

# Combining ecological momentary assessment, wrist-based eating detection, and dietary assessment to characterize dietary lapse: A multi-method study protocol

Digital Health  
Volume 7: 1–16  
© The Author(s) 2021  
Article reuse guidelines:  
sagepub.com/journals-  
permissions  
DOI: 10.1177/2055207620988212  
journals.sagepub.com/home/dhj



Stephanie P Goldstein<sup>1,2</sup> , Adam Hoover<sup>3</sup>, E Whitney Evans<sup>1,2</sup> and J Graham Thomas<sup>1,2</sup>

## Abstract

**Objectives:** Behavioral obesity treatment (BOT) produces clinically significant weight loss and health benefits for many individuals with overweight/obesity. Yet, many individuals in BOT do not achieve clinically significant weight loss and/or experience weight regain. Lapses (i.e., eating that deviates from the BOT prescribed diet) could explain poor outcomes, but the behavior is understudied because it can be difficult to assess. We propose to study lapses using a multi-method approach, which allows us to identify objectively-measured characteristics of lapse behavior (e.g., eating rate, duration), examine the association between lapse and weight change, and estimate nutrition composition of lapse.

**Method:** We are recruiting participants (n = 40) with overweight/obesity to enroll in a 24-week BOT. Participants complete biweekly 7-day ecological momentary assessment (EMA) to self-report on eating behavior, including dietary lapses. Participants continuously wear the wrist-worn ActiGraph Link to characterize eating behavior. Participants complete 24-hour dietary recalls via structured interview at 6-week intervals to measure the composition of all food and beverages consumed.

**Results:** While data collection for this trial is still ongoing, we present data from three pilot participants who completed EMA and wore the ActiGraph to illustrate the feasibility, benefits, and challenges of this work.

**Conclusion:** This protocol will be the first multi-method study of dietary lapses in BOT. Upon completion, this will be one of the largest published studies of passive eating detection and EMA-reported lapse. The integration of EMA and passive sensing to characterize eating provides contextually rich data that will ultimately inform a nuanced understanding of lapse behavior and enable novel interventions.

**Trial registration:** Registered clinical trial NCT03739151; URL: <https://clinicaltrials.gov/ct2/show/NCT03739151>

## Keywords

Dietary lapse, dietary assessment, weight loss, passive sensing, ecological momentary assessment

Submission date: 18 June 2020; Acceptance date: 22 December 2020

## Background

Overweight and obesity affect two-thirds of Americans and contribute to the development of multiple cardiovascular disease risk factors (e.g., physical inactivity, hypertension, hyperlipidemia, sleep-disordered breathing, and diabetes mellitus), in addition to being independent risk factors for cardiovascular disease.<sup>1–4</sup> Behavioral obesity treatment (BOT) produces clinically

<sup>1</sup>The Miriam Hospital Weight Control and Diabetes Research Center, Providence, USA

<sup>2</sup>Department of Psychiatry and Human Behavior, Warren Alpert Medical School of Brown University, Providence, USA

<sup>3</sup>Holcombe Department of Electrical and Computer Engineering, Clemson University, Clemson, USA

### Corresponding author:

Stephanie P Goldstein, The Miriam Hospital Weight Control and Diabetes Research Center, 196 Richmond St., Providence, RI 02903, USA.  
Email: [stephanie\\_goldstein@brown.edu](mailto:stephanie_goldstein@brown.edu)



significant weight losses through decreased caloric intake and increased physical activity.<sup>5</sup> Weight loss and maintenance are achieved through adherence to the recommended diet in BOT and nonadherence can compromise outcomes.<sup>6</sup> Dietary lapses (i.e., specific instances of nonadherence to one or more BOT dietary goals or eating behaviors) are frequent (~3-4 times per week) and have been shown to undermine weight loss outcomes.<sup>7,8</sup> Moreover, the ability to cope with temptation, and thus prevent lapses, is associated with BOT success.<sup>9-11</sup>

Despite the clear potential for dietary lapses to influence weight loss outcomes, they remain understudied as traditional measurement tools for assessing eating are inadequate; dietary lapses are difficult to simulate in laboratory settings and are vulnerable to inaccuracies and inherent bias when studied via retrospective self-report questionnaires.<sup>12,13</sup> Ecological momentary assessment (EMA) has enhanced the reliability and validity of lapse measurement through real-time assessment.<sup>13</sup> EMA employs short surveys delivered (typically via smartphone) repeatedly over the course of the day to self-report behaviors, cognitive/emotional states, and environmental contexts.<sup>14</sup> While EMA has been used in prior studies to strengthen lapse measurement,<sup>7,15,16</sup> EMA is still dependent on participants' awareness of their behavior and their willingness and ability to report on it accurately. This limitation precludes a complete understanding of lapse characteristics, the nutritional composition of lapses compared to non-lapse eating, and the role of lapses in impeding weight outcomes.

Thus, an important advancement to improve lapse measurement would be to reduce reliance on self-report by integrating objective monitoring of eating behavior in real-time assessment. For example, rate of eating,<sup>17</sup> number of bites,<sup>18</sup> meal frequency,<sup>19</sup> meal duration,<sup>20</sup> meal timing,<sup>21</sup> and meal size<sup>19</sup> are objective measurements of eating that have been associated with obesity. As such, these same characteristics may serve as markers of eating behaviors that jeopardize weight loss, including dietary lapses. Objective monitoring of eating behaviors is an ever-evolving field that capitalizes on several passive sensing methodologies (e.g., wrist devices, chewing sensors, neck-worn cameras), one of which involves wrist-based devices to detect wrist-roll motion as food is being brought to the mouth. While this methodology is being developed by many research groups,<sup>22-25</sup> our team has previously used wrist-based passive sensing to infer bites from eating episodes that were bounded by the participant manually pressing a button at the start and end of eating.<sup>26,27</sup> The manual button press was helpful in ensuring reports of ground truth, but participants in our prior work have anecdotally reported that it

incurred additional burden, suffered from user error (e.g., forgetting to press start and/or stop), and prompted reactivity (e.g., not eating because they do not want to press the button). As such, our recent work has focused on developing and refining algorithms that can infer eating episodes from continuously collected wrist data (thus eliminating the need for a button press).<sup>27-31</sup>

Our methods for eating inference and free-living bite counting have been developed and refined over five studies (each using differing wrist-based sensors, e.g., InertiaCube3 sensor, iPhone 4, Bite Counter device, STMicroelectronics MEMS sensors, Shimmer3 device). In the first, we counted bites taken during a meal with 86% sensitivity among 47 participants wearing a wrist device during an uncontrolled eating setting in which they brought their own meal to a lab.<sup>28</sup> In the second, we achieved 81% accuracy of inferring the occurrence of eating (ground truth measured via manual event logs) among 43 participants yielding 449 hours of data with total 116 meals and snacks.<sup>29</sup> Third, we found that wrist-inferred bite count was moderately correlated with kilocalories as measured by a 24-hour food recall ( $R^2=0.44$ ,  $p<.001$ ) among 77 free-living participants over 14 days with a total of 2,975 meals.<sup>18,31</sup> Fourth, we studied 271 participants eating a single uncontrolled meal in a cafeteria setting (yielding 24,088 bites) and achieved 75% sensitivity with a positive predictive value of 89% for inferring bites without a button press.<sup>27</sup> Fifth, our most recent study evaluated 4,680 hours of data across 351 participants who wore a wrist device for 1 day during free-living and used an activity log to self-report ground truth of eating instances and secondary activities (i.e., walking and resting).<sup>30</sup> This work capitalized on accounting for secondary activities detected in the wrist data to achieve 77% weighted accuracy of inferring eating episodes. Further analysis showed that the improvements over our prior methods (which achieved 74% weighted accuracy in this sample)<sup>29</sup> due to accounting for walking and rest were significant ( $p<.001$ ) and could be attributable to a 23% reduction in false negatives.

While continued work is required to enhance accuracy and ensure generalizability of our results across different samples and contexts, these studies do illustrate that wrist-based passive eating inference and characterization in free-living settings (i.e., outside of the lab) can be executed without a button press. This method has the potential to supplement EMA by providing information on eating behavior that can enhance accuracy of lapse detection and its characterization.<sup>32</sup> Improving lapse detection and characterization is important because it can dramatically enhance our ability to intervene with personalized intervention

tools and enable more complex conceptual models of the behavior.<sup>33</sup>

Despite the potential for a combination of EMA and passive sensing to advance the study of eating behavior, few studies have employed both together.<sup>34,35</sup> To date, there are no studies that integrate EMA, passive sensing of eating, and direct assessment of dietary intake to understand a clinically relevant problematic eating behavior, such as dietary lapse, in the context of a health behavior change intervention. Given the importance of lapse and the promise of passive sensing technology to improve our understanding of the behavior, we are conducting an NIH-funded study (F32HL143954) with the following aims: 1) identify objectively-measured characteristics of lapse behavior by monitoring device predicted timing, duration, bite count, and rate of eating amongst EMA-reported lapse episodes and comparing to non-lapse eating episodes, 2) test the association between dietary lapse frequency and weight change, and 3) estimate nutrition composition of dietary lapses via 24-hour food recalls.

