trend for a slightly faster speed but this did not quite reach statistical significance. Consequently, whilst fatigue
likely did become a factor in the longer test, its effects were small, and those most affected were those who performed worst in the 6MW test anyway. Thus, the prediction of the 30MWD from the 6MWD was high.

12. There was a significant difference in BMI and physical activity scores (Table 1) between the 2 groups. These parameters may be a possible confounding factors to explain difference in walking performance in 30MW and 6MW tests in the 2 groups. The authors should therefore discuss this important point in the Discussion.

We agree that the higher BMI and lower physical activity scores are possible confounders in walking performance of the TKA cohort. However, even when the higher BMI is taken into account by using published reference equations that factor in height and weight, the TKA cohort was still substantially below their predicted performance. This suggests that in the TKA cohort, the contribution of the higher BMI to poorer performance is relatively small.

Regarding physical activity levels, we cannot determine the degree to which the lower physical activity level in the TKA cohort contributed to their poor performance. However, it is very likely that the low physical activity levels is closely associated with, or a manifestation of, the comorbidities, other joint disease and obesity that are commonly present in the TKA recipient. Hence, the effect of lower reported physical activity on performance can be attributed to the characteristics of a typical TKA recipient.

We have discussed the effect of significantly different BMI and reported physical activity on performance between the health and TKA cohort in the Discussion session.

“Lower physical activity levels reported by the TKA cohort, which are a likely manifestation of the co-morbidities, joint disease and obesity associated with patients requiring TKA, may also be a contributing factor to their poorer performance.”

“Using the reference equation published by Troosters et al. [16], our healthy cohort performed very similarly to their predicted distance (582 m vs predicted 576 m), while the performance of our TKA cohort was much lower than their predicted distance (424 m vs predicted 538 m). As this reference equation takes into account age, gender, height and weight (the constituents of BMI), it demonstrates that the higher BMI of the TKA cohort only makes a minor contribution to the large difference in walking performance compared to the healthy cohort.”

13. The authors should add a comparison between the 6MW distances of the two groups and the standard values calculated with the reference equations for the 6MW test (Enright et al., American Journal of Respiratory and Critical Care Medicine 1998; Troosters et al., The
“The observed residual impairment in mobility amongst our TKA cohort is further supported by comparisons between their 6MW test performance with published standard values of healthy elderly populations. Using the reference equation published by Troosters et al. [16], which takes into account age, height, weight and gender, our healthy cohort performed very similarly to their predicted distance (582 m vs predicted 576 m), while the performance of our TKA cohort was much lower than their predicted distance (424 m vs 538 m). Using reference equations published by Enright et al. [32], which have been considered by some to underestimate normal walking distance [33, 34], the performance of our TKA cohort was still inferior, with the predicted distance for TKA group of 464 m.”
5. P. 6, second paragraph: add the numbers for each specific aim. “[...] The specific aims of the study were multiple: 1) [...] surgery; 2) [...]”.

The specific aims have been numbered.

6. After the title of each paragraph, the authors should add a dot (e.g. “Protocol development.”, p. 6 first paragraph of the “Methods”).

A full stop has been added after the headings in Methods section.

7. P. 10, paragraph “Statistical analysis”, last line. Please change the sentence to “Data from the first testing session of 30 MW were used in the regression analyses.”.

“of the 30MW test” had been added to the last sentence of Statistical Analysis section.

8. P.11, last paragraph, l. 4. Add “)” after “3”: “(Table 3)”.

“)” has been inserted after Table 3.

9. Why was the distance walked significantly higher in the second 30MW test only in controls? Please briefly discuss.

Discussion of the possible reason for significant better performance in the second 30MW test in the control group has been added to the Discussion section.

“The 30MW test was observed to be highly repeatable. The healthy cohort walked significantly further by 50m in the second test. This difference may be due to a practice effect, which has been described by some authors in the 6MW test [16]. It is likely that a significant increase was not seen in the TKA cohort because there was greater variability in the performance, as indicated by a larger standard deviation of the difference between the first and second test (TKA SD 215 m vs Controls SD 93 m).”

10. Table 1 (p. 24). The authors should define all abbreviations used in the tables and in the Figures in their Legends (e.g., Table 1, define: CI, IQR, n, WOMAC and TKA). Please check and add throughout the manuscript.

Definitions of abbreviations have been added to all tables and Figure legends.
11. *Table 1* (p. 24, l. 18). Change “IRQ” to “IQR”.

Typographical error corrected.

