TECHNICAL UNIVERSITY OF MUNICH

HABILITATION THESIS

Improving Weight Loss in Lifestyle Interventions

Dr. sc. med. Christoph Höchsmann

Venia Legendi for

HEALTH SCIENCES

with a focus on

Physical Activity, Ingestive Behavior, and Health

Submitted to the

Department Health and Sport Sciences

TUM School of Medicine and Health

Technical University of Munich

September 2024

Mentors:

Prof. Dr. Karsten Köhler, Technical University of Munich (Chair)

Prof. Dr. Dr. Anja Bosy-Westphal, Christian-Albrechts-University Kiel

Prof. Dr. Anne Kerstin Reimers, Friedrich-Alexander-University Erlangen-Nuremberg

Table of Contents

DECLARATION

Dr. Christoph Höchsmann was accepted as a habilitation candidate at the Department of Sport and Health Sciences (since October 1, 2023 TUM School of Medicine and Health) of the Technical University of Munich in accordance with the habilitation regulations of December 9, 2003 (in the version of the amendment statutes of December 13, 2005) by the Department Council on February 1, 2022.

According to §9 para. 2 of the habilitation regulations, the habilitation procedure intends to determine the "ability to conduct independent research based on [...] several publications with the weight corresponding to a habilitation thesis". The requirements for this cumulative habilitation were defined and agreed upon between the habilitation candidate and the mentors (*Zielvereinbarung*) and approved by the Department Council on June 1, 2022. The interim evaluation took place on July 25, 2023, and was approved by the mentors and the Department Council. The cumulative habilitation thesis was submitted to the mentors for final approval on April 11, 2024. After positive evaluation, the School Council gave the final approval of the completed habilitation procedure on September 18, 2024 and thereby granted the Venia Legendi for *Health Sciences with a focus on Physical Activity, Ingestive Behavior, and Health*.

ACKNOWLEDGMENTS

I would like to thank my mentors, Prof. Dr. Karsten Köhler, Prof. Dr. Dr. Anja Bosy-Westphal, and Prof. Dr. Anne Kerstin Reimers, for their support and excellent supervision of my habilitation. I would particularly like to thank Prof. Dr. Karsten Köhler for giving me the opportunity to join his research group and gain further expertise in grant writing, establishing research collaborations, and helping me develop my own independent line of research. I very much appreciate your motivation, ever-constructive input, excellent guidance, and the overall positive work environment in your lab. I also would like to thank all current and former lab members of the Exercise, Nutrition and Health Professorship. I am particularly grateful to Dr. Paulina Wasserfurth-Grzybowska for the many lively discussions, helpful feedback, and pleasant work atmosphere in our office and to Julian Brummer for his outstanding work during his student assistant position, but also to Dr. Hande Hofmann, Helena Engel, Sieglinde Hechenbichler Figueroa, Valentin Nußer, Lucas Thurnherr, Christina Glasbrenner, Vanessa Corrêa Mascarenhas, and Elisabeth Bäurle as well as all student assistants and interns for the excellent and always supportive work environment.

I would also like to acknowledge my research funding and thank the National Institute of Diabetes and Digestive and Kidney Diseases for supporting my research at Pennington Biomedical Research Center for two years through a T32 Research Service Award (06/2019 – 05/2021) as well as the Alexander von Humboldt Foundation for supporting my research for two years at the Technical University of Munich via a Humboldt Postdoctoral Research Fellowship (04/2022 – 03/2024).

Finally, I would like to thank my collaborators and co-authors, without whom my habilitation research would not have been possible. First and foremost, I would like to thank my current mentor, Prof. Dr. Karsten Köhler, as well as my mentors at Pennington Biomedical Research Center, Dr. Corby Martin, Dr. Peter Katzmarzyk, and Dr. John Apolzan, who have played a crucial role in my habilitation publications. Other collaborators and co-authors I would like to acknowledge are: Dr. James Dorling, Dr. Robert Newton Jr., Dr. Neil Johannsen, Dr. Kara Denstel, Emily Mire, Dr. Daniel Hsia, Dr. William Johnson, Dr. Vivian Fonseca, Dr. Terry Davis, Dr. Nicole Fearnbach Viverito, Dachuan Zhang, Dr. Connie Arnold, Dr. Carl Lavie, Dr. Candice Myers, Dr. Tim Church, Dr. Tera Fazzino, Dr. Susan Racette, Shengping Yang, Dr. Michelle Cardel, Dr. José Ordovás, Dr. Gary Foster, Dr. Frank Greenway, Dr. Catherine Champagne, Dr. William Kraus, Dr. Wentao Cao, Dr. Sai Krupa Das, Dr. Qingzhao Yu, Dr. Manjushri Bhapkar, Dr. Leanne Redman, Dr. Eric Ravussin, Dr. Diana Thomas, Dr. Carl Pieper, Dr. Julie Boron, Dr. Jeffrey Stevens, Dr. Jeffrey French, Safiya Beckford, Dr. Tina Thethi, Dr. Eboni Price-Haywood, Dr. Benjamin Springgate, Dr. Bernard Fuemmeler, Dr. Caroline Cobb, Gabrielle Maldonado, Vineela Rachagiri, Dr. Thokozeni Lipato, Dr. Rabia Imran, Karissa Neubig, Dr. Andrew Barnes, Akansha Anbil, Julian Brummer, Sieglinde Hechenbichler Figueroa, and Christina Glasbrenner.

SYNOPSIS

This publication-based habilitation thesis is divided into three parts representing the main research areas of my habilitation work. Part I focuses on determining factors that influence the degree of weight loss and energy compensation in lifestyle interventions aiming to improve weight and health-related outcomes. Part II focuses on technology-assisted methods for food intake assessment aiming to enhance the remote measurement and monitoring of dietary intake and eating behaviors. Part III highlights further work of my habilitation research, examining the health effects of large lifestyle interventions beyond weight loss and addressing the challenges in defining adherence to calorie restriction goals.

PART I – Predictors of Energy Compensation and Weight Loss in Lifestyle Interventions

According to the World Health Organization, 2.5 billion adults (43%) worldwide were overweight in 2022, and 890 million of these were living with obesity. Obesity is associated with an increased risk for numerous severe health conditions and premature mortality, presenting a substantial public health risk and economic burden worldwide. Intensive lifestyle interventions, combining a calorie restriction diet with physical activity prescription and behavioral therapy, are the current non-medical gold-standard approach to promote weight loss and effectively treat obesity and manage associated health risks. However, there is considerable variability in the response to lifestyle-based weight loss treatment, and many participants do not achieve clinically significant weight loss of more than 5%, especially in the medium to long term. Identifying these individuals early can help improve longer-term weight loss success. Exercise interventions are particularly prone to considerable variability in the weight loss response, and most individuals (>75%) lose significantly less weight than expected based on the energy expended in exercise. The discrepancy between the expected and actual energy deficit (or weight loss) is called compensation. Energy (or weight) compensation is primarily caused by increases in energy intake in response to the exercise regimen, which partially or even fully negate the energy deficit created by exercise and thereby impede weight loss. While regular exercise yields many health benefits independent of weight loss, these benefits can be maximized when exercise results in sustained weight loss. And because the prospect of weight loss is the primary reason and motivator for many individuals to engage in any exercise- or lifestyle-based weight loss intervention, absent weight loss often leads to discontinuation of the program as it is perceived as unsuccessful. Therefore, identifying predictors of weight loss success is important to guide treatment and identify patients needing increased support for weight loss, such as a dietary intervention as an add-on to an exercise intervention or increased contact with interventionists and individualized behavioral strategies. As part of my habilitation work, we comprehensively analyzed baseline and dynamic, intervention-specific predictors of energy compensation and weight loss, ranging from acute exercise bouts to exercise interventions and a pragmatic lifestyle intervention. Finally, in an industry-funded project, we went beyond classic predictor analyses and examined whether specific genotype patterns modify diet effects on weight loss.

In Publication 1, we aimed to identify baseline predictors of post-exercise energy intake and compensation after acute exercise. For this, we utilized data from the EAT-FC study, a randomized crossover study comparing the effects of 45 min aerobic exercise vs. rest on post-exercise energy intake in a laboratory-based test meal. Biological (sex, body composition, appetite hormones) and behavioral (habitual exercise via prospective exercise log, eating behavior traits) characteristics were considered as potential predictors. We found that biological and behavioral characteristics differentially affect total and relative (compensatory) post-exercise energy intake in men and women. In men, fasting concentrations of appetite-regulating hormones (peptide YY and adiponectin) predicted total post-exercise energy intake, explaining 78% and 44% of the variance in post-exercise energy intake, while in women, only habitual exercise predicted total post-exercise energy intake, with more exercise protecting against compensatory eating. These results can help identify individuals who are more likely to (over-) compensate for the energy expended in exercise via increased post-exercise energy intake, allowing to deploy targeted countermeasures ahead of time.

In Publication 2, we went beyond acute exercise and aimed to determine whether habitual physical activity behavior predicts weight change, weight compensation, and changes in energy intake during a 24-week supervised aerobic exercise intervention. We utilized data from the Examination of Mechanisms of Exercise-induced Weight Compensation (E-MECHANIC) trial for these analyses. The primary aim of E-MECHANIC was to identify mechanisms of exercise-induced weight compensation by examining the effect of the two different doses of exercise training (moderate dose: 8 kcal/kg of body weight per week [KKW] of exercise-induced energy expenditure; high-dose: 20 KKW) on energy intake over a 24-week intervention period compared to a no-exercise control condition. In E-MECHANIC, most exercisers (82.6%) compensated (less weight loss than expected). Our results showed that, while there was substantial variability in the data, on average, lower baseline levels of moderate-to-vigorous physical activity (MVPA) were associated with less weight loss from exercise, higher compensation, and increased energy intake. Specifically, for every 15 min more of habitual MVPA at baseline, participants lost 0.23 kg more weight, compensated 0.20 kg less, and decreased daily energy intake by 22 kcal per day from baseline to follow-up. Consequently, individuals with lower habitual MVPA levels at baseline could benefit from an additional dietary intervention when participating in an exercise intervention for weight loss to prevent compensatory mechanisms such as increased energy intake.

