



# HHS Public Access

Author manuscript

*Obesity (Silver Spring)*. Author manuscript; available in PMC 2018 February 01.

Published in final edited form as:

*Obesity (Silver Spring)*. 2017 February ; 25(2): 363–369. doi:10.1002/oby.21737.

## Validation of Clinic Weights from Electronic Health Records against Standardized Weight Measurements in Weight Loss Trials

Lan Xiao<sup>a</sup>, Nan Lv<sup>a</sup>, Lisa G. Rosas<sup>a,b</sup>, David Au<sup>c</sup>, and Jun Ma<sup>a,d</sup>

<sup>a</sup>Palo Alto Medical Foundation Research Institute, 795 El Camino Real, Ames Building, Palo Alto, CA 94301, U.S.A

<sup>b</sup>Stanford University School of Medicine, 291 Campus Drive, Stanford, CA 94305

<sup>c</sup>VA Puget Sound Health Care System, HSR&D, University of Washington, Division of Pulmonary and Critical Care Medicine, 1660 S. Columbian Way, Seattle WA 98101

<sup>d</sup>School of Public Health and College of Medicine, University of Illinois at Chicago, Chicago, IL 60608, U.S.A

### Abstract

**Objective**—To validate clinic weights in electronic health records against researcher-measured weights for outcome assessment in weight loss trials.

**Methods**—Clinic and researcher-measured weights from a published trial (“BE WELL”) were compared using Lin’s concordance correlation coefficient (CCC), Bland and Altman’s limits-of-agreement, and polynomial regression model. Changes of clinic and researcher-measured weights in BE WELL and another trial, “E-LITE,” were analyzed using growth curve modeling.

**Results**—Among BE WELL (n=330) and E-LITE (n=241) participants, 96% and 90% had clinic weights (mean (SD) of 5.8 (6.1) and 3.7 (3.9) records) over 12 and 15 months of follow-up, respectively. The CCC was 0.99 and limits-of-agreement plots showed no pattern between or within treatment groups, suggesting overall good agreement between researcher-measured and nearest-in-time clinic weights up to 3 months. The 95% confidence intervals for predicted percent differences fell within +/-3% for clinic weights within 3 months of the researcher-measured weights. Furthermore, the growth curve slopes for clinic and researcher-measured weights by treatment group didn’t differ significantly, suggesting similar inferences about treatment effects over time, in both trials.

---

**Corresponding author:** Jun Ma, MD, PhD, Institute for Health Research and Policy (MC275), 5th Floor, West Research Office Building, 1747 West Roosevelt Road, Chicago, Illinois 60608, maj2015@uic.edu, Phone: (312) 413-9830; (312) 413-4295, Fax: (312) 413-9835.

Clinical trial registration

E-LITE trial registration: NCT00842426

BE WELL trial registration: NCT00901095

### Disclosure

The authors declare no conflict of interest.

**Conclusions**—Compared with researcher-measured weights, close-in-time clinic weights showed high agreement and inference validity. Clinic weights can be a valid pragmatic outcome measure in weight loss studies.

### Keywords

research weights; electronic health records; clinic weights; weight loss; obesity

---

### Introduction

Obesity is a major public health challenge facing policy makers, researchers, the healthcare system, providers, and patients. The annual health care costs of obesity are estimated at \$190.2 billion or nearly 21% of annual medical spending in the US.(1) Despite significant efforts across multiple sectors, the prevalence of obesity continues to rise, reaching 38% of US adults in 2014.(2) To address this costly and pervasive health challenge, it is imperative to disseminate and implement evidence-based weight loss interventions into practice, and equally importantly, to develop and test new interventions in research and non-research settings. One of the fastest growing interests among researchers, policy makers, healthcare systems, governmental agencies, funders, and other key stakeholders is large pragmatic trials and real-world implementation studies, which call for new experimental designs and cost-efficient measurement methods.(3–5)

Rigorous evaluation of obesity interventions typically involves collecting weight data per standardized research protocols at baseline and one or more follow-up time points over an extended period of months to years and perhaps for hundreds to even thousands of participants as determined by power analyses.(6) A critical barrier to innovation in pragmatic obesity prevention and treatment research is the high cost of in-person data collection both in terms of research staff time and participant burden. Furthermore, this barrier can be a significant challenge to recruitment as well as retention. This may adversely affect the internal and external validity of a trial. Self-reported weight as a surrogate of directly measured weight has been suggested to overcome this challenge but is prone to reporting biases and errors.(7, 8) Thus, alternative approaches to obtaining objectively measured weight data in pragmatic and low-cost ways are needed, and the validity of these methods needs to be established to substantiate their substitution for “gold standard” researcher-measured weights.

