Research Article

Measurement of Weight in Clinical Trials: Is One Day Enough?

Nisa M. Maruthur,¹ William M. Vollmer,² Jeanne M. Clark,^{1,3,4} Gerald J. Jerome,^{1,5}
Lillian F. Lien,⁶ Catherine M. Loria,⁷ and Lawrence J. Appel^{1,3,4}

¹ Division of General Internal Medicine, The Johns Hopkins University School of Medicine, 2024 E Monument, Room 2-518, Baltimore, MD 21287, USA

² Center for Health Research, Kaiser Permanente, Portland, OR 97227, USA

³ Welch Center for Prevention, Epidemiology, and Clinical Research, The Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA

⁴ Department of Epidemiology, The Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD 21287, USA

⁵ Department of Kinesiology, Towson University, Towson, MD 21252, USA

⁶ Division of Endocrinology, Metabolism, and Nutrition, Duke University Medical Center, Durham, NC 27710, USA

⁷ National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20824, USA

Correspondence should be addressed to Nisa M. Maruthur, maruthur@jhmi.edu

Received 27 May 2010; Accepted 9 July 2010

Academic Editor: Jack Adam Yanovski

Copyright © 2010 Nisa M. Maruthur et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Weight is typically measured on a single day in research studies. This practice assumes negligible day-to-day weight variability, although little evidence exists to support this assumption. We compared the precision of measuring weight on one versus two days among control participants in the Weight Loss Maintenance trial. **Methods.** Trained staff measured weight on two separate days at baseline, 12 months, and 30 months (2004–2007). We calculated the standard deviation (SD) of mean weight change from baseline to the 12- and 30-month visits using (a) the first and (b) both daily weights from each visit and conducted a variance components analysis (2009). **Results.** Of the 316 participants with follow-up measurements, mean (SD) age was 55.8 (8.5) years, BMI was 30.8 (4.5) kg/m², 64% were women, 36% were black, and 50% were obese. At 12 months, the SD of mean weight change was 5.1 versus 5.0 kg using one versus two days of weight measurements ($P = .76$), while at 30 months the corresponding SDs were 6.3 and 6.3 kg ($P = .98$). We observed similar findings within subgroups of BMI, sex, and race. Day-to-day variability within individuals accounted for <1% of variability in weight. **Conclusions.** Measurement of weight on two separate days has no advantage over measurement on a single day in studies with well-standardized weight measurement protocols.

1. Introduction

Short- and long-term variability in outcomes can affect power in research studies. Weight is usually measured on a single day per study visit (i.e., time point in study: baseline and followup) in clinical trials of weight loss or maintenance. However, day-to-day variability in weight within an individual may decrease precision, thus impacting necessary sample sizes. Because of this type of variability in blood pressure, sample size is optimized by using at least 3 sets of daily blood pressure measurements per study visit [1].

We aimed to evaluate the hypothesis that measurement of weight on one day is sufficiently precise compared to

two days by (1) comparing the standard deviation of weight change based on one versus two days of measurements and (2) evaluating components of weight variability using data from the Weight Loss Maintenance (WLM) Trial [2].

2. Methods

2.1. Study Design. WLM was an NHLBI-supported, 4-site, and randomized controlled trial comparing 2 weight loss maintenance interventions to a self-directed control group in two phases: an initial 6-month weight-loss phase common to all participants (phase 1) and a randomized, 30-month weight-maintenance phase (phase 2). This study uses data

from phase 2. Data collection occurred between 2004 and 2007. Methods and results have been published previously [2–4]. Participants had a body mass index (BMI) between 25 and 45 kg/m², treated hypertension, hyperlipidemia, or both, access to a telephone and the Internet, no active cardiovascular disease, diabetes, or medical condition(s) precluding participation, absence of >9 kg weight loss in the prior 3 months, no use of weight loss medication, and no history of weight loss surgery [2]. Participants provided informed consent, and institutional review boards at each study site approved the protocol [2]. Participants were eligible for phase 2 if they lost ≥ 4 kg during the initial weight loss phase [2].

The current study is limited to the 316 (out of 342) control group participants who attended either the 12- or 30-month follow-up visits in phase 2. This group received printed lifestyle guidelines (for diet and physical activity) while meeting with an interventionist at randomization [3] and after the 12-month visit [2].

2.2. Study Measurements. Weight was measured during screening, at the conclusion of phase 1 (baseline), and at 12- and 30-month postrandomization into phase 2. Height was measured once during screening. Trained and certified staff used high-quality calibrated digital scales and calibrated stadiometers, and participants wore light clothes and no shoes [2].