In Publication 3, we shifted the focus away from classic baseline (participant) characteristics as predictors. Instead, we focused on dynamic interventions-specific factors as predictors of weight loss in an exercise intervention. While Publication 2 showed that habitual MVPA at baseline predicts weight loss, weight compensation, and changes in energy intake after six months of supervised exercise, the effect was only moderate, with habitual MVPA explaining 12%, 13%, and 21% of the variance in these outcomes. This leaves a substantial portion of the variance unexplained, as is often the case for behavioral baseline predictors. Dynamic, intervention-specific factors that are modifiable have the potential to be more reliable predictors of medium-to-long-term weight loss success, as has been shown for restrictive

dietary interventions, including large trials such as Look AHEAD (Action for Health in Diabetes). We used data from two supervised 6-month exercise intervention studies: E-MECHANIC and the Dose-Response to Exercise in Postmenopausal Women (DREW) study. Similar to E-MECHANIC, DREW compared the effects of different exercise doses (4 KKW, 8 KKW, and 12 KKW) on weight loss and compensation and further examined the dose-response relations between exercise training and cardiorespiratory fitness and cardiovascular disease risk factors. For our analyses, participants in both studies were divided into tertiles based on percent weight change from baseline to week 4, with tertiles 1 and 3 exhibiting the least and most initial weight loss, respectively. At month 6, weight loss was lower, and compensation was greater in tertile 1 than in tertile 3 in both studies. Further, changes in triglycerides and high-density lipoprotein (HDL) cholesterol were less favorable in tertile 1 compared to tertile 3. These results show that less initial weight loss was associated with longer-term attenuated weight loss and greater compensation during aerobic exercise training. Individuals with less initial weight loss during exercise may require additional intervention measures early on to decrease compensation and facilitate weight loss.

In Publication 4, an invited review article, we aimed to answer the question of why exercise by itself is a relatively ineffective method for weight loss despite playing an undisputedly important role in the prevention and management of obesity, summarizing our and others' previous work. We discuss energy-related aspects of weight loss as well as exercise effects on eating behavior, illustrating that a greater calorie deficit can be achieved more efficiently through calorie restriction compared to exercise and that exercise is often accompanied by compensatory eating, which slows or even negates weight loss. We also show that the preservation of fat-free mass through exercise, which contributes to slower weight loss compared to calorie restriction, is, in fact, beneficial as it supports better long-term appetite and energy balance regulation and can thereby prevent weight regain.

In Publication 5, we tested whether initial weight loss can predict longer-term weight loss in a 2-year pragmatic weight loss intervention using data from the Promoting Successful Weight Loss in Primary Care in Louisiana (PROPEL) trial. Contrary to previous large intensive lifestyle interventions typically conducted in academic health centers, PROPEL used a more pragmatic approach, delivering the intervention content to participants via trained health coaches embedded in primary-care clinics in weekly sessions initially and monthly sessions in months 7 through 24. A key intervention component in PROPEL was daily self-weighing and the incorporation of a personalized weight graph, which automatically plotted participants' weight data in relation to the expected individualized weight loss trajectory (10% at six months with lower [7.5%] and upper [12.5%] bounds). In addition to initial weight loss, we utilized these self-weighing and weight graph data for our prediction analyses. Specifically, we tested whether initial weight loss, the number of daily weights, and the number of adherent weights (i.e., on the expected weight loss trajectory) at 2, 4, and 8 weeks predicted medium-to-long-term weight loss at 6, 12, and 24 months of the intervention. Our results show that greater initial weight loss, daily self-weighing adherence, and adherence to the expected weight loss trajectory predicted weight loss at all time points, while initial weight loss remained the best

predictor in multivariable models. As expected, a longer initial timeframe and the shortest outcome timeframe (i.e., six months) generally yielded the highest predictive value; 2-week weight loss explained 15%, 11%, and 9% of the variance in weight loss at 6, 12, and 24 months, whereas 8-week weight loss explained 50%, 32%, and 16%. These results highlight the importance of initial weight loss for predicting long-term success and show that self-weighing and adherence to the expected weight loss trajectory can improve weight loss prediction beyond initial weight loss alone.

In Publication 6, we went beyond predictor analyses and aimed to identify the mediators of weight change during the PROPEL intervention. Understanding what factors drive weight loss in an intensive lifestyle intervention is important, as strategies and behaviors can be targeted and tested in future interventions to enhance the efficacy of weight-management programs. Specifically, we assessed whether self-reported eating behaviors (restraint, disinhibition), dietary intake (percentage fat intake, fruit/vegetable intake), physical activity, and weight- and health-related quality of life constructs mediated between-group (intervention group vs. control group) variations in weight change from baseline to month 12 and from month 12 to month 24. At 12 months, the intervention group lost 7.2 kg more weight compared to the control group, and improvements in disinhibition, percentage fat intake, physical activity, and subjective fatigue in the intervention group partially explained this between-group difference. Weight loss at 24 months was 5.4 kg greater in the intervention compared to the control group, showing that the intervention group (re-)gained 2.2 kg from months 12 to 24 compared to the control group. Change in fruit and vegetable intake partially explained this response, and no variables attenuated the weight regain of the intervention group. These results show that several psychological and behavioral variables mediated weight change during a 2-year pragmatic weight loss intervention, which could help refine weight management regimens in underserved patients with obesity.

In Publication 7, we examined genotype-diet interactions on weight loss. Instead of conducting a classic predictor analysis, we designed a personalized nutrition intervention, testing the effects of diets with different macronutrient compositions on weight loss among individuals with different macronutrient-responsive genotypes. Specifically, we identified participants as fat-responders or carbohydrate-responders via a genetic algorithm that considered their combined genotypes at ten genetic variants. The algorithm was based on the most recent literature, suggesting that participants with carbohydrate-responsive polymorphisms lose more weight on high-carbohydrate vs. high-fat diets and vice versa for those with fat-responsive polymorphisms. After identifying participants as fat-responders or carbohydrate-responders, they were randomized to either a calorie-restricted (daily deficit of ~750 kcal) high-fat or high-carbohydrate diet for 12 weeks. Dietitians delivered the weight loss intervention via 12 weekly diet-specific small group sessions and participants received daily diet-specific meal plans to self-prepare meals during the intervention period. Weight loss at 12 weeks was ~5 kg (5.5%) and did not differ between the genotype-concordant and genotype-discordant diets for the whole sample or when analyzing carbohydrate- and fast-responders separately. Similarly, results for secondary endpoints such as percent body fat, waist and hip circumference, or

blood pressure did not differ between the genotype-concordant and genotype-discordant diets. These results show that with the current ability to genotype participants as fat- or carbohydrate-responders, evidence does not support greater weight loss on genotype-concordant diets. This manuscript has been selected as an Editor's Highlight, which features the 50 best publications in *Nature Communications*.

Taken together, these results indicate that individual characteristics such as sex, fasting hormone levels, and habitual exercise play significant roles in post-exercise energy intake. The sex-dependent effects of fasting appetite hormone levels on acute post-exercise energy intake were particularly novel and should be further explored more longitudinally in future studies. Our results also highlight the importance of higher (pre-intervention) habitual MVPA for achieving weight loss during an exercise intervention, consistent with findings showing "unregulated" energy intake levels in individuals with low physical activity levels. Further, incorporating dynamic, intervention-specific predictors (i.e., initial weight loss), in addition to baseline characteristics, can substantially improve prediction models and identify individuals who may struggle with weight loss and compensation during both exercise- and diet-based lifestyle interventions to allow targeted countermeasures early on. These countermeasures should focus on psychological and other behavioral factors, such as disinhibition, dietary intake, and physical activity, to improve weight loss and prevent weight regain. Finally, specific genotypes that might enhance weight loss via a high-carbohydrate vs. high-fat diet are currently not supported by evidence, and the variability in weight loss is more likely explained by dietary adherence and other participant characteristics.

PART II – Technology-Assisted Food Intake Assessment

Accurate measurement of food intake is essential for assessing diet-health interactions in observational studies, examining and monitoring the effects of dietary changes (e.g., adherence to calorie restriction targets) on obesity treatment and health, and informing public health policies based on empirical data. To date, self-report methods (e.g., food records, food recalls, and food frequency questionnaires) are still commonly used in epidemiology and clinical research settings despite the evident inaccuracy of these methods in assessing energy and nutrient intake caused by (un-)intentional under- or over-reporting of foods, portion size estimation errors, and reactivity due to awareness of being measured. Technology- and particularly image-assisted methods of food intake assessment that quantify food intake via active or passive image capture and automated or semi-automated analysis have gained popularity in recent years, and these methods have addressed many of the limitations of traditional self-report. Another emerging approach for remote meal detection is continuous glucose monitoring (CGM), which measures glucose concentrations in the interstitial fluid as a proxy for blood glucose levels and offers the advantage of being completely unobtrusive. As part of my habilitation work, we comprehensively reviewed image-assisted methods of food intake assessment, including their accuracy, feasibility, and acceptability in different research settings and for day-to-day monitoring of dietary intake. We further reviewed CGM-based approaches regarding their accuracy in meal detection as standalone and combined methods.

Finally, we evaluated the applicability of these image- and sensor-based methods for Just-In-Time Adaptive Interventions (JITAIs) that aim to detect dietary lapses and non-adherence in (near) real-time and respond with intervention content delivery in the moment when it is most needed and the patient is likely to be (most) receptive.

In Publication 8, we conducted a literature review of the validity and feasibility of image-assisted methods for dietary assessment. Our review showed that in free-living conditions, active smartphone-based image capture of food selection and plate waste can produce accurate energy intake estimates, though accuracy is not guaranteed and relies heavily on the quality of analysis by trained human raters. A limitation of active image capture that remains is intentional and/or unintentional under-reporting of foods due to social desirability or forgetfulness and reactivity, similar to traditional self-report. When an accurate assessment of habitual food intake is not the main objective, the reactivity (or self-monitoring) effect can, in fact, be beneficial, for example, when the goal is to change dietary behaviors, and many weight loss studies have used food recording as an effective behavior change strategy. Passive image capture via wearable cameras is promising and aims to reduce user burden. However, only pilot data with limited validity were available at the time the review was conducted, and these methods remained obtrusive and cumbersome. Further, the technology required for automated and semi-automated food recognition and portion size estimation was still in its infancy at that point, and fully automated food intake assessment with acceptable precision was not a reality. In general, analysis by human raters was more accurate and less variable than (semi-)automated image analysis, not least because existing nutrient databases could be utilized. With recent advances in artificial intelligence and machine-learning technology, methods with passive image capture and automated analysis have promise to dramatically improve in the future, allowing accurate automated in-the-moment feedback on food intake data to patients and researchers.

In Publication 9, we conducted an online survey (n=1959) to examine participants' preference, expected burden, and willingness to use image-based compared to traditional pen-and-paper self-report methods for food and alcohol intake assessment. While validity, reproducibility, usability, and feasibility of the method selected for food intake assessment are essential for collecting high-quality data, participants' preference, acceptability, and perceived burden of a specific method also play an important role. These factors directly influence compliance and, thereby, data quality, which is particularly important in dietary weight loss interventions. We assessed participants' preference, expected burden, and willingness to use the following four methods to record their food/drink intake for three days as part of a clinical or study setting: food/drink record, 24h recall, Remote Food Photography Method© (RFPM), and a novel app (PortionSize®) that allows the in-app portion size estimation of foods/drinks by the user. For both food intake and alcohol consumption, the greatest percentage of participants rated the expected burden of the RFPM as low, followed by PortionSize®, food/drink records, and the 24h recall. Preference for these methods mirrored the ratings of expected burden, and correlations between low expected burden and high preference were strong for all methods. The willingness to use the RFPM was greater than that of PortionSize® and the 24h recalls,

but it was not different from food records. Because preference for and expected burden of a specific food intake assessment method likely affect participants' compliance over time and, thereby, data quality, these results can help inform the method selection in conjunction with the reliability, validity, and usability/feasibility data of these methods.

