PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Bicycling to school improves the cardiometabolic risk factor profile: a randomised controlled trial
AUTHORS	Østergaard, Lars ; Børrestad, Line; Tarp, Jakob; Andersen, Lars

VERSION 1 - REVIEW

REVIEWER	Elling Bere Professor University of Agder Norway I don't have any competing interests
REVIEW RETURNED	05-Jun-2012

GENERAL COMMENTS	This is an innovative study and a well written paper on an important topic. I have only a few comments: - Page 5, line 10. Why was only 10 of the 36 schools visited. Please elaborate. - Page 5, line 14. Please describe how the assessment of cycling the last three months was assesses. Also, please state what months of the year these three months were. - Page 5, line 49. Unclear what you mean by "matching test date". Please elaborate. - Page 6, line 41. Please write out RER and HR first time. - Page 8, line 48. VO2peak is included in the score and also analysed as a separate dependent variable. Please describe why. - Page 10. For VO2peak you see an association between km cycled, but no intervention effect, and for the score you see the opposite (an intervention effect, but no association with km cycled). Please discuss.

REVIEWER	Dr Charlie Foster BHF HPRG Department of Public Health University of Oxford UK
REVIEW RETURNED	11-Jun-2012

GENERAL COMMENTS	This is a well conducted, written and elegant study of a RCT to promote cycle use in young adolescents. The paper is clear and the scientific methods are appropriate. This is a challenging study performed with a challenging group. It should act as a pilot for a larger study to counter the statistical limitations outlines by the authors. Further suggestions, as possible suggested changes, are offered for each section.

	Introduction The rationale for the paper might be further supported by the addition of relevant conclusions of a recent systematic review. The review reported weak evidence from observational studies that fitness may be higher in children who undertake active travel to schools. It also called for the need for experimental studies with objective measures to tease out the real health effects of active travel confounded within cross sectional studies. Lubans D, Boreham C, Kelly P, Foster C. The relationship between active travel to school and health-related fitness in children and adolescents: a systematic review, International Journal of Behavioral Nutrition and Physical Activity, 2011, 8:5. It might be helpful for international readers to know the age range of children in an elementary school. Methods The study team should be commended for their excellent effort to maintain study quality, i.e. the blinding to allocation status of participants to the measurement team at baseline and follow up. Did any of the children who withdrew have friends within the study elsewhere? This may explain withdrawals (approximately 20% of eligible participants). The research team made considerable efforts to recruit participants and it would be useful to know how long the recruitment period lasted, and the most effective approach for getting eligible subjects as study numbers are a limit to the overall power of their results. Perhaps the authors could comments on this in their discussion section? Non-broadcast media methods might be helpful here. Have the authors considered any possible seasonal effect of winter/spring on cycling rates or indeed willingness to re-start cycling in relation to recruitment. Please could the authors comment or respond about their reasons for using the ITOF BMI ranges? Results It would be useful for the authors to comment upon the possible measurement error in the VO2 assessment? How might this relate to the amount of change achieved by the intervention? Could the authors discuss the relationship between the assessed intensity of the cycle journeys in relation to estimates of the VO2 capacities of children, at an individual level? The author are in the position to cross check the intensities of children's journeys with their performance in the VO2 test to see if the cycling was sufficiently hard enough to change fitness, rather than a mean of the groups. Discussion

	Please could the authors expand on their rationale for the definition of compliance by % of journeys? This is a key factor that drives their analysis. Please could the authors describe their views on possible lag between intervention effects and changes in VO2 max over an 8 week intervention period? The authors might consider other suggestions for objective outcome measures of the behaviour could include GPS and digital visual systems, both used with this age group. These also corroborate journey mode, duration and possible road conditions. Please could authors consider the possible impact of different trip profiles to and from school? A number of UK studies have suggested that these are very different in children – see Project PEACH papers

REVIEWER	Siew-Pang Chan Senior Lecturer La Trobe University Australia
REVIEW RETURNED	27-Jul-2012

THE STUDY	1) There should be a table showing the ANCOVA results based on regression analyses. 2) ANCOVA is not necessarily univariate. With the help of multiple linear regression more than one covariates could be analyzed. 3) There should a detailed discussion about the goodness of fit of the ANCOVA/regression models (e.g., R squares), and whether the crucial assumptions are violated (presence of outliers and influential observations). 4) It is unclear how QQ plots and the Shapiro-Wilks tests were applied. Were they applied to ascertain the residulas of the models or the outcomes? Linear regression estimated with ordinary least squares (OLS) does not require the outcomes be normally distributed. 5) It is unclear which regression model was applied. OLS has a zero breakdown point when outliers are present.
RESULTS & CONCLUSIONS	1) A table showing the regression analyses should be included.

VERSION 1 – AUTHOR RESPONSE

Elling Bere

Point #2:
Page 5, line 10. Why was only 10 of the 36 schools visited. Please elaborate.
Response: We initially sent invitation letters to approximately 3000 school children. Due to low expression of interest we personally visited as many schools as possible for a period of 2 weeks. A prolonged recruitment period would have implied that the intervention period would have been shortened since it had to end before the summer vacation at the elementary schools started. However, since the included children were randomized afterwards, a selection of schools should not affect intervention results.

