Author's response to reviews

Title: Randomised controlled trial of an automated, interactive telephone intervention (TLC Diabetes) to improve type 2 diabetes management: Baseline findings and six-month outcomes

Authors:

Emily D Williams (emily.d.williams@monash.edu)
Dominique Bird (dbird@coh.uq.edu.au)
Andrew W Forbes (andrew.forbes@monash.edu)
Anthony Russell (Anthony_Russell@health.qld.gov.au)
Susan Ash (s.ash@qut.edu.au)
Robert Friedman (Robert.Friedman@bmc.org)
Paul A Scuffham (p.scuffham@griffith.edu.au)
Brian Oldenburg (brian.oldenburg@monash.edu)

Version: 4 Date: 1 May 2012

Author's response to reviews: see over
Dear Professor Lotufo

Thank you very much for providing us the opportunity to respond to the reviewers’ comments and to provide a revised manuscript that addresses the issues raised by the reviewers.

Reviewer 1:
The authors need to discuss the limitations of the study as the discussion is very positive towards the intervention. No mention is made of the impact of the smaller sample size and the impact of the methodological adjustments which took place on the reliability of the findings for example. There are limitations of the intervention also such as those associated with socio-economic factors and the fact that to participate you need reliable access to a telephone.

In line with Reviewer 1’s comment, we have developed our discussion of the limitations of our study on page 15. This section now reads as follows:

“Although only glomerular filtration rate significantly varied across the study arms at baseline, other baseline characteristics (Table 2) showed some differences. Separate analyses tested the impact of the inclusion of these variables on the main results and the main outcome results did not change. As with most research, it is possible that a selection bias operated in this study, with people willing to participate being more likely to prioritise their health and/or have the social, educational, and economic resources to accommodate participation. The study requirement of access to a telephone meant that there may have been a socioeconomic selection bias; however in the geographic area from which we recruited, over 96% of households have a fixed phone connection, so we are confident that this criterion did not appreciably influence participation. It is also possible that the reduced sample size and some of the challenges associated with trial recruitment may limit generalizability. More research is required to investigate generalizability and to explore uptake by others with diabetes.”

Reviewer 2:
1. I am most concerned about the balance of characteristics between the two arms. Even though there may be no statistical differences except in one or two variables, there are clinically important differences in major confounding factors such as age, levels of education, self care etc which may affect the results such that it explains the apparent benefit of the self management arm. It is essential that a statistical reviewer comment on this and that these factors are explored in adjusted analyses to see if they affect the results.

We appreciate the reviewer’s concern about the balance of characteristics between the two study arms. Strict randomization procedures were applied and therefore any differences between the two arms were due solely to chance from the random allocation. We tabulated a total of 45 baseline characteristics and therefore some imbalances due to chance in a few factors, as observed, are to be expected. The importance of imbalances in baseline characteristics is a combination of the size of the imbalance together with how prognostic the particular factor is. Although it is debatable as to the clinical importance of the observed baseline differences (e.g. a 2 year difference in average age), we have now performed additional analyses adjusting for age, education, adherence to blood glucose testing recommendations and daily insulin/diabetes medications, and foot inspections. This adjustment did not affect the group differences in follow-up.
outcomes observed previously, with group differences at follow-up remaining for HbA\textsubscript{1c} ($p = 0.007$) and mental health-related quality of life (HRQL) ($p = 0.006$), and no differences between study arms in follow-up physical HRQL ($p = 0.76$). As shown above, in response to the reviewer’s first comment, discussion of this additional adjustment has been added to the Limitation section of the Discussion on page 15.

2. Although the text states that there were no significant differences between the 2 arms except the glomerular filtration rate, the values given in the table do appear to show material differences in many variables e.g. education, levels of depression, physical activity, adherence to guidelines etc which may not be significantly different, but which show a trend to the intervention group being more healthy and better at adherence etc. This section should be looked at by a statistician and a decision made about whether these differences may affect the results, and what measures should be taken to try to adjust for them.

As per the response to point 1, we have now adjusted for many of the characteristics that show the greatest differences between the study arms at baseline and the intervention effects on the main outcome variable were unchanged.

3. This should also be described and discussed appropriately in the text of the results and discussion. It is dealt with only briefly and inadequately in the text of the results.

The following additional sentences have been included in the document to address the non-significant baseline differences between the study arms.