In Publication 10, we conducted a scoping review on the accuracy of CGM for automatic meal detection. CGM is a particularly attractive option for passive, objective meal detection as it is much less obtrusive than current passive image-based approaches and consequently less prone to reactivity. Further, user and analysis burden are reduced compared to food photography, leading to better long-term compliance and remote data capture. Our review found several promising CGM-based approaches for automatic meal detection, with some approaches showing close to 100% sensitivity. Meal detection times ranged between 9 and 45 min, with the majority being between 20 and 40 min. These detection times are likely too long for true in-the-moment intervention delivery (i.e., while the participant is still eating); however, this is likely not the goal. More realistically, intervention strategies would aim at altering behavior at a subsequent meal for which the current detection times are more than sufficient. However, the overall heterogeneity of studies (e.g., age range, controlled vs. free-living settings, participants with vs. without type 1 diabetes), algorithms (e.g., input data being CGM only vs. additional data such as accelerometer, body core temperature, heart rate), performance (outcome) metrics, and methodological issues (e.g., self-report as ground truth method) prevent clear recommendations of a single approach and specific cases will dictate the most suitable approach.

Collectively, these results show that image-based approaches with human rater-based analysis remain the state-of-the-art approach for the objective assessment of food intake in many settings, particularly of food items and portion size. These methods have excellent validity in estimating energy and macro-/micronutrient intake, and they are at the same time less burdensome and show higher patient acceptability than traditional pen-and-paper methods. Passive image capture, while promising, has, to date, limited validity, and the current methods are too obtrusive and cumbersome. While not yet accurate enough to serve as a standalone tool for food intake assessment, including timing, identification, and quantification, CGM can help further minimize under-reporting by reducing missing data, and it can be used to detect meal timings more accurately. Going forward, pairing active image-based approaches to assess *what* individuals eat with objective, passive methods such as CGM to determine *when* they eat is a promising prospect. Using mathematical modeling to integrate the multi-sensor data will likely yield further improvements in the accuracy of food intake assessment and monitoring, with automated in-the-moment feedback to patients and researchers being a realistic vision for the near future.

PART III – Other Work

Part III highlights further research of my habilitation work and postdoctoral training, such as examining the health effects of large lifestyle interventions beyond weight loss, discussing

challenges in defining adherence to calorie restriction goals in weight loss interventions, and gaining further expertise in food intake research.

In Publication 11, we showed that the PROPEL intervention, as a patient-centered, pragmatic, and scalable obesity treatment program delivered by health coaches to over 800 patients in primary care, can elicit clinically meaningful improvements in cardiometabolic health. We found that fasting glucose decreased, and HDL cholesterol increased in the intervention group compared with the control group, with clinically significant improvements in both parameters. In addition to traditional cardiometabolic risk markers, we applied a novel, validated continuous metabolic syndrome severity score to better detect a worsening or improving condition over time compared to the conventional binary metabolic syndrome classification. Metabolic syndrome severity decreased significantly in the intervention group compared to control at 12 and 24 months. Blood pressure and other blood lipids did not change throughout the intervention. These results show that by removing barriers to participation, PROPEL offered a viable and more successful primary care-based treatment option for obesity-related comorbidities compared to the existing model in the United States, which is supported by Medicare/Medicaid and offers insufficient access, particularly to underserved populations. During the last six months of the trial, I oversaw the intervention sessions at 6 of the 18 participating clinics and led weekly meetings with the health coach. This manuscript was selected as *Circulation*'s paper of the month (02/2021) by the American Heart Association.

In Publication 12, we delved deeper into the health effects of structured exercise. Specifically, using the E-MECHANIC dataset, we examined the impact of exercise dose on changes in central adiposity and its relationship with exercise-induced compensation. Exercise has been shown to decrease central adiposity, though the influence of exercise dose is poorly understood. Because central adiposity is more strongly linked to metabolic diseases than fat stored in other regions, reductions in central adiposity are particularly important to improve metabolic health. Our results showed that both the 8 KKW exercise dose, which is recommended for health, and the 20 KKW exercise dose, which is recommended for weight loss and maintenance, led to similar, negligible changes in waist circumference and visceral adipose tissue (via dual-energy X-ray absorptiometry). However, exercisers who compensated (82.6% of all participants) exhibited small (unfavorable) increases in waist circumference and visceral adipose tissue compared to those who did not compensate, and greater subjective desire to eat (via visual analog scale) directly predicted changes in visceral adipose tissue during exercise. To enhance reductions in central adiposity and improve metabolic health, exercise-induced compensation should be prevented and treated potentially through strategies managing appetite and compensatory food behaviors. This manuscript was selected as the *Journal of Clinical Endocrinology & Metabolism*'s featured article of the week in calendar week 11 of 2024.

In Publication 13, we aimed to tackle a common challenge of dietary weight loss interventions: defining successful adherence to calorie restriction goals using data from the Comprehensive Assessment of Long-term Effects of Reducing Intake of Energy phase 2 trial (CALERIE 2). A recent approach for estimating dietary adherence in real-time is using body weight as a proxy.

For this, the participant's actual (daily) weights are compared to expected weights over time, and if the actual body weight reflects the expected weight, adherence to the calorie restriction goal can be inferred. Because there are natural day-to-day fluctuations in body weight independent of calorie intake, participants are provided with a range (i.e., bounds around the goal weight; cf. Publication 5) of acceptable body weights ("zone of adherence") that reflects calorie restriction adherence. Weights above the zone are considered non-adherent, and additional intervention strategies are needed to help participants better restrict their energy intake. Weights below the zone indicate a too-restrictive eating behavior, and efforts are needed to increase energy intake. In our analyses, we aimed to determine the level of calorie restriction associated with the zone of adherence in CALERIE 2 by utilizing a validated weight loss calculator (NIDDK Body Weight Planner) and to determine if participants' actual level of calorie restriction was within the zone of adherence by using the intake-balance method. Our results show that the upper bound of the zone of adherence reflected a percent calorie restriction that was well below CALERIE 2's 25% calorie restriction goal at months 12 (13.7%) and 24 (10.4%) though the average level of calorie restriction achieved by participants (determined post-hoc via the intake-balance method) was within the zone at months 12 (15.2%) and 24 (11.9%). The lower bound of the zone nearly reflected 25% calorie restriction only at month 12, and by month 24, the lower bound of the zone reflected ~19% calorie restriction. The fact that the zone of adherence in CALERIE 2 considered calorie restriction far less than the 25% goal as being adherent should be considered in designing future calorie restriction interventions and strategies to promote adherence. For example, lower adherence zones should be used to achieve higher levels of calorie restriction when participants' weights are in the zone of adherence.

In Publication 14, as part of my postdoctoral training in food intake research, I designed a randomized crossover study measuring food intake during a test meal with acute e-cigarette use ("vaping") as a manipulator. Because nicotine consumption (by itself or via cigarettes) has been shown to reduce appetite and, ultimately, body weight via brain and hormonal mechanisms, weight control is a common motive to start smoking, and the fear of post-cessation weight gain is frequently cited as a barrier to smoking cessation. It is to date unclear if e-cigarettes have similar effects on appetite and energy intake, despite being actively marketed as methods of weight control and consumed by individuals with the intention to lose weight. In our study, participants completed two randomly ordered 20-min clinical lab sessions (20 puffs of an e-cigarette [5% nicotine] vs. access to an uncharged e-cigarette) and a buffet-style meal with 21 food/drink items about 40 min after product administration. Our results show that while acute e-cigarette use increased subjective feelings of satiety and decreased subjective feelings of hunger, these subjective effects did not translate into reductions in acute energy intake. These findings are inconsistent with the smoking-related energy intake suppression observed in previous studies, and greater variability in energy intake and greater diversity in race/ethnicity and sex in our study compared to other studies may have influenced the effects. Nevertheless, while further research is needed, these findings could inform actionable approaches toward enhancing e-cigarette cessation efforts, with a particular focus on addressing perceptions of e-cigarette-related appetite control and/or weight management.

PART I

Predictors of Energy Compensation and Weight Loss in Lifestyle Interventions

Publication 1

Biological and behavioral predictors of relative energy intake after acute exercise.

Christoph Höchsmann [1], Safiya E Beckford [2], Jeffrey A French [3], Julie B Boron [4], Jeffrey R Stevens [5], Karsten Koehler [1]

[1] Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany

[2] Department of Nutrition and Health Sciences, University of Nebraska-Lincoln, Lincoln, NE, USA

[3] Program in Neuroscience and Behavior, University of Nebraska at Omaha, Omaha, NE, USA

[4] Department of Gerontology, University of Nebraska at Omaha, Omaha, NE, USA

[5] Department of Psychology, University of Nebraska-Lincoln, Lincoln, NE, USA

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
Appetite, 184: 106520	2023	10.1016/j.appet.2023.106520	5.4 (2022)	Q1	Scopus: 0 Google Scholar: 0

Author contribution:

First author; conducted the statistical analyses, drafted the manuscript, and created tables and figures.

Appetite 184 (2023) 106520

Contents lists available at ScienceDirect

Appetite

journal homepage: www.elsevier.com/locate/appet

Biological and behavioral predictors of relative energy intake after acute exercise

Christoph Höchsmann [a], Safiya E. Beckford [b], Jeffrey A. French [c], Julie B. Boron [d], Jeffrey R. Stevens [e], Karsten Koehler [a,*]

[a] Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany
[b] Department of Nutrition and Health Sciences, University of Nebraska-Lincoln, Lincoln, NE, USA
[c] Program in Neuroscience and Behavior, University of Nebraska Omaha, Omaha, NE, USA
[d] Department of Gerontology, University of Nebraska Omaha, Omaha, NE, USA
[e] Department of Psychology, University of Nebraska-Lincoln, Lincoln, NE, USA

ARTICLE INFO

Handling Editor: Dr. S. Higgs

ABSTRACT

Energy intake in the post-exercise state is highly variable and compensatory eating – i.e., (over-) compensation of the expended energy via increased post-exercise energy intake – occurs in some individuals but not others. We aimed to identify predictors of post-exercise energy intake and compensation. In a randomized crossover design, 57 healthy participants (21.7 [SD = 2.5] years; 23.7 [SD = 2.3] kg/m^2, 75% White, 54% female) completed two laboratory-based test-meals following (1) 45-min exercise and (2) 45-min rest (control). We assessed associations between biological (sex, body composition, appetite hormones) and behavioral (habitual exercise via prospective exercise log, eating behavior traits) characteristics at baseline and total energy intake, relative energy intake (intake – exercise expenditure), and the difference between post-exercise and post-rest intake. We found a differential impact of biological and behavioral characteristics on total post-exercise energy intake in men and women. In men, only fasting (baseline) concentrations of appetite-regulating hormones (peptide YY [PYY, $\beta = 0.88$, $P < 0.001$] and adiponectin [$\beta = 0.66$, $P = 0.005$] predicted total post-exercise energy intake, while in women, only habitual exercise ($\beta = -0.44$, $P = 0.017$) predicted total post-exercise energy intake. Predictors of relative intake were almost identical to those of total intake. The difference in energy intake between exercise and rest was associated with VO_{2peak} ($\beta = -0.45$, $P = 0.020$), fasting PYY ($\beta = 0.53$, $P = 0.036$), and fasting adiponectin ($\beta = 0.57$, $P = 0.021$) in men but not women (all $P > 0.51$). Our results show that biological and behavioral characteristics differentially affect total and relative post-exercise energy intake in men and women. This may help identify individuals who are more likely to compensate for the energy expended in exercise. Targeted countermeasures to prevent compensatory energy intake after exercise should take the demonstrated sex differences into account.

