©Journal of Sports Science and Medicine (2019) **18,** 830-833
http://www.jssm.org

**Letter to editor**

# Caution with Conclusions Required: A Response to the Paper "Objectively Measured Aerobic Fitness is not related to Vascular Health Outcomes and Cardiovascular Disease Risk in 9-10 Year Old Children"

**Dear Editor-in-chief**

In this issue Farr et al. (2019) examined whether aerobic fitness in 9-10 year old children was related to macro and microvascular health and cardiovascular disease (CVD) risk. The authors should be congratulated for publishing null findings, an under-appreciated and often unfairly difficult endeavor. The study had many important strengths including direct determination of aerobic fitness, normalization for body size and body composition, and multiple measures of macro and microvascular function. However, the study as designed and presented also has several clear limitations, some of which are discussed below.

First, it is unfortunate that the authors did not report all 'statistically non-significant' test statistics, especially for analyses that were central to the aim of the paper (e.g., correlations between aerobic fitness and vascular outcomes). This rather common practice within our field diminishes the transparency of the research and can serve as an avenue for biases to present themselves. Notably, ANOVAs and ANCOVAs were used to examine mean differences between low and high aerobic fitness groups on vascular health outcomes and markers of CVD risk, yet none of these model statistics were reported either in-text or as supplementary material. This is especially problematic since the authors used the results of these tests to draw their primary conclusion, which happens to serve as the title of their manuscript – "objectively measured aerobic fitness is not related to vascular health outcomes and cardiovascular disease risk in 9-10 year old children."

Second, information pertaining to the log-transformed variables is also absent from the article. Given that many of the variables (e.g., ACh AUC, SNP AUC, Specific insulin) have a variance greatly exceeding the mean, particularly the variables with proportion values (PAI, β- cell function, HOMA-IS), some descriptives on the log-transformed variables would help reassure the reader that the analyses performed were appropriate. It is highly encouraging that the authors properly screened the distribution of their variables – a practice too often ignored in quantitative science; however, in our experience with skewed data, conducting log-transformations rarely brings the data into conformity with a normal distribution. In the event that the log-transformations did not significantly improve the variables' distribution, using ANOVAs and ANCOVAs would be questionable. In this scenario, using a model that can handle skewed variables (e.g., negative binomial, Poisson) would be more appropriate. As others have recommended (Feng et al., 2014), we believe that using models that can handle skewed data is preferable to data transformation.

Third, an obvious limitation of this study was the small sample size of 96 children, and even smaller in some analyses ($n$ = 69) due to loss of data. We suspected that this study was underpowered and therefore performed ANOVA (fixed effects, omnibus, one-way) post-hoc power analyses, using an effect size of 0.25 and alpha level of 0.05. Our analysis showed that a sample size of 70 generated a 46% probability of Type II error (power = .54), whereas a sample size of 96 generated a 32% probability of Type II error (power = .68). This suggests the study was underpowered, increasing the likelihood that a researcher will fail to reject the null hypothesis when it should, in fact, be rejected. Indeed, it might be argued – somewhat crudely – that when the power is 54%, committing a Type II error boils down to a coin toss.

Finally, the authors – like many other scientists (including ourselves at times!) – seem to have misinterpreted (and misused) the $p$-value (Greenland et al., 2016). The authors conclude that aerobic fitness was not related to macro and microvascular health outcomes and CVD risk factors, and base these conclusions on $p$-values >0.05. However, a $p$-value >0.05 *does not* mean that an absence of an association was demonstrated, but rather that the null hypothesis is among *many* other hypotheses that are compatible with the data (Greenland et al., 2016). Therefore, the conclusion that aerobic fitness was not related to macro and microvascular health outcomes is likely unfounded. Moreover, the authors appear to make scientific inferences based solely on $p$-values >0.05, without any apparent consideration given to contextual factors such as study design, sample size, effect size and intra-subject measurement reliability (Wasserstein & Lazar, 2016). For example, on page 517, they report "[w]hen assessed as a whole group and in both boys and girls separately there was a negative correlation between BMI and carotid to ankle PWV in the whole group ($r$ = -0.24; p = 0.020) and boys ($r$ = -0.30; $p$ = 0.038) but not girls ($r$ = -0.24; $p$ = 0.13)." Here the authors dismiss the negative coefficient in the subsample of girls ($r$ = -0.24) simply because the $p$-value exceeded the threshold of 0.05. It is important to remember that a label of 'statistical non-significance' does not imply that "an association or effect is improbable, absent, false, or unimportant" (Wasserstein, Schirm, & Lazar, 2019, p. 2). Put another way, the absence of evidence is not evidence of absence (Altman & Bland, 1995).