On any given day, weight was measured as the average of two independent measurements. For the baseline and for the 12- and 30-month visits, this process was repeated on two separate days (day 1 and day 2). At baseline, these repeated measurements could have been separated by as much as 32 days, and for purposes of this analysis, we restricted participants to those individuals whose baseline weight measures were within 21 days apart from each other. Only 5 (1.6%) out of a possible 314 participants were excluded from the 12-month analysis for this reason, and only 4 (1.4%) out of a possible 285 participants were excluded from the 30-month analysis. Because the baseline weights were taken during the period of active weight loss intervention when the cohort as a whole was still losing weight, we adjusted the initial (day 1) weights to have the same mean as the final weights (day 2) by adding to each person's day-1 weight the mean change in weight across all participants between day 1 and 2. For the 12- and 30-month visits, adjustment for temporal drift between days 1 and 2 was not needed since participants were not undergoing an active weight loss intervention. The mean differences in the two weights at 12- and 30-months were -0.02 and -0.03 kg, respectively.

Participants were not required to fast, and weight measurement did not occur at a specific time of day. However, each participant was required to present in a fasting state for at least one day per visit which included laboratory testing; this visit would likely occur in the morning.

2.3. Statistical Methods. Analyses were completed using SAS 9.2. Using (a) the first and (b) both daily weights from each visit, we calculated the mean and variance in weight

change from baseline to 12 months and from baseline to 30 months. We defined the variance in weight change as the standard deviation of weight change and used F-tests to compare them. We conducted similar analyses stratified by screening BMI (25–29.9 kg/m² and ≥ 30 kg/m²), sex, and race (African American and non-African American). We also conducted a nested analysis of variance to estimate between-person, between-visit, and day-to-day (within-visit) variability in weight. Within-day variability was not assessed, but was negligible. A P value $< .05$ was considered statistically significant. Analyses were performed in 2009.

3. Results

Mean (SD) age was 55.8 (8.5) years, and mean (SD) BMI was 30.8 (4.5) kg/m²; 64% of participants were women; 36% were African American; 50% were obese (Table 1). The mean number of days between day 1 and day 2 of weight measurement for at the baseline, 12-month, and 30-month visits was 8.9, 9.1, and 10.4, respectively.

Although the standard deviation of weight change differed for measurements taken 12 and 30 months apart, it was essentially unchanged regardless of whether weight was measured as the mean of one or two days of measurements (Table 1). For example, the standard deviation for 30-month weight change for all participants was 6.3 kg when weight was measured on a single day (day 1) and 6.3 kg when weight was measured on two separate days (days 1 and 2, $P = .98$). We found no statistically significant differences overall or in any of the subgroups studied. Day-to-day variability (i.e., of mean change in weight from a single day compared to that using both day 1 and 2 of the visits) within individuals accounted for $< 1\%$ of total variability in weight, while between-visit variability within individuals accounted for 6–17% of the variability in weight change depending on the subgroup (Table 1).

4. Discussion

In the context of this study being conducted with a rigorous weight measurement protocol, the use of one versus two measurement days per study visit (time point) had essentially no impact on the variability of weight change in adults measured over 12 and 30 months, and day-to-day variability in weight correspondingly accounted for less than 1% of the total variability in weight.

A limitation of the study is that the baseline weights were measured while the participants were still engaged in an active weight-loss intervention. We minimized the added variability in weight caused by this by adjusting for temporal drift and by only including participants with 21 or fewer days between day 1 and 2 measurements at the baseline visit. Also, we evaluated the contribution of day-to-day variability in weight to variability in weight change over 12 and 30 months. Day-to-day variability in weight may be of more significance in shorter studies thus limiting the generalizability of our results to studies brief in duration.

TABLE 1: Mean weight change (standard deviation) and weight variability among control participants in weight loss maintenance^a.