1. Introduction

Regular physical activity (PA) and exercise are recommended as methods of weight control; however, the effectiveness of exercise by itself as a method for weight loss is highly variable (Church et al., 2009; Martin et al., 2019). This is primarily due to the substantial variability in acute and chronic post-exercise energy intake. While some individuals show reduced energy intake post-exercise, allowing for an exercise-induced energy deficit, others show a compensatory increase in energy intake which negates the potential for exercise to promote negative energy balance and subsequent weight loss (Dorling et al., 2018; King et al., 2017). Post-exercise energy intake is affected by a variety of behavioral and biological variables.

For example, higher levels of habitual PA are associated with improved energy balance regulation and maintenance of healthy body weight (Church & Martin, 2018; Edholm, 1956; Mayer et al., 1956). Further, higher levels of habitual moderate-to-vigorous PA before an exercise-based weight loss intervention are associated with greater weight loss and importantly less weight compensation (i.e., less weight loss than expected based on measured energy expenditure) (Höchsmann

* Corresponding author. Technical University of Munich Connollystraße, 32 80809 Munich, Germany.
E-mail address: Karsten.Koehler@tum.de (K. Koehler).

https://doi.org/10.1016/j.appet.2023.106520
Received 2 March 2022; Received in revised form 17 January 2023; Accepted 28 February 2023
Available online 1 March 2023

et al., 2020a). The effects of habitual exercise on appetite and acute energy intake are, however, less clear and conflicting. Some studies have shown increases in subjective appetite following exercise training, whereas others have reported negligible effects or even reductions in appetite and energy intake after training (Dorling et al., 2018). The effects of habitual PA on acute exercise-induced energy compensation are to date largely unknown.

Eating behavior traits such as cognitive restraint, uncontrolled eating, and emotional eating are associated with unintentional overeating in the presence of food (Feig et al., 2018; French et al., 2012). Over time, these eating behavior traits can lead to a continued positive energy balance and consequently contribute to weight gain (Feig et al., 2018; French et al., 2012). It is unknown how these eating behavior traits affect energy intake (and compensation) after acute exercise and if they override the acute and transient exercise-induced suppression in appetite (exercise-induced anorexia) (Dorling et al., 2018; King et al., 1994).

Appetite-regulating hormones such as acylated ghrelin (appetite-stimulating), peptide YY (PYY), and glucagon-like peptide 1 (GLP-1) – the latter two are appetite-suppressing – modulate feelings of satiation on a meal-by-meal basis (Dorling et al., 2018; Murphy & Bloom, 2006). Acute bouts of moderate-to-vigorous aerobic exercise have been shown to suppress circulating concentrations of acylated ghrelin and simultaneously elevate concentrations of PYY and GLP-1, (Broom et al., 2007; Broom et al., 2009; Martins et al., 2007; Balaguera-Cortes et al., 2011) driving the aforementioned transient exercise-induced anorexia. In the fasted state, PYY concentrations are inversely correlated with markers of adiposity such as BMI and body fat (Guo et al., 2006), and fasting PYY concentrations have been shown to increase after exercise training-related weight (fat mass) loss (Jones et al., 2009), with suggested benefits for satiety regulation. Similarly, fasting GLP-1 concentrations have been reported to be lower in individuals with obesity compared to healthy individuals (Alssema et al., 2013; Ranganath et al., 1996), and greater fasting GLP-1 concentrations are associated with reduced carbohydrate but not total energy intake during an *ad libitum* test-meal (Basolo et al., 2019). It is to date unknown if fasting concentrations of appetite-regulating hormones (measured before an acute exercise bout) affect energy intake and compensation after acute exercise. While the effects of these various behavioral and biological factors on post-exercise food intake have been examined, to our knowledge it has not yet been investigated whether it is possible to predict a person's post-exercise food intake based on these (and other) participant characteristics at baseline. Therefore, the overall objective of the present study was to identify predictors of post-exercise energy intake from a wide range of behavioral and biological characteristics that can be relatively easily assessed in a clinical setting before individuals participate in exercise-based weight loss interventions. If successful, this approach would allow to identify possible compensators before these individuals engage in exercise programs for weight loss and to deliver targeted behavioral add-ons to the exercise intervention (e.g., cognitive-behavioral strategies, additional calorie restriction, preparation of a post-exercise meal ahead of time, or consumption of a small meal before exercise) to help prevent compensatory eating ahead of time and maximize the weight loss potential of exercise. To achieve this objective, we aimed to assess predictors of (1) total energy intake (kcal) during an *ad libitum* test-meal following a 45-min aerobic exercise bout and (2) relative energy intake (kcal), i.e., test-meal energy intake relative to the energy expended during the exercise bout. Additionally, as control, we aimed to identify predictors of *ad libitum* test-meal intake following 45 min of rest as well as of the difference in energy intake between the post-exercise test-meal and test-meal following the rest condition. We investigated the contributions of sex, anthropometrics, behavioral characteristics, and physiological/endocrine factors at baseline to compensatory *ad libitum* post-exercise energy intake. We hypothesized that greater levels of habitual physical activity and exercise would be associated with lower compensatory post-exercise energy intake. In explorative analyses, we additionally examined the predictive effects of eating behavior traits and fasting concentrations of appetite-regulating hormones on compensatory post-exercise energy intake.

2. Materials and methods

2.1. Study design and participants

This report is a secondary analysis of the EAT-FC (Exercise, Appetite, and Temporal Food Choices) randomized crossover study (Koehler et al., 2021). In the EAT-FC study, following two preliminary assessment visits, two study conditions were completed in random order on two separate days at least 3 days apart: (1) one 45-min exercise bout and (2) one rest period of identical duration. The study used block randomization (block size of 4) to allocate participants to the study conditions. Volunteers for this study were recruited from the University of Nebraska and its surrounding communities via fliers and word-of-mouth. Men and women were eligible if they were 19–29 years old, had a body mass index (BMI) of 18.5–29.9 kg/m^2, were weight stable ($\leq$2.5 kg weight change during the past six months), and exercised regularly ($\geq$1 bout/week). Exclusion criteria included pregnancy, smoking, any medical condition or use of medication that could affect appetite or present any contraindications to exercise, a history of or current eating disorder, or a self-reported inability to exercise at a moderate intensity for 45 min. Further, participants who were allergic to or strongly disliked the food offered during the *ad libitum* test meal were excluded (Koehler et al., 2021). All study procedures were approved by the University of Nebraska-Lincoln's Institutional Review Board (project number 17239) and written informed consent was obtained from all participants before participation in the study.

2.2. Measures

Anthropometric data, PA behavior, cardiorespiratory fitness, and eating behavior traits were assessed during two preliminary assessment visits that occurred before participation in the study conditions, and appetite-regulating hormones were measured at each study condition visit immediately after arrival at the laboratory.

2.2.1. Anthropometric data

Weight and height were measured using a digital scale and stadiometer (Seca, Hamburg, Germany). Total body fat (%) and fat-free mass (FFM, kg) were estimated via a 7-site skinfold assessment.

2.2.2. Physical activity behavior

PA behavior, and specifically moderate-to-vigorous PA (MVPA, min/week, cut-points according to Freedson (Freedson et al., 1998)), was assessed over seven days using accelerometry (hip-worn GT3X+, Actigraph, Pensacola, FL; epoch length 10 s). Participants were instructed to wear the GT3X + continuously throughout the 7-day monitoring period and to only remove it for swimming or taking a shower. Only full days of data, defined as a wear time of $\geq$95% (equating to 22 h and 48 min or 1368 min), were included in the analyses. Additionally, as it has been reported before that Actigraph devices are inaccurate at recording activities such as strength training and cycling (Berntsen et al., 2010; Herman Hansen et al., 2014; Höchsmann et al., 2020b), we instructed participants to prospectively record their habitual exercise (min/week and days/week) over the same period using an exercise log.

2.2.3. Cardiorespiratory fitness

Peak oxygen uptake (VO_{2peak}) was measured using an incremental all-out exercise test on a bicycle ergometer (LC6, Monark, Vansbro, Sweden). Participants began cycling at a resistance of 60 W for 3 min, and the work rate was increased by 35 W every 3 min until exhaustion (Achten & Jeukendrup, 2003). Maximal exhaustion was accepted when

at least two of the following were met: (1) Heart rate of ≥90% of age-predicted maximal heart rate, (2) a respiratory exchange ratio ≥1.1, (3) rating of perceived exertion ≥19,[28] (4) a plateau in oxygen uptake despite the increasing workload. Throughout the test, respiratory gas parameters were analyzed breath by breath (Quark CPET, COSMED, Rome, Italy) and heart rate was monitored through telemetry (Polar, Kempele, Finland).

2.2.4. Eating behavior traits

Cognitive restraint, uncontrolled eating, and emotional eating were assessed with the revised 18-item Three-Factor Eating Questionnaire (TFEQ-R18v2). The TFEQ-R18v2 is a shortened version of the original well-validated 51-item TFEQ by Stunkard and Messick (Stunkard & Messick, 1985), which has demonstrated improved psychometric properties, minimized floor and ceiling effects in the emotional eating domain, and improved internal consistency in the cognitive restraint domain compared to the earlier shortened versions of the TFEQ (TFEQ-R18 and TFEQ-R21), with an overall robust factor structure and good reliability in two large North American samples (Cappelleri et al., 2009; Karlsson et al., 2000).