The limitations must be considered with the study's strengths. As scientists it is our responsibility to accurately interpret and communicate our findings, practice transparency, critically evaluate our study's shortcomings and meaningfulness, and situate our findings within a broader context by considering previous research and contextual factors. We must all be vigilant and critical of our own

Received: 09 September 2019 / Accepted: 11 October 2019 / Published (online): 19 November 2019

work to collectively improve the scientific rigor of our field.

**Michelle D. Guerrero ✉, Joel D. Barnes and Mark S. Tremblay**
Healthy Active Living and Obesity Research Group, Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Road, Ottawa, Ontario, Canada, K1H 8L1

✉ **Michelle D. Guerrero, PhD**
**E-mail:** mguerrero@cheo.on.ca

## References

Altman, D. G. and Bland, J. M. (1995) Statistics notes: Absence of evidence is not evidence of absence. *BMJ* **311,** 485.

Farr, C., Middlebrooke, A. R., Armstrong, N., Barker, A. R., Fulford, J., Mawson, M. and Mcmanus, A. M. (2019) Objectively Measured Aerobic Fitness is Not Related to Vascular Health Outcomes and Cardiovascular Disease Risk In 9-10 Year Old Children. *Journal of Sports Science & Medicine* **18,** 513–522.

Feng, C., Wang, H., Lu, N., Chen, T., He, H., Lu, Y. and Tu, X. M. (2014) Log-transformation and its implications for data analysis. *Shanghai Archives of Psychiatry* **26,** 105–109.

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N. and Altman, D. G. (2016) Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology* **31,** 337–350.

Wasserstein, R. L., and Lazar, N. A. (2016) The ASA Statement on p-Values: Context, Process, and Purpose. *The American Statistician* **70,** 129–133.

Wasserstein, R. L., Schirm, A. L. and Lazar, N. A. (2019) Moving to a World Beyond "p < 0.05." *The American Statistician* **73,** 1–19.

---

# Authors' response

**Dear Editor-in-chief**

We would like to thank Dr. Guerrero and colleagues for their interest in our recently published manuscript entitled "Objectively Measured Aerobic Fitness is Not Related to Vascular Health Outcomes and Cardiovascular Disease Risk In 9-10 Year Old Children". We appreciate the acknowledgement by the authors of the difficulties that can be encountered in publishing null findings. Indeed, we have an ethical duty in science to disseminate our findings, irrespective of whether these are null or support our hypotheses, and as suspected by Guerrero and colleagues it did require more than a little patience before publication was finally achieved!

The main strengths of our paper, as acknowledged by Guerrero and colleagues, was the direct determination of aerobic fitness (peak V̇O2) using a cycle based protocol, and the use of valid scaling techniques to control for body size and composition when expressing peak V̇O2. Our study not only reported commonly used markers of cardiovascular disease (CVD) risk, but also determined a number of important vascular outcomes (at the micro and macrovascular levels), which are poorly explored in the paediatric literature, particularly in younger children.

