



# HHS Public Access

Author manuscript

*Obesity (Silver Spring)*. Author manuscript; available in PMC 2021 November 01.

Published in final edited form as:

*Obesity (Silver Spring)*. 2020 November ; 28(11): 2107–2115. doi:10.1002/oby.22958.

## Can Electronic Health Records Validly Estimate the Effects of Health System Interventions Aimed at Controlling Body Weight?

John A. Gallis, ScM<sup>1,2</sup>, Kristie Kusibab, MSc<sup>3,\*</sup>, Joseph R. Egger, PhD<sup>2</sup>, Maren K. Olsen, PhD<sup>1,4</sup>, Sandy Askew, MPH<sup>2,5</sup>, Dori M. Steinberg, PhD<sup>2,5,6</sup>, Gary G. Bennett, PhD<sup>2,5,7,\*</sup>

\*During the study, Ms. Kusibab was a Master of Science student in the Department of Biostatistics & Bioinformatics at Duke University.

<sup>1</sup>Department of Biostatistics & Bioinformatics, Duke University, Durham, NC, United States

<sup>2</sup>Duke Global Health Institute, Duke University, Durham, NC, United States

<sup>3</sup>PharPoint Research, Inc., Durham, NC, United States

<sup>4</sup>Center for Health Services Research in Primary Care, Durham VA Medical Center, Durham, NC, United States

<sup>5</sup>Duke Global Digital Health Science Center, Duke University, Durham, NC, United States

<sup>6</sup>Duke School of Nursing, Duke University, Durham, NC, United States

<sup>7</sup>Department of Psychology and Neuroscience, Duke University, Durham, NC, United States

### Abstract

**Objective:** To compare weight collected at clinics and recorded in the electronic health record (EHR) to primary study-collected trial weights to assess the validity of utilizing EHR data in future pragmatic weight loss or weight gain prevention trials.

**Methods:** For both the Track and Shape obesity intervention randomized trials, we compared clinic EHR weight data to primary trial weight data over the same time period. In analyzing the EHR weights, we estimated intervention effects on the primary outcome of weight (in kilograms) with EHR data, using linear mixed effects models.

---

\* **Corresponding Author Contact Info:** Gary G. Bennett, gary.bennett@duke.edu; 919-668-3420; 116 Allen Building, Box 90024, Durham NC 27708.

**Author Contributions:** Dr. Bennett had full access to all the data in the study and takes responsibility for the integrity of the data and accuracy of the data analysis. Concept and design: GB, DS. Acquisition, analysis, or interpretation of data: All authors. Drafting of the manuscript: JG, JE. Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: JG, JE, MO. Obtained funding: GB

**Clinical Trial Registration:** The Track Trial is registered at [www.clinicaltrials.gov](http://www.clinicaltrials.gov) NCT01827800. The Shape Trial is registered at [www.clinicaltrials.gov](http://www.clinicaltrials.gov) NCT00938535.

**Data Sharing:** Deidentified data related specifically to the analyses described in this paper will be shared upon reasonable request with researchers who provide a methodologically sound proposal for analyzing the data. Proposals should be directed to gary.bennett@duke.edu.

**Disclosure:** Dr. Bennett is currently on the scientific advisory board of WW (Weight Watchers) and has equity in Coeus Health. Dr. Steinberg is on the medical advisory board for Omada Health. No other disclosures were reported.

**Results:** EHR weight measurements were higher on average and more variable than trial weight measurements. The mean differences and 95% confidence intervals are similar at all time points between the estimates using EHR and study-collected weights.

**Conclusions:** The results of this study can be used to help guide the planning of future pragmatic weight-related trials. This study provides evidence that body weight measurement abstracted from EHRs can provide valid, efficient, and cost-effective data to estimate treatment effects from randomized clinical weight loss and weight management trials. However, care should be made to properly understand the data generating process and any mechanisms that may affect the validity of these estimates.

### Keywords

Electronic health records (EHR); pragmatic trials; comparative effectiveness research; obesity

---

## INTRODUCTION

Electronic health record (EHR) data are increasingly used in comparative effectiveness research, particularly with pragmatic trial designs.<sup>1-4</sup> Electronic health record data are used to enroll and characterize study participants, as well as to assess study endpoints.<sup>2,3,5</sup> However, considerable debate remains regarding the ability of EHR data to validly estimate intervention effects on trial outcomes, due to issues of misclassification and sparse, irregularly measured data.<sup>3,6</sup> Body weight is a commonly collected data point measured in clinical settings, however, there is little empirical research assessing the validity of body weights measured in the EHR.<sup>7,8</sup>

Using passively-collected EHR data in prospective studies of weight loss interventions can reduce research costs and lower participant burden, potentially influencing the number and quality of pragmatic obesity trials. Obesity-related pragmatic trials that rely solely on the EHR to collect outcome data are beginning to emerge.<sup>1</sup> However, before valid inference can be drawn from results of these studies, we need additional empirical evidence supporting the use of routinely collected body weight ascertained from the EHR for helping to plan and for validly assessing the effectiveness of obesity treatment trials.

The purpose of this paper is multi-fold. First, we compare intervention effects estimated using primary study-collected data from two randomized-controlled trials (RCTs)—one assessing a health system-based digital obesity weight loss intervention<sup>9,10</sup> (the “Track” study), the other an interactive obesity prevention approach<sup>11,12</sup> (the “Shape” study)—to intervention effects estimated using EHR-derived body weight data on the same individuals measured at routine clinical visits over the same study period. A comparison of outcome measures allows us to assess the validity of utilizing EHR data when it is infeasible or inappropriate to collect primary data on study participants. Second, we use these EHR data of body weight in the Track study to examine the impact of the intervention six months after the intervention ended, to assess durability effects of the intervention solely with EHR data. Finally, we compare variability and measurement frequency between the two data sources to help inform study planning and sample size estimation for obesity-related pragmatic trials.

