Published in final edited form as:

J Women Aging. 2018 ; 30(1): 2–5. doi:10.1080/08952841.2017.1407575.

Failing to Account for Regression to the Mean Results in Unjustified Conclusions

Tanya M. Halliday¹, Diana M. Thomas², Cynthia O. Siu³, and David B. Allison^{4,*}

¹Division of Endocrinology, Metabolism and Diabetes, Department of Medicine, University of Colorado-Denver, Anschutz Medical Campus, Aurora, CO 80045

²Department of Mathematical Sciences, United States Military Academy, West Point, NY 10996.

³COS & Associates Ltd., Central, Hong Kong.

⁴Department of Epidemiology and Biostatistics, School of Public Health, Indiana University-Bloomington, Bloomington, IN 47405.

A paper recently published by Tomisek et al. in this journal reported on preliminary findings from, the “Strong. Healthy. Energized (SHE) program¹. SHE addressed an important topic, testing a 12-week ‘healthy weight intervention’ designed specifically for lesbian women 60 years of age and above, with overweight/obesity. The stated goals of this program were to 1) decrease the average waist circumference by 5%; 2) increase the average number of steps per day by 2,000; 3) decrease the percentage of participants consuming sugar-sweetened beverages daily by 25%; and 4) increase the percentage of participants meeting fruit and vegetable intake guidelines by 25%. The authors conclude the SHE program should be adopted by senior centers and lesbian, gay, bisexual, and transgender (LGBT) centers as it demonstrated “*effectiveness in improving health behaviors and short-term health outcomes in the target population*”¹. Specifically, the authors make this conclusion based upon a “*marked step increase*”¹ for participants in the lowest tertile-defined category of baseline step count. However, the analysis does not support this conclusion. This is because regression to the mean (RTM), rather than treatments effectiveness, explains, in part, the arrived conclusion.

RTM is a statistical phenomenon that describes the tendency for extreme values observed on initial assessment to be less extreme, and closer to the population mean with repeated measurement when the correlation coefficient is less than 1.0². First described in regards to height of children in comparison to their parents and the population by Francis Galton as “regression to mediocrity”³, RTM is a concept that has often been ignored and misunderstood in health and obesity-related research^{4–8}. Failure to account for RTM often

*Corresponding Author: David B. Allison, Indiana University School of Public Health-Bloomington, 1025 E 7th St, PH 111, Bloomington, IN 47405. Allison@IU.edu.

Competing interests

None pertinent. Dr. Allison’s prior University received numerous grants and gifts from for-profit and non-profit organizations with interests in obesity and health, including food, beverage, and pharmaceutical companies to support research and educational activities. Dr. Siu serves as a consultant for Sunovion, Sumitomo Dainippon Pharma, Centre for Addiction and Mental Health (Schizophrenia Program), University of Toronto, and the Chinese University of Hong Kong (Psychology, Biomedical Sciences).

leads to errors in interpretation of results and unjustified conclusions. In pre-post study designs that lack a comparator control group, neglecting RTM can lead to the inaccurate conclusion that an intervention was effective in improving a health outcome in a group of participants ^{4,5}.

In the results section of this paper the authors state, “*The SHE program was most effective for participants with low levels of physical activity and steps*”¹. As shown in Table 8, subjects whose baseline values were below the bottom tertile averaged ~3,900 steps/day at baseline, and ~5,540 steps at the end of the SHE program. The expected analytical approach of evaluating change in step count for the entire sample was not reported (beyond the authors stating that the goal of an increase of 2,000 steps/day on average was not reached). Interestingly, in regards to waist circumference and body weight, Tomisek et al. first present results of the change from pre- to post-intervention in the full sample. In the case of body weight, no significant difference was detected between baseline and post-program measurement ($p=0.22$). The authors go on to show that participants with higher baseline body weights lost more weight than those in the bottom or middle tertile. They comment that this result was “unsurprising”; which it should be, given that RTM occurs when selecting a subgroup of participants with higher baseline value of body weight from the full sample. Therefore, given the acknowledgement that a decrease in body weight in the group with highest initial body weights was expected, the same logic should have been applied to the outcome of change in step counts. Thus, the conclusion that there was evidence for effectiveness is not justified given that the results are likely due to RTM and not specific intervention effects attributable to the SHE program. This does not mean that the SHE program is not effective, only that it was not convincingly shown to be effective by ordinary scientific standards in this study.

Interventions that are evaluated without the use of a control group are susceptible to the reliance on results that may be a consequence of RTM. Greater vigilance regarding RTM is necessary throughout the research and publication process.

Acknowledgements:

Supported in part by NIH grants R25DK099080; R25HL124208; R25GM116167; T32DK07658–26. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or any other organization.

References:

1. Tomisek A, Flinn B, Balsky T, Gruman C, Rizer AM. Strong, healthy, energized: Striving for a healthy weight in an older lesbian population. *J Women Aging*. 2017;29(3):230–242. [PubMed: 27484950]
2. Bland JM, Altman DG. Regression towards the mean. *BMJ (Clinical research ed)*. 1994;308(6942):1499.
3. Galton F Regression Towards Mediocrity in Hereditary Stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*. 1886;15:246–263.
4. George BJ, Beasley TM, Brown AW, et al. Common scientific and statistical errors in obesity research. *Obesity (Silver Spring, Md)*. 2016;24(4):781–790.
5. Linden A Assessing regression to the mean effects in health care initiatives. *BMC Medical Research Methodology*. 2013;13(1):119. [PubMed: 24073634]

6. Skinner AC, Heymsfield SB, Pietrobelli A, Faith MS, Allison DB. Ignoring regression to the mean leads to unsupported conclusion about obesity. *The international journal of behavioral nutrition and physical activity*. 2015;12:56. [PubMed: 25948534]
7. Cockrell Skinner A, Goldsby TU, Allison DB. Regression to the Mean: A Commonly Overlooked and Misunderstood Factor Leading to Unjustified Conclusions in Pediatric Obesity Research. *Childhood obesity (Print)*. 2016;12(2):155–158. [PubMed: 26974388]
8. Allison DB, Loebel AD, Lombardo I, Romano SJ, Siu CO. Understanding the relationship between baseline BMI and subsequent weight change in antipsychotic trials: Effect modification or regression to the mean? *Psychiatry Research*. 2009;170(2):172–176. [PubMed: 19897253]