As with any study, there were a number of limitations in our paper which we highlighted within the discussion of our study results. In addition, we thank the editor for the opportunity to address each of the concerns of Guerrero and colleagues, as highlighted below:

1) Reporting of 'statistically non-significant' test statistics: We agree with Guerrero and colleagues that the reporting of all results, irrespective of their significance, is important in scientific research to ensure transparency. Through the revision process, all data and test results were submitted to the journal, but two tables containing the correlations between aerobic fitness and both vascular outcomes and CVD risk were removed through the editorial process, presumably due to space limitations and readability of the article. In hindsight, we should have questioned this decision before publication. However, we have requested that these tables be attached as an online supplementary file for further interest (attached as Tables 1 and 2 below).

2) Use of log-transformed variables: Despite the non-normal distribution of many of our results we felt a parametric approach would reduce the likelihood of a type II error, (especially given our reduced final sample size). Indeed, given the problems associated with over transforming data (using reciprocal, reciprocal root etc.) we thus incorporated a commonly published method of data transformation (i.e. log transformation), before affirming normality through further analyses.

3) Sample size: The authors raise an important point regarding the sample size of our data. The novelty and technical nature of the methods employed in the study, including numerous vascular outcomes and the use of magnetic resonance imaging to assess adiposity, meant data collection was very time consuming and labour intensive. Originally, the intention was to recruit 128 children for the study but due to a delay in the commencement of data collection, the time window within which to collect data was impacted, and subsequently the final number of children recruited was reduced to 100. A further four participants were lost due to withdrawal or the use of vasoactive medication leaving a total sample of 96 children. Despite our best efforts, the number of children with a complete dataset was unfortunately further reduced to 69 due to issues surrounding iontophoresis analysis (i.e. technical error, poor laser signal, or a child requesting we stop due to discomfort). Studies of young participants are difficult and data loss is unfortunately often a consequence. As such this limitation was reported within our discussion.

4) P-value: We would like to thank Guerrero and colleagues for highlighting that interpretation of the p-value alone is fraught with complications, and it is thus important for it to be considered alongside other statistical factors such as the size of the observed effect. We would contend in the case of the correlational analyses such as those we present in the paper and in Tables 1 and 2 below, that the magnitude of the relationships between aerobic fitness and the CVD outcomes were mostly trivial or small in

magnitude, according to the thresholds proposed by Cohen (1988).

We would to thank Dr Guerrero and colleagues for providing some important reflections on our study and its findings. We trust our responses take into account the strengths of the study, as well as further explaining the limitations, restrictions and challenges we encountered while researching a fascinating group of young participants. The identification of future CVD risk in childhood using clinical techniques is a fascinating and complex topic, and one which clearly warrants further consideration. Early identification of individuals 'at-risk' of future CVD is paramount so that suitable primary prevention strategies can be implemented, and it is only through rigorous research in children and young adults that such a vision can be fully realised.

**Colin Farr [1]✉, Andrew R. Middlebrooke [1], Neil Armstrong [1], Alan R. Barker [1], Jon Fulford [3], David M. Mawson [2] and Ali M. McManus [4]**

[1] Children's Health and Exercise Research Centre, Sport and Health Sciences, College of Life and Environmental Sciences, University of Exeter, Exeter, UK; [2] Diabetes and Vascular Medicine, NIHR Exeter Clinical Research Facility, University of Exeter Medical School, RD&E (Wonford), Barrack Road, Exeter, UK; [3] NIHR Exeter Clinical Research Facility, University of Exeter Medical School, University of Exeter, Exeter, UK; [4] Centre for Heart, Lung and Vascular Health, School of Health and Exercise Sciences, University of British Columbia, Kelowna, Canada

## Reference

Cohen, J. (1988). Statistical power analysis for the behavioural sciences (2nd ed.). New Jersey: Lawrence Erlbaum.