Authorized abstraction of clinically measured weights from patients’ electronic health records (EHR) is one such promising approach due to the relative low cost of extracting EHR data and the ability to eliminate patient burden as a result of separate longitudinal research data collection. Many adults, especially those with or at high risk for obesity-related comorbidities, routinely visit their health care providers and have their weight objectively measured at the clinic visits.(9–11) Several studies found good agreement between clinic and research weight measurements at one point in time.(12–14) The utility of clinic weights for evaluating intervention effects on weight changes over time requires further research. Leveraging repeated-measures weight data from 2 published behavioral

weight loss treatment trials, (15–18) this study aimed to validate clinic weights from EHR against research weights measured according to standardized protocols.

## Methods

### Study Participants

The 2 randomized controlled clinical trials (RCTs), BE WELL (The Breathe Easier through Weight Loss Lifestyle) and E-LITE (Evaluation of Lifestyle Interventions to Treat Elevated Cardiometabolic Risk in Primary Care), were conducted separately in 2 health systems in the San Francisco Bay Area, California. BE WELL was a 2-arm RCT that compared the effectiveness of a coach-led, face-to-face group lifestyle intervention (n=165) and usual care (n=165) on asthma control among primary care patients with obesity and uncontrolled asthma (recruitment period: 04/2010–08/2012).(17) Although the primary asthma control outcome was null, BE WELL intervention participants achieved significantly greater weight loss than controls at 12-month follow-up. E-LITE was a 3-arm RCT also based in primary care that demonstrated the effectiveness of 2 adapted Diabetes Prevention Program lifestyle interventions for weight loss among overweight or obese adults with prediabetes, metabolic syndrome, or both.(15, 16) Participants recruited from 05/2009–06/2010 were randomized to either of the intervention groups, one similar to the BE WELL coach-led group intervention (n=79) or a self-directed DVD intervention (n=81), or to the usual care control group (n=81), and were followed for 15 months.

The BE WELL trial was approved by an Institutional Review Board of the Kaiser Foundation Research Institute in Northern California. The E-LITE trial was approved by the Palo Alto Medical Foundation's Institutional Review Board. All participants in both studies provided written informed consent. Complete protocols of both trials were published previously.(16, 18)

### Research weights

Following standardized protocols, trained research staff measured body weight (“research weight”) at 2 baseline visits and multiple follow-up visits in both BE WELL (6 and 12 months post randomization) and E-LITE (3, 6 and 15 months post randomization). At each visit weight was measured twice in light clothes without shoes using either a balance beam scale or a high-quality digital scale. The 2 measurements were averaged for analysis.

### Clinic weights

Body weight measurements were abstracted from the EHR (“clinic weight”) for participants from 3 months before randomization to 3 months after the primary endpoint of each trial which was 12 months in BE WELL and 15 months in E-LITE, post randomization – namely, the abstraction period.

### Statistical analyses

We examined agreement between research weights and clinic weights using 3 approaches (with data from the BE WELL study in the first 2 approaches and data from both studies in the third approach). E-LITE was not amenable to the first 2 approaches because of the

limited available clinic weights within the defined index window for each follow-up time point of the study.

The first approach compared BE WELL participants' research weights at 6 and 12 months follow-up with clinic weights available within an index window of up to 3 months before and after the date of each participant's randomization date plus 6 and 12 months. The index window for baseline, however, only included up to 3 months before the date of each participant's randomization date because the first 3 months post randomization were the intensive treatment phase for participants randomized to the intervention group. If multiple clinic weight measurements were available within the index window, the closest weight in time was chosen. We examined 2 agreement indices for continuous measures, Lin's concordance correlation coefficient (CCC)(19) and Bland and Altman's limits of agreement. (20) CCC assesses the agreement of two measurements, ranging from -1 to 1, with perfect agreement at 1. It is an adaptation of the Pearson correlation coefficient that incorporates the departure from a 45 degree line through the origin representing perfect agreement. Bland and Altman's limits of agreement are often used to determine whether the new measurement (e.g., clinic weight) can replace the established one (e.g., research weight), using graphics to plot the difference of the 2 measurements and 95% limits of agreement (i.e., mean of the difference  $\pm 1.96 \times$  standard deviation of the difference) against the mean of the 2 measurements for each subject.

The second approach took advantage of the availability of multiple clinic weights over the study period to compare with research weights in BE WELL. We used a polynomial regression model to fit percentage difference ((clinic weight-research weight)/research weight \* 100%) between research weights and all available clinic weights within the index window specified above versus days from research weight at each time point. The resulting predicted percentage differences and 95% confidence intervals (CI) were plotted against days between clinic and research weights. Reference lines mark differences of  $\pm 3\%$  because of the established definition of weight maintenance as weight changes within these limits. (21) Hence, the application of this definition helps identify differences between clinic and research weights that were outside these limits, which may be more likely to be true differences and less confounded by factors such as natural weight fluctuations within individuals.