## METHODS

Data come from two separate trials, conducted in the community health center systems of Piedmont Health Services (PHS), Inc., serving central North Carolina. Design characteristics have been described in detail elsewhere.<sup>9,10</sup> In brief, the Track trial was an individually-randomized two-arm parallel-group RCT of a racially/ethnically diverse study population of 351 adults. Track included a 12-month weight loss intervention for individuals with obesity (BMI: 30.0–44.9 kg/m<sup>2</sup>) diagnosed with hypertension, diabetes, and/or hyperlipidemia. Additional inclusion and exclusion criteria are described in the protocol paper.<sup>9</sup> The trial was conducted in four community health centers within the PHS network. Participants were allocated to intervention or usual care in a 1:1 ratio using a minimization algorithm accounting for community health center, gender, and ethnicity in order to balance these characteristics across arms. Participants randomized to the intervention arm received several components of an intervention, which included behavioral goals, self-weighing, and counseling; while the participants randomized to usual care received the standard care currently offered by their healthcare providers.<sup>9</sup> Participants were followed for 12 months, and primary trial data were collected at 6- and 12-month endpoints. Data were collected between the years 2013 and 2015. The Track study received ethical approval from the Duke University IRB (#2017–1282).

The Shape trial was an individually-randomized two-arm parallel-group RCT with 185 premenopausal black women with overweight or class 1 obesity (25–34.9 kg/m<sup>2</sup>).<sup>11,12</sup> Patients were recruited from six community health centers within PHS. Study participants were randomly allocated equally (1:1) to one of the two treatment arms. The length of the intervention was 12 months and focused on implementing behavior change goals that would lead to prevention of weight gain.<sup>11</sup> Those in the usual care arm received general wellness newsletters every six months, but otherwise were given standard care at each visit. Participants were followed for 18 months, and primary trial data were collected at 6-, 12-, and 18-month post-randomization endpoints. Data were collected between the years 2009 and 2012. Notably, while the goal of the Track intervention was weight loss, in Shape the goal was weight gain prevention. Additional study details are described in the protocol paper and main outcomes paper.<sup>11,12</sup> The Shape study received ethical approval from the Duke University IRB (#2017–1169).

In both trials, data were obtained from Piedmont Health's electronic medical record on eligible participants who consented to EHR data abstraction. In the Track study, these EHR data were obtained from between 12 months prior to enrollment and 24 months post-enrollment. Of the 337/351 (96%) eligible participants from the primary trial, 307 (91%; 154 intervention, 153 usual care) both consented to EHR data extraction (n=24 did not consent) and had weight records in the EHR within the study period (n=6 consented but did not have EHR weight records within the 24-month period; Figure S1). Of the 3,881 visits recorded in the EHR across the 337 participants, only 61 (1.6%) visits across 41 participants did not include a weight measurement. In the Shape study, the EHR data were obtained on participants from enrollment to 24 months post-enrollment. Of the 185 eligible participants in the primary trial, 139 (75%) both consented to EHR data abstraction (n=15 did not consent) and had weight records in the EHR within the study period (n=31

consented but did not have EHR weight records within the 24-month period; Figure S2). Of the 778 visits recorded in the EHR across the 139 participants, only 4 (0.5%) visits across 3 participants did not include a weight measurement. Of note, research staff calibrated scales at the clinics and trained health system staff at the start of both trials.

The primary outcome for both studies is change in body weight, in kilograms; however, the goal in Track was weight loss whereas the goal in Shape was weight maintenance. Although the EHR data extend to 24 months post-randomization in both trials, there are relatively few (<50%) participants with weight data collected in the EHR after 18 months, causing estimates after 18 months to be imprecise and potentially invalid. In both studies separately, we estimated and compared the mean weight at baseline and across time in the EHR data and trial data. We defined baseline in the EHR data as within 30 days of the baseline trial visit (includes 30 days before and after in Track and 30 days after in Shape). To estimate the variability around visit timing, we also calculated—overall and by intervention arm—the number and range of visits per participant, the mean time between visits for both EHR and trial data, as well as the percentage of participants with only one visit recorded in the EHR during the EHR abstraction period.

In both studies, we estimated intervention effects on the primary outcome with EHR data, using a constrained<sup>13</sup> linear mixed effects model, with non-linearity in the weight change trajectories estimated using either quadratic terms or linear splines. The number and location of knots for linear splines were determined using a combination of theory and fit criteria. Fit criteria (BIC) using Shape data indicated that a quadratic term captured the non-linearity well, whereas in the Track data linear splines with knots at 6, 12, and 18 months fit the data well. Random effects parameters included a participant-level random intercept and a random linear and quadratic slope (Shape) or linear slope for each segment of the spline curve (Track). From these models, we computed the estimated mean weight change by trial arm and the intervention effect at 6, 12, and 18 months in both trials. The Track analysis was adjusted for health center, gender, and race/ethnicity to account for the minimization randomization design. The Shape analysis was unadjusted, to match the main study analyses. We compared these results to the original trial findings by extracting the reported regression results directly from the Track and Shape main trial papers.<sup>10,12</sup> For both studies, to account for potential selection bias, we compared those with EHR data to those who declined EHR data extraction, and adjusted for any baseline variables indicated as differential ( $p < 0.10$ ) by EHR extraction status as a sensitivity analysis.

All analyses were conducted in Stata version 16.1 (StataCorp, College Station, TX).