✉ **Colin Farr**
Research Assistant, MRC Epidemiology Unit, Princess of Wales Hospital, Lynn Road, Ely, Cambridgeshire, CB6 1DN, UK
**E-mail:** colin.farr@mrc-epid.cam.ac.uk

**Table 1. Correlations between aerobic fitness and vascular outcomes. Data presented are the Pearson's correlation coefficient (r), probability (p) and sample size (n).**

| | Peak oxygen uptake ($mL \cdot kg^{-0.61} \cdot min^{-1}$) | | | Peak oxygen uptake ($mL \cdot FFM^{-1} \cdot min^{-1}$) | | |
|---|---|---|---|---|---|---|
| | **Whole group** | **Boys** | **Girls** | **Whole group** | **Boys** | **Girls** |
| | | | **Macrovascular** | | | |
| **PAI (%)** | r=-0.20; p=0.06 n=93 | r=-0.09; p=0.53 n=50 | r=-0.25; p=0.11 n=43 | r=0.04; p=0.73 n=83 | r=0.01; p=0.96 n=47 | r=0.11; p=0.52 n=36 |
| **Carotid to ankle PWV ($m.s^{-1}$)** | r=0.11; p=0.30 n=90 | r=0.21; p=0.15 n=50 | r=0.11; p=0.49 n=40 | r=0.15; p=0.17 n=81 | r=0.20; p=0.18 n=47 | r=0.13; p=0.47 n=34 |
| **Carotid to radial PWV ($m.s^{-1}$)** | r=-0.01; p=0.90 n=94 | r=-0.10; p=0.49 n=52 | r=0.16; p=0.32 n=42 | r=0.05; p=0.68 n=84 | r=<0.01; p=0.99 n=49 | r=0.14; p=0.43 n=35 |
| | | | **Microvascular** | | | |
| **FMR (V)** | r=-0.03; p=0.76 n=96 | r=0.09; p=0.52 n=53 | r=-0.14; p=0.37 n=43 | r=-0.03; p=0.81 n=86 | r=0.08; p=0.57 n=50 | r=-0.14; p=0.43 n=36 |
| **Peak ACh response (V)** | r=-0.05; p=0.61 n=93 | r=0.11; p=0.44 n=51 | r=-0.05; p=0.76 n=42 | r=0.15; p=0.16 n=83 | r=0.19; p=0.19 n=48 | r=0.08; p=0.66 n=35 |
| **ACh AUC (V x time)** | r=0.06; p=0.60 n=92 | r=0.10; p=0.51 n=50 | r=-0.08; p=0.62 n=42 | r=0.14; p=0.22 n=82 | r=0.19; p=0.20 n=47 | r=0.01; p=0.98 n=35 |
| **Peak SNP response (V)** | r=-0.01; p=0.92 n=69 | r=0.03; p=0.86 n=38 | r=-0.11; p=0.56 n=31 | r=0.09; p=0.50 n=63 | r=0.09; p=0.59 n=37 | r=0.06; p=0.77 n=26 |
| **SNP AUC (V x time)** | r=-0.02; p=0.86 n=69 | r=0.01; p=0.93 n=38 | r=-0.12; p=0.52 n=31 | r=0.10; p=0.43 n=63 | r=0.09; p=0.59 n=37 | r=0.08; p=0.70 n=26 |

FFM – fat free mass; PAI - peripheral non-transformed augmentation index; PWV - pulse wave velocity; FMR - functional microvascular reserve; ACh - acetylcholine; AUC - area under the curve; SNP - sodium nitroprusside.

**Table 2. Correlations between aerobic fitness and markers of CVD risk. Data presented are the Pearson's correlation coefficient (r), probability (p) and sample size (n).**