The third approach used all available clinic and research weights from BE WELL and E-LITE separately. We used mixed model growth curve analysis to compare slopes and quadratics of weight change by treatment group within each trial as determined by research weights at baseline and follow-up and by all available clinic weights during the abstraction period. BE WELL and E-LITE data were analyzed in 2 separate models; each included treatment group (*Group*), source of weight data (*Source*; clinic vs. research), days to randomization (*Days*), and interaction terms for *Group*  $\times$  *Days* and *Source*  $\times$  *Group*  $\times$  *Days*. We fitted linear and quadratic terms for *Days* for the growth curves, and tested them for the significance of differences between sources in the same group. The quadratic terms were removed in the final models because their main effects and interactions were nonsignificant. All analyses were conducted using SAS, version 9.3 (SAS Institute Inc., Cary, North Carolina).

## Results

### Participant characteristics

This study included the original trial participants in BE WELL (n=330) and E-LITE (n=241). (15, 17) In BE WELL, participants were racially and ethnically diverse and predominantly middle-aged, female, with some college or more education, and obese. In E-LITE, participants were mostly middle-aged, non-Hispanic White, overweight/obese men and women with a college degree or higher (Table 1).

### Clinic weights availability

The mean (SD) number of available clinic weights was 5.8 (6.1) for the 96% of BE WELL participants who had a least one clinic weight during the EHR abstraction period for a 12-month trial, and 3.7 (3.9) for the 90% of E-LITE participants who had a least one clinic weight during the EHR abstraction period for a 15-month trial. The mean numbers of clinic weights differed significantly between the usual care control group and the self-directed intervention group in E-LITE, 4.2 (4.2) vs. 2.8 (2.1), but were otherwise similar between treatment groups of each trial (Table 2).

### Lin's CCC and Bland and Altman's limits of agreement

The CCC (95% CI; number of subjects) for the nearest-in-time clinic weights within 1 month of research weights in BE WELL were 0.996 (95% CI: 0.994, 0.997; n=98) at baseline, 0.996 (95% CI: 0.995, 0.998; n=116) at 6 months, and 0.991 (95% CI: 0.987, 0.994; n=101) at 12 months. The corresponding CCC for the nearest clinic weights within 3 months were 0.988 (95% CI: 0.984, 0.992; n=166), 0.992 (95% CI: 0.990, 0.994; n=195), and 0.990 (95% CI: 0.987, 0.993; n=167).

The 95% limits of agreement for weights at baseline, 6- and 12-month are presented in Figure 1. There were less than 5% outliers (i.e., the difference between research weights and the closest clinic weights in time in the index window of 1 or 3 months > the 95% limits of agreement) at each time point. Overall, no pattern was detected in the variability of the differences between and within the treatment groups in these Bland-Altman plots. The distributions of the differences appeared random around the no difference line, suggesting that there was no systematic difference between clinic and research weights.

### Percent weight difference

The predicted percent weight differences and the 95% CIs resulting from polynomial regression line fell within +/- 3% reference lines comparing clinic weights in the index window of 3 months and research weights at baseline, 6 and 12months, respectively (Figure 2).

### Weight change trajectories

The slopes of weight change trajectories did not differ significantly between clinic and research weights for each treatment group in both studies (difference in slopes [95% CI]: -0.03 [-0.09, 0.04] for coach-led intervention and 0.03 [-0.04, 0.09] for usual care in BE

WELL, and 0.06 [−0.03, 0.15] for self-directed intervention, 0.01 [−0.07, 0.08] for coached intervention, and −0.01 [−0.09, 0.06] for usual care in E-LITE) (Table 3).

## Discussion

The main findings of this study suggest that weights measured in routine clinical care and recorded in the EHR can be a valid alternative to gold-standard researcher-measured weights for longitudinal obesity research studies. We used 3 different approaches to examine agreement (Lin's CCC and Bland-Altman plots, polynomial regression model-based predicted percent differences, and mixed-model growth curve analysis of weight change trajectories) using clinic and research weight data from 2 randomized behavioral weight loss treatment trials. The BE WELL and E-LITE trials were conducted with different patient population samples in different health systems. The consistent findings of good agreement across the 3 approaches based on these trial data despite their heterogeneity confirmed that the weights from clinical care could be used to replace the weight collected by research staff.