## RESULTS

Baseline characteristics of the participants who consented to EHR data extraction, stratified by intervention arm, are shown in Table 1 for Track and Table 2 for Shape, with more details on the study populations available in the associated main outcomes papers.<sup>10,12</sup> The participants who consented to EHR data extraction in Track were 70% female and an average (standard deviation [SD]) age of 50.5 (9.0) years, with 30% below the poverty line. The average body mass index (BMI) at enrollment was approximately 36 kg/m<sup>2</sup>. In Shape,

which consisted entirely of black women, the average (SD) age was 35.7 (5.6) years, with 35% below the poverty line. The average BMI at enrollment was approximately 31 kg/m<sup>2</sup>.

Table 3 displays the comparison of EHR and study-collected trial data on weight and visit timing for both Track and Shape. In Track, the mean (SD) weight across all time points was 98.0 (14.9) kg in the trial data and 98.9 (15.4) kg in the EHR data. In Shape, the mean (SD) weight across all time points was 81.1 (9.45) kg in the trial data and 83.2 (10.6) kg in the EHR data. Thus, in both cases, the average EHR weight was larger than the average trial-collected weight (mean [SE] weight difference 0.89 [0.61] kg in Track and 2.18 [0.47] kg in Shape), and the EHR data also had greater variability.

We also defined a 30-day window around study baseline visit in both studies (30 days before and after in Track and 30 days after in Shape) to compare baseline weight measurements. In Track, 137/307 (44.6%) participants had weight measurements in this window. The mean (SD) weight at baseline among these participants was 99.2 (13.4) kg in the trial data and 100.3 (13.6) kg in the EHR data. In Shape, 48/139 (34.5%) had weight measurements in the window. The mean (SD) weight at baseline among these participants was 80.6 (8.2) kg in the trial data and 82.4 (8.4) kg in the EHR data. In both cases, the baseline comparisons also showed higher average weight and greater variability in the EHR data versus the trial data, although it is important to note that in both cases this result is based on <50% of those with EHR data.

In both studies, the trial visits were scheduled to be six months (i.e., 180 days) apart. In Track, the average (SD) days between any two visits in the trial data was 186.6 (28.4) days and 86.0 (84.9) days in the EHR data (Table 3). 75.6% had visits recorded in the EHR within the first 6 months, 71.7% in the 6 to 12 month time frame, 66.8% in the 12 to 18 month time frame, and 97.7% had more than one visit. Table 4 displays summaries of the number of visits across intervention arms, both for the entire study period as well as in 6 month windows around the primary time points. The median number of visits in the EHR across the 24 months post-baseline was 8.0 (Q1, Q3 = 6.0, 12.0) with a minimum of 1.0 and a maximum of 30.0. The mean and median number of visits does not significantly differ by intervention arm, either across the entire EHR abstraction period or in the windows around the primary time points (Table 4).

Similarly, in Shape, the average (SD) days between any two visits in the trial data was 188.8 (30.6) days and 110.0 (142.7) days in the EHR data (Table 3). 77.0% had visits recorded in the EHR within the first 6 months, 66.9% in the 6 to 12 month time frame, 60.4% in the 12 to 18 month time frame, and 87.1% had more than one visit. The median number of visits in the EHR across the 24 months was 4.0 (Q1, Q3 = 2.0, 8.0) with a minimum of 1.0 and a maximum of 23.0. The mean and median number of visits does not significantly differ by intervention arm, either across the entire EHR abstraction period or in the windows around the primary time points (Table 4).

For Track, mean weight change was estimated separately for both EHR and trial data at 6 and 12 months; and 18 months using EHR data only (Figure 1, Panel A). The average treatment effect is reported as a mean difference with 95% confidence interval (CI) using

an intent-to-treat analysis, and is reported separately using trial data and EHR data (Figure 1, Panel B). Comparing the EHR- and trial estimates, the mean differences and 95% CIs are similar at both 6 and 12 months (Table 5), and lead to the same conclusion regarding intervention effectiveness. Estimates of a durability effect of the intervention using only EHR data suggest a continued effect of the intervention at 18-months post-enrollment (mean difference =  $-2.8$  kg, 95% CI:  $-4.5, -1.1$ ). The sensitivity analysis, additionally adjusting for insurance status and total cholesterol, did not significantly alter the results (18-month mean difference =  $-2.9$  kg, 95% CI:  $-4.6, -1.1$ ).

For Shape, mean weight change was estimated separately for both EHR and trial data at 6, 12, and 18 months (Figure 2, Panel A). The average treatment effect is reported as a mean difference with 95% confidence interval (CI) using an intent-to-treat analysis, and is reported separately using trial data and EHR data (Figure 2, Panel B). Comparing the EHR- and trial estimates, the mean differences and 95% CIs are similar at 6, 12, and 18 months (Table 5), and lead to similar conclusions regarding intervention effectiveness. The sensitivity analysis, additionally adjusting for waist circumference, glucose, education level, and income, did not significantly alter the results, although the EHR effect estimates became larger (Table 5).

## DISCUSSION

Using EHR data, we were able to validly and precisely reproduce the original study trial treatment effect of two separate obesity interventions. Our analyses utilized linear mixed effects models which can readily handle longitudinal data where participants have a varying number of measurements at varying intervals, a characteristic and challenge of EHR-ascertained weights. For each study, the treatment effects over time were estimated by determining the best fitting functional form via both fixed and random effects components. The two trials were conducted in different populations in central North Carolina ranging from 2009 through 2015, showing some durability to temporal impacts. Our analytic approach and results have positive implications for use of EHR weights in obesity-related clinical trials at all stages on the translational continuum, both to estimate treatment effects at primary endpoints, as well as to estimate intervention durability effects past the primary endpoint.

The results of this paper can also be used to help guide the planning and sample size calculation for future pragmatic weight-related trials. Information on timing and availability of EHR measurements, as well as the higher average weight and variability in the EHR data versus the trial data, can provide crucial information for sample size calculations. Research trials frequently employ stronger weight measurement protocols than health care providers. Both Track and Shape employed protocols intended to optimize precision of weight measurements. Participants were asked to change into hospital gowns and paper booties prior to their weight assessments, and to use the restroom, if needed. Additionally, in order to collect fasting blood glucose, participants in both trials were asked to follow a fasting protocol prior to their study visits, with one exception of the 6-month Track visit. We found that EHR weights, on average, were heavier and more variable than protocolized, study-collected weights for both trials; however, this measurement bias was not found to

be differential by treatment arm. Furthermore, we observed differences in the timing and frequency of EHR weight measurements across the study populations. Researchers planning pragmatic weight loss or weight gain prevention trials that will solely utilize EHR data should expect higher variability in weight—leading to a larger sample size needed for the same amount of statistical power as a traditional RCT with trial-collected weights—and note that less healthy populations will likely have more frequent visits to their healthcare providers.

Researchers should also abstract EHR data both before and after the trial study period to avoid issues with validity and precision of the treatment effects. Specifically, intervention effect estimates based on EHR data are less precise the fewer measurements there are near the time point of interest. Thus, in planning pragmatic trials, it is important that investigators extract EHR data several months after the final endpoint of interest. If EHR data collection is ended right at this endpoint, clinical visits near but after the endpoint will be missed. Furthermore, our results showed stronger concordance in treatment effects comparing EHR data with trial data for the Track study than for the Shape study. This was likely due to our lack of pre-baseline EHR data in the Shape study. As a result of this, our model-predicted ‘baseline’ weight value was not a true baseline weight and was affected, to some degree, by the intervention (e.g., if the intervention successfully reduced body weight, this would result in a downward bias in the ‘baseline’ weight prediction). This highlights the need for collection of EHR data prior to the study enrollment period to validly estimate baseline body weight.

A prior study by Xiao and colleagues conducted similar longitudinal analyses on study populations that were primarily white women with obesity receiving care at health systems in the San Francisco Bay area of California.<sup>8</sup> They also found high levels of agreement between study and EHR weights measured closely in time and between estimated slopes of weight change between the two modes of measurement. A similar in-progress analysis among weight loss trial and EHR data among veterans in Durham, North Carolina also shows high levels of agreement between study and EHR weights longitudinally.<sup>14</sup> As noted in Xiao and colleagues,<sup>8</sup> three other studies have found good concordance between EHR and study data at one time point,<sup>15–17</sup> but to our knowledge there are no additional published studies that examine concordance of weight measurements between EHR and study data longitudinally, although similar concordance studies have been performed for other outcomes, such as chronic obstructive pulmonary disease exacerbation episodes.<sup>18</sup> In many obesity trials the primary outcome is either achieving a certain percentage weight loss (e.g., 3%) or preventing a certain percentage weight gain. While this is generally straightforward to determine for study-collected weights, which are collected approximately at the time point of interest, it is less straightforward for EHR weights. An interesting area of future research would be to explore methods of validly and reliably estimating such binary weight loss or weight gain prevention outcomes using EHR data.

There are several limitations of using EHR weights as measurements for RCTs. First, because we do not have observed weights at common time points, all treatment effects at specific points in time must be model estimated. These predicted weight estimates may be sensitive to model selection. Second, variations in equipment and calibration protocols

can contribute to variability in EHR data, particularly when data is drawn from multiple sites. Third, there is no standardized protocol for measuring weights. Depending on the visit day and the healthcare professional, a patient may or may not be asked to remove his or her shoes and jacket when having weight measured. This lack of standardized protocol is, in part, what leads to higher variability in EHR weights compared to trial weights. For example, our results showed EHR-measured body weight was routinely estimated to be higher and more variable on average compared to trial body weight. As noted previously, this may lead to a larger sample of participants needed in an RCT using EHR weights versus an RCT using study-collected weights for the same level of statistical power. To induce a bias in the intervention effect, the amount the EHR measurement overestimates the study-collected weight would need to be differential (e.g., informative) by intervention assignment. In fact, there is evidence of slightly higher overestimation in control than intervention in Shape (2.7 kg vs. 1.7 kg), which may have contributed to some of the differences in estimates obtained comparing the EHR and trial data. Generally, however, we do not expect EHR weight mismeasurement to be significantly differential by treatment arm with baseline randomization of study participants. Fourth, in trials using EHR data only, there may be a higher rate of clinical encounters among heavier participants or less healthy participants. However, unless this is differential by intervention arm (e.g., the intervention leads to health improvements resulting in fewer doctor visits in the intervention arm), it will not bias the intervention effect estimate.<sup>19</sup> Finally, in the Shape and Track trials, the patient populations were relatively stable, with little movement of participants out of the health system or study area. This likely improved the validity of the EHR data to estimate treatment effects at both primary endpoints, as well as in durability effects. In populations with a larger movement of patients out of the health system, use of EHR data may negatively affect both the precision and validity of these treatment estimates if the sparse data mechanism is not completely random.

There is growing interest in utilizing EHR data to plan and evaluate clinical interventions.<sup>2,4,5</sup> Pragmatic trials of weight loss or weight gain prevention using EHR data should report information on timing and frequency of visits, as well as the variability of EHR weights across time, in order to provide crucial information for researchers planning future pragmatic trials. Because body weight is routinely and reliably collected during clinical encounters in the United States, use of these data from the EHR is attractive for evaluating weight loss and weight management trials. Additionally, its collection should generally lead to relatively little bias in treatment effect estimates in RCTs with weight as the outcome.<sup>19</sup> The present study provides evidence that body weight measurement abstracted from EHRs can provide valid, efficient, and cost-effective data to estimate treatment effects from randomized clinical weight loss and weight management trials. However, care should be made to properly understand the data generating process and any mechanisms that may affect the validity of these estimates.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.