2.2.5. Appetite-regulating hormones

Fasting plasma concentrations of total GLP-1 (1-37a), acylated ghrelin, PYY (3-36), and adiponectin were measured at each study condition visit immediately after arrival to the laboratory before participants received a standardized breakfast and continued with the initial 30-min rest period. Whole-blood samples were collected into ethylenediaminetetraacetic acid (EDTA) tubes from participants in a seated position. A protease inhibitor (aprotinin; Sigma Aldrich, St. Louis, MO) was added to PYY and ghrelin samples. Immediately after collection, the EDTA tubes were placed on ice for 15 min and then centrifuged at 1800×*g* for 10 min at +4 °C. Subsequently, plasma fractions were aliquoted and stored at −80 °C until analysis. Enzyme-linked immunosorbent assays (ELISAs) were used to measure concentrations of PYY (Millipore Sigma, Burlington, MA; inter-assay coefficient of variation [CV]: 6%; intra-assay CV: 7%), GLP-1 (Invitrogen™, Carlsbad, CA; inter-assay CV: <12%; intra-assay CV: <10%), ghrelin (Invitrogen™, Carlsbad, CA; inter-assay CV: 8.5%; intra-assay CV: 6%), and adiponectin (Invitrogen™, Carlsbad, CA; inter-assay CV: 3.1%; intra-assay CV: 4.2%).

2.3. Study conditions

On the day of each study condition, participants arrived at the lab between 06:30 and 10:00 (appointments scheduled 30–60 min after habitual wake-up time, identical time at each visit), following an overnight fast and abstinence from alcohol for at least 24 h. Participants further refrained from exercise and strenuous physical activity the day before and the morning of their visits, with compliance monitored via accelerometry (GT3X+, Actigraph, Pensacola, FL). During their first study condition visit, participants completed a 24-h diet recall using an Automated Self-Administered 24-h Dietary Assessment Tool (ASA24, National Cancer Institute, Bethesda, MD, USA). Participants received a copy of their recall after the visit and they were instructed to replicate the diet as closely as possible on the day before their second study condition visit. At each study condition visit, participants were provided with a small standardized breakfast (commercially available cereal bar [240 kcal] and 8 ounces of bottled water) upon arrival and instructed to rest for 30 min in a seated position.

2.3.1. Exercise condition

Following the initial 30-min rest, participants exercised on a bicycle ergometer (LC6, Monark, Vansbro, Sweden) for 45 min at an intensity of 60% of their VO_{2peak}. Heart rate and ratings of perceived exertion (Borg, 1982) were monitored at regular intervals throughout the exercise bout. After completion of the exercise bout, participants rested for another 30 min before being offered the test meal.

2.3.2. Rest condition

For the resting condition, participants were instructed to sit quietly in a chair for 45 min, following the initial 30-min rest period. To ensure an overall identical timing and sequence of the two study condition visits, participants rested for an additional 30 min after the 45-min rest condition before being offered the test meal. Throughout both visits, participants were allowed to listen to music or watch pre-approved TV programs that did not contain any images of or references to food.

2.3.3. Ad libitum test-meal

Thirty minutes after each study condition (exercise or rest), participants were offered an identical single-item *ad libitum* test meal. The test meal (frozen family-size cheese pizza, Hy-Vee, West Des Moines, IA) was prepared by the study staff, and participants were offered the entire pizza (~3200 kcal, above energy needs) at once. The test meal was consumed in a separate room and under supervision, and cell phone use was restricted. Participants were instructed to eat as much or as little of the test meal as they would like and to make sure to eat the pre-cut pizza slices evenly (i.e., not leave/discard the crust or take the cheese off, etc.). Pre- and post-meal weights (grams) were recorded, with the difference in weight representing food intake. Gram weights were converted to energy intake (kcal) using the pizza's nutrition label.

2.4. Statistical analyses

The distribution of variables was verified by visual inspection of histograms and quantile-quantile plots of the residuals. The exclusion of outliers (≤2 outliers for all models) did not change the results meaningfully; therefore, the models including outliers are reported. Descriptive data are reported as mean and standard deviation (SD). We used simple linear regression models to estimate the association between anthropometric characteristics (weight, BMI, FFM, percent body fat), physiological characteristics (VO_{2peak}, maximal power, and fasting concentrations of appetite-regulating hormones such as GLP-1, ghrelin, PYY, and adiponectin), and behavioral characteristics (habitual exercise min/week and days/week, MVPA, and eating behavior traits) and energy intake during an *ad libitum* single-item test meal. Specifically, we used the following four variables as dependent variables in our models: (1) post-exercise energy intake (kcal), (2) relative energy intake, which was defined as post-exercise energy intake [kcal] – energy expenditure during exercise session [kcal], (3) energy intake following the rest condition (kcal), and (4) the difference in energy intake between the post-exercise and the post-rest test meal (post-exercise energy intake [kcal] – post-rest energy intake [kcal]). Because it has been demonstrated in several studies that FFM is a predictor of meal size and single meal food intake (Blundell et al., 2012; Hopkins et al., 2016; Weise et al., 2014), we included FFM as a covariate in sensitivity analyses; however, results did not differ meaningfully and we consequently report the results without FFM as a covariate. In additional analyses, we performed multiple regression analyses using the significant single predictors for each outcome variable. Further, results for eating behavior traits (cognitive restraint, uncontrolled eating, and emotional eating) based on the TFEQ-R18 and TFEQ-R21 did not differ meaningfully from the results of the TFEQ-R18v2 presented herein. Because the TFEQ-R18v2 has been validated in North American samples and improved internal consistency has been reported, as described above, only these results are reported. We used SPSS Statistics for Windows, version 27 (IBM Corp., Armonk, NY) for our analyses, and results were considered significant at $P < 0.05$.

3. Results

3.1. Participant characteristics and energy intake following each study condition

Sixty-five participants were enrolled in the study. Eight participants were excluded (intensity of exercise session not at 60% VO_{2peak} [n = 6], no exercise data [n = 2]); hence 57 participants were included in our analyses. Baseline characteristics of all included participants (mean age 21.7 [SD = 2.5] years, mean BMI 23.7 [SD = 2.3], 75% White, 54% female), as well as average energy expenditure during the exercise session and energy intake following each study condition, are presented in Table 1. Energy expenditure during the exercise session was greater in men (400 [SD = 85] kcal) compared to women (296 [SD = 46] kcal, $P < 0.001$).

3.2. Predictors of total post-exercise energy intake

Total post-exercise energy intake was inversely associated with habitual exercise minutes ($\beta = -0.29$, $P = 0.032$; Table 2, Fig. 1A) and positively associated with FFM ($\beta = 0.30$, $P = 0.025$; Table 2) and fasting concentrations of PYY ($\beta = 0.39$, $P = 0.015$; Table 2, Fig. 1D). We also found a sex effect, as men consumed on average 261.9 kcal more than women ($P = 0.015$). After stratifying by sex, PYY ($\beta = 0.88$, $P < 0.001$) and additionally adiponectin ($\beta = 0.66$, $P = 0.005$, Table 2, Fig. 1G) were significant predictors of total post-exercise energy intake only in men, while habitual exercise minutes ($\beta = -0.44$, $P = 0.017$) were a significant predictor of total post-exercise energy intake only in women.

3.3. Predictors of relative energy intake

Similar to total post-exercise energy intake, relative energy intake was inversely associated with habitual exercise minutes ($\beta = -0.31$, $P = 0.024$; Table 3, Fig. 1B) and positively associated with fasting concentrations of PYY ($\beta = 0.37$, $P = 0.021$; Table 3, Fig. 1E). Similar to total post-exercise energy intake, PYY ($\beta = 0.85$, $P < 0.001$) and additionally adiponectin ($\beta = 0.69$, $P = 0.003$; Table 3, Fig. 1H) were significant predictors of relative energy intake only in men, while habitual exercise (min/week: $\beta = -0.44$, $P = 0.016$; days/week: $\beta = -0.39$, $P = 0.032$) and additionally VO_{2peak} (relative: $\beta = -0.36$, $P = 0.044$; absolute: $\beta = -0.42$; $P = 0.020$; Table 3) were significant predictors of relative energy intake only in women.

3.4. Predictors of post-rest energy intake

Energy intake after the rest condition was positively associated with weight ($\beta = 0.35$, $P = 0.008$), FFM ($\beta = 0.38$, $P = 0.004$), and aerobic fitness as measured by absolute VO_{2peak} ($\beta = 0.43$, $P < 0.001$) and maximal power during the exercise test ($\beta = 0.43$, $P < 0.001$; Supplemental Table 1). Similar to post-exercise, post-rest energy intake differed by sex ($P < 0.001$). Absolute VO_{2peak} ($\beta = 0.53$, $P = 0.006$) and maximal power ($\beta = 0.54$, $P = 0.005$) were only associated with *ad*libitum energy intake in men, while in women, only habitual exercise minutes per week were a significant predictor of *adlibitum* energy intake ($\beta = -0.37$, P=0.048).

3.5. Predictors of the difference between post-exercise and post-rest energy intake

The difference in total energy intake between exercise and rest was inversely associated with aerobic fitness as measured by relative ($\beta = -0.31$, $P = 0.020$) and absolute ($\beta = -0.35$, $P = 0.008$) VO_{2peak}. The difference between exercise and rest was also positively associated with fasting PYY concentrations ($\beta = 0.33$, $P = 0.038$; Supplemental Table 2, Fig. 1F). Notably, significant associations were driven by men, and they were not significant for women (Supplemental Table 2). In men, above a VO_{2peak} cut point of 40.9 mL/kg/min (3.0 L/min), post-rest energy intake was greater than post-exercise energy intake, while below the cut point, post-exercise energy intake was greater than post-rest energy intake. For PYY in men, post-exercise energy intake was greater than post-rest energy intake above the cut point of 118.6 pg/mL, while below the cut point, post-rest energy intake was greater than post-exercise

Table 1
Participant characteristics.