| | Peak oxygen uptake ($mL \cdot kg^{-0.61} \cdot min^{-1}$) | | | Peak oxygen uptake ($mL \cdot FFM^{-1} \cdot min^{-1}$) | | |
|---|---|---|---|---|---|---|
| | **Whole group** | **Boys** | **Girls** | **Whole group** | **Boys** | **Girls** |
| **BMI ($kg \cdot m^{-2}$)** | r=-0.12; p=0.25 n=96 | r=-0.16; p=0.25 n=53 | r=-0.08; p=0.62 n=43 | r=-0.11; p=0.34 n=86 | r=-0.03; p=0.82 n=50 | r=-0.17; p=0.33 n=36 |
| **TBF%** | **r=-0.34; p=0.001** n=86 | r=-0.19; p=0.19 n=50 | **r=-0.34; p=0.044** n=36 | r=0.04; p=0.75 n=86 | r=0.23; p=0.17 n=50 | r=-0.04; p=0.84 n=36 |
| **VAT ($cm^3$)** | r=-0.07; p=0.56 n=86 | r=-0.25; p=0.09 n=49 | r=0.06; p=0.72 n=37 | r=-0.04; p=0.73 n=78 | r=-0.06; p=0.69 n=46 | r=0.03; p=0.87 n=32 |
| **Total cholesterol ($mmol.L^{-1}$)** | r=-0.11; p=0.30 n=79 | r=-0.09; p=0.54 n=49 | r=-0.10; p=0.60 n=30 | r=0.06; p=0.61 n=71 | r=0.05; p=0.75 n=46 | r=0.02; p=0.93 n=25 |
| **HDL-C ($mmol.L^{-1}$)** | r=-0.05; p=0.67 n=78 | r=-0.11; p=0.44 n=49 | r=-0.07; p=0.74 n=29 | r=0.10; p=0.43 n=70 | r=0.05; p=0.74 n=46 | r=0.15; p=0.48 n=24 |
| **LDL-C ($mmol.L^{-1}$)** | r=-0.06; p=0.63 n=78 | r=-0.05; p=0.75 n=49 | r=-0.08; p=0.67 n=29 | r<0.01; p=0.99 n=70 | r=-0.01; p=0.95 n=46 | r=-0.03; p=0.91 n=24 |
| **TG ($mmol.L^{-1}$)** | r=0.02; p=0.88 n=79 | r=0.15; p=0.31 n=49 | r=-0.07; p=0.73 n=30 | r=0.04; p=0.75 n=71 | r=0.10; p=0.50 n=46 | r=-0.04; p=0.86 n=25 |
| **Glucose ($mmol.L^{-1}$)** | r=-0.10; p=0.39 n=80 | r=-0.09; p=0.55 n=49 | r=-0.10; p=0.58 n=31 | r=-0.14; p=0.25 n=72 | r=-0.12; p=0.42 n=46 | r=-0.14; p=0.49 n=26 |
| **Specific insulin ($pmol.L^{-1}$)** | r=0.07; p=0.55 n=77 | r=0.12; p=0.44 n=46 | r=-0.17; p=0.36 n=31 | r=0.10; p=0.44 n=69 | **r=0.34; p=0.028** n=43 | r=-0.21; p=0.30 n=26 |
| **β-cell function (%)** | r=-0.05; p=0.69 n=77 | r=0.13; p=0.39 n=46 | r=-0.12; p=0.52 n=31 | r=0.16; p=0.20 n=69 | **r=0.38; p=0.011** n=43 | r=-0.12; p=0.55 n=26 |
| **HOMA-IS (%)** | r=0.07; p=0.52 n=77 | r=-0.13; p=0.40 n=46 | r=0.19; p=0.30 n=31 | r=0.09; p=0.47 n=69 | **r=-0.32; p=0.036** n=43 | r=0.21; p=0.31 n=26 |
| **HOMA-IR** | r=-0.08; p=0.50 n=77 | r=0.10; p=0.50 n=46 | r=-0.17; p=0.35 n=31 | r=0.08; p=0.52 n=69 | **r=0.31; p=0.042** n=43 | r=-0.21; p=0.31 n=26 |

BMI - body mass index; TBF% - total body fat percentage; VAT - visceral adipose tissue; HDL-C – high density lipoprotein cholesterol; LDL-C – low density lipoprotein cholesterol; TG -Triglyceride; HOMA – Homeostatic model assessment; IS – Insulin sensitivity; IR – Insulin resistance.