Prior studies have shown good agreement for clinic weights within a short time frame (e.g., 2 weeks, 1 month, or 3 months) of standardized weight measurements taken by researchers at a single time point.<sup>(12–14)</sup> By leveraging extant data from 2 prospective RCTs, this study extends the available literature by showing agreement consistently between repeated clinic and research weight measures within and between treatment groups at multiple time points. We also demonstrated that as opposed to using weights measured by researchers at fixed follow-up times across participants (e.g., 6 and 12 months), longitudinal clinic weight data collected at varying times among patients resulted in similar inferences about treatment effects on weight change over time. This finding suggests the statistical inference validity of clinic weights relative to research weights. It provides a pragmatic demonstration that the conclusions of those studies about the treatment effect on weight loss trajectories over time would have been the same had clinic weights been used instead of research weights.

The findings have several significant implications for pragmatic weight loss treatment trials. First, abstracting weight data from the EHR as opposed to research staff measuring participants' weight may improve recruitment and retention rates because it decreases participant burden. Second, relying on EHR data can lower the costs of research and potentially enable studies of larger sample sizes. Additionally, the findings inform pragmatic outcome measure choices in longitudinal obesity research studies. The availability of EHR weights continuously over time allows for longitudinal analysis of weight trajectories over time at marginal costs relative to collecting primary weight data separately for the research. If needed, point estimates can also be made from the trajectories for comparison of predicted weight changes between groups at a specific time point. Thus, relying on EHR data for weight measures in weight loss treatment trials offers numerous advantages for pragmatic approaches to research.

There are several practical issues to consider with using the EHR approach to weight measurement in obesity research. First, for this analysis we used data from 2 trials that were implemented in primary care settings within large integrated healthcare systems. These healthcare systems have robust centralized EHR databases that capture all visit data

including weights measured at visits in primary and specialty care. The findings indicate that utilizing EHR weight data works well in this type of setting. Strategies that may improve the availability and quality of data from the EHR in these settings include: (1) Standardize training of health care staff to accurately and consistently measure patients' weights and record them in EHR at every visit; and (2) During the course of a weight loss study, reinforce the importance of obtaining an accurate weight measure at every clinic visit with both staff and participants. These improvements may result in more reliable estimates of weight change over time. Utilizing EHR data for weight measurements may not be suitable for studies conducted in community settings or small healthcare practices given such challenges as the necessity of contacting multiple providers in different health systems or the lack of adequate infrastructure for EHR data management and abstractions for research purposes.

The findings from this study have important implications for the movement towards establishing learning healthcare systems. Learning healthcare systems are those that generate and utilize quality evidence to improve the safety, effectiveness, and value of healthcare for patients, their families, and their providers.(22) The promise of learning healthcare systems is predicated on the quality of available data. The findings from this study demonstrate that weight measurements from the EHR are available to support research and innovation in a learning healthcare system. However, the outliers identified in our analysis indicate opportunities for improving data quality. The movement toward continuous learning in healthcare systems is propelling the experts from a wide range of disciplines including medicine, public health, informatics, clinical research to explore the data quality issues and strategies central to clinical data improvement. The suggested strategies include minimizing the burden imposed on data collectors, maintaining the provenance and context of the data, and developing more standardized digital health data definitions and representations to improve data quality.(22)

This study had some potential limitations. We used the data from 2 clinical trials of behavioral weight loss treatments that were not designed to evaluate the agreement between clinic and research weights. Baseline clinic weights were not required when participants were enrolled in these studies. Availability of baseline clinic weights would make findings even more robust due to more clinic weight data at baseline. The 2 trials were behavioral weight loss treatment studies for overweight or obese adults with specific comorbid conditions, thus our findings may not generalize to patients with other comorbid conditions or healthy individuals. Furthermore, participants with more and evenly time-spaced weight measurements throughout the study period may be systematically different than those with fewer records (e.g., those with more comorbidities may visit the clinic more frequently than those with fewer comorbidities). Randomization or properly matched control techniques that take these factors into account can reduce bias in this regard.

## Conclusion

Our study responds to the need for low-cost and valid objective weight measurements, especially given the growing interest among funders and researchers in large, pragmatic, and patient-centered weight loss studies. We demonstrated that albeit less rigorously measured

than research weights according to standardized protocols, close-in-time weight data (up to 3 months) abstracted from EHR showed high agreement as well as statistical inference validity in evaluating the treatment effect on weight loss trajectories over time. Thus, clinic weights can be a valid pragmatic outcome measure in weight loss studies. Use of clinic weights instead of research weights could significantly reduce study cost from directly measuring participants' weights, decrease participant burden, and improve both recruitment and retention rates in obesity research studies.