	All (N = 57)		Men (n = 26)		Women (n = 31)	
Race/Ethinicity, n (%)						
White	43	(75.4)	18	(69.3)	25	(80.6)
African American	9	(15.8)	6	(23.1)	3	(9.7)
Asian	4	(7.0)	1	(3.8)	3	(9.7)
Other	1	(1.8)	1	(3.8)	0	(0.0)
	Mean	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**
Age, years	21.7	(2.5)	21.4	(2.4)	21.9	(2.6)
Weight, kg	68.7	(10.2)	73.6	(11.3)	64.6	(7.0)
BMI, kg/m^2	23.7	(2.3)	23.8	(2.7)	23.5	(2.1)
Fat-free mass, kg	59.6	(9.0)	66.0	(8.9)	54.2	(4.6)
Total body fat, %	13.2	(6.0)	9.9	(5.4)	15.9	(5.1)
Physical activity behavior and cardiorespiratory fitness						
Total habitual exercise, min/week [a]	245.9	(181.2)	236.0	(137.7)	254.4	(213.9)
Habitual exercise days, days/week	3.4	(1.9)	3.7	(1.9)	3.3	(2.0)
MVPA, min/week	332.4	(145.7)	350.4	(159.9)	317.3	(133.4)
Relative VO_{2peak}, mL/kg/min	37.4	(6.2)	40.6	(5.8)	34.7	(5.2)
Absolute VO_{2peak}, L/min	2.6	(0.6)	2.3	(0.6)	2.2	(0.3)
Maximal power, W [b]	220.6	(48.5)	248.5	(47.6)	196.5	(34.8)
Eating behavior traits						
Cognitive Restraint via TFEQ-R18v2	5.9	(2.2)	5.7	(2.6)	6.0	(1.8)
Uncontrolled Eating via TFEQ-R18v2	17.3	(4.7)	17.6	(4.9)	17.0	(4.5)
Emotional Eating via TFEQ-R18v2	9.8	(3.3)	9.2	(2.7)	10.2	(3.7)
Appetite-regulating hormones [c]						
GLP-1, pg/mL [d]	10.9	(5.2)	10.0	(4.4)	11.4	(5.8)
PYY, pg/mL [d]	110.8	(47.0)	110.4	(46.5)	111.1	(48.4)
Ghrelin, pg/mL [e]	865.2	(393.0)	807.3	(435.0)	906.5	(365.3)
Adiponectin, ng/mL [f]	11.7	(7.1)	8.9	(6.4)	13.8	(6.9)
Exercise session						
Energy expenditure, kcal	343	(85)	400	(85)	296	(46)
Test meal						
Energy intake (post exercise), kcal	867	(411)	1010	(478)	748	(304)
Energy Intake (post rest), kcal	821	(383)	999	(376)	672	(326)
Difference in energy intake, kcal [g]	46	(303)	11	(398)	75	(193)
Relative energy intake, kcal [h]	526	(406)	609	(482)	452	(320)

Data are mean (standard deviation) unless stated otherwise.
Abbreviations: BMI, body mass index; GLP-1, Glucagon-like Peptide 1; MVPA, moderate-to-vigorous physical activity; PYY, peptide YY; SD, standard deviation; TFEQ-R18v2, revised 18-item Three-Factor Eating Questionnaire.
[a] Data available for 54/57 participants (25/26 men and 29/31 women).
[b] Data available for 56/57 participants (26/26 men and 30/31 women).
[c] Hormone concentrations are reported as means between pre-exercise and pre-rest. Fasting concentrations before the two study conditions did not differ (all $P \geq 0.08$).
[d] Data available for 39/57 participants (16/26 men and 23/31 women).
[e] Data available for 36/57 participants (15/26 men and 21/31 women).
[f] Data available for 38/57 participants (16/26 men and 22/31 women).
[g] Post-exercise *ad libitum* energy intake (kcal) – energy intake following the rest condition (kcal).
[h] Energy intake during test meal (kcal) – energy expenditure during exercise session (kcal).

Table 2
Linear regression analysis for the association between anthropometrics, physiological and behavioral baseline characteristics and total energy intake during the post-exercise *ad libitum* test meal.

	All participants					Men					Women				
	Energy intake (kcal)					Energy intake (kcal)					Energy intake (kcal)				
	R-squared values	B	SE	β	*P*	R-squared values	B	SE	β	*P*	R-squared values	B	SE	β	*P*
Sex [a]	0.103	261.9	104.5	0.32	**0.015**										
Age, years	0.000	−2.3	22.3	−0.01	0.920	0.002	−8.2	40.0	−0.04	0.840	0.008	10.7	22.0	0.090	0.632
Weight, kg	0.052	9.2	5.3	0.23	0.088	0.022	6.2	8.5	0.15	0.471	0.000	0.2	8.1	0.00	0.979
BMI, kg/m^2	0.009	16.8	23.5	0.10	0.477	0.000	2.9	36.4	0.02	0.937	0.034	27.0	26.9	0.18	0.323
Fat-free mass, kg	0.088	13.5	5.9	0.30	**0.025**	0.059	13.1	10.7	0.24	0.233	0.031	−11.6	12.0	−0.18	0.340
Total body fat, %	0.018	−9.4	9.2	−0.14	0.315	0.016	−11.2	18.1	−0.13	0.542	0.060	14.7	10.8	0.25	0.183
Physical activity behavior and cardiorespiratory fitness															
Habitual exercise, min/week	0.085	−0.7	0.3	−0.29	**0.032**	0.032	−0.6	0.7	−0.18	0.395	0.194	−0.6	0.3	−0.44	**0.017**
Habitual exercise days/week	0.024	−33.0	28.1	−0.16	0.246	0.013	−28.9	50.6	−0.12	0.572	0.103	−49.3	27.0	−0.32	0.078
MVPA, min/week	0.012	0.3	0.4	0.11	0.418	0.019	0.4	0.6	0.14	0.498	0.000	0.0	0.4	−0.02	0.924
Relative VO_{2peak}, mL/kg/min	0.000	1.2	9.0	0.02	0.896	0.009	−7.8	16.8	−0.09	0.648	0.074	−16.0	10.5	−0.27	0.138
Absolute VO_{2peak}, L/min	0.021	97.5	90.2	0.14	0.284	0.000	9.4	158.1	0.01	0.953	0.083	−270.1	166.2	−0.29	0.115
Maximal power, W	0.049	1.9	1.1	0.22	0.101	0.015	1.2	2.0	0.12	0.558	0.003	−0.5	1.7	−0.05	0.781
Eating behavior traits															
Cognitive Restraint	0.013	−21.4	25.3	−0.11	0.401	0.058	−44.1	36.3	−0.24	0.237	0.040	34.5	31.5	0.20	0.282
Uncontrolled Eating	0.010	8.9	11.9	0.10	0.455	0.004	−6.2	19.7	−0.06	0.757	0.091	20.6	12.1	0.30	0.098
Emotional Eating	0.002	5.8	16.9	0.05	0.732	0.011	−18.0	35.5	−0.10	0.618	0.097	26.0	14.7	0.31	0.087
Appetite-regulating hormones															
GLP-1, pg/mL	0.000	0.7	12.8	0.09	0.957	0.108	38.4	29.5	0.33	0.214	0.036	−8.4	9.5	−0.19	0.388
PYY, pg/mL	0.149	3.3	1.3	0.39	**0.015**	0.775	9.6	1.4	0.88	**<0.001**	0.013	−0.6	1.1	−0.12	0.600
Ghrelin, pg/mL	0.005	0.1	0.2	0.07	0.681	0.087	0.3	0.3	0.30	0.285	0.021	−0.1	0.2	−0.14	0.535
Adiponectin, ng/mL	0.043	12.0	9.5	0.21	0.213	0.438	52.9	16.0	0.66	**0.005**	0.010	3.8	8.4	0.10	0.659

Bold font indicates statistical significance ($P < 0.05$). Dependent variable in all models: Total energy intake during the post-exercise *ad libitum* test meal (kcal).
Abbreviations: B, unstandardized regression coefficient; β, standardized regression coefficient; BMI, body mass index; GLP-1, Glucagon-like Peptide 1; MVPA, moderate-to-vigorous physical activity; PYY, peptide YY; SE, standard error.
[a] Female = 0, male = 1.

energy intake.

3.6. Multiple regression analyses

For total post-exercise energy intake, the combination of the significant single predictors (FFM, habitual exercise minutes, and PYY) increased the coefficient of determination ($R^2 = 0.279$; $P = 0.013$, Supplemental Table 3) compared to the single predictors alone (all $R^2 \leq 0.149$, Table 2). For relative post-exercise energy intake and post-rest energy intake, the combination of the significant single predictors did not change the coefficient of determination meaningfully compared to the single predictors alone (Table 3, Supplemental Table 1, and Supplemental Table 3). For men, the combination of the significant single predictors did not change the coefficient of determination meaningfully compared to the single predictors alone for total and relative post-exercise energy intake as well as post-rest intake (Table 2, Table 3, Supplemental Table 1, and Supplemental Table 3). For women, the combination of the significant single predictors (habitual exercise minutes, habitual exercise days, and VO_{2peak}) only had an impact on relative post-exercise energy intake, with an increased coefficient of determination ($R^2 = 0.311$; $P = 0.009$, Supplemental Table 3) compared to the single predictors alone (all $R^2 \leq 0.197$, Table 3).

4. Discussion

The present study aimed to identify predictors of post-exercise energy intake and compensation in healthy adults following a single 45-min aerobic exercise bout. Our results show that individuals with lower habitual exercise and/or higher fasting concentrations of PYY eat more after an acute exercise bout, even after accounting for the energy expended during the exercise bout. Notably, these biological and behavioral characteristics differentially affected post-exercise energy intake in men and women; habitual exercise behavior was only predictive of post-exercise energy intake in women whereas fasting PYY concentrations were only a significant predictor of post-exercise intake in men. For habitual exercise in women, every 30 min/week increase was associated with a decrease in post-exercise energy intake and compensation of ~20 kcal. In men, albeit not significant, the trend in the association between habitual exercise and post-exercise energy intake was similar to that in women. Of note, in women, habitual exercise was also a

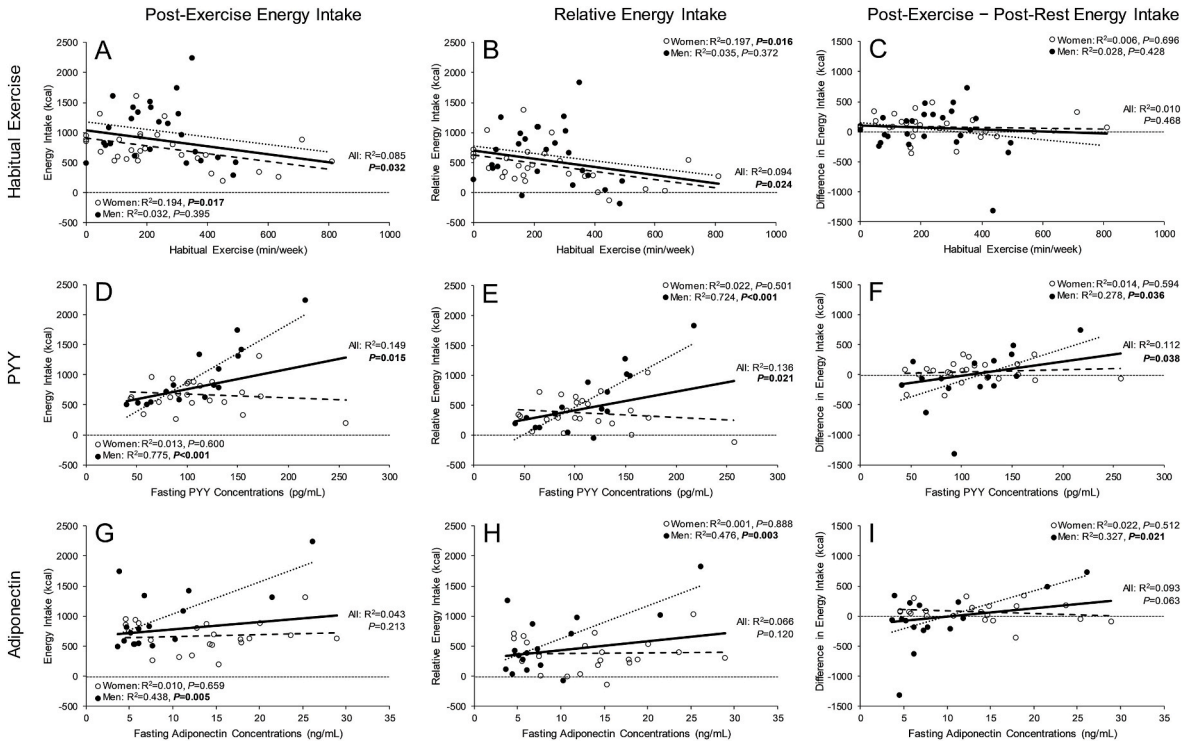


Fig. 1. Associations between habitual exercise, fasting PYY concentrations, and fasting adiponectin concentrations and post-exercise energy intake, relative energy intake (post-exercise energy intake − exercise energy expenditure), and the difference between post-exercise and post-rest energy intake. Regression lines are displayed for the entire sample (solid line), for men (dotted line), and for women (dashed line).