To accomplish these aims, we utilize three assessment methods concurrently during a behavioral obesity treatment: EMA (the traditional gold-standard method for measuring dietary lapse), a wrist-based device that continuously detects wrist motion from which we can infer eating episodes and relevant characteristics of eating (e.g., rate of eating, duration of eating, bite count), and telephone-based 24-hour food recalls (to assess macro-nutrient and energy intake). The following sections detail the study protocol for this trial. While data collection is ongoing, we present an illustration of pilot data collected from three participants to underscore the strengths, challenges, and potential impact of this work as it evolves. This protocol is informative it represents one of the first applications of this type of multi-method assessment protocol combining gold-standard and cutting-edge measurement strategies. Moreover, the study of lapse is widely applicable to the management of any chronic health condition requiring dietary adherence to a set of recommendations (e.g., Type 2 diabetes, cardiovascular disease, gastrointestinal disorders).

## Methods

### Overview

This project adds continuous passive eating inference and periodic dietary assessment to EMA to study dietary lapses in 40 adults with overweight/obesity and cardiovascular disease risk throughout a 12-week in-person BOT program and a 12-week period of weight loss maintenance. Participants self-report all eating via EMA, including dietary lapses. The ActiGraph GT9X

Link (ActiGraph, LLC, Pensacola, FL, USA), a wrist-based device that is typically used to measure physical activity and sleep, is being used to detect the wrist-motion of food being brought to the mouth. Prior work, completed with other types of wrist-mounted devices, has developed and refined algorithms that can infer eating episodes and characteristics from wrist motion and wrist velocity data.<sup>28–31</sup> We are now extending these algorithms to infer eating from wrist data collected by the ActiGraph. Because the composition of foods consumed during lapses is an important and understudied facet of lapse behavior, participants also complete periodic 24-hour dietary recalls via structured telephone interview to measure the composition of all food and beverages consumed. Data from the ActiGraph and 24-hour food recalls are being used to assess objective eating characteristics (i.e., timing, duration, bites, rate), estimated caloric content, and nutritional composition of EMA-reported dietary lapses compared to non-lapses. Weight is measured at every treatment session.

### Setting

This study takes place at The Miriam Hospital (located in Providence, Rhode Island, USA).

### Participants

We are recruiting 40 participants on a rolling basis. Eligible participants are men and women with overweight or obesity (body mass index 25–50 kg/m<sup>2</sup>), between the ages of 18 and 70, and who have been diagnosed by a physician with one or more cardiovascular risk factors (Type 2 diabetes, hypercholesterolemia, or hypertension). Individuals are excluded if they report health problems that are contraindication for weight loss or physical activity, are currently pregnant or breastfeeding, are currently enrolled in another weight loss program, have lost  $\geq 5\%$  of their initial body weight in the last 6 months, are currently taking weight loss medication, have had a surgical procedure for weight loss, or report a history of a clinically diagnosed eating disorder excluding Binge Eating Disorder.

### Procedure

See Figure 1 for a study schematic. Participants are being recruited on a rolling basis using advertisements in local newspapers, the research center's website, advertising resources available through the Miriam Hospital (e.g., email newsletters circulated to patients and employees), and through physician referrals from local primary care practices. Interested individuals are contacted via telephone to complete an initial screening in which the study procedures are summarized.

	Study week																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Treatment session	X	X	X	X	X	X	X	X	X	X	X	X			X				X					X
Data download session													X				X				X			
Wear ActiGraph device	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Complete EMA	X		X		X		X		X		X		X		X		X		X		X		X	
24-hr food recalls	X				X						X						X						X	
Final assessment & questionnaire																								X

**Figure 1.** Study schematic and timeline.

Individuals who appear eligible are scheduled for an in-person appointment for additional screening and orientation. Should the individual consent to participate during the orientation visit, this appointment also includes baseline assessment procedures. At the orientation/baseline visit, participant height and weight are measured by trained research staff, and they complete questionnaires and are shown how to wear and maintain the ActiGraph. Participants are scheduled for their initial treatment visit approximately one week after the baseline visit. As a behavioral run-in, participants are asked to wear the ActiGraph and record dietary intake (either via smartphone app or paper-based diary) for one week in between the baseline and orientation visit. Participants are only eligible to continue with treatment if they wear the ActiGraph for a minimum of  $\geq 10$  hours on  $\geq 5$  days (including  $\geq 1$  weekend day) and record at least 2 meals a day for 7 days during the run-in phase. Finally, participants are required to provide physician permission to participate in the study procedures given their cardiovascular risk factors.

During the participants' initial treatment visit (60 minutes), ActiGraph and dietary record data are checked to confirm eligibility. Participants are then instructed to wear the ActiGraph every day during waking hours for the remainder of the 24-week study period. Next, participants are trained in how to complete EMA surveys (and in particular oriented to the definition of a dietary lapse) and phone-based 24-hr food recalls. Per Figure 1, participants are asked to complete 7 days of EMA on a biweekly basis for 24 weeks. A biweekly assessment schedule was chosen to minimize participant burden while ensuring the adequacy of the data. Participants with smartphones are encouraged to use their personal devices for EMA. If a

participant does not own a smartphone device or wish to use their personal device, they are provided with a study device for EMA. Per Figure 1, dietary intake is assessed every 6 weeks via phone-based 24-hr food recalls in intervals that align with EMA data collection. Participants begin their first week ("baseline") of EMA and food recalls after the initial treatment visit is completed.

Participants attend in-person treatment sessions on a weekly basis for 12 weeks, and then transition to monthly booster sessions in the subsequent 12 weeks (referred to here as the maintenance phase). During the maintenance phase, participants are asked to attend an additional appointment once per month, separate from the monthly treatment visit, to initiate downloading data from the ActiGraph when it approaches the maximum storage capacity. Participants receive \$20 after attending each of the three data download appointments (during which time they receive no intervention) for a total of \$60 over the 24-week study period. Participants are weighed by trained study staff at every treatment and data download session. At the final study visit, participants complete questionnaires regarding feasibility and acceptability of the program and the technology.

### Behavioral obesity treatment

Session structure and content of the BOT are based on the Diabetes Prevention Program and LookAHEAD trials,<sup>36,37</sup> which produce an average clinically significant weight loss of 7% of initial body weight. Sessions were held on an individual basis and were 30 minutes in length. Per national guidelines for weight loss, a daily goal of 1200-1800 kcal/day is prescribed depending on



baseline weight.<sup>38</sup> A Mediterranean diet is prescribed to meet the recommended calorie goal and be consistent with dietary recommendations for patients with cardiovascular disease risk.<sup>39</sup> Participants are encouraged to achieve 200 minutes/week of moderate intensity activity by gradually increasing activity level to achieve approximately 40 minutes of activity over five days per week.<sup>40</sup> Increasing lifestyle activity (e.g., parking farther away from a building entrance or taking the stairs) to reduce sedentary time is also recommended. Participants are taught strategies to assist them in meeting their dietary and physical activity goals. To increase self-regulatory skills and enhance awareness, participants are instructed to track dietary intake and daily physical activity, as well as weigh themselves regularly. Participants are provided with instructions for using My Fitness Pal, a commercially available smartphone app that is commonly used in BOT studies because it is based in social cognitive theory (the theoretical foundation also underlying BOT), for tracking diet, activity, and weight.<sup>41</sup> In addition to facilitating self-monitoring, My Fitness Pal contains brief messages and information (e.g., recipes, exercise suggestions) that are consistent with evidence-based BOT, which makes it ideal for use in conjunction with an in-person BOT program that provides more substantial support.<sup>41,42</sup> Participants are provided with personalized feedback regarding dietary intake and activity by their clinicians at the outset of each treatment session. Participants are also taught skills for goal-setting, eating mindfully, behavioral reinforcement, stimulus control, and problem-solving.

## Measures

*Objective eating characteristics.* Eating behavior is objectively quantified using algorithms for inferring eating episodes (previously validated using other wrist-mounted devices) that are applied to ActiGraph data.<sup>28–31</sup> Participants are asked to wear the ActiGraph on the wrist of their dominant hand (or that used for eating) during all waking hours, exclusive of bathing and swimming, for the duration of the 24-week study. This ActiGraph is equipped with a wear time sensor and the face of the device is an LCD display that provides date, time, steps, and battery life. When the ActiGraph battery is depleted it requires re-initialization via ActiLife software, as such, participants are asked to charge the ActiGraph nightly (during sleeping hours) throughout the study period. Participants are also instructed to monitor the battery via the LCD display, which alerts them when the battery is at 10% capacity, and place on the charger as needed to prevent depletion.

