Published in final edited form as:

Int J Obes (Lond). 2018 December ; 42(12): 2064–2065. doi:10.1038/s41366-018-0251-6.

In response to: ‘Information bias in measures of self-reported physical activity’

Yann C. Klimentidis^{1,*}, David A. Raichlen², Jennifer Bea^{3,4}, David O. Garcia⁵, Nathan E. Wineinger⁶, Lawrence J. Mandarin⁷, Gene E. Alexander⁸, Zhao Chen¹, and Scott B. Going⁴

¹Department of Epidemiology and Biostatistics, Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, Arizona, USA

²School of Anthropology, University of Arizona, Tucson, Arizona, USA

³Department of Medicine, University of Arizona, Tucson, Arizona, USA

⁴Department of Nutritional Sciences, University of Arizona, Tucson, Arizona, USA

⁵Department of Health Promotion Sciences, Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, Arizona, USA

⁶Scripps Translational Science Institute, La Jolla, CA, USA

⁷Center for Disparities in Diabetes, Obesity and Metabolism, Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, University of Arizona, Tucson, Arizona, USA

⁸Departments of Psychology and Psychiatry, Neuroscience and Physiological Sciences Interdisciplinary Programs, BIO5 Institute, and Evelyn F. McKnight Brain Institute, University of Arizona, Tucson, Arizona, USA; Arizona Alzheimer’s Consortium, Phoenix, Arizona, USA

We thank the editors for the opportunity to respond to the issues raised by Folley et al. [1]. We also thank Folley et al. for their interest in our paper and for helping to provide further insight into our findings.

In our paper, we sought to identify genetic correlates of habitual physical activity (PA) by examining five PA measures: three based on self-report, and two based on accelerometry data [2]. Folley et al. hypothesized that individuals with poor cognitive function are more likely to over-estimate and over-report their level of physical activity. They support their hypothesis by showing that those with lower memory, lower education, and older age tend to self-report higher levels of PA, but generally exhibit lower levels of accelerometry-measured PA – an objective and presumably more accurate measure of actual PA (although not without its own limitations). The trends that they show are relatively consistent, and do seem to suggest that there is over-reporting of PA by these individuals. This could suggest that some of the loci identified for moderate-to-vigorous PA (MVPA), and perhaps for some of the

*Corresponding Author: Department of Epidemiology and Biostatistics, Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, Arizona, 85724, 520-621-0147, yann@email.arizona.edu.

This research was conducted using the UK Biobank Resource under Application Numbers 15678 and 21259.

The authors declare no conflict of interest.

other self-report measures, may be related, at least in part, to relatively poorer cognitive performance rather than to MVPA.

To determine whether the associations between the nine SNPs and self-reported MVPA described in Klimentidis et al. [2] are unrelated to education and/or memory, we performed analyses with and without adjustment for educational attainment (College/University degree, other qualification, no qualification), and then for both educational attainment and memory (pairs-matching test), both categorized in a similar fashion as Folley et al. We find that the SNP associations with MVPA remain relatively unchanged after these adjustments (see Table 1), suggesting that these genetic associations are not mediated by educational attainment or memory.

Folley et al. also mention that this information bias may be particularly exemplified by the fact that the Alzheimer's (and heart disease) $\epsilon 4$ risk allele at *APOE* was found to be associated with significantly greater levels of MVPA. In our paper, we suggested that this may be the result of selection/survival bias due to an enrichment, at older ages, of *APOE* $\epsilon 4$ positive individuals with healthy lifestyles (individuals with risk allele and an unhealthy lifestyle are less likely to be alive or healthy enough to participate). We reported that the frequency of this allele in the sample decreases with age, and its strength of association with MVPA increases with age. Folley et al. mention there was not a significant association of *APOE* with accelerometry measures of PA, which would argue against the selection bias hypothesis. However, as we mentioned in our paper, although not genome-wide significant, the *APOE* risk allele was nominally and positively associated with acceleration average ($p=0.008$). Furthermore, in another study in the UK Biobank, Jones et al. [3] found that $\epsilon 4$ carriers exhibited a higher quality of sleep (lower nocturnal sleep episodes), based on accelerometry. Additionally, Lyall et al. [4] did not find a significant association of the *APOE* variant with cognitive ability in the UK Biobank. Thus, given the analyses we present here, and the aforementioned findings, we suspect that the *APOE* finding, if not representing a causal/direct effect on MVPA, is more likely the result of selection bias rather than information bias.