significant predictor of energy intake after rest, suggesting that in women the predictive value of habitual exercise behavior is not restricted to post-exercise energy intake, but applies to all habitual energy intake. Generally, the association between greater habitual exercise and lower (post-exercise) energy (compensation) is in line with our hypothesis and previous research. It has been shown that energy balance is better regulated at higher levels of PA-related energy expenditure due to better satiety signaling and the fact that exercise-induced food rewards and cravings play a less important role (Beaulieu et al., 2017, 2020; Church & Martin, 2018; Mayer et al., 1956). It is noteworthy that accelerometer-measured MVPA did not confirm these results, as MVPA was not a significant predictor of post-exercise energy intake. Overall, our sample showed relatively high average MVPA levels (~330 min/-week) at baseline, with 89% of participants above the established weekly recommendations of 150 min of MVPA and 54% even above 300 min/week (Bull et al., 2020). It can be speculated whether self-reported exercise behavior (prospectively recorded) included certain types of (even high-intensity) exercise such as strength training, cycling, or swimming that were not (accurately) captured by the hip-worn Actigraph devices, as demonstrated before (Berntsen et al., 2010; Herman Hansen et al., 2014; Höchsmann et al., 2020b), and whether this contributed to exercise behavior (min/week) being a better predictor of post-exercise energy intake and compensation than overall MVPA. The generally high MVPA level suggests that most participants would be in the regulated zone of Mayer's curve, in which energy expenditure and energy intake are in balance (Church & Martin, 2018; Mayer et al., 1956). Nevertheless, despite the overall high MVPA levels, there was substantial variability in the post-exercise energy intake, demonstrating that the single exercise session evoked a greater relative (compensatory) energy intake response in some participants than others.

The associations between fasting concentrations of appetite-regulating hormones and post-exercise energy intake are striking. Particularly the sex differences with a strong predictive value of PYY and adiponectin in men, explaining 78% and 44% of the variance in post-exercise energy intake, respectively (72% and 48% for energy compensation) but no significant associations with post-exercise energy intake in women were unexpected. In men, every ten pg/mL increase in fasting PYY concentrations and every one ng/mL increase in fasting adiponectin concentrations was associated with an increase in post-exercise energy intake of 96 kcal and 53 kcal, respectively (energy compensation: 94 kcal and 56 kcal, respectively). We are not aware of previous findings of similar sex differences in the association between fasting appetite-regulating hormones and (post-exercise) energy intake. In our study, PYY concentrations did not differ by sex at baseline, which is in line with the literature (Cooper, 2014). In the fasted state, circulating concentrations of PYY are usually low with rapid increases upon nutrient ingestion (Cooper, 2014; Druce et al., 2004). During aerobic exercise, PYY has been shown to increase and remain elevated for up to 5h after exercise (Stensel, 2010), with the appetite-suppressing effects usually diminishing within 30–60 min of exercise cessation, however (Dorling et al., 2018). It has been reported that sex differences exist in the PYY response to moderate-intensity exercise (bike ergometer at 65% VO_{2max}, similar to our study), with greater increases and a greater subsequent post-exercise appetite suppression in men compared to women (Hazell et al., 2017). Importantly, however, the exercise-induced increases in PYY and associated appetite suppression do not always translate into de-facto decreases in post-exercise energy intake. In fact, the majority of exercise studies show no change in energy intake after acute bouts of exercise (Stensel, 2010), and it has been found in a review that 19% of such studies even report increases in energy intake while 16% show a decrease (65% no change) (Blundell & King, 1999). Nevertheless, the strong associations between greater fasting concentrations of PYY (and adiponectin) and greater post-exercise energy intake and compensation as shown in the present study are still somewhat unexpected. Further research is needed to confirm our findings and examine why fasting PYY may affect energy intake and compensation after exercise but not after a no-exercise rest condition; also of interest is why PYY concentrations affect post-exercise energy

Table 3
Linear regression analysis for the association between anthropometrics, physiological and behavioral baseline characteristics and relative post-exercise *ad libitum* energy intake (energy intake [kcal] – energy expenditure [kcal]).

	All participants					Men					Women				
	Relative energy intake (kcal)					Relative energy intake (kcal)					Relative energy intake (kcal)				
	R-squared values	B	SE	β	*P*	R-squared values	B	SE	β	*P*	R-squared values	B	SE	β	*P*
Sex [a]	0.038	157.4	106.9	0.19	0.147										
Age, years	0.000	−2.0	22.0	−0.01	0.929	0.007	−16.1	40.3	−0.08	0.693	0.012	13.8	23.1	0.11	0.554
Weight, kg	0.007	3.4	5.4	0.09	0.528	0.000	0.9	8.7	0.02	0.923	0.002	−2.0	8.5	−0.04	0.816
BMI, kg/m^2	0.000	3.0	23.3	0.02	0.899	0.007	−14.5	36.6	−0.08	0.695	0.020	21.8	28.5	0.14	0.449
Fat-free mass, kg	0.019	6.1	6.0	0.14	0.312	0.013	6.1	11.0	0.11	0.583	0.054	−16.0	12.5	−0.23	0.209
Total body fat, %	0.009	−6.4	9.2	−0.09	0.490	0.027	−14.9	18.1	−0.17	0.419	0.053	14.6	11.4	0.23	0.211
Physical activity behavior and cardiorespiratory fitness															
Habitual exercise, min/week	0.094	−0.7	0.3	−0.31	**0.024**	0.035	−0.6	0.7	−0.19	0.372	0.197	−0.7	0.3	−0.44	**0.016**
Habitual exercise days/week	0.054	−48.7	27.4	−0.23	0.081	0.028	−41.9	50.6	−0.17	0.416	0.149	−62.4	27.6	−0.39	**0.032**
MVPA, min/week	0.012	0.3	0.4	0.11	0.419	0.027	0.5	0.6	0.16	0.421	0.001	−0.1	0.4	−0.02	0.900
Relative VO_{2peak}, mL/kg/min	0.019	−9.1	8.8	−0.14	0.306	0.043	−17.3	16.7	−0.21	0.312	0.133	−22.6	10.7	−0.36	**0.044**
Absolute VO_{2peak}, L/min	0.004	−40.1	89.9	−0.06	0.657	0.026	−127.0	157.4	−0.16	0.428	0.173	−408.6	166.0	−0.42	**0.020**
Maximal power, W	0.001	0.3	1.1	0.04	0.781	0.001	−0.4	2.1	−0.04	0.853	0.027	−1.5	1.7	−0.16	0.386
Eating behavior traits															
Cognitive Restraint	0.012	−20.4	25.0	−0.11	0.420	0.050	−41.2	36.8	−0.22	0.274	0.021	26.4	33.4	0.15	0.436
Uncontrolled Eating	0.008	7.7	11.7	0.09	0.517	0.008	−8.5	19.9	−0.09	0.671	0.094	22.0	12.7	0.31	0.093
Emotional Eating	0.004	7.4	16.7	0.06	0.660	0.016	−22.6	35.7	−0.13	0.532	0.095	27.0	15.5	0.31	0.092
Appetite-regulating hormones															
GLP-1, pg/mL	0.002	3.0	12.5	0.04	0.811	0.108	38.8	29.8	0.33	0.214	0.024	−7.5	10.3	−0.16	0.477
PYY, pg/mL	0.136	3.1	1.3	0.37	**0.021**	0.724	9.4	1.6	0.85	**<0.001**	0.022	−0.8	1.2	−0.15	0.501
Ghrelin, pg/mL	0.018	0.1	0.2	0.14	0.431	0.090	0.4	0.3	0.30	0.278	0.001	−0.0	0.2	−0.03	0.911
Adiponectin, ng/mL	0.066	14.6	9.2	0.26	0.120	0.476	55.8	15.6	0.69	**0.003**	0.001	1.3	9.0	0.03	0.888

Bold font indicates statistical significance ($P < 0.05$). Dependent variable in all models: Relative post-exercise *ad libitum* energy intake (energy intake during test meal [kcal] – energy expenditure during exercise session [kcal]).
Abbreviations: B, unstandardized regression coefficient; β, standardized regression coefficient; BMI, body mass index; GLP-1, Glucagon-like Peptide 1; MVPA, moderate-to-vigorous physical activity; PYY, peptide YY; SE, standard error.
[a] Female = 0, male = 1.

intake and compensation differently in men than women, despite similar fasting concentrations.

Contrary to previous findings showing a general association between eating behavior traits, and particularly disinhibition or uncontrolled eating, with overeating in the presence of food (Brunner et al., 2021; Garcia-Garcia et al., 2022; Vainik et al., 2019), eating behavior traits (cognitive restraint, uncontrolled eating, and emotional eating) were not significant predictors of energy intake. This is similar to recent findings in adolescents in whom the three trait measures as assessed via the TFEQ-R18v2 were also not associated with *ad libitum* post-exercise energy intake (Fearnbach et al., 2022). Further, physiological processes may override these eating behavior traits in the post-exercise state; however, in our study, cognitive restraint, uncontrolled eating, or emotional eating were also not associated with *ad libitum* energy intake following the control condition involving rest.