12. *Legends* (p. 23). Why are the numbers of subjects (n) different for TUG and 6MW? Clarify and justify in the text of the manuscript. Moreover, add the “n” for 30MW.

The number for 30MW test has been added. The different number for TUG and 6MW test was due to missing data from administrative error, and explained in the following sentence in the Results section.

“In the data set, there were three missing values of TUG in the TKA cohort and one in the healthy cohort due to administration error.”

13. *Figure Legends* (p. 23). Please add “r”, “p” and the equations of the two correlations.

The r and P values, and equations of the two correlations have been added to the figure legend.


A full stop has been added after “et al”.

15. *P. 15*, first paragraph, l. 10. Please add a reference after “[...] actual physical activity.”.

The following reference has been added.


16. *P. 15*, the last paragraph, l. 7. Change the quoting format of reference #33 in the text.

The quoting format has been corrected.
Referee 2:

Major compulsory revisions:

1. Throughout the paper, a main piece that is missing is any discussion of pain in the TKA population. To begin, in the results section, it is mentioned that the TKA group found the test difficult, I think a reader would appreciate knowing if that was due to pain, endurance or muscle fatigue, or other reasons. Further, in the discussion section, the argument is made that pain may be one reason for inferior performance in the TKA group, however this was not measured in this study; additionally, pain is rarely an issue 12+ months after surgery. If possible, could you include reported pain levels during these tests? Also please include a discussion about the effect of pain on these tests.

Pain levels were not measured during the walk tests. We agree that at 12+ months after surgery pain in the index knee is less likely to limit mobility. Pain of the index joint was only reported by 5 participants as the limiting factor in the 30MW test. Other factors limiting performance in the 30MW test reported by TKA recipients included back and other lower limb joint pain, and poor aerobic fitness. To clarify this, the number of participants reporting these factors has been added to the Results section.

“TKA recipients reported factors which limited their performance included pain in the index knee (n = 5), back and other lower limb joint pain (n=13), and poor aerobic fitness (n=10).”

Minor essential revisions: None

Discretionary Revisions:

1. Methods Section/Sample size: While this explanation of choosing a sample size appears logical, it is hard to determine if you were underpowered, overpowered or sufficiently powered to detect difference in this study. If no data was available for estimating a sample size, have you calculated power from the results of this study to determine power level?

Using the results of this study, (overall standard deviation of 735m for the 30MW distance), we have calculated that with a sample of 75, we can detect a difference of 591m, at 90% power. As the difference between the healthy and TKA cohort was 978m, we were sufficiently powered to detect this difference. We have not included this information in our manuscript as this revision was discretionary.

2. Methods Section/testing protocol/procedure: the final sentence regarding 30MW test was performed after HR returned to resting levels should be moved earlier in the paragraph for clarification. Also, the notion of whether fatigue played a role in testing for either population could be addressed in the discussion section.
The Method section has been reorganised and the statement regarding performing the 30MW test after a resting period is now earlier in the section.

The ideal method to eliminate the effect of fatigue would be to randomise the order of testing. However, if the 30MW test was completed first, participants would need a significantly longer resting period (hours) before they could complete the 6MW test. This would not have been practical. Nevertheless, we do not believe the 6 MW test in itself had a fatigue effect on the 30 MW test. If it had, we should have seen a better 30 MW test on the 2nd test amongst the TKA cohort because that test was not preceded by a 6 MW test unlike it was in the first test. In our experience with a trial of 250 TKA recipients in which the 6MW test was conducted at one year after surgery, while performance was less than normal, participants did not report fatigue as a limiting factor, unlike for walk tests we have done with OA populations.

In response to this comment we have added the following in our discussion.

“Not randomizing the order of testing to eliminate the effect of fatigue during testing could be viewed as another potential limitation in this study. However, the amount of time required after the 30MW test to ensure sufficient rest made this impractical. Furthermore, it was unlikely that fatigue affected testing because amongst the TKA cohort, in whom fatigue would have a larger potential impact, those who returned for a repeat 30MW test (which did not follow a 6MW test) did not perform significantly better in the second test.”

3. Methods Section/statistical analysis: Each model was limited to 3 independent/explanatory which were decided upon apriori. These variables could be mentioned here.

The 3 variables are stated in the Statistical Analysis section, “To compare the predictive value of the three walk tests, three models were generated for each dependent variable using gender, body mass index (BMI), and one of the walk tests as independent variables”

Thank you for considering our revised manuscript.

Yours sincerely,

Victoria Ko (corresponding author), Justine Naylor, Ian Harris, Jack Crosbie, and Tony Yeo.