Point #3:
Page 5, line 14. Please describe how the assessment of cycling the last three months was assesses. Also, please state what months of the year these three months were.
Response: Mode of transport was self-reported at the time of registration for the study.
More detailed information has been added to the manuscript. The paragraph is now:
“Participants were included if they at the time of registration stated that they had not bicycled regularly to and from school for at least 3 months (i.e. at least from January onwards) prior to the intervention, and if willing to be randomised to one of the two study groups (i.e. control group or bicycling group)”

Point #4:
Page 5, line 49. Unclear what you mean by "matching test date". Please elaborate.
Response:
The paragraph is now:
“Matching baseline test date with follow-up test (in order to achieve similar intervention duration for all subjects) was strived for, but logistically not possible for all participants.”

Point #5:
Page 6, line 41. Please write out RER and HR first time.
Response:
The requested has been added to the manuscript.
Point #6:
Page 8, line 48. VO2peak is included in the score and also analysed as a separate dependent variable. Please describe why.
Response: We aimed at, in a controlled design, to document that regular cycling to school cause higher cardiorespiratory fitness and thus analysed this variable separately. We chose to integrate VO2peak in the clustered score because it has been shown to be is associated with a clustered risk score in children (Hong et al.,Pediatr Exerc Sci. 2011 May;23(2):270-80) and adults (Hassinen et al., DIABETES CARE, VOLUME 31, NUMBER 6, JUNE 2008) cardiorespiratory is likely to be part of the complex causing the metabolic syndrome. Furthermore VO2peak has been included as a part of the risk score in previous investigations of clustering of risk factors among children and adolescents (Andersen et al. Lancet 2006; 368: 299–304).

Point #7:
Page 10. For VO2peak you see an association between km cycled, but no intervention effect, and for the score you see the opposite (an intervention effect, but no association with km cycled). Please discuss.

Response:
Bicycling to school did not have the expected effect on cardiorespiratory fitness. This might be due to a relatively short intervention period in the present study and perhaps also because of too short distance to school. The observed intervention effect on clustering of risk factors could be due to inclusion of additional variables which adapt at a lower degree of exposure. It is noteworthy that the two dose-response analyses – in contrast to the main analyses (intervention effect) only comprise subjects from the intervention group (i.e. statistical significance in main analyses also dependent on change in the control group). Finally a statistically significant dose-response association between km cycled and delta VO2peak but not delta mean of z-scores could be due to the fact that the latter include various variables with different dose-response associations whereas the relationship between km and delta VO2 peak is more straight forward.

Charlie Foster
Point #8:
The rationale for the paper might be further supported by the addition of relevant conclusions of a recent systematic review. The review reported weak evidence from observational studies that fitness may be higher in children who undertake active travel to schools. It also called for the need for experimental studies with objective measures to tease out the real health effects of active travel confounded within cross sectional studies.
Lubans D, Boreham C, Kelly P, Foster C. The relationship between active travel to school and health-related fitness in children and adolescents: a systematic review, International Journal of Behavioral Nutrition and Physical Activity, 2011, 8:5.
Response: This relevant reference has been added (page 4, bottom).

Point #9:
It might be helpful for international readers to know the age range of children in an elementary school.
Response: The recruitment was directed at pupils in the 4th and 5th grade aged 11-12 yrs. We believe that addition of the age span for the entire elementary school (5-17 years) would confuse the reader. Mean age and sd for the participants is stated in table 1.

Point #10:
The study team should be commended for their excellent effort to maintain study quality, i.e. the blinding to allocation status of participants to the measurement team at baseline and follow up.
Did any of the children who withdrew have friends within the study elsewhere? This may explain withdrawals (approximately 20% of eligible participants). The research team made considerable efforts to recruit participants and it would be useful to know how long the recruitment period lasted, and the most effective approach for getting eligible subjects as study numbers are a limit to the overall power of their results. Perhaps the authors could comments on this in their discussion section? Non-broadcast media methods might be helpful here. Have the authors considered any possible seasonal effect of winter/spring on cycling rates or indeed willingness to re-start cycling in relation to recruitment.
Response: We do not have registrations on “friend-associations”. We do, however, have information on the various reasons of withdrawal and this seem not to be caused by a Lemming effect but rather family factors (e.g. unwillingness, relocation, divorce, expulsion from school).
The first recruitment letter was sent through the central school administration to all public elementary school in the city of Odense in january 2011 (approximately 3 months before baseline measurements). Potential study participants was enticed by money, bicycle gear and competition for a bicycle. This approach gave us only few elegible subjects. Our perception was that the study material had reached far from all intended pupils and with great variances in teacher enthusiasm. We therefore produced more appealing recruitment material which we personally distributed in the school classes. Hundreds of children then contacted us in order to participate. Unfortunately many of these did not meet the inclusion criteria – mainly because they were already bicycling to school.

On page 13 we have added: “We experienced that direct personal contact to school pupils was the most efficient way to recruit participants.”

Also see the original comments on page 12 regarding difficulties in relation to participant acceptance of randomisation.