In their letter, Folley et al. focus on one of the three self-report variables that we considered. Although we did find that age is positively associated with MVPA, it is negatively associated with the two other self-reported PA variables (and the accelerometry-based phenotypes, as also shown by Folley et al.). Similarly, the overall genetic correlations with intelligence and years of schooling are generally positive for the two other self-report PA measures. Thus, we suspect that among the self-report variables, MVPA may be most prone to the influence of potential information bias related to relatively poorer cognitive performance, especially, and perhaps principally, among older individuals. However, as mentioned above, the top loci identified for MVPA do not seem to be subject to this information bias.

We thank Folley et al. for their contribution towards better understanding our results, and for drawing attention to the potential for information bias in self-reported PA. Although our findings do not support a cognitively-based information bias explaining our top genetic results for self-reported MVPA, we do agree that the loci we identified will require further

scrutiny through replication and detailed follow-up efforts in order to better understand if and how they may alter the propensity for habitual physical activity.

References

1. Folley S, Zhou A, Hypponen E. Information bias in measures of self-reported physical activity. *Int J Obes*. 2018 In press.
2. Klimentidis YC, Raichlen DA, Bea J, Garcia DO, Wineinger NE, Mandarino LJ, Alexander GE, Chen Z, Going SB. Genome-wide association study of habitual physical activity in over 377,000 UK Biobank participants identifies multiple variants including *CADM2* and *APOE*. *Int J Obes*. 2018; 42: 1161–1176.
3. Jones SE, van Hees VT, Mazzotti DR et al. Genetic studies of accelerometer-based sleep measures in 85,670 individuals yield new insights into human sleep behavior. *bioRxiv*. 2018; 10.1101/303925
4. Lyall DM, Ward J, Ritchie SJ et al. Alzheimer disease genetic risk factor APOE e4 and cognitive abilities in 111,739 UK Biobank participants. *Age and Ageing*. 2016; 45: 511–517. [PubMed: 27103599]

Table 1:

Association of nine self-report MVPA-associated SNPs before and after adjusting for educational attainment and memory.

SNP	Without Education or Memory			With Education			With Education and Memory		
	β (SE)	p-value	n	β (SE)	p-value	n	β (SE)	p-value	n
rs429358	0.0216 (0.0031)	1.87E-12	377,234	0.0221 (0.0031)	5.97E-13	375,064	0.0215 (0.0031)	2.38E-12	373,895
rs7804463	-0.015 (0.0022)	1.64E-11	373,605	-0.0142 (0.0022)	1.94E-10	371,451	-0.0139 (0.0022)	5.05E-10	370,291
rs2854277	-0.0325 (0.0088)	2.11E-04	293,874	-0.033 (0.0088)	1.73E-04	292,207	-0.0318 (0.0088)	2.99E-04	291,285
rs7791992	-0.0134 (0.0023)	2.93E-09	374,272	-0.0135 (0.0023)	2.53E-09	372,114	-0.0133 (0.0023)	3.97E-09	370,957
rs3094622	-0.0195 (0.0032)	1.75E-09	377,234	-0.0193 (0.0032)	2.33E-09	375,064	-0.0191 (0.0032)	3.62E-09	373,895
rs149943	-0.0187 (0.0031)	2.02E-09	377,234	-0.019 (0.0031)	1.32E-09	375,064	-0.019 (0.0031)	1.44E-09	373,895
rs2035562	-0.014 (0.0024)	3.77E-09	372,923	-0.014 (0.0024)	3.95E-09	370,773	-0.0139 (0.0024)	5.55E-09	369,615
rs2988004	0.0123 (0.0023)	7.01E-08	359,818	0.0125 (0.0023)	4.71E-08	357,754	0.0124 (0.0023)	5.89E-08	356,650
rs1043595	-0.0142 (0.0025)	8.85E-09	372,622	-0.0134 (0.0025)	5.96E-08	370,486	-0.0135 (0.0025)	5.26E-08	369,329

The results shown here differ slightly from those shown in the original paper, because we used Plink to extract the relevant genotypes (sets missing to some imputed dosages), and we did not use a linear mixed model as implemented in BOLT-LMM. All associations are performed using Model 1, as specified in the paper, and with education and memory as three-level categorical variables.