## Acknowledgments

### Funding

The E-LITE study was supported by grant R34DK080878 from the National Institute of Diabetes and Digestive and Kidney Diseases, a Scientist Development Grant award (0830362N) from the American Heart Association, and the BE WELL study was supported by grant R01HL094466 from the National Heart, Lung, and Blood Institute. This study was also supported by internal funding from the Palo Alto Medical Foundation Research Institute. No sponsor or funding source had a role in the design or conduct of the study; collection, management, analysis or interpretation of the data; or preparation, review or approval of the manuscript.

## References

1. Cawley J, Meyerhoefer C. The medical care costs of obesity: an instrumental variables approach. *J Health Econ.* 2012; 31(1):219–30. [PubMed: 22094013]
2. Ogden, C., Carroll, M., Fryar, C., Flegal, K. Prevalence of Obesity Among Adults and Youth: United States, 2011–2014. Hyattsville, MD: National Center for Health Statistics; 2015.
3. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ.* 2015; 350:h2147. [PubMed: 25956159]
4. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care.* 2012; 50(3):217–26. [PubMed: 22310560]
5. Concannon TW, Guise JM, Dolor RJ, Meissner P, Tunis S, Krishnan JA, et al. A national strategy to develop pragmatic clinical trials infrastructure. *Clin Transl Sci.* 2014; 7(2):164–71. [PubMed: 24472114]
6. Wolfenden L, Wiggers J. Strengthening the rigour of population-wide, community-based obesity prevention evaluations. *Public health nutrition.* 2014; 17(02):407–21. [PubMed: 23157970]
7. Dalton WT 3rd, Wang L, Southerland JL, Schetzina KE, Slawson DL. Self-reported versus actual weight and height data contribute to different weight misperception classifications. *Southern medical journal.* 2014; 107(6):348–55. [PubMed: 24945166]
8. Murphy JK, Bruce BK, Williamson DA. A comparison of measured and self-reported weights in a 4-year follow-up of spouse involvement in obesity treatment. *Behavior Therapy.* 1985; 16(5):524–30.
9. Waring ME, Roberts MB, Parker DR, Eaton CB. Documentation and management of overweight and obesity in primary care. *J Am Board Fam Med.* 2009; 22(5):544–52. [PubMed: 19734401]
10. Ma J, Xiao L, Stafford RS. Adult obesity and office-based quality of care in the United States. *Obesity (Silver Spring).* 2009; 17(5):1077–85. [PubMed: 19197264]
11. Kraschnewski JL, Sciamanna CN, Pollak KI, Stuckey HL, Sherwood NE. The epidemiology of weight counseling for adults in the United States: a case of positive deviance. *Int J Obes.* 2013; 37(5):751–3.
12. Stevens VJ, Wagner EL, Rossner J, Craddick S, Greenlick MR. Validity and usefulness of medical chart weights in the long-term evaluation of weight loss programs. *Addict Behav.* 1988; 13(2): 171–5. [PubMed: 3369325]
13. Arterburn D, Ichikawa L, Ludman EJ, Operskalski B, Linde JA, Anderson E, et al. Validity of Clinical Body Weight Measures as Substitutes for Missing Data in a Randomized Trial. *Obes Res Clin Pract.* 2008; 2(4):277–81. [PubMed: 19956347]