Point #11:
Please could the authors comment or respond about their reasons for using the ITOF BMI ranges?
Response: We used IOTF / Cole´s cut-off values solely in order to describe the study sample. We chose this approach since it is a simple and well known way to take into account gender and age based on international reference population.

Point #12:
It would be useful for the authors to comment upon the possible measurement error in the VO2 assessment? How might this relate to the amount of change achieved by the intervention?
Response: The metabolic cart was prior to both baseline and endline measurements found to produce precise values as compared to steady state VO2 measured with the douglas bag method. We are aware that a large variation would diminish the possibility of detecting an effect on VO2peak.

Point #13:
Could the authors discuss the relationship between the assessed intensity of the cycle journeys in relation to estimates of the VO2 capacities of children, at an individual level? The author are in the position to cross check the intensities of children's journeys with their performance in the VO2 test to see if the cycling was sufficiently hard enough to change fitness, rather than a mean of the groups.
Response: Good point.
1) We first cross checked the relative average commuter intensity (self-chosen average commuter intensity vs. mean of laboratory assessed maximal heart rate at baseline and endline) with VO2peak changes. 2) Then we cross checked the relative maximal commuter intensity (maximal HR during commuter intensity vs. mean of laboratory assessed maximal heart rate at baseline and endline) with VO2peak changes. If possible we would like to present this interesting relationship as data/online supplement. Graphs are added to the re-submission as TIFF files.

The following has been added to page 12.
"Interestingly post-hoc linear regression showed that both the relative average and the relative maximal intensity during commuter bicycling was positively associated (see supplement 1).with cardiorespiratory fitness improvements (p=0.005 and p=0.002 respectively)."

Point #14:
Please could the authors expand on their rationale for the definition of compliance by % of journeys? This is a key factor that drives their analysis.
Response: This factor drives only the supplementary efficacy analyses, whereas in the intention to treat analyses compliance is neglected. We acknowledge that the efficacy analyses are based on an arbitrary compliance cutoff level. We found a cut-off allowing bicyclists in the intervention group to miss one school day, walk or passively commute one day/week to be appropriate. Essentially this matter is a trade-off between a correct classification of commuter mode versus loss of statistical power and potential dilution of intervention effect.

Point #15:
Please could the authors describe their views on possible lag between intervention effects and changes in VO2 max over an 8 week intervention period?
Response: There is evidence that 8 weeks of commuter bicycling should be sufficient time to improve maximal cardiorespiratory fitness in adults (Møller et. al. 2011, The effect on cardiorespiratory fitness after an 8-week period of commuter cycling — A randomized controlled study in adults, preventive medicine). Also 12 weeks have been shown to be sufficient in children (Børrestad et al., 2012, Experiences from a randomised, controlled trial on cycling to school: Does cycling increase cardiorespiratory fitness?, Scandinavian Journal of Public Health, 2012; 40: 245–252). The explanation for a possible lag could be a combination of a short intervention period and too low intensity in some subjects.

Point #16:
The authors might consider other suggestions for objective outcome measures of the behaviour could include GPS and digital visual systems, both used with this age group. These also corroborate journey mode, duration and possible road conditions.
Response: This is taken ad notam.

Point #17:
Please could authors consider the possible impact of different trip profiles to and from school? A number of UK studies have suggested that these are very different in children – see Project PEACH papers
Response:
Though we have no data to document the degree of trip profile variation most children and parents gave us the impression that the participants (unless when visiting a friend on the way home or attending leisure time activities after school) took the same route to and from school. This matter is highly affected by the age of the children with a tendency that the youngest commute the same route because it is perceived to be safe.

Siew-Pang Chan

Point #18:
There should be a table showing the ANCOVA results based on regression analyses.
Response: If possible we would like to present these results as data/online supplement. Tables are added to the re-submission as docx files.

Point #19:
ANCOVA is not necessarily univariate. With the help of multiple linear regression more than one covariates could be analyzed.
Response: univariate (page 8, mid) deleted.

Point #20:
There should a detailed discussion about the goodness of fit of the ANCOVA/regression models (e.g., R squares), and whether the crucial assumptions are violated (presence of outliers and influential observations).
It is unclear how QQ plots and the Shapiro-Wilks tests were applied. Were they applied to ascertain the residulas of the models or the outcomes? Linear regression estimated with ordinary least squares (OLS) does not require the outcomes be normally distributed.

Response:
The following has been added (page 11):
“Standardized residuals were plotted against the predicted values and no systematic patterns were observed which confirmed variance-homogeneity. QQ plots and Shapiro Wilks tests of the standardized residuals of the model expressed normality. The goodness of fit as indicated by r-squared values in the regression modeling of change in the standardised composite Z score were 0.16 and 0.19 for ITT and efficacy analyses respectively. Goodness of fit for the modeling of change in $\dot{V}O_2peak$ was 0.17 and 0.18 for ITT and efficacy analyses respectively (for additional results from the regression analyses see supplement 2).”

Point #21:
A table showing the regression analyses should be included.
Response: see point #18

Formalities
- Previous ref #9: Østergaard et al. has now been published and the reference has been changed accordingly.