On page 9-10, in the Statistical Analyses section of the Methods, the following sentence was added: “The sensitivity of conclusions to imbalances in baseline characteristics was assessed via additional ANCOVA analyses adjusting for all characteristics exhibiting any potentially important imbalances.”

The Results section on page 10 now reads: “Comparison of the baseline characteristics across usual care and intervention arms revealed important differences in e-GFR, which showed a significantly greater impairment in renal function in the intervention compared with usual care arm, and creatinine. Other differences observed were in age, education, and self-care behaviours (adherence to blood glucose testing recommendations and daily insulin/diabetes medications, and foot inspections). Adjustments were made for these variables in sensitivity analyses”.

4. The clinical reasoning behind adjustment for glomerular filtration rate is not evident.

The decision to adjust for eGFR (and creatinine) was based on the large and clearly clinically important difference between study arms at baseline.

5. METHODS: The allocation ratio, method of allocation and allocation sequence generation are not stated. Even if these were in the protocol publication they should also be in the report.

The allocation ratio was 1:1 and the allocation sequence was computer-generated. The arm allocation was conducted using a 4x4 block randomised block design with the participant as the unit of randomisation. This information has been included on page 5 of the Methods section.

6. Table 2 provides a comparison of the 2 arms. The text provides a description of the overall sample which is not reflected in the table. The table should also have a column which describes the overall sample, and the text should describe any differences between the two arms.

As requested by the reviewer, Table 2 now includes a column that describes the overall TLC sample. As discussed in response to the reviewer’s third comment (above), the differences between the two arms are stated on page 10.
METHODS

7. Blinding: were their usual care providers blinded to their allocation?

The treating physicians were not blinded. The following comment was added on page 5 in study arms: “The treating physicians were not blinded to the allocation.”

8. Please make the intervention section clearer by including appropriate sub-headings for each topic eg Overall aims of the intervention; baseline visit; weekly telephone calls; topics/responses; follow-up etc.

We thank the reviewer for this suggestion and have now added sub-headings to pages 5-7.

9. Outcome variables: please provide more information about the blood tests. Were these distinct from the weekly blood glucose testing fed back to the TLC? Where were they undertaken and by whom? Were they fasting?

The blood tests taken as part of the clinical assessments at baseline and follow-up were distinct from the weekly blood glucose testing that the intervention participants completed and uploaded to the TLC system. The usual care group did not provide these weekly data, but did participate in the clinical assessments.

The blood samples were taken at the two hospitals where appointments were held and blood tests were conducted by Queensland Pathology using standardised assays. This information has been clarified in the manuscript on page 4, with the following text. “At that appointment, baseline questionnaires were completed and fasting blood specimens were taken, along with other clinical data (blood pressure, weight, height and waist circumference). Blood tests were conducted by Queensland Pathology using standardised assays.”

Blood tests were taken when participants were fasting. This information is available in the final sentence on page 4 and 7.

10. In the statistical analysis section there is a sentence describing Table 1 which should be in the results section and also labelled Table 2.

We thank the reviewer for highlighting this error; this has been amended.

RESULTS

11. In table 3 and the text it must be made clear that this is an ITT analysis and give the numbers analysed for each group.

A sentence has been added to page 11 (Study Outcomes section): “These analyses were based on intention-to-treat”. “n=60” was added to the headers of each appropriate column in Table 3 to indicate that all randomized patients were included in the analyses, i.e. intention-to-treat analysis was performed as specified in the methods section.

12. Why is the difference presented as a ratio? And should this have a %? This is confusing.

The difference is presented as a ratio because logarithmic transformation was applied to reduce skewness for the analysis. The analysis produces differences in averages on the logarithmic scale, which then translate to ratios of geometric means on the HbA1c scale. The reviewer is correct that there should not be a “%” sign and we apologise for that error and have corrected it. The ratio of 0.91 means that the geometric mean HbA1c at six months in the TLC arm is 0.91 of the value in the usual care arm after adjustment for baseline covariates. This final sentence has been added to the results on page 11.

13. We need p values and confidence intervals in the text to describe % achieving recommended HbA1c levels etc.

These values have been added to page 11.
14. Comparing with AUSDiab: There is no table 4 – do you mean table 2? Please correct this.
The reviewer is correct. The text should read table 2. We apologize for this. This has been corrected in the manuscript on page 12.

15. The abstract needs to include the fact that diabetes patients were required to have HbA1c > 7.5% and also in the results that the physical HRQL was unaffected.
The Abstract now contains mention of the HbA1c selection criterion and the physical HRQL result.

Reviewer 3:

1. I am not as familiar with seeing differences in geometric means over time, rather than raw A1c values. Perhaps the arithmetic mean by arm could be included (at least in the text, as this was used as inclusion criterion), even if the regression analyses log-transform these values due to the skewed nature of the data. In addition, it is not clear to me that the Table 1 A1c values are arithmetic or geometric means.
The arithmetic means have now been included in parentheses to the text in the Results section on page 11. As for all of the other variables with data that were not normally distributed, Table 2 presents HbA1c values as medians and interquartile ranges.

2. The results on page 10 state that there is “little evidence of difference between study arms varied with baseline A1c”, but p-value for the interaction is marginally significant at 0.09 (especially given the limited power of this sample size). I think this wording should be less strongly stated. I believe the TLC arm started the trial with somewhat lower A1c values and still made more significant improvement compared to usual care.
The text on page 11 has now been altered to read:
“There was slight evidence that the difference in HbA1c at six months between study arms increased with baseline HbA1c (p=0.09 for the interaction term in regression model). This suggested that the difference in six-month HbA1c between TLC and usual care patients was greater in patients with high baseline HbA1c values than in patients with low values.”

In addition, the following statement has been added to the discussion on page 15:
“Although there was a suggestion of an increasing effect of intervention with increasing baseline HbA1c values (from the interaction test), this did not reach conventional levels of statistical significance and should be reassessed in future studies.”

3. I wanted clarification about the AusDiab comparison sample, especially since the sample size seemed small. If there were over 10,000 individuals in that study (I thought all of whom had diabetes based on the Methods description?), how was this limited to the 156 for the comparison group? If this smaller comparison group is still the most representative of the diabetes population in Australia, then I think it could be fine, but I was wondering about the statistical power to detect a difference in sample characteristics between the survey and the trial participants – that could be a limitation.
The description of the AusDiab study has been amended on page 7 and 11-12, to better reflect the AusDiab study’s aim to measure diabetes prevalence. The comparison group taken from the overall AusDiab sample was the total number of people with diagnosed diabetes, who fit the TLC inclusion criteria. These criteria were applied to provide the closest comparison to the TLC sample.

Page 7: “To examine the representativeness of the Australian TLC Diabetes sample, the baseline characteristics were compared with data from the Australian Diabetes, Obesity and Lifestyle (AusDiab) study [29], the largest national, population-based diabetes prevalence study in Australia. This nationally representative study’s subsample of individuals identified as having diabetes (and based on TLC inclusion criteria) provides the best comparison for the TLC study sample.”

Page 11-12: “To determine the representativeness of the TLC sample at baseline, we used a comparable subsample of individuals from the AusDiab study, obtained from applying the Australian
TLC Diabetes criteria for age range and HbA$_{1c}$ levels ($\geq 7.5\%$) to the subsample ($n=643$) of those classified in AusDiab as having diabetes. 156 AusDiab participants were identified for comparison with the Australian TLC Diabetes sample.”

4. The Discussion states on page 11 that this is the first study to examine an automated, interactive telephone intervention in the world. To be more specific, I think the statement that this is the first TLC intervention without any healthcare provider assistance/ follow-up might be more accurate (that is, differentiating from the Piette et al. trial, etc.).

An amendment has been made on page 12: “As far as we are aware, this is one of the first studies in the world to formally evaluate an automated telephone system for diabetes management that involves tailoring to individual needs”.

5. The conclusions compare a 0.8% reduction in A1c to the UKPDS study, but I was unsure if they also used geometric means? I was wondering about the comparability of the different mean measurements.

This is an important point raised by the reviewer. The UKPDS did not use geometric means, however the comparability is still appropriate, especially now that we have provided the arithmetic means to the results section (which also show an HbA$_{1c}$ reduction of 0.8%).

Yours Sincerely,

Dr. Emily Williams
School of Public Health and Preventive Medicine
Monash University, Alfred Hospital
89 Commercial Road, Melbourne, Australia 3004
Phone +61 3 9903 0255
emily.d.williams@monash.edu