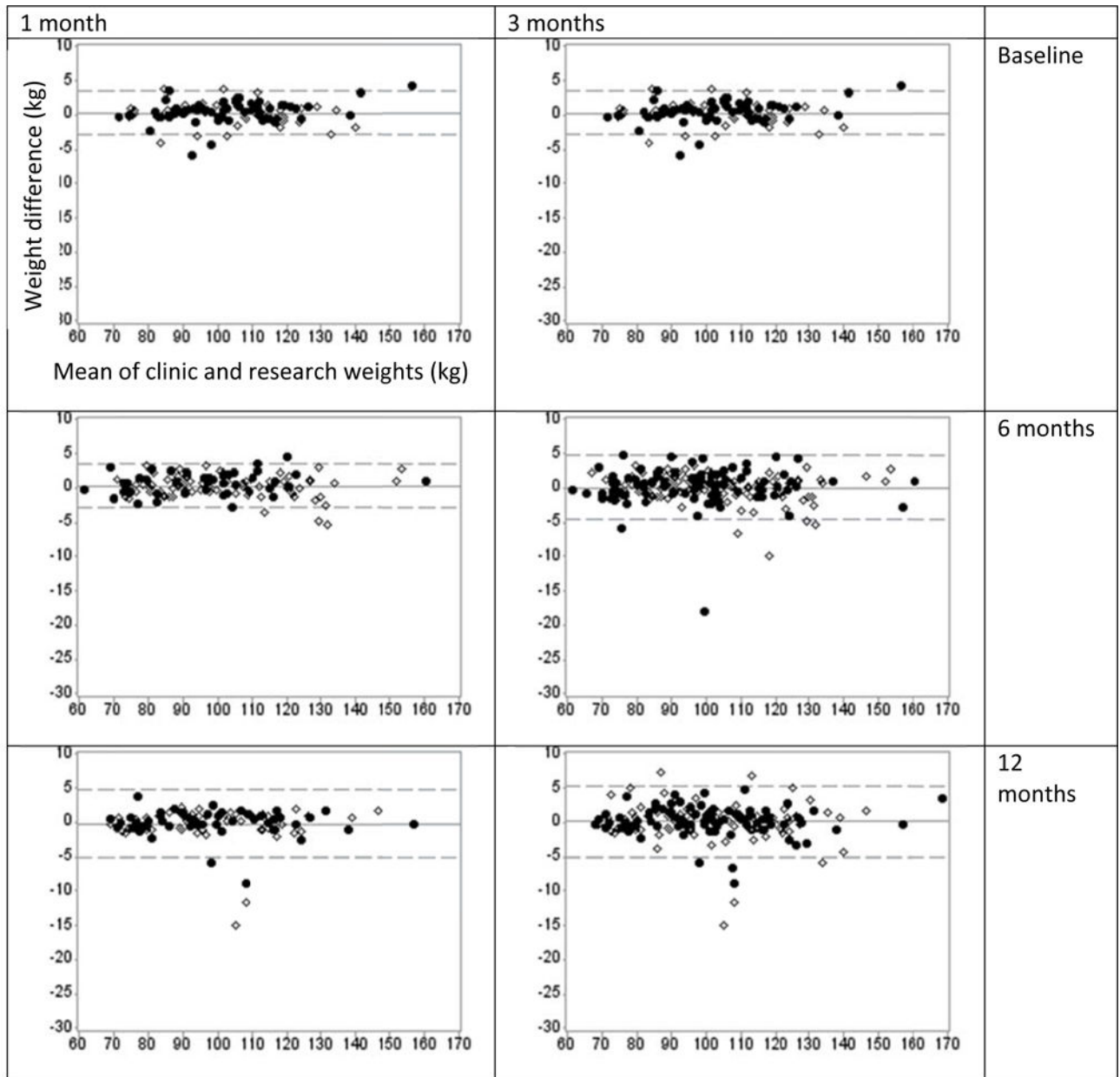
14. Leo MC, Lindberg NM, Vesco KK, Stevens VJ. Validity of medical chart weights and heights for obese pregnant women. *EGEMS (Wash DC)*. 2014; 2(1):1051. [PubMed: 25848585]
15. Ma J, Yank V, Xiao L, Lavori PW, Wilson SR, Rosas LG, et al. Translating the Diabetes Prevention Program lifestyle intervention for weight loss into primary care: a randomized trial. *JAMA Intern Med*. 2013; 173(2):113–21. [PubMed: 23229846]
16. Ma J, King AC, Wilson SR, Xiao L, Stafford RS. Evaluation of lifestyle interventions to treat elevated cardiometabolic risk in primary care (E-LITE): a randomized controlled trial. *BMC Fam Pract*. 2009; 10:71. [PubMed: 19909549]
17. Ma J, Strub P, Xiao L, Lavori PW, Camargo CA Jr, Wilson SR, et al. Behavioral weight loss and physical activity intervention in obese adults with asthma. A randomized trial. *Ann Am Thorac Soc*. 2015; 12(1):1–11. [PubMed: 25496399]
18. Ma J, Strub P, Camargo CA Jr, Xiao L, Ayala E, Gardner CD, et al. The Breathe Easier through Weight Loss Lifestyle (BE WELL) Intervention: a randomized controlled trial. *BMC Pulm Med*. 2010; 10:16. [PubMed: 20334686]
19. Lawrence IKL. A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics*. 1989; 45(1):255–68. [PubMed: 2720055]
20. Altman DG, Bland JM. Measurement in Medicine: The Analysis of Method Comparison Studies. *Journal of the Royal Statistical Society Series D (The Statistician)*. 1983; 32(3):307–17.
21. Stevens J, Truesdale KP, McClain JE, Cai J. The definition of weight maintenance. *Int J Obes (Lond)*. 2006; 30(3):391–9. [PubMed: 16302013]
22. Digital Data Improvement Priorities for Continuous Learning in Health and Health Care: Workshop Summary. Washington (DC): 2013.

**What is already known about this subject**

- A critical barrier to pragmatic research on obesity prevention and control is the high cost of in-person weight data collection both in terms of research staff time and participant burden.
- Authorized abstraction of clinically measured weights from patients' electronic health records (EHR) is one promising solution considering the relative low cost of EHR data abstraction and the ability to eliminate patient burden as a result of separate data collection for research purposes only.
- Prior studies showed good agreement between clinic and research weight measurements at one point in time, but we could not identify any papers that examined agreement using repeated measures of clinic and research weights over time or validity in inferences about weight loss intervention effects.

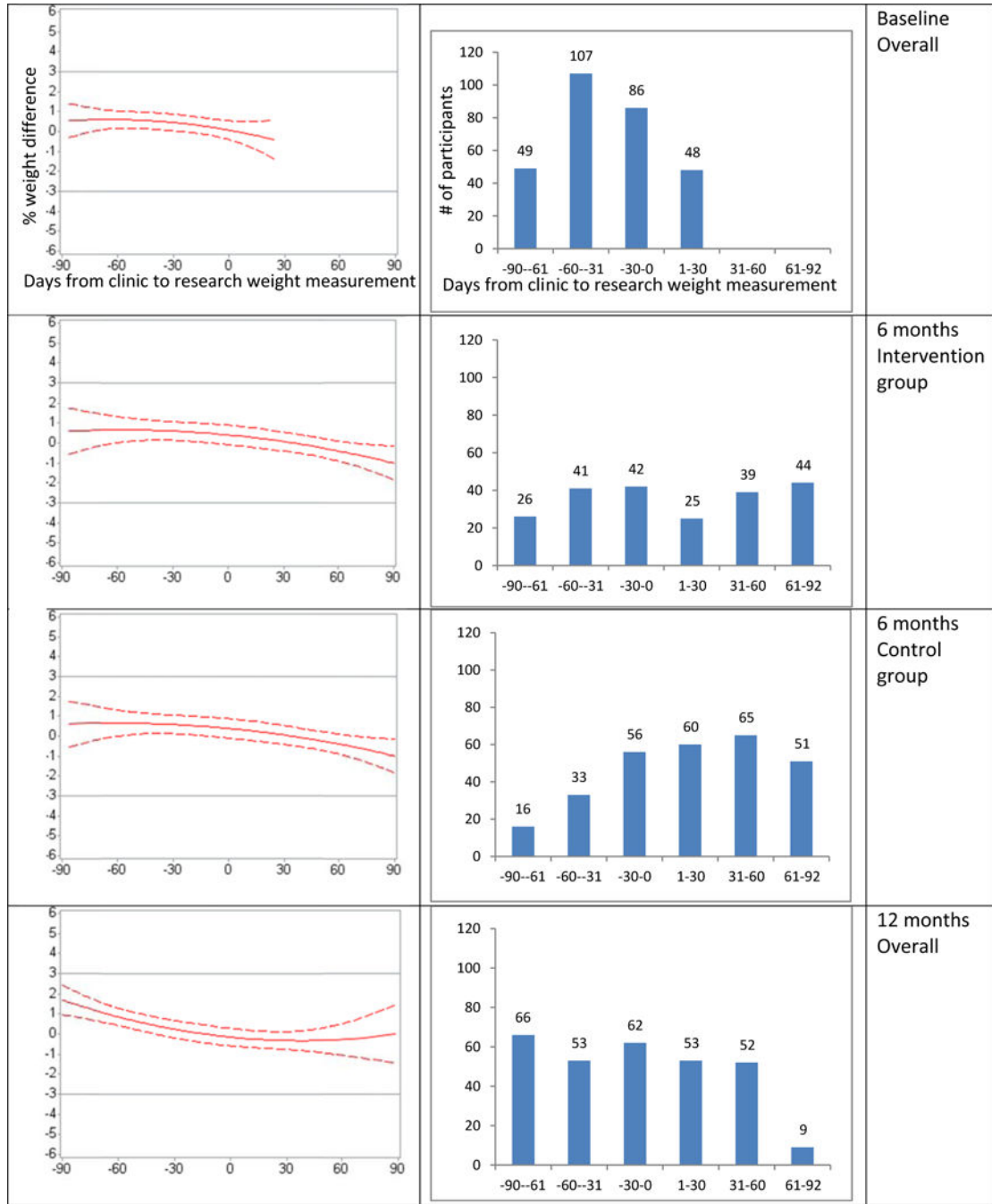
**What this study adds**

- We demonstrated high agreement between research weights according to standardized measurement protocols and close-in-time weight data (up to 3 months) abstracted from EHR.
- We also showed statistical inference validity in evaluating the treatment effect on weight loss trajectories over time using clinic weights vs. research weights.
- Clinic weights can be a valid pragmatic outcome measure in weight loss studies, which may significantly reduce participant burden and research costs associated with direct research weight measurement and improve recruitment and retention rates in obesity research studies.



● Intervention ◊ Control

**Figure 1.** Bland and Altman’s limits of agreement plots. The plots illustrate the differences between clinic and research weights (y-axis, in kg) and the lower and upper bounds for the 95% limits of agreement (horizontal dashed lines, mean of differences  $\pm$  1.96 standard deviation) against the means of these measurements (x-axis, in kg) in the BE WELL trial. Two index windows, 1 and 3 months, were applied to capture the closest clinic weight in time in comparison to research weights measured at baseline, 6 and 12 months.



**Figure 2.** Polynomial regression model results. The graphs show the predicted percentage weight differences, (clinic weight-research weight)/research weight, in kg, and the 95% CIs (left) or the numbers of participants with available clinic and research weight (right) versus days between clinic and research weight measurements by study time point. The coefficients for the treatment group-by-days interaction term were not statistically significant at baseline and 12 months. Thus, separate regression lines by treatment group are shown at 6 months only.

**Table 1**

Baseline characteristics of participants in the BE WELL and E-LITE trials

<b>Characteristic</b>	<b>BE WELL n=330</b>	<b>E-LITE n=241</b>
Age, years, mean $\pm$ SD	47.6 $\pm$ 12.4	52.9 $\pm$ 10.6
Female, %	70.6	46.5
Race/ethnicity, %		
Non-Hispanic White	49.7	78.0
Non-Hispanic Black	20.0	0.4
Asian/Pacific Islander	8.2	17.0
Latino/Hispanic	20.3	4.1
Education, %		
High school graduate or less	16.2	2.5
Some college	41.6	15.1
College level or above	42.2	72.4
Weight, kg, mean $\pm$ SD	104.2 $\pm$ 19.6	93.8 $\pm$ 17.7
Body mass index, kg/m <sup>2</sup> , mean $\pm$ SD	37.5 $\pm$ 5.9	32.0 $\pm$ 5.4

SD, standard deviation.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

Availability of clinic weights during the electronic health record abstraction period<sup>\*</sup>, by trial and treatment group

	Participants with clinic weights %	# of clinic weights per participant mean $\pm$ SD
<b>BE WELL, 12 months follow-up<sup>*</sup></b>		
Coach-led intervention (n=165)	95%	5.8 $\pm$ 5.5
Usual care control (n=165)	96%	5.8 $\pm$ 6.7
All participants (n=330)	96%	5.8 $\pm$ 6.1
<b>E-LITE, 15 months follow-up<sup>*</sup></b>		
Self-directed intervention (n=81)	86%	2.8 $\pm$ 2.1 <sup>†</sup>
Coach-led intervention (n=79)	90%	4.0 $\pm$ 4.8
Usual care control (n=81)	95%	4.2 $\pm$ 4.2 <sup>†</sup>
All participants (n=241)	90%	3.7 $\pm$ 3.9

SD, standard deviation.

<sup>\*</sup> EHR abstraction period was defined as the time frame of participants from 3 months before randomization to 3 months after the primary endpoint of each trial.

<sup>†</sup> Indicates statistically significant difference ( $p < 0.05$ ) between the self-directed intervention and usual care control groups in the E-LITE trial.

**Table 3**

Differences in slope of change based on clinic and research weights over the course of follow-up in the BE WELL and E-LITE trials

	Clinic weights slope $\pm$ SE	Research weights slope $\pm$ SE	Difference in slope (95% CI)	P-value
<b>BE WELL, 12 months follow-up</b>				
Coach-led intervention (n=165)	-0.10 $\pm$ 0.02	-0.07 $\pm$ 0.03	-0.03 (-0.09, 0.04)	0.41
Usual care control (n=165)	0.03 $\pm$ 0.02	-0.01 $\pm$ 0.03	0.03 (-0.04, 0.09)	0.43
<b>E-LITE, 15 months follow-up</b>				
Self-directed intervention (n=81)	-0.04 $\pm$ 0.03	-0.10 $\pm$ 0.03	0.06 (-0.03, 0.15)	0.17
Coach-led intervention (n=79)	-0.10 $\pm$ 0.03	-0.11 $\pm$ 0.03	0.01 (-0.07, 0.08)	0.85
Usual care control (n=81)	0.01 $\pm$ 0.03	0.02 $\pm$ 0.03	-0.01 (-0.09, 0.06)	0.74

SE, standard error; CI, confidence interval.