	All	Overweight ^b	Obese ^c	Men ^d	Women ^d	African american ^d	Non-African american ^d
<i>N</i>	316	134	158	113	203	113	203
<i>Baseline characteristics</i>							
Mean age (SD), years	55.8 (8.5)	57.3 (8.2)	54.9 (8.4)	56.6 (8.6)	55.4 (8.5)	53.5 (9.0)	57.1 (8.0)
% Female	64.2	61.9	67.1	0	100	71.7	60.1
% African American	35.8	26.9	44.9	28.3	39.9	100	0
% Overweight	42.4	100	0	45.1	40.9	31.9	48.3
% Obese	50	0	100	46	52.2	62.8	42.9
Mean BMI (SD), kg/m ²	30.8 (4.5)	27.7 (1.4)	34.5 (3.2)	30.3 (4.1)	31.2 (4.7)	32.1 (4.9)	30.1 (4.1)
Mean weight, kg	86.9 (15.0)	78.4 (8.8)	96.6 (13.4)	94.7 (15.1)	82.5 (13.1)	89.5 (15.9)	85.4 (14.3)
<i>N for 12-month analysis</i>	309	131	155	109	200	110	199
<i>12-Month weight Δ (SD), kg</i>							
Single day	3.9 (5.1)	3.5 (4.3)	4.5 (5.7)	3.9 (5.8)	3.9 (4.6)	3.6 (4.4)	4.0 (5.4)
Two days	3.9 (5.0)	3.5 (4.2)	4.5 (5.7)	3.9 (5.8)	3.9 (4.5)	3.6 (4.4)	4.1 (5.3)
<i>N for 30-month analysis</i>	281	116	141	104	177	101	180
<i>30-Month weight Δ (SD), kg</i>							
Single day	6.0 (6.3)	5.7 (5.1)	6.4 (7.4)	6.7 (8.0)	5.5 (5.0)	5.4 (5.0)	6.2 (6.9)
Two days	6.0 (6.3)	5.8 (5.0)	6.4 (7.4)	6.6 (8.1)	5.6 (4.9)	5.4 (5.0)	6.3 (6.9)
<i>% Total variance</i>							
Between-individual, %	91.3	82.4	86.7	88.9	90.7	93.5	89.9
Between-visit, %	8.5	17.1	13	10.9	9	6.3	9.9
Within-visit, %	0.2	0.4	0.3	0.2	0.2	0.2	0.2

SD: standard deviation; kg: kilograms; BMI: body mass index

^a*P* values > .70 for all *F*-tests comparing standard deviations within each group (e.g., for all participants, comparing SD of mean weight change at 30 months using a single day (6.3 kg) to that using 2 days (6.3 kg), *P* value = .98).

^bBMI 25–29.9 kg/m²

^cBMI ≥ 30 kg/m²

^dBMI ≥ 25 kg/m².

To our knowledge, this is the first study of its kind to evaluate day-to-day variability in weight as it applies to measurement in clinical studies. Especially pertinent is our subgroup analysis showing the stability of day-to-day weight variability across subgroups defined by BMI, sex, and race. Finally, WLM used a stringent weight measurement protocol that included calibrated digital scales, trained staff, and duplicate within-day measurements to minimize the probability of measurement or transcription error [2]. Observed day-to-day variability could be expected to be greater in studies with less stringent weight measurement protocols, though our findings would suggest such increased variability would likely be due to measurement error rather than to true day-to-day variability.

In conclusion, we show that in the setting of a well-standardized weight measurement protocol with negligible measurement error, there is little day-to-day variability in weight. Weight measurement on a single day is sufficiently precise for use in a clinical trial.

Acknowledgments

Dr. Maruthur was supported by Grant no. 1KL2 RR025006-01, National Center for Research Resources at the National Institutes of Health and National Institutes of Health

Roadmap for Medical Research and Grant no. 5 T32 HL007180, National Heart, Lung, and Blood Institute at the National Institutes of Health. Weight Loss Maintenance was supported by Grant nos. 5-UO1 HL68734, 5-UO1 HL686676, 5-UO1 HL68790; 5-UO1 HL68920, and 5-UO1 HL68955, National Heart, Lung, and Blood Institute at the National Institutes of Health.

References

- [1] W. M. Vollmer, L. J. Appel, L. P. Svetkey et al., "Comparing office-based and ambulatory blood pressure monitoring in clinical trials," *Journal of Human Hypertension*, vol. 19, no. 1, pp. 77–82, 2005.
- [2] L. P. Svetkey, V. J. Stevens, P. J. Brantley et al., "Comparison of strategies for sustaining weight loss: the weight loss maintenance randomized controlled trial," *Journal of the American Medical Association*, vol. 299, no. 10, pp. 1139–1148, 2008.
- [3] P. Brantley, L. Appel, J. Hollis et al., "Design considerations and rationale of a multi-center trial to sustain weight loss: the weight loss maintenance trial," *Clinical Trials*, vol. 5, no. 5, pp. 546–556, 2008.
- [4] J. F. Hollis, C. M. Gullion, V. J. Stevens et al., "Weight loss during the intensive intervention phase of the weight-loss maintenance trial," *American Journal of Preventive Medicine*, vol. 35, no. 2, pp. 118–126, 2008.