Energy intake after rest was associated with weight and FFM as well as cardiorespiratory fitness (VO_{2peak} and maximal power). Of note, only absolute but not relative VO_{2peak} were significant predictors of post-rest energy intake, suggesting that this association was driven by the significant predictors of weight, FFM, and sex, which have repeatedly been shown to predict meal size and single meal food intake (Blundell et al., 2012; Hopkins et al., 2016; Weise et al., 2014). When assessing predictors of the difference in energy intake between after exercise and after rest, we found significant associations with cardiorespiratory fitness, in this case, both relative and absolute VO_{2peak}, for men but not women. Further, the associations revealed a turning point at 40.9 mL/kg/min or 3.0 L/min, respectively. Individuals with average-to-above-average fitness seem to have reduced total energy intake post-exercise compared to post-rest while those with below-average fitness seem to have increased total energy intake post-exercise compared to post-rest. The finding of greater post-exercise energy intake in individuals with lower fitness compared to those with higher fitness is supported by previous findings showing that individuals with higher compared to lower fitness (mean VO_{2peak} of 51.6 vs. 37.0) compensated after an acute exercise session compared to after rest (Charlot & Chapelot, 2013). Additionally, this relationship may be influenced by higher body fat accumulation, which often co-occurs with lower fitness and has been shown to predict greater post-exercise energy

compensation due to a progressively increasing impairment of energy balance (Careau et al., 2021).

Overall, our results may help identify individuals who are likely to show post-exercise energy compensation and help explain why this adverse response to exercise occurs in some individuals but not others. To prevent increased post-exercise energy intake, countermeasures such as the selection and preparation of the post-exercise meal ahead of the exercise session may be beneficial, as demonstrated previously (Gustafson et al., 2018; Koehler et al., 2021). Strengths of the present study include the crossover design, the relatively balanced sample of men and women, and the comprehensive analysis of the contribution of anthropometric, behavioral, cognitive, and endocrine factors at baseline to post-exercise energy intake. Limitations include the relatively small sample, particularly for the appetite-regulating hormone data, and the relative lack of racial and age-related diversity. However, we specifically chose to recruit a convenience sample of young adults to minimize the impact of age, as it has been shown that exercise-induced consequences of hunger, satiety, and compensation differentially affect adults aged 60+ years (Hubner et al., 2021).

Nevertheless, future studies with larger and more diverse samples should examine potential differences in predictors of post-exercise energy intake by race/ethnicity and age. Further, it would be interesting to assess how different exercise modalities (type, intensity, duration) with similar energy expenditure affect post-exercise food intake and compensation and whether the found associations hold for repeated exercise bouts.

5. Conclusions

Biological and behavioral characteristics differentially affect post-exercise energy intake in men and women. In women, habitual exercise behavior predicts post-exercise energy intake, with more exercise protecting against compensatory eating. In men, appetite-regulating hormones, specifically PYY and adiponectin, play a role in the energy intake response to acute exercise, even when measured before exercise and after controlling for post-rest energy intake. Our results can help identify individuals who are more likely to (over-) compensate for the energy expended in exercise via increased post-exercise energy intake allowing to deploy targeted countermeasures ahead of time.

Financial support

This research was funded by the University of Nebraska Food for Health Collaboration Initiative. The funding source had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.

Ethical statement

The study was approved by the University of Nebraska-Lincoln's Institutional Review Board (project number 17239) and written informed consent was obtained from all participants before participation in the study. All study procedures were conducted in accordance with the Declaration of Helsinki.

Author contributions

K.K., J.R.S., J.B.B., and J.A.F. acquired funding. K.K. and J.R.S. designed the study and S.E.B., J.A.F., J.B.B., and K.K. collected data. C. H. conducted statistical analyses, drafted the manuscript, and created tables and figures. K.K., S.E.B., J.A.F., J.R.S., and J.B.B. provided critical revision of the manuscript for important intellectual content. All authors have read and agreed to the published version of the manuscript.

Declaration of competing interest

The authors have no conflicts of interest to declare.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.appet.2023.106520.

References

Achten, J., & Jeukendrup, A. E. (2003). Maximal fat oxidation during exercise in trained men. *Int J Sports Med, 24*, 603–608.

Alssema, M., Rijkelijkhuizen, J. M., Holst, J. J., Teerlink, T., Scheffer, P. G., Eekhoff, E. M. W., Gastaldelli, A., Mari, A., Hart, L. M., Nijpels, G., & Dekker, J. M. (2013). Preserved GLP-1 and exaggerated GIP secretion in type 2 diabetes and relationships with triglycerides and ALT. *European Journal of Endocrinology, 169*, 421–430.

Balaguera-Cortes, L., Wallman, K. E., Fairchild, T. J., & Guelfi, K. J. (2011). Energy intake and appetite-related hormones following acute aerobic and resistance exercise. *Applied Physiology, Nutrition, and Metabolism, 36*, 958–966.

Basolo, A., Heinitz, S., Stinson, E. J., Begaye, B., Hohenadel, M., Piaggi, P., Krakoff, J., & Votruba, S. B. (2019). Fasting glucagon-like peptide 1 concentration is associated with lower carbohydrate intake and increases with overeating. *Journal of Endocrinological Investigation, 42*, 557–566.

Beaulieu, K., Hopkins, M., Long, C., Blundell, J., & Finlayson, G. (2017). High habitual physical activity improves acute energy compensation in nonobese adults. *Medicine & Science in Sports & Exercise, 49*, 2268–2275.

Beaulieu, K., Oustric, P., & Finlayson, G. (2020). The impact of physical activity on food reward: Review and conceptual synthesis of evidence from observational, acute, and chronic exercise training studies. *Current Obesity Reports, 9*, 63–80.

Berntsen, S., Hageberg, R., Aandstad, A., Mowinckel, P., Anderssen, S. A., Carlsen, K.-H., & Andersen, L. B. (2010). Validity of physical activity monitors in adults participating in free-living activities. *British Journal of Sports Medicine, 44*, 657–664.

Blundell, J. E., Caudwell, P., Gibbons, C., Hopkins, M., Näslund, E., King, N. A., & Finlayson, G. (2012). Body composition and appetite: Fat-free mass (but not fat mass or BMI) is positively associated with self-determined meal size and daily energy intake in humans. *British Journal of Nutrition, 107*, 445–449.

Blundell, J. E., & King, N. A. (1999). Physical activity and regulation of food intake: Current evidence. *Medicine & Science in Sports & Exercise, 31*, S573.

Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Medicine & Science in Sports & Exercise, 14*, 377–381.

Broom, D. R., Batterham, R. L., King, J. A., & Stensel, D. J. (2009). Influence of resistance and aerobic exercise on hunger, circulating levels of acylated ghrelin, and peptide YY in healthy males. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology, 296*, R29–R35.

Broom, D. R., Stensel, D. J., Bishop, N. C., Burns, S. F., & Miyashita, M. (2007). Exercise-induced suppression of acylated ghrelin in humans. *Journal of Applied Physiology, 102*, 2165–2171.

Brunner, E. J., Maruyama, K., Shipley, M., Cable, N., Iso, H., Hiyoshi, A., … Llewellyn, C. (2021). Appetite disinhibition rather than hunger explains genetic effects on adult BMI trajectory. *International Journal of Obesity, 45*, 758–765.

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J.-P., Chastin, S., Chou, R., Dempsey, P. C., DiPietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., … Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine, 54*, 1451–1462.

Cappelleri, J. C., Bushmakin, A. G., Gerber, R. A., Leidy, N. K., Sexton, C. C., Lowe, M. R., & Karlsson, J. (2009). Psychometric analysis of the three-factor eating questionnaire-R21: Results from a large diverse sample of obese and non-obese participants. *International Journal of Obesity, 33*, 611–620.

Careau, V., Halsey, L. G., Pontzer, H., Ainslie, P. N., Andersen, L. F., Anderson, L. J., Arab, L., Baddou, I., Bedu-Addo, K., Blaak, E. E., Blanc, S., Bonomi, A. G., Bouten, C. V. C., Buchowski, M. S., Butte, N. F., Camps, S. G. J. A., Close, G. L., Cooper, J. A., Das, S. K., … Speakman, J. R. (2021). Energy compensation and adiposity in humans. *Current Biology, 31*, 4659–4666. e2.

Charlot, K., & Chapelot, D. (2013). Energy compensation after an aerobic exercise session in high-fat/low-fit and low-fat/high-fit young male subjects. *British Journal of Nutrition, 110*, 1133–1142.

Church, T., & Martin, C. K. (2018). The obesity epidemic: A consequence of reduced energy expenditure and the uncoupling of energy intake? *Obesity, 26*, 14–16.

Church, T. S., Martin, C. K., Thompson, A. M., Earnest, C. P., Mikus, C. R., & Blair, S. N. (2009). Changes in weight, waist circumference and compensatory responses with different doses of exercise among sedentary, overweight postmenopausal women. *PLoS ONE, 4*, Article e4515.

Cooper, J. A. (2014). Factors affecting circulating levels of peptide YY in humans: A comprehensive review. *Nutrition Research Reviews, 27*, 186–197.

Dorling, J., Broom, D. R., Burns, S. F., Clayton, D. J., Deighton, K., James, L. J., King, J. A., Miyashita, M., Thackray, A. E., Batterham, R. L., & Stensel, D. J. (2018). Acute and chronic effects of exercise on appetite, energy intake, and appetite-related hormones: The modulating effect of adiposity, sex, and habitual physical activity. *Nutrients, 10*.

Druce, M. R., Small, C. J., & Bloom, S. R. (2004). Minireview: Gut peptides regulating satiety. *Endocrinology, 145*, 2660–2665.

Edholm, O. G. (1956). Energy expenditure in relation to nutrition. *Proceedings of the Nutrition Society, 15*, 80–83.

Fearnbach, N., Staiano, A. E., Johannsen, N. M., Hsia, D. S., Beyl, R. A., Carmichael, O. T., & Martin, C. K. (2022). Predictors of post-exercise energy intake in adolescents ranging in weight status from overweight to severe obesity. *Nutrients, 14*, 223.

Feig, E. H., Piers, A. D., Kral, T. V. E., & Lowe, M. R. (2018). Eating in the absence of hunger is related to loss-of-control eating, hedonic hunger, and short-term weight gain in normal-weight women. *Appetite, 123*, 317–324.

Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the computer science and applications, inc. accelerometer. *Medicine & Science in Sports & Exercise, 30*, 777–781.

French, S. A., Epstein, L. H., Jeffery, R. W., Blundell, J. E., & Wardle, J. (2012). Eating behavior dimensions: Associations with energy intake and body weight: A review. *Appetite, 59*, 541–549.

Garcia-Garcia, I., Neseliler, S., Morys, F., Dadar, M., Yau, Y. H. C., Scala, S. G., … Dagher, A. (2022). Relationship between impulsivity, uncontrolled eating and body mass index: A hierarchical model. *International Journal of Obesity, 46*, 129–136.

Guo, Y., Ma, L., Enriori, P. J., Koska, J., Franks, P. W., Brookshire, T., … Tataranni, P. A. (2006). Physiological evidence for the involvement of peptide YY in the regulation of energy homeostasis in humans. *Obesity, 14*, 1562–1570.

Gustafson, C. R., Rakhmatullaeva, N., Beckford, S. E., Ammachathram, A., Cristobal, A., & Koehler, K. (2018). Exercise and the timing of snