Reviewer’s report

Title: Translational research: Are community-based child obesity treatment program scalable?

Version: 1 Date: 5 May 2015

Reviewer: João Valente-dos-Santos

Reviewer’s report:

1. GENERAL COMMENT

This paper has the potential to be of interest to readers. This is another output from the Go4Fun® program; details have been previously published on BMC Public Health (2014, 14:140 doi:10.1186/1471-2458-14-140). The present paper seeks to report the outcomes of the program on children’s weight and weight-related behaviours. This is an interesting question in an important population, thus the authors should be commended for their effort. However, I believe that this paper can be improved on some counts. I have mainly focused my feedback on Major Compulsory Revisions and occasionally in Minor Essential Revisions and Discretionary Revisions. That said these limitations should be easy for the authors to address.

2. MAJOR COMPULSORY REVISIONS

INTRODUCTION

2.1.

Translational research includes “several” areas of translation. One of the most interesting, from where I stand, concerns research aimed at enhancing the adoption of best practices in the community. Therefore, cost-effectiveness of prevention and treatment strategies are definitively an important part of translational science. However, the first paragraph of the introduction section could be rewritten to make it more focused in the topic under discussion. Be concise without indulging on general reasoning around the impact of Australian national policies.

2.2.

Second paragraph: External and internal validity are not independent types of validity. The second helps to ensure confidence in the casual linkage between the programs and the outputs. There is probably the need to highlight these questions taking into account that several constraints are faced in community-based obesity programs.

2.3.

Fourth paragraph: It would be informative if you stated your hypotheses.
METHODS

2.4.
In the methods, under the topic “outcome measures”, in addition to “… July 2009 and October 2012”, let the readers know if the design was controlled for seasonal variation. For example, sedentary behaviour may range from higher levels in the Fall to lower values in the Spring. If seasonal variation was not controlled, the limitations should be assumed in the manuscript which continues a very valuable contribution.

2.5.
How was stature, body mass and waist circumference assessed? I would be delighted to have some technical error of measurement on anthropometry. Meantime, the relevance of WtHtr should be part of the introduction, not in the methods section.

2.6.
The well-known variation in the tempo and timing of biological maturity for boys and girls of the same chronological age necessitates the use of an accurate measure of maturation in research involving children and adolescents.

The biological maturity of the child will confound the results and has a large independent impact on adiposity, fitness and cardio-metabolic risk. I would argue that it is essential to control for these influences. You can include a measure of biological maturation, parallel to the anthropometric description of the sample (e.g., Moore et al., 2014).


Maturity-related effects should also be briefly mentioned in the discussion.

If the authors cannot do this, then it needs to be acknowledge as a limitation. Maturity-related effects should also be briefly mentioned in the discussion.

2.7.
I recommend that the authors provide a little more detail regarding the process of survey procedures and multilevel modelling. Although I am quite familiar with some of these statistical techniques, and believe that the authors have used this technique appropriately, I think that those unfamiliar with these methods would struggle to understand or replicate the analytical process.

For example, when you account for the clustered design of the program in the multilevel models, they accommodate within-individual variance {#ij~N[0, var(#ij)]} for the ith assessment in the jth individual; the between-individuals intercept variance {μj~N[0, var(μ)]}; and, also the between-individuals slope variance {#ijxij~N[0, var(#ijxij)]}. So, in multilevel modelling each individual has their own trajectory, with intercepts and slope coefficients varying between individuals.
Therefore, we controlled for independent inter-group effects while simultaneously controlling for the effects of age, sex, area level of socio-economic disadvantage, area level of remoteness, and the number of sessions attended, within each individual. If this was the way, or not, please clearly state that.

2.8.
Multilevel model technique allows the number of observations and temporal spacing between measurements to vary among subjects, thus using all available data. It is assumed that the probability of data being missing is independent of any of the random variables in the model. As long as a full information estimation procedure is used, such as maximum likelihood for normal data, the actual missing mechanism can be ignored. Was maximum likelihood estimation used for the models?

For discrete response multilevel models, maximum likelihood estimation is computationally intensive, and therefore quasi-likelihood methods were implemented in several statistical packages. If so please clearly state that.

2.9.
How did you checked for heteroscedasticity? Heteroscedasticity can result in biased parameter estimates and standard error estimates that are either too large or too small. Consequently, heteroscedasticity can increase Type I-error (be overly optimistic about parameters being statistically different from zero) or increase Type II-error (be overly pessimistic about parameters being statistically different from zero). One approach to detecting violations of homoscedasticity is to examine plots of residuals versus predicted values for evidence that residuals increase as a function of the predicted value. However, heteroscedasticity can also be a result of a violation of the linearity or normality assumptions, and it may also be remedied by addressing those violations.

RESULTS

2.10
The authors make a point, in Figure 1, that pre and post BMI distribution of all children attending #1 session shows a slight left shift in the distribution curves following participation in the program. The authors should make a clearer case as to whether or not the left shift is or not statistical significant.

2.11.
Table 1, 2 and 3: Reporting only the P-values for an analysis is not adequate for readers to fully understand the results. Magnitude-based inference is far superior to the traditional approach of null-hypothesis significance testing. This approach allows for conclusions about sample-based effects (e.g., Batterham & Hopkins, 2006; Hopkins et al., 2009).

I am fully aware of the criticism of Welsh and Knight (2015) to the use of magnitude-based Inferences, but I am also aware of the mistakes of that statistical review. However, if the authors do not believe that magnitude-based inference can be specifically valuable in follow-up studies as it is a less conservative method, there are other alternatives. Instead of magnitude-based inference you can chose to report (partial) eta-squared statistics as a measure of effect-size, along with the traditional statistical significance and P-values. By doing so, you will provide a measure of effect-magnitude, but still within the traditional statistical framework. Moreover, you can do the square root of eta-squared and interpret it as a correlation. Something like:

The effect size correlations (ES-r) were estimated using the square root of the ratio of the F-value squared and the difference between the F-value squared and degrees of freedom (Rosnow and Rosenthal, 1996). Coefficients were interpreted as follows: trivial (r < 0.1), small (0.1 < r < 0.3) moderate (0.3 < r < 0.5), large (0.5 < r < 0.7), very large (0.7 < r < 0.9), nearly perfect (r < 0.9) and perfect (r = 1) (Hopkins et al., 2009).

DISCUSSION

2.12.

The discussion section, in general, is good and provides straight forward and transparent account of the findings of this study. A considerable effort to link the results of the currently study in relation to existing literature was performed. My recommendation would be that the authors spend a little more time discussing the potential practical implications of this research.

3. MINOR ESSENTIAL REVISIONS

INTRODUCTION

3.1.

NSW (i.e., New South Wales) - Not previously defined.

3.2.

At the end of the second paragraph “in the real world” – correct to “in the real world.”
METHODS

3.3.
“Outcome measures” subheading is repeated twice under the methods section.

RESULTS

3.4.
I would present the median number of participants per program, who completed the program, etc. and not the average.

3.5.
Table 1 is easy to follow but you do not explain what the “*” stands for.

DISCUSSION

3.6.
Fourth paragraph: I am not a native English speaker but I found this sentence somewhat confusing: “This may reflect the current status of the range of different nutrition labels suggest that policy makers need to ensure nutrition labels are understandable by community members”.

Conclusions

3.7.
In this section the authors should address more possible take home message/s and avoid replicating statements already stated in other parts of the paper.

REFERENCE LIST

3.8.
Check the guidelines for the reference style. There are several inconsistencies.

4. DISCRETIONARY REVISIONS

4.1.
Keywords: for indexing purposes, repeating a word used in the title (i.e., translational research) does not seams a robust approach.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.
Declaration of competing interests:

I declare that I have no competing interests