Most importantly, the ActiGraph has an integrated Inertial Measurement Unit (IMU) that contains triaxial gyroscope, magnetometer, and accelerometer sensors. The IMU provides information about movement and rotation which are necessary for inferring eating. Eating characteristics that are being inferred in the current study are: timing, duration, number of bites taken, and rate of eating. Our eating inference algorithms segment the IMU data at peaks of wrist motion,<sup>29</sup> and then these segments are classified as eating, walking, resting, or other, using a Bayesian classifier.<sup>43</sup> Classification of eating is further described in the Statistical Analysis section. Once eating episodes are inferred, we calculate duration and timing of eating. An algorithm based on wrist-roll motion is used to estimate the number of bites taken during an eating episode (also described further in Statistical Analysis<sup>28</sup>). In future studies, bite estimates can be used to provide estimates of energy intake.<sup>44</sup> Lastly, rate of eating is calculated as seconds per bite using number of bites and duration of the meal.

*Ecological momentary assessment.* Participants are asked to complete seven days of EMA on a biweekly basis for 24 weeks. EMA is conducted using the secure LifeData, platform for EMA. Participants download the LifeData app (available for iPhone or Android) on their smartphone to engage with the EMA protocol. Once enrolled, participants are prompted through the app via vibration and audible tone to complete self-report ratings, which include occurrence of dietary lapse and meals/snacks (i.e., non-lapses). The first survey of each day also assesses whether any lapses or eating occurred in the period between the last survey of the previous day and the first survey of the next day. Ratings are prompted semi-randomly around five anchor times throughout the day (9:00am, 11:00am, 2:00pm, 5:00pm, and 8:00pm). Semi-random prompting has been used in prior EMA studies of eating behavior to prevent survey notifications from arriving too close to one another, thus protecting against user burden and ensuring experiences throughout the day are sampled.<sup>45,46</sup> LifeData randomly prompts users to complete an EMA survey within +/- 1 hour of the anchor time, which minimizes reactivity by ensuring that participants cannot anticipate the exact timing of the assessment. Anchor times were selected to ensure 1) feasible completion by all participants (i.e., not too early or late in the day), and 2) that participants could not receive two surveys at once (i.e., that one survey would expire before another could be delivered). Participants are given 60 minutes to respond to an EMA survey after each prompt, with a reminder at 30 minutes. Participants are also encouraged to self-initiate ratings of dietary lapses and meals/snacks as they occur. De-identified participant data is

transmitted to a secure, password-protected server that is only accessible to the research team. This platform is accessible via a Web-based interface to control the start and end date of EMA protocol, view summaries of data in real-time, and download the complete dataset of participant responses.

Each EMA survey assesses the following: Dietary lapse is assessed by asking participants to report if they have experienced a lapse since the last survey. Consistent with prior work,<sup>7</sup> lapse is defined as any “eating or drinking likely to cause weight gain, and/or put weight loss/maintenance at risk”. Participants who report lapses are asked to record the time and location of the lapse. Participants are asked “how would you describe the lapse?” and select all that apply (“I ate a larger portion of a meal or snack than I intended”, “I ate when I hadn’t intended to eat”, “I ate a type of food that I intended to avoid”, “I ate too quickly”). This method of describing the lapse *type* aids in achieving more accurate self-reports and will later improve our ability to classify lapse behaviors.<sup>47</sup> Non-lapse eating is assessed by asking participants to report if they had eaten a meal or snack since the last survey and to note the approximate times of each. These data will distinguish lapse from non-lapse eating episodes throughout the day. Grazing is assessed by asking “Did you engage in grazing since the last survey (e.g., eating small amounts of food within a period of one hour in an unplanned and repetitious way without feeling hungry?)” Previous testing revealed that it is difficult for the ActiGraph to passively detect very small eating instances (such as one or two bites of food). While participants in the study were discouraged from grazing, knowledge of grazing will be valuable when analyzing and aligning data from the ActiGraph and EMA. Eating pace is assessed by asking “How quickly were you eating?”, with response options of “slower than my normal pace”, “my normal pace”, and “faster than my normal pace”. This question is designed to assist in classifying and understanding rates of eating as detected via the ActiGraph. Eating duration is assessed by asking “How long, in minutes, did the eating episode last?”. This question is designed to assist in validating objective assessment of eating duration estimated by the ActiGraph.

*Dietary composition.* Dietary intake is assessed via telephone-based 24-hour dietary recalls at 6-week intervals that align with EMA data collection. The 24-hour recall method was chosen because it is a gold-standard dietary assessment tool capable of providing rich data on dietary quality of lapses as compared to other non-lapse eating.<sup>48,49</sup> Recalls are collected by trained research staff over a series of three random, non-consecutive days of the week (2 weekdays and 1 weekend day)<sup>49</sup> Given that lapses occur approximately 3–4

times per week, we estimated that the standard series of three recalls would likely capture at least one dietary lapse (if not more). The Nutrition Data System for Research (NDSR), a computer-based software application developed at the University of Minnesota Nutrition Coordinating Center facilitates the collection of dietary recalls in a standardized fashion.<sup>50</sup> Dietary intake data is collected via a multiple-pass method which collects information on eating occasion name and time as well as each food and drink consumed.<sup>51</sup> NDSR supports a dietary quality analysis of recalled intake using a robust food and nutrient database that pulls from the USDA’s Food and Nutrient Database for Dietary Studies. Recalled food intake and independently observed food intake are highly correlated, indicating the 24-hour dietary recall is highly valid and is thus considered the gold-standard for dietary assessment.<sup>52</sup> Estimated caloric intake, as well as macronutrient content (i.e., protein, fiber, carbohydrates, fat, sodium) and an overall dietary quality score (as determined by the Healthy Eating Index-2015), will be calculated for each eating episode using NDSR software. Twenty-four-hour recall data will inform a greater understanding of which dietary factors may characterize to lapse. These data can be used in future studies to inform more targeted EMA-based dietary recalls employed at the same frequency of lapse assessment, thus providing more granular data on the dietary quality of lapses with reduced user burden.<sup>53</sup>

*Questionnaires.* Participants demographics and a rating of comfort and familiarity with using technology are assessed via questionnaire at baseline. At the final assessment, participants complete questionnaires regarding their level of satisfaction with the treatment and assessment protocol.

*Weight and height.* Participants’ body weight is measured to the nearest 0.1 kg using a calibrated digital scale at each study appointment (baseline assessment, weekly, monthly boosters, and at ActiGraph download appointments). Height is measured at baseline in millimeters using a wall-mounted stadiometer. Body Mass Index (BMI; kg/m<sup>2</sup>) is calculated from height and weight measures. Weight change is assessed in kilograms (kg) and percentage change from baseline.

### Statistical analysis plan

*Inferring eating episodes.* Periods of eating are inferred using algorithms described in prior publications.<sup>29,30</sup> The following briefly reviews the method in addition to how eating characteristics are extrapolated from these data. First, data are smoothed using a Gaussian kernel to reduce the effect of noise. Wrist motion energy is calculated as the sum of linear acceleration across the X, Y and Z axes. Peaks in wrist motion

energy have been found to occur before and after eating episodes as a person prepares a meal and cleans up after finishing. Peaks are found in the wrist motion energy using a custom hysteresis threshold and used to segment the data.<sup>29</sup> Each segment is classified using a naïve Bayes classifier into one of four classes: eating, walking, resting or other.<sup>30</sup> Timing of eating is calculated as the time of day that an eating episode is inferred; these data will be categorized into six categories (breakfast, mid-morning, afternoon, mid-afternoon, dinner, evening) for ease of interpretation, as in our prior work.<sup>54</sup> The duration of eating is calculated as the time between the start and end peaks used to segment the data. The number of bites for each eating episode is calculated using a bite detection algorithm that utilizes the velocity and timing of successive wrist-roll motions (described in more detail in our previous publications).<sup>28,31</sup> Using bites and duration (in seconds), rate of eating is calculated as seconds per bite.<sup>55</sup>

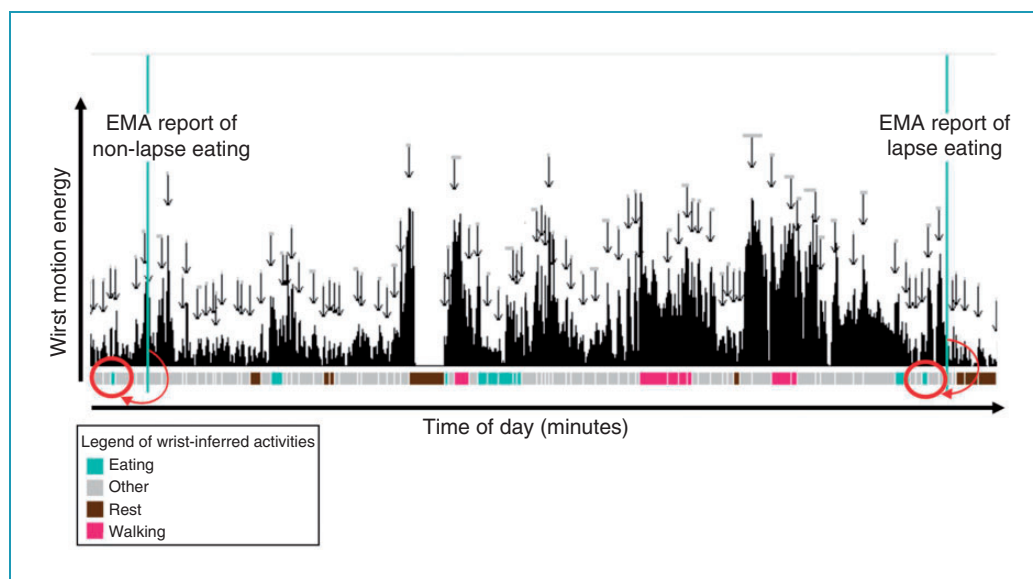
*Matching eating episodes.* It is of note that all methods of assessment in this study (EMA, ActiGraph, and 24-hour food recall) incur some degree of error due to mis-reporting and algorithm-detected false positives. To the best of our knowledge, there are also no precedents from prior research that can be used to inform matching decision rules. As such, error in this matching process is anticipated and the following procedures seek to maximize our data while removing cases of *obvious* error. In other words, as this work is in its very early stages, attempts are made to retain matches wherever possible.

After eating episodes are inferred from the ActiGraph data, they are matched with EMA-reported lapse and non-lapse eating episodes using the time of day eating was detected/reported. See Figure 2 for example output from a custom program that our team built to map EMA-reported eating episodes onto wrist-inferred eating episodes (and secondary activities) throughout a given day. Given that EMA was administered biweekly, eating episode matching procedures are conducted on 12 weeks of the 24-week protocol. Days on which the ActiGraph was not worn are removed from the matching process. To identify possible error in EMA-reported eating prior to matching, we first note eating episodes in which the EMA survey completion timestamp fell greater than 30 minutes *before* the EMA-reported eating episode. This time difference would imply that the participant forecasted eating episodes when EMA surveys clearly only ask about eating in the previous hours, not future eating. Thirty minutes was chosen as the cut-off to account for possible rounding errors when reporting eating (e.g., participant completes an EMA survey at 12:15 when they are currently eating but reports the

eating episode at 12:30pm). Eating episodes that were forecasted > 30 minutes into the future were identified as likely reporting errors and removed from analysis.

Each EMA-reported eating episode is then matched to its closest wrist-inferred eating episode and the time difference in hours between the two eating instances is calculated. We define “successful” matches as those in which the EMA-reported eating and wrist-inferred eating were within  $\pm 2$  hours of one another. In other words, if an EMA-reported eating episode matches to a nearest wrist-inferred eating episode that is over two hours away in either direction, the match is considered “unsuccessful”. It is well known that participants have difficulty accurately self-reporting their dietary intake (even in controlled settings in which they are given the food items).<sup>56,57</sup> While there are no studies that explore bias in self-reported eating times, we assumed that the evidence of misreporting and bias in other areas of self-reported dietary assessment (e.g., amount of food consumed, number of foods consumed) might also extend to reports of eating time.<sup>58</sup> As such, a two-hour window for successful matches was chosen to retain as much data as possible by allowing for generous margin for human error in self-reporting the timing of eating episodes. During preliminary analyses to decide the eating match criterion, we discovered that instituting a two-hour match rule eliminates 20.9% of matches, a 90-minute match rule eliminates 25.3% of matches, and a one-hour match rule eliminates 30.8% of matches. Given that a one-hour match rule eliminated almost 1/3 of our matches, we chose the rule that allowed us to retain the most data for analysis. Upon trial completion, we plan to conduct analyses using the data from more stringent to less stringent match rules to examine differential patterns in the data (see Discussion section for more details). Eating occasions assessed via 24-hour food recall will be matched to EMA- and wrist-detected eating using a similar approach (i.e., using a two-hour window and examining differential patterns in stringency of match rules).

*Aim 1: Objective characteristics of lapse.* A supervised classification approach will be employed to determine which objective eating characteristics distinguish EMA-reported lapse from non-lapse eating (as such, EMA-reported data act as ground truth for the purposes of model building and validation).<sup>59</sup> We propose to employ several different supervised classification models including support vector machine, random forests, and several types of C4.5 decision trees from the R WEKA package (e.g., Logit.Boost, Bagging, Random Subspace, Bayes Net<sup>60</sup>). Model performance will be evaluated by calculating an Area Under the Curve (AUC) estimate, which represents the ability of the classifier to correctly discriminate lapse from non-lapse cases.<sup>61</sup> Model validation will be determined via



**Figure 2.** Example of output from wrist-inferred eating and secondary activities using our algorithms (depicts one participant's data over one day). Downward arrows represent algorithm-identified peaks in wrist motion energy (y axis): colored boxes represent activities classified from wrist motion data and length of time performed in minutes (x axis): blue vertical lines represent EMA-reported lapses and non-lapse eating; red circles enclose the wrist-inferred eating episodes that would be matched to EMA-reported eating.

leave-one-out cross validation (2/3 participants' data are used to build the algorithm and the remaining 1/3 participants' data are left out as a blinded test set), which ensures that the model is unbiased and generalizable to new individuals.<sup>29,62</sup> Accuracy, sensitivity, and specificity of the model for classifying EMA-reported lapse from non-lapse will be calculated using the blinded test set. Predictive validity of each eating characteristic will be assessed via least absolute selection and shrinkage operator regression, which we have used in our prior work on lapse classification due to its variable selection capabilities and reduced risk of overfitting.<sup>63–65</sup>

**Aims 2 & 3: Lapses and weight change & nutritional composition of lapses.** General linear mixed effect models will be used to evaluate the association between weight change (dependent variable) and monthly lapse frequency (time-varying predictor). Models will also be run with the time-varying objective eating characteristics described in Measures (timing of eating, duration of eating, number of bites taken, and rate of eating) as predictors, irrespective of lapse. To estimate the nutrient composition of lapse compared to non-lapses, general longitudinal linear mixed effect models will be used to evaluate the association between lapse classification (dependent variable) and time-varying macro- and micro-nutrition components as measured via dietary recall (i.e., Healthy Eating Index score, protein, fiber, carbohydrates, fat, sodium, and total caloric intake). Both sets of models will be evaluated three times; for

lapses identified by ActiGraph, for lapses identified by EMA self-report, and for all lapses identified by either method. Models will also control for sex, race, ethnicity, age, EMA compliance (where applicable), and weekday/weekend day (where applicable).

**Justification for sample size.** Data from our prior work studying lapses in the context of a 10-week mobile BOT<sup>54</sup> was used to estimate power for Aim 2 (lapses and weight change) via a Monte Carlo simulation with 5,000 iterations. With  $N=40$ ,  $\alpha=.05$ , two-tailed tests, mean (SD) lapses per month of 18.9 (11.6), a typical mean (SD) weight loss of 1.8 (1.8) kg during weight loss and 0 (1.8) kg during maintenance, a conservative intraclass correlation coefficient (ICC) of 0.5 for both lapse and weight loss, power will be .80 to detect an effect size of Cohen's  $d=0.48$  (a medium effect) for the association between the time-varying covariates (e.g., monthly lapse frequency) and weight loss.

## Results

Prior to the current trial, we conducted a pilot test of our EMA and wrist-based eating inference methods, and evaluated the feasibility of the eating episode matching process among three participants who underwent behavioral obesity treatment. We present data from the eating inference methods and matching steps of the above-described analytical plan to serve as a demonstration of our methods. Due to our small



**Table 1.** Descriptive Information for Individual Participants and Mean Calculation.

Variable (unit)	Pt 1	Pt 2	Pt 3	Mean
Age (years)	43	56	56	51.7
BMI (kg/m <sup>2</sup> )	30.4	28.9	39.3	32.9
24-week weight loss (%)*	3.5	4.2	-5.8	0.63
EMA surveys complete (%)	59.9	76.8	48.6	61.8
Days of ActiGraph Wear (n)	126	170	137	144.3
ActiGraph wear time per day (hours)	7.7	11.13	10.16	9.8
Total EMA-reported eating episodes (n)	103	236	154	164.3
Total EMA-reported lapses (n)	44	27	30	33.7
Total EMA-reported non-lapses (n)	59	209	124	130.7
Total wrist-detected eating episodes (n)	659	1,783	970	1,137.3
Average daily bite totals (# of bites)	232.4	334.3	211.8	259.5
Average bites per meal (# of bites)	24.5	28.3	30.8	27.8
Average daily eating duration (minutes)	71.7	100.6	62.4	78.2
Average duration of eating episodes (minutes)	7.5	8.54	9.1	8.4
Successful EMA-reported and wrist-detected eating matches (n)	29	174	102	101.7

\*Note: Positive value indicative of weight loss, negative indicative of weight gain.

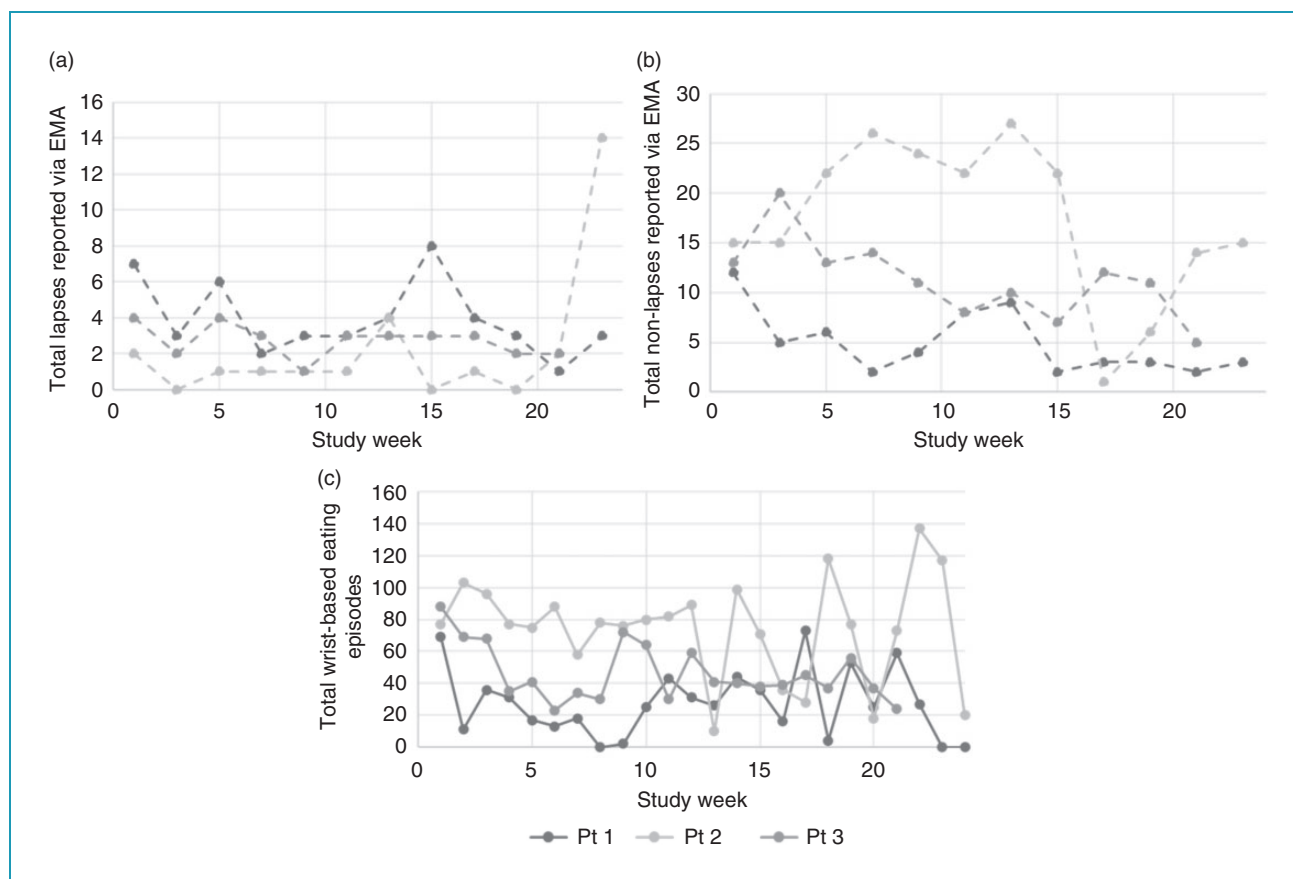
sample, we present means and individual participant data in Table 1. The results from the inferential analyses described for Aims 1-3 will be presented in the primary outcomes paper of this trial with the full study sample and 24-hour food recall assessment (as dietary assessment was not conducted with these pilot participants).

Participants were three women who were middle-aged with an average baseline BMI in the obese range. Two participants identified as White, and one participant identified as Other. During the 24-week behavioral obesity treatment, two participants lost weight and the third gained weight. See Table 1 for demographic information and weight change across participants.

### EMA-reported eating

As displayed in Table 1, participants were moderately compliant with the EMA protocol with 639 (60.7%) signal-contingent prompts out of 1052 total prompted surveys completed. Additionally, participants

completed 37 event-contingent prompts, thus totaling to 676 EMA surveys across 3 subjects. Participants reported 493 eating episodes via EMA. An average of 164.3 ( $SD = 54.8$ ) eating episodes were reported per participant across the 12-week period, which corresponded to approximately 13.7 eating episodes per week per participant. Of these, participants reported 101 total lapses ( $M = 33.7$  lapses,  $SD = 7.4$ ; 20.5% of eating episodes), which corresponded to an average of 2.8 lapses per week per participant. The remainder were 392 total non-lapse eating episodes ( $M = 130.7$ ,  $SD = 61.4$ ; 79.5% of eating episodes), which corresponded to an average of 10.9 non-lapse eating occasions per week per participant. See Table 1 for descriptive information on EMA-reported eating episodes and Figure 3 (panels a and b) for a depiction of lapse and non-lapse eating over time. Figure 3 confirms that lapse and non-lapse eating are imbalanced (i.e., more non-lapse episodes are reported than lapse episodes)<sup>59</sup> and demonstrates that lapse and non-lapse eating are time-varying across individuals. We also see that eating episodes appear to trend down over



**Figure 3.** Eating episodes across the 24-week study period by participant: (a) depicts total weekly dietary lapses per participant. (b) depicts total weekly non-lapses per participant, and (c) depicts total weekly wrist-detected eating episodes per participant. Note: Dashed lines are indicative of non-continuous assessment weeks (EMA) and solid lines are indicative of continuous assessment weeks (ActiGraph).

time, although this is likely confounded with compliance with EMA which is also known to decrease over time.<sup>14</sup>

### Wrist-based eating inference

Across three participants, there were 433 total days with ActiGraph wear time. Per Table 1, participants averaged approximately 85.9% of study days with wear time ( $M = 144.3$ ,  $SD = 22.9$ ). On days with wear time, participants averaged 9.7 hours of wear time per day ( $SD = 1.7$ ). Across participants, there were 47 total days on which the ActiGraph battery was fully depleted ( $M = 15.7$  days per participant,  $SD = 6.4$ ), thus leading to partial days of wear time. Because the ActiGraph requires re-initialization at the research center after the battery is fully depleted, this led to an additional 66 days with no wear time ( $M = 22.0$  days per participant,  $SD = 13.9$ ).

See Table 1 for descriptive information on wrist-inferred eating and eating characteristics. Participants

evidenced 3,412 wrist-inferred eating episodes over 24 weeks, which corresponded to an average of 47.4 eating episodes per week per participant. See Figure 3 (panel c) for wrist-inferred eating episodes over time across participants. While the pattern of wrist-inferred eating episodes over time is likely influenced by compliance with wearing the ActiGraph, these data indicate that there could be individual variability in eating patterns throughout a weight loss program. Average total daily bites per participant was 259.5 ( $SD = 53.5$ ), and participants averaged 27.8 bites per meal ( $SD = 2.6$ ). Average total daily eating duration was 4694.8 seconds ( $SD = 976.5$ ), which corresponds to approximately 79 minutes per day spent eating. Per eating episode, participants evidenced an average eating duration of 504.9 seconds ( $SD = 40.4$ ), which corresponds to approximately 8 minutes per eating episode. Of note, these eating characteristics are generally comparable to those found in previous lab studies,<sup>66</sup> and lower than those found in previous free-living studies.<sup>31</sup> While results from the full sample will be needed

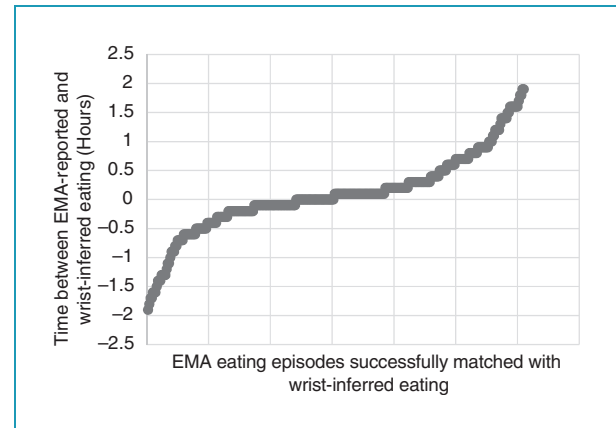
to draw definitive conclusions, discrepancies from previous studies could be due to small sample size, improvements in eating inference and bite counting algorithms over the lifetime of this research, and the removal of the button press from the assessment protocol.

### Matching

We identified 78 total EMA-reported eating events that could not be matched because the ActiGraph was not worn when participants completed EMA ( $M=26.0$  across participants,  $SD=10.0$ ). Seven total EMA-reported events were marked as errors (i.e., eating episodes reported into the future) and were removed. Of note, all seven episodes were from one participant. Using the two-hour window rule, 108 total matches were removed ( $M=36.0$  across participants,  $SD=4.3$ ) as “unsuccessful” matches. Thus, 61.9% ( $n=305$ ) EMA-reported eating events were considered to have “successfully” matched to wrist-inferred eating events ( $M=101.7$  episodes matched,  $SD=59.2$ ). Table 1 depicts descriptive information for the frequency of successfully matched eating episodes across participants. Figure 4 demonstrates the time difference, in hours, between the EMA-reported eating and its wrist-inferred eating match (e.g., an EMA-reported eating episode at 2:00pm that was matched with a wrist-inferred eating episode detected at 1:30pm would have a time difference of -0.5 hours). Negative time lag indicates that the EMA-reported episode occurred after the wrist-detected match, positive time lag indicates that EMA-reported episode occurred before the wrist-detected match. As seen in Figure 4, matches are roughly split with regarding to negative and positive times. The average absolute time between EMA- and wrist-eating matches was 0.5 hours ( $SD=0.5$ ).

### Discussion

The current study will be the first to capitalize on the *integration* of EMA and passive sensing to characterize dietary lapse during a behavioral obesity treatment. In line with the NIH foci on improving measurement and integrating approaches to capture dietary data, utilizing EMA, passive sensing tools (such as the wrist-based eating inference), and dietary assessment together provides contextually rich data that can subsequently inform a nuanced understanding of lapse behavior and ultimately enable sophisticated “in-the-moment” interventions for lapse.<sup>33,67–69</sup> Our initial pilot data provide evidence that combining EMA and wrist-based eating inference methods to understand dietary lapse is feasible and informative.



**Figure 4.** Hours between matched EMA-reported and wrist-inferred eating episodes. Negative values indicate EMA eating episode was reported after its wrist-inferred match. Positive values indicate the EMA eating episode was reported before its wrist-inferred match.

A major strength of our approach is combining previously validated dietary assessment tools (i.e., EMA and 24-hour food recalls) with novel sensing technologies (i.e., wrist-based eating inference), which enhances the scientific rigor of the project.<sup>32</sup> The method of wrist-based eating inference that we used has been developed and validated over decades of research, beginning with using a wrist watch device requiring a button press for bite counting and now inferring eating episodes and characteristics from continuous wrist motion data.<sup>28,29,43</sup> We supplement the wrist-based eating inference with EMA, which is one of the most valid and reliable tools for assessing individuals’ behaviors throughout the course of their daily lives and has been used to study dietary lapses across approximately seven other research studies.<sup>7,15,16,54,59,70,71</sup> Finally, we use a well-established method of dietary assessment, the 24-hour food recall.<sup>52</sup> The combination of these tools is paramount because each method has the potential to overcome weaknesses of the others. For example, wrist-based eating inference could be used in future studies to identify underreporting of eating that may occur via dietary recall or EMA.

Another strength is the size of this proposed work; upon completion, this project will be one of the largest (both with regards to sample size and duration) published studies of objectively inferred eating *and* EMA-reported lapses in free-living adults.<sup>32</sup> Moreover, the temporal granularity of our methods affords an unprecedented amount of data that will allow us to explore both within-person and time-varying associations between EMA-reported lapses and wrist-inferred eating characteristics, as well as lapse behavior (both subjectively reported and objectively detected) and weight change.<sup>33</sup>

There have also been general technical challenges that are normative when using novel technology-based assessment methods. The process of conducting our pilot study and preliminary analysis helped us to identify important considerations that will ultimately lead to improvements to this work. Although we would caution against using these data as indicators of true population values given the small sample size ( $n=3$ ), we were able to identify challenges with the compliance with EMA and ActiGraph measurement protocols and the EMA-wrist data matching process.

We observed moderate compliance with EMA and good compliance with wearing the ActiGraph across the study period. There could be several reasons for missed EMA surveys, which likely contributed to fewer lapse or meal recordings than we would expect. To limit common obstacles in completing EMA, such as device-related problems (e.g., low battery) and situational barriers (e.g., driving),<sup>72</sup> the main trial now includes more participant training in using the LifeData app to report on eating occasions (e.g., a Frequently Asked Questions document, periodic brief refreshers on using LifeData and reporting on lapses at treatment visits, specific instruction for using the ‘user-initiated surveys’ function to report lapses and meals if EMA notifications are missed due to situational barriers). In addition to procedures to limit barriers for completing EMA, monetary compensation has been used in prior work to enhance participants’ motivation to respond to surveys. Due to limited resources, we did not provide compensation that was contingent on completing EMA or wearing the ActiGraph. Other studies have combined compensation with frequent detailed reports of progress (e.g., percentage of completed surveys or hours of wear time) to enhance compliance.<sup>7,13,71</sup> Given the importance of good data quality at this phase of the research, future studies should strongly consider using these methods for enhancing assessment compliance. Moreover, this study protocol requires frequent and high-levels of EMA, which has been supported in our prior work but nonetheless may have been difficult for participants to complete.<sup>54,59</sup> Future studies might also consider the trade-off between more continuous assessment and participant burden (e.g., completing EMA/wearing ActiGraph every 2 or 3 weeks instead of every week/biweekly), which might mitigate the risk of non-compliance over the long-term and ultimately enhance data quality.

With regards to the wearing the ActiGraph device, there were two major technical barriers that likely impacted participant wear time. First, the ActiGraph battery does not last a full 24-hour period when collecting data via IMU. Further, the ActiGraph battery capacity tends to deplete after repeated days of IMU use and re-initialization. As such, participants reported

having to start charging the ActiGraph prior to finishing their day to prevent the battery from fully depleting. Second, the ActiGraph requires re-initialization with proprietary software only available to the research team if the device battery fully depletes. Thus, when the battery died, participants were required to return the device to the research center in-person as soon as possible, which often fell outside of their planned treatment sessions. Together, these two factors likely led to an inflation of non-compliance with ActiGraph device wear and missing data on wrist-inferred eating episodes. Future studies may consider utilizing other commercially available smart watches (e.g., Samsung Galaxy Smart Watch) that have better battery life and do not require reactivation by the research team after battery depletion.

Developing the methods for matching the EMA-reported and wrist-inferred eating posed another interesting challenge. In regards to matching, our pilot data provided numerous important insights. As depicted in Table 1 and Figure 3, wrist-inferred eating episodes far out-number EMA-reported eating. This result is likely due to respective error in both wrist-based eating inference and EMA-reported eating. After transitioning away from the button press method (which incurred burden but also helped to assess accuracy of ground truth), our continuous wrist-based eating inference algorithms currently make an estimated three to five false positives per actual eating episode,<sup>29</sup> thus overestimating eating. Conversely, EMA-reported eating indicated approximately two eating episodes per day, which is a likely indicator of underreporting. These initial data show that substantial work is still required to actualize the potential of combining EMA and wrist-based passive sensing to infer eating; this research would benefit from improving the accuracy of each assessment tool individually and the sophistication of methods for combining the data. With regard to the former, we are continuously refining our algorithms for wrist-based eating inference (in other work not discussed here) which will ultimately enhance accuracy and improve our ability to match with EMA in future studies. With regard to the latter, we illustrate below how our pilot data can begin to inform EMA-wrist data matching approaches.

With our pilot data, we chose to use a simple algorithm to match EMA and wrist-inferred eating episodes that maximized data availability. Based on these data, a “greedy” matching rule (i.e., matching each EMA-reported eating episode to its nearest wrist-inferred eating episode by time) was chosen for now because it allows us to retain many matches. Of note, we hope this will serve as a starting point for matching rules that the larger trial can build from with greater levels of sophistication. The notable drawback is that



the two-hour window for matching EMA and wrist-inferred eating episodes is still quite large. When data collection for the main trial is complete, we plan to formally evaluate the impact of this two-hour matching rule (as compared to 90, 60, or 30-minute matching rules) by comparing patterns of results across the different matching rules. Another strategy, which would be ideal for larger datasets, would be to use more robust analytical tools to fine-tune our simple decision rule for matching EMA and wrist eating episodes by searching for optimal matching algorithm parameters such as thresholds, smoothing factors, or averaging window sizes (genetic algorithms are one such tool; they are often applied to biomedical data and timetable problems to find the set of parameters that best approximates the expected result).<sup>73,74</sup>

As evidenced by the EMA-wrist matches reported in Table 1, another important insight from this work was that the process of matching wrist- and EMA-eating was more successful for some participants than others. Such results may indicate that there will be individual variability in the accuracy of EMA-reported eating and/or wrist-inferred eating, thus making the matching of the two data streams more difficult. Given the challenges that arose with matching differing data streams, future applications of this work would benefit substantially from linking the wrist-based eating inference and the EMA, such that detecting a wrist episode triggers a survey to be sent to the participants.<sup>34</sup> While this technology has not yet been developed for our wrist-based passive sensing tool, it is well-within reach as it would require a program that 1) uses previously-developed algorithms in real-time to infer eating episodes, 2) pushes a notification to the participant, and 3) stores the survey response. This innovation would eliminate the need for matching EMA-reported and wrist-based eating episodes altogether. In the meantime, future studies in this area could consider utilizing both continuous eating inference and the manual button press method to ensure access to ground truth information (although this would require use of a smartwatch for wrist-based sensing, as the ActiGraph does not contain the manual button press feature).

## Conclusion

The project is designed to accumulate large quantities of longitudinal data regarding dietary lapse and passively sensed eating behavior throughout weight loss and weight loss maintenance phases of behavioral obesity treatment. Examples of potential research questions that will be tested using these data include: distinguishing lapse from non-lapse eating episodes using objective eating characteristics; investigating

types of eating behaviors that are associated with weight change (subjective lapses, objective lapses, or some combination of subjective and objective information); evaluating the potential impact of completing EMA on eating behavior by comparing objective eating characteristics on EMA weeks to non-EMA weeks; investigating individual differences with regards to objective eating characteristics of lapse; and examining the time-varying nature of the above-described associations (e.g., evaluating if objective eating characteristics of a lapse differ between weight loss and weight loss maintenance treatment phases). Additionally, the use of a traditional dietary assessment tool, the 24-hr food recall, will provide valuable information related to the dietary composition of a lapse.<sup>32,75</sup> Overall, this study incorporates several cutting-edge methodologies that are still being heavily refined due to some important limitations, but we can already observe the tremendous potential of this work. Data and methods from the current study are expected to provide a foundation for subsequent NIH-funded trials using these technology-based assessment methods to understand dietary lapses and other types of problematic eating. Eventually, this research line will lead to the development of novel technology-based interventions for targeting problematic eating.

**Acknowledgements:** We would like to thank Dr. Rena Wing for her assistance and guidance in this research. We also wish to acknowledge the pilot participants for their time and dedication.


**Contributorship:** All authors assisted in the conceptualization of this study and are named contributors on the NIH F32 award. All authors were involved in protocol development. SG and JGT were responsible for gaining ethical approval, patient recruitment, and data collection. SG and AH conceptualized the analytical plan and conducted analyses. SG wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version.

**Declaration of conflicting interests:** The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: SG, EWE, and JGT declare that they have no conflicts of interest. AH is the co-creator of the Bite Counter. While this study did not use Bite Counters, it benefited from use of the Bite Counter algorithms.

**Ethical approval:** The Institutional Review Board of The Miriam Hospital approved this study (IRB record 1089492-21).

**Funding:** The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Heart Lung and Blood Institute of the National Institutes of Health [grant number: F32HL143954; Principal Investigator: SG].

**Guarantor:** SG.

**ORCID iD:** Stephanie P Goldstein  <https://orcid.org/0000-0002-9656-1675>

**Peer review:** Dr. Francesco Sartor and Dr. Laura König, University of Konstanz, have reviewed this manuscript.

## References

- Hruby A, Manson JE, Qi L, et al. Determinants and consequences of obesity. *Am J Public Health* 2016; 106: 1656–1662.
- Van Gaal LF, Mertens IL and Christophe E. Mechanisms linking obesity with cardiovascular disease. *Nature* 2006; 444: 875–880.
- Fryar CD, Carroll MD and Ogden CL. Prevalence of Overweight, Obesity, and Severe Obesity Among Adults Aged 20 and Over: United States, 1960–1962 Through 2015–2016. Hyattsville, MD: NCHS Health EStats, CDC Stacks Public Health Publications; 2018. Available at: <https://stacks.cdc.gov/view/cdc/58670> (accessed 6 January 2021).
- Poirier P, Giles TD, Bray GA, et al. Obesity and cardiovascular disease: pathophysiology, evaluation, and effect of weight loss: an update of the 1997 American Heart Association Scientific Statement on Obesity and Heart Disease from the Obesity Committee of the Council on Nutrition, Physical Activity, and Metabolism. *Circulation* 2006; 113: 898–918.
- Wadden TA and Foster GD. Behavioral treatment of obesity. *Medical Clinics of North America* 2000; 84: 441–461.
- Alhassan S, Kim S, Bersamin A, et al. Dietary adherence and weight loss success among overweight women: results from the a to Z weight loss study. *Int J Obes (Lond)* 2008; 32: 985–991.
- Forman EM, Schumacher LM, Crosby R, et al. Ecological momentary assessment of dietary lapses across behavioral weight loss treatment: characteristics, predictors, and relationships with weight change. *Ann Behav Med* 2017; 51: 741–753.
- Wadden TA, Crerand CE and Brock J. Behavioral treatment of obesity. *Psychiatr Clin North Am* 2005; 28: 151–170.
- Drapkin RG, Wing RR and Shiffman S. Responses to hypothetical high risk situations: do they predict weight loss in a behavioral treatment program or the context of dietary lapses? *Health Psychol* 1995; 14: 427–434.
- McGuire MT, Wing RR, Klem ML, et al. Behavioral strategies of individuals who have maintained long-term weight losses. *Obes Res* 1999; 7: 334–341.
- Grilo CM, Shiffman S and Wing RR. Coping with dietary relapse crises and their aftermath. *Addict Behav* 1993; 18: 89–102.
- Grilo CM, Shiffman S and Wing RR. Relapse crises and coping among dieters. *J Consult Clin Psychol* 1989; 57: 488–495.
- Stone AA and Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. *Ann Behav Med* 1994; 16: 199–202.
- Shiffman S, Stone AA and Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol* 2008; 4: 1–32.
- McKee HC, Ntoumanis N and Taylor IM. An ecological momentary assessment of lapse occurrences in dieters. *Ann Behav Med* 2014; 48: 300–310.
- Carels RA, Hoffman J, Collins A, et al. Ecological momentary assessment of temptation and lapse in dieting. *Eat Behav* 2001; 2: 307–321.
- Hill SW and McCutcheon NB. Contributions of obesity, gender, hunger, food preference, and body size to bite size, bite speed, and rate of eating. *Appetite* 1984; 5: 73–83.
- Alex J, Turner D, Thomas DM, et al. Bite count rates in free-living individuals: new insights from a portable sensor. *BMC Nutr* 2018; 4: 23.
- Spiegel TA, Kaplan JM, Tomassini A, et al. Bite size, ingestion rate, and meal size in lean and obese women. *Appetite* 1993; 21: 131–145.
- Ma Y, Bertone ER, Stanek IIE, et al. Association between eating patterns and obesity in a free-living US adult population. *Am J Epidemiol* 2003; 158: 85–92.
- Forslund HB, Lindroos A, Sjöström L, et al. Meal patterns and obesity in Swedish women – a simple instrument describing usual meal types, frequency and temporal distribution. *Eur J Clin Nutr* 2002; 56: 740–747.
- Kyritsis K, Diou C and Delopoulos A. Modeling wrist micromovements to measure in-meal eating behavior from inertial sensor data. *IEEE J Biomed Health Inform* 2019; 23: 2325–2334.
- Fortuna C, Giraud-Carrier C and West J. Hand-to-mouth motion tracking in free-living conditions for improved weight control. In: *IEEE international conference on healthcare informatics (ICHI)*, 2016, pp.341–348. USA: IEEE.
- Farooq M and Sazonov E. Accelerometer-based detection of food intake in free-living individuals. *IEEE Sens J* 2018; 18: 3752–3758.
- Thomaz E, Essa I and Abowd GD. A practical approach for recognizing eating moments with wrist-mounted inertial sensing. In: *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, 2015, pp.1029–1040.
- Scisco JL, Muth ER, Dong Y, et al. Slowing bite-rate reduces energy intake: an application of the bite counter device. *J Am Diet Assoc* 2011; 111: 1231–1235.
- Shen Y, Salley J, Muth E, et al. Assessing the accuracy of a wrist motion tracking method for counting bites across demographic and food variables. *IEEE J Biomed Health Inform* 2016; 21: 599–606.
- Dong Y, Hoover A, Scisco J, et al. A new method for measuring meal intake in humans via automated wrist motion tracking. *Appl Psychophysiol Biofeedback* 2012; 37: 205–215.

29. Dong Y, Scisco J, Wilson M, et al. Detecting periods of eating during free-living by tracking wrist motion. *IEEE J Biomed Health Inform* 2014; 18: 1253–1260.
30. Sharma S, Jasper P, Muth E, et al. The impact of walking and resting on wrist motion for automated detection of meals. *ACM Trans Comput Healthcare* 2020; 1: 1–19.
31. Scisco JL, Muth ER and Hoover AW. Examining the utility of a bite-count–based measure of eating activity in free-living human beings. *J Acad Nutr Diet* 2014; 114: 464–469.
32. Bell BM, Alam R, Alshurafa N, et al. Automatic, wearable-based, in-field eating detection approaches for public health research: a scoping review. *NPJ Digit Med* 2020; 3: 38.
33. Riley WT. Behavioral and social sciences at the national institutes of health: Methods, measures, and data infrastructures as a scientific priority. *Health Psychol* 2017; 36: 5–7.
34. Ye X, Chen G, Gao Y, et al. Assisting food journaling with automatic eating detection. In: *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems*, (CHI EA '16) San Jose, CA, USA, 7-12 May 2016; pp. 3255–3262.
35. Gomes D and Sousa I. Real-Time drink trigger detection in free-living conditions using inertial sensors. *Sensors* 2019; 19: 2145.
36. Diabetes Prevention Program Research Group. The diabetes prevention program (DPP): description of lifestyle intervention. *Diabetes Care* 2002; 25: 2165–2171.
37. Look AHEAD Research Group. Look AHEAD (action for health in diabetes): design and methods for a clinical trial of weight loss for the prevention of cardiovascular disease in type 2 diabetes. *Control Clin Trials* 2003; 24: 610–628.
38. NHLBI Obesity Education Initiative. *The practical guide: identification, evaluation, and treatment of overweight and obesity in adults*. USA: National Institutes of Health, National Heart, Lung, and Blood Institute, 2000.
39. Estruch R, Ros E, Salas-Salvadó J, et al. Primary prevention of cardiovascular disease with a Mediterranean diet. *N Engl J Med* 2013; 368: 1279–1290.
40. Piercy KL, Troiano RP, Ballard RM, et al. The physical activity guidelines for Americans. *JAMA* 2018; 320: 2020–2028.
41. Laing BY, Mangione CM, Tseng C-H, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. *Ann Intern Med* 2014; 161: S5–S12.
42. Pagoto S, Schneider K, Jovic M, et al. Evidence-based strategies in weight-loss mobile apps. *Am J Prev Med* 2013; 45: 576–582.
43. *The impact of walking and resting on wrist motion for automated detection of meals* *ACM Trans Comput Healthcare* 2020; 1: 4. Available: <https://doi.org/10.1145/3407623>.
44. Salley JN, Hoover AW and Muth ER. Between-and within-Subjects predictors of the kilocalorie content of bites of food. *J Acad Nutr Diet* 2019; 119: 1109–1117.
45. Smyth JM, Wonderlich SA, Heron KE, et al. Daily and momentary mood and stress are associated with binge eating and vomiting in bulimia nervosa patients in the natural environment. *J Consult Clin Psychol* 2007; 75: 629–638.
46. Thomas JG, Doshi S, Crosby RD, et al. Ecological momentary assessment of obesogenic eating behavior: combining person-specific and environmental predictors. *Obesity (Silver Spring)* 2011; 19: 1574–1579.
47. Goldstein SP, Dochat C, Schumacher LM, et al. Using ecological momentary assessment to better understand dietary lapse types. *Appetite* 2018; 129: 198–206.
48. Steinfeldt L, Anand J and Murayi T. Food reporting patterns in the USDA automated multiple-pass method. *Procedia Food Sci* 2013; 2: 145–156.
49. Conway JM, Ingwersen LA and Moshfegh AJ. Accuracy of dietary recall using the USDA five-step multiple-pass method in men: an observational validation study. *J Am Diet Assoc* 2004; 104: 595–603.
50. Feskanich D, Sielaff BH, Chong K, et al. Computerized collection and analysis of dietary intake information. *Comput Methods Programs Biomed* 1989; 30: 47–57.
51. Johnson RK, Driscoll P and Goran MI. Comparison of multiple-pass 24-hour recall estimates of energy intake with total energy expenditure determined by the doubly labeled water method in young children. *J Am Diet Assoc* 1996; 96: 1140–1144.
52. Blanton CA, Moshfegh AJ, Baer DJ, et al. The USDA automated multiple-pass method accurately estimates group total energy and nutrient intake. *J Nutr* 2006; 136: 2594–2599.
53. Schembre SM, Liao Y, O'Connor SG, et al. Mobile ecological momentary diet assessment methods for behavioral research: systematic review. *JMIR mHealth Uhealth* 2018; 6: e11170.
54. Goldstein SP, Thomas JG, Foster GD, et al. Refining an algorithm-powered just-in-time adaptive weight control intervention: a randomized controlled trial evaluating model performance and behavioral outcomes. *Health Informatics J* 2020; 26: 2315–2331.
55. Salley JN, Hoover AW, Wilson ML, et al. Comparison between human and bite-based methods of estimating caloric intake. *J Acad Nutr Diet* 2016; 116: 1568–1577.
56. Lichtman SW, Pisarska K, Berman ER, et al. Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. *N Engl J Med* 1992; 327: 1893–1898.
57. Lafrenière J, Lamarche B, Laramée C, et al. Validation of a newly automated web-based 24-hour dietary recall using fully controlled feeding studies. *BMC Nutr* 2017; 3: 34.
58. Gibson RS. *Principles of nutritional assessment*. Oxford: Oxford University Press, 2005.
59. Goldstein SP, Zhang F, Thomas JG, et al. Application of machine learning to predict dietary lapses during weight loss. *J Diabetes Sci Technol* 2018; 12: 1045–1052.
60. Hornik K, Buchta C and Zeileis A. Open-source machine learning: R meets weka. *Comput Stat* 2009; 24: 225–232.

61. Cortes C and Mohri M. Confidence intervals for the area under the ROC curve. *Adv Neural Inform Process Syst* 2005; 18: 305–312.
  62. Wong T-T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition* 2015; 48: 2839–2846.
  63. Goldstein SP. *A preliminary investigation of a personalized risk alert system for weight control lapses*. Philadelphia, PA: Drexel University, 2016.
  64. Goldstein SP. *Comparing effectiveness and user behaviors of two versions of a just-in-time adaptive weight loss smartphone app*. Philadelphia, PA: Drexel University, 2018.
  65. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B (Methodol)* 1996; 58: 267–288.
  66. Lin B and Hoover A. A comparison of finger and wrist motion tracking to detect bites during food consumption. In: *IEEE 16th international conference on wearable and implantable body sensor networks (BSN)*, 2019, pp.1–4. USA: IEEE.
  67. Riley WT, Rivera DE, Atienza AA, et al. Health behavior models in the age of mobile interventions: are our theories up to the task? *Transl Behav Med* 2011; 1: 53–71.
  68. Hekler EB, Michie S, Pavel M, et al. Advancing models and theories for digital behavior change interventions. *Am J Prev Med* 2016; 51: 825–832.
  69. US Department of Health & Human Services; National Institutes of Health. 2020-2030 Strategic Plan for NIH Nutrition Research; 2020. Available at: <https://www.niddk.nih.gov/about-niddk/strategic-plans-reports/strategic-plan-nih-nutrition-research> (accessed 6 January 2021).
  70. Carels RA, Douglass OM, Cacciapaglia HM, et al. An ecological momentary assessment of relapse crises in dieting. *J Consult Clin Psychol* 2004; 72: 341–348.
  71. Forman EM, Goldstein SP, Zhang F, et al. OnTrack: development and feasibility of a smartphone app designed to predict and prevent dietary lapses. *Transl Behav Med* 2019; 9: 236–245.
  72. Ziesemer K, König LM, Boushey CJ, et al. Occurrence of and reasons for “missing events” in mobile dietary assessments: results from three event-based ecological momentary assessment studies. *JMIR mHealth Uhealth* 2020; 8: e15430.
  73. Ross P and Corne D. Comparing genetic algorithms, simulated annealing, and stochastic hillclimbing on timetabling problems. In: *AISB workshop on evolutionary computing*, 1995, pp.94–102. Berlin: Springer.
  74. Levin M. Use of genetic algorithms to solve biomedical problems. *MD Comput* 1995; 12: 193–199.
  75. Westerterp KR and Goris AH. Validity of the assessment of dietary intake: problems of misreporting. *Curr Opin Clin Nutr Metab Care* 2002; 5: 489–493.
-