## Acknowledgements

We express deep gratitude to the administration and staff of Piedmont Health for their collaboration and participation in both the Track and Shape trials. We would also like to especially thank the men and women who participated in the trials.

**Funding:** The Track and Shape studies were funded by the National Institute of Diabetes and Digestive and Kidney Diseases (Track: R01DK093829; Shape: R01DK078798). Work on this paper also helped in planning for the analysis of the Balance pragmatic trial, funded by the National Institute of Diabetes and Digestive and Kidney Diseases (R01DK109518). The research presented in this paper is that of the authors and does not reflect the official policy of the NIH.

## REFERENCES

1. Berger MB, Steinberg DM, Askew S, et al. The Balance protocol: a pragmatic weight gain prevention randomized controlled trial for medically vulnerable patients within primary care. *BMC Public Health* 2019;19(1):596. [PubMed: 31101037]
2. University of Pennsylvania. Electronic health records (EHR) in randomized clinical trials: challenges and opportunities Paper presented at: 12th Annual Conference on Statistical Issues in Clinical Trials2019; Philadelphia, PA.
3. Hersh WR, Weiner MG, Embi PJ, et al. Caveats for the use of operational electronic health record data in comparative effectiveness research. *Med Care* 2013;51(8 Suppl 3):S30–S37. [PubMed: 23774517]
4. Rockhold FW, Goldstein BA. Pragmatic Randomized Trials Using Claims or Electronic Health Record Data. In: Piantadosi S, Meinert CL, eds. *Principles and Practice of Clinical Trials* Switzerland: Springer International Publishing; 2019.
5. Bhasin S, Gill TM, Reuben DB, et al. Strategies to Reduce Injuries and Develop Confidence in Elders (STRIDE): a cluster-randomized pragmatic trial of a multifactorial fall injury prevention strategy: design and methods. *The Journals of Gerontology: Series A* 2018;73(8):1053–1061.
6. Rockhold FW, Tenenbaum JD, Richesson R, Marsolo KA, O'Brien EC. Design and analytic considerations for using patient-reported health data in pragmatic clinical trials: report from an NIH Collaboratory roundtable. *J Am Med Inform Assoc* 2020.
7. Concannon TW, Guise J-M, Dolor RJ, et al. A National Strategy to Develop Pragmatic Clinical Trials Infrastructure. *Clin Transl Sci* 2014;7(2):164–171. [PubMed: 24472114]
8. Xiao L, Lv N, Rosas LG, Au D, Ma J. Validation of clinic weights from electronic health records against standardized weight measurements in weight loss trials. *Obesity* 2017;25(2):363–369. [PubMed: 28059466]
9. Foley P, Steinberg D, Levine E, et al. Track: a randomized controlled trial of a digital health obesity treatment intervention for medically vulnerable primary care patients. *Contemp Clin Trials* 2016;48:12–20. [PubMed: 26995281]
10. Bennett GG, Steinberg D, Askew S, et al. Effectiveness of an app and provider counseling for obesity treatment in primary care. *Am J Prev Med* 2018;55(6):777–786. [PubMed: 30361140]
11. Foley P, Levine E, Askew S, et al. Weight gain prevention among black women in the rural community health center setting: the Shape Program. *BMC Public Health* 2012;12(1):305. [PubMed: 22537222]
12. Bennett GG, Foley P, Levine E, et al. Behavioral treatment for weight gain prevention among black women in primary care practice: A randomized clinical trial. *JAMA internal medicine* 2013;173(19):1770–1777. [PubMed: 23979005]
13. Lu K. On Efficiency of Constrained Longitudinal Data Analysis Versus Longitudinal Analysis of Covariance. *Biometrics* 2010;66(3):891–896. [PubMed: 19764951]
14. Olsen MK, Voils CI. Understanding Selection bias and Missing-data Mechanisms of Weight Data in Electronic Health Records-based Research International Conference in Health Policy Statistics; 2018; Charleston, SC.

15. 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):171–175. [PubMed: 3369325]
16. Arterburn D, Ichikawa L, Ludman EJ, et al. Validity of clinical body weight measures as substitutes for missing data in a randomized trial. *Obes Res Clin Pract* 2008;2(4):277–281. [PubMed: 19956347]
17. Leo MC, Lindberg NM, Vesco KK, Stevens VJ. Validity of medical chart weights and heights for obese pregnant women. *eGEMs* 2014;2(1).
18. Sperrin M, Webb DJ, Patel P, et al. Chronic obstructive pulmonary disease exacerbation episodes derived from electronic health record data validated using clinical trial data. *Pharmacoepidemiol Drug Saf* 2019;28(10):1369–1376. [PubMed: 31385428]
19. Goldstein BA, Phelan M, Pagidipati NJ, Peskoe SB. How and when informative visit processes can bias inference when using electronic health records data for clinical research. *J Am Med Inform Assoc* 2019;26(12):1609–1617. [PubMed: 31553474]

## STUDY IMPORTANCE

### What is already known about this subject?

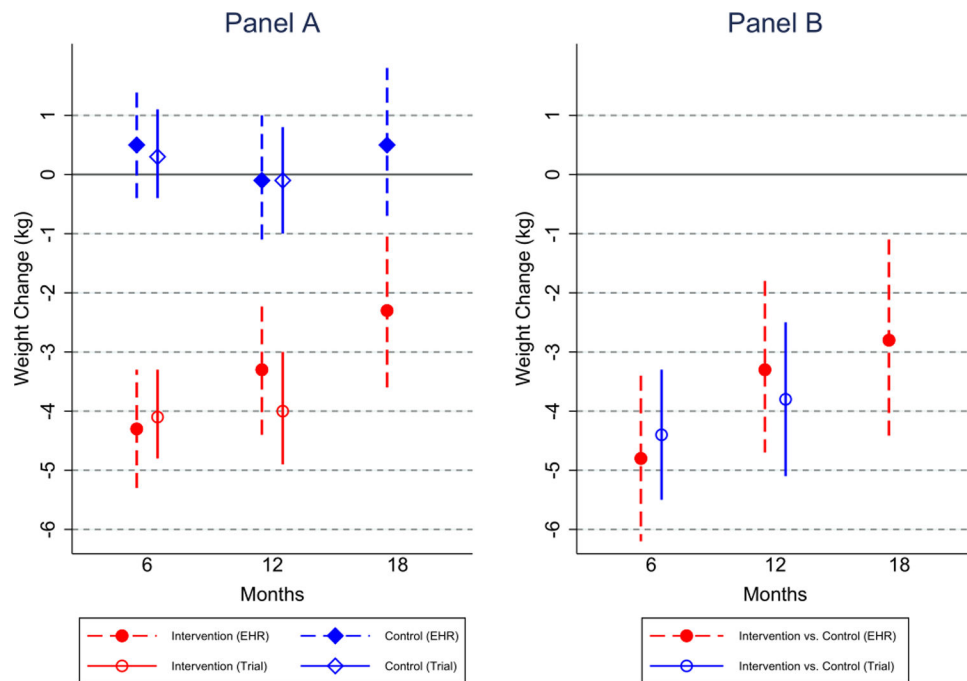
- Electronic health record (EHR) data are increasingly used in comparative effectiveness research, particularly with pragmatic trial designs.
- Using passively-collected EHR data in prospective studies of weight loss interventions can reduce research costs and lower participant burden.

### What are the new findings?

- Using EHR data, we were able to validly and precisely reproduce the original study trial treatment effect of two separate obesity interventions.

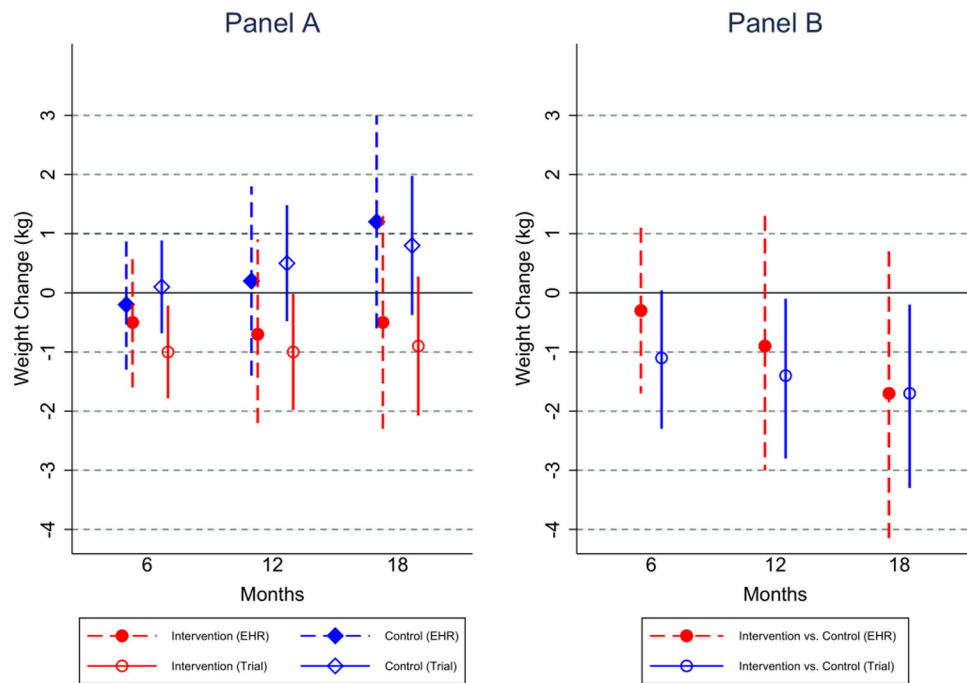
### How might these results change the direction of research or the focus of clinical practice?

- The number of randomized trials leveraging electronic health record data is increasing. The results of our study can be used to help guide the planning and sample size calculation for future pragmatic weight-related trials.
- Our analytic approach and results have positive implications for use of EHR weights in obesity-related clinical trials at all stages on the translational continuum.



**Figure 1. Comparison of EHR data and trial data estimated mean weight loss and intervention effect, and EHR-estimated long-term durability effects, from the Track Trial.**

Panel A shows the estimated mean weight change separately in intervention and control, for the original Track trial data and the EHR data separately, along with the long-term EHR-estimated weight changes. Panel B displays the difference in weight change between intervention and control, estimating using separate mixed effects models for the original trial data (extracted from the Track main results paper) and the EHR data, along with the long-term (18 month) EHR-estimated treatment effect.



**Figure 2. Comparison of EHR data and trial data estimated mean weight loss and intervention effect from the Shape Trial.**

Panel A shows the estimated mean weight change separately in intervention and control, for the original Shape trial data and the EHR data separately. Panel B displays the difference in weight change between intervention and control, estimating using separate mixed effects models for the original trial data (extracted from the Shape main results paper) and the EHR data.

**Table 1.**

Characteristics of participants who consented to EHR data extraction in the Track study.

	Usual Care (N=153)	Intervention (N=154)	Total (N=307)
<b>Gender, Female</b>	110 (71.9%)	106 (68.8%)	216 (70.4%)
<b>Race/Ethnicity</b>			
Non-Hispanic White	44 (28.8%)	43 (27.9%)	87 (28.3%)
Non-Hispanic Black	80 (52.3%)	84 (54.5%)	164 (53.4%)
Hispanic (all races)	19 (12.4%)	21 (13.6%)	40 (13.0%)
Non-Hispanic other/unreported	10 (6.5%)	6 (3.9%)	16 (5.2%)
<b>Education</b>			
Less than high school graduate	26 (17.0%)	21 (13.6%)	47 (15.3%)
High school graduate	46 (30.1%)	62 (40.3%)	108 (35.2%)
Some college or vocational/trade school	60 (39.2%)	60 (39.0%)	120 (39.1%)
4-year college degree or higher	21 (13.7%)	11 (7.1%)	32 (10.4%)
<b>Poverty Status</b>			
Below	49 (32.0%)	43 (27.9%)	92 (30.0%)
Borderline	23 (15.0%)	28 (18.2%)	51 (16.6%)
Above	55 (35.9%)	70 (45.5%)	125 (40.7%)
Unknown	26 (17.0%)	13 (8.4%)	39 (12.7%)
<b>Marital Status</b>			
Married or Living with Partner	71 (46.4%)	74 (48.1%)	145 (47.2%)
Not Married or Living with Partner	81 (52.9%)	80 (51.9%)	161 (52.4%)
Unreported	1 (0.7%)	0 (0.0%)	1 (0.3%)
<b>Has Insurance</b>	80 (52.3%)	78 (50.6%)	158 (51.5%)
<b>Age, Mean (SD)</b>	50.40 (8.71)	50.60 (9.25)	50.50 (8.97)
<b>Weight (in kg), Mean (SD)</b>	101.37 (15.40)	100.90 (14.82)	101.13 (15.09)
<b>Body Mass Index (kg/m<sup>2</sup>), Mean (SD)</b>	35.85 (3.67)	36.07 (4.02)	35.96 (3.84)
<b>Waist circumference (cm), Mean (SD)</b>	114.98 (10.39)	114.55 (9.81)	114.76 (10.09)
<b>Systolic Blood Pressure (mmHg), Mean (SD)</b>	129.47 (17.15)	129.84 (17.18)	129.65 (17.14)
<b>Diastolic Blood Pressure (mm Hg), Mean (SD)</b>	81.58 (11.58)	82.19 (11.84)	81.89 (11.69)
<b>Total Cholesterol, Mean (SD)</b>	187.85 (39.07)	183.00 (33.60)	185.42 (36.45)

Abbreviations: EHR – Electronic Health Records; mmHg – millimeters of mercury; cm – centimeters; kg – kilograms; m – meters; SD – Standard Deviation

**Table 2.**

Characteristics of participants who consented to EHR data extraction in the Shape study.

	Usual Care (N=71)	Intervention (N=68)	Total (N=139)
<b>Education</b>			
Less than high school graduate	10 (14.1%)	9 (13.2%)	19 (13.7%)
High school graduate	18 (25.4%)	18 (26.5%)	36 (25.9%)
Some college or vocational/trade school	39 (54.9%)	38 (55.9%)	77 (55.4%)
4-year college degree or higher	4 (5.6%)	3 (4.4%)	7 (5.0%)
<b>Poverty Status</b>			
Below	29 (40.8%)	20 (29.4%)	49 (35.3%)
Borderline	20 (28.2%)	21 (30.9%)	41 (29.5%)
Above	21 (29.6%)	26 (38.2%)	47 (33.8%)
Unknown	1 (1.4%)	1 (1.5%)	2 (1.4%)
<b>Marital Status</b>			
Married	22 (31.0%)	14 (20.6%)	36 (25.9%)
Other	46 (64.8%)	53 (77.9%)	99 (71.2%)
Unreported	3 (4.2%)	1 (1.5%)	4 (2.9%)
<b>Age, Mean (SD)</b>	35.38 (5.66)	35.93 (5.47)	35.65 (5.56)
<b>Weight (in kg), Mean (SD)</b>	81.65 (9.08)	82.01 (9.00)	81.83 (9.01)
<b>Body Mass Index (kg/m<sup>2</sup>), Mean (SD)</b>	30.23 (2.35)	30.49 (2.64)	30.35 (2.49)
<b>Waist circumference (cm), Mean (SD)</b>	97.87 (7.85)	99.07 (8.11)	98.46 (7.97)
<b>Systolic Blood Pressure (mmHg), Mean (SD)</b>	123.31 (14.61)	123.44 (15.57)	123.37 (15.03)
<b>Diastolic Blood Pressure (mm Hg), Mean (SD)</b>	80.71 (11.06)	81.19 (10.31)	80.94 (10.66)
<b>Total Cholesterol, Mean (SD)</b>	242.31 (511.19)	174.47 (37.57)	208.88 (365.32)

Note: All participants are black women

Abbreviations: EHR – Electronic Health Records; mmHg – millimeters of mercury; cm – centimeters; kg – kilograms; m – meters; SD – Standard Deviation

**Table 3.**

Comparison of EHR and trial data

	Track (n=307)		Shape (n=139)	
	Trial Data*	EHR Data	Trial Data*	EHR Data
Weight across time, kg, mean (SD)	98.0 (14.9)	98.9 (15.4)	81.1 (9.5)	83.2 (10.6)
Weight at baseline <sup>**</sup> , kg, mean (SD)	99.2 (13.4)	100.3 (13.6)	80.6 (8.2)	82.4 (8.4)
Time between visits, days, mean (SD)	186.6 (28.4)	86.0 (84.9)	188.8 (30.6)	110.0 (142.7)

\* Trial data summarized is of those who also had EHR data.

\*\* Baseline is for EHR data is determined as 30 days within the first trial visit (before and after for Track; after *only* for Shape). In Track, 137/307 (44.6%) had EHR data in this window; in Shape, 48/139 (34.5%) had EHR data in this window.

Abbreviations: kg – kilograms; EHR – Electronic Health Records; SD – Standard Deviation



**Table 4.**

Summaries of the number of visits recorded in the EHR where weight was collected

	Track			Shape		
	Control (n=153) 2 (1.3%)	Intervention (n=154) 5 (3.2%)	Overall (n=307) 7 (2.3%)	Control (n=71) 10 (14.1%)	Intervention (n=68) 8 (11.8%)	Overall (n=139) 18 (12.9%)
<b>Number with only one visit recorded in the EHR, N (%)</b>						
<b>Number of visits</b>						
Mean (SD)	9.7 (5.5)	9.0 (5.6)	9.3 (5.5)	5.6 (4.5)	5.5 (4.4)	5.6 (4.4)
Median (Q1, Q3)	9.0 (6.0, 12.0)	8.0 (5.0, 12.0)	8.0 (6.0, 12.0)	4.0 (2.0, 8.0)	4.0 (2.0, 7.0)	4.0 (2.0, 8.0)
Min, Max	1.0, 30.0	1.0, 30.0	1.0, 30.0	1.0, 23.0	1.0, 22.0	1.0, 23.0
<b>Number of visits in 3–9 months</b>						
Mean (SD)	1.5 (1.5)	1.4 (1.3)	1.4 (1.4)	1.5 (1.5)	1.6 (1.4)	1.6 (1.5)
Median (Q1, Q3)	1.0 (1.0, 2.0)	1.0 (0.0, 2.0)	1.0 (0.0, 2.0)	1.0 (1.0, 2.0)	1.0 (0.5, 2.0)	1.0 (1.0, 2.0)
Min, Max	0.0, 10.0	0.0, 7.0	0.0, 10.0	0.0, 7.0	0.0, 6.0	0.0, 7.0
<b>Number of visits in 9–15 months</b>						
Mean (SD)	1.6 (1.5)	1.4 (1.4)	1.5 (1.4)	1.5 (1.5)	1.4 (1.6)	1.4 (1.6)
Median (Q1, Q3)	1.0 (0.0, 2.0)	1.0 (0.0, 2.0)	1.0 (0.0, 2.0)	1.0 (0.0, 2.0)	1.0 (0.0, 2.0)	1.0 (0.0, 2.0)
Min, Max	0.0, 7.0	0.0, 6.0	0.0, 7.0	0.0, 7.0	0.0, 7.0	0.0, 7.0
<b>Number of visits in 15–21 months</b>						
Mean (SD)	1.3 (1.4)	1.5 (1.6)	1.4 (1.5)	1.0 (1.2)	0.8 (1.0)	0.9 (1.1)
Median (Q1, Q3)	1.0 (0.0, 2.0)	1.0 (0.0, 2.0)	1.0 (0.0, 2.0)	1.0 (0.0, 2.0)	0.0 (0.0, 1.0)	1.0 (0.0, 1.0)
Min, Max	0.0, 5.0	0.0, 9.0	0.0, 9.0	0.0, 4.0	0.0, 5.0	0.0, 5.0

Abbreviations: kg – kilograms; EHR – Electronic Health Records; SD – Standard Deviation;

Q - Quartile

**Table 5.**

Comparison of EHR data and primary Track and Shape trial data.

Study Population	Time Point	Data Type	Estimated mean change in weight* from baseline, intervention group	Estimated mean change in weight from baseline, usual care group	Estimated mean difference in change from baseline, intervention vs. usual care	Estimated mean difference in change from baseline, intervention vs. usual care, sensitivity analysis
<b>6 Months</b>	<b>EHR data</b>		-4.3 (-5.3, -3.3)	0.5 (-0.4, 1.5)	-4.8 (-6.2, -3.5)	-4.6 (-6.0, -3.3)
	<b>Trial Data</b>		-4.1 (-4.8, -3.3)	0.3 (-0.4, 1.1)	-4.4 (-5.5, -3.3)	
<b>12 Months</b>	<b>EHR data</b>		-3.3 (-4.4, -2.2)	-0.05 (-1.1, 1.0)	-3.3 (-4.7, -1.8)	-3.2 (-4.6, -1.7)
	<b>Trial Data</b>		-4.0 (-4.9, -3.0)	-0.1 (-1.0, 0.8)	-3.8 (-5.1, -2.5)	
<b>18 Months</b>	<b>EHR Data</b>		-2.3 (-3.6, -1.0)	0.5 (-0.7, 1.8)	-2.8 (-4.5, -1.1)	-2.9 (-4.6, -1.1)
<b>6 Months</b>	<b>EHR data</b>		-0.5 (-1.6, 0.6)	-0.2 (-1.3, 0.9)	-0.3 (-1.7, 1.1)	-0.7 (-2.1, 0.6)
	<b>Trial Data</b>		-1.0 (-1.8, -0.2)	0.1 (-0.7, 0.9)	-1.1 (-2.3, 0.04)	
<b>12 Months</b>	<b>EHR data</b>		-0.7 (-2.2, 0.9)	0.2 (-1.4, 1.8)	-0.9 (-3.0, 1.3)	-1.5 (-3.4, 0.5)
	<b>Trial Data</b>		-1.0 (-2.0, -0.02)	0.5 (-0.5, 1.5)	-1.4 (-2.8, -0.1)	
<b>18 Months</b>	<b>EHR Data</b>		-0.5 (-2.3, 1.3)	1.2 (-0.6, 3.0)	-1.7 (-4.2, 0.7)	-2.2 (-4.6, 0.2)
	<b>Trial Data</b>		-0.9 (-2.1, 0.3)	0.8 (-0.4, 2.0)	-1.7 (-3.3, -0.2)	

\* Notes: Trial data estimates are extracted directly from the Track and Shape main results paper. All estimates are in kilograms (kg). Estimates for both types of data are from mixed effects models, and in Track they are adjusted for health center, gender, and ethnicity. Numbers in parentheses are lower and upper ends of 95% confidence intervals. Abbreviation: EHR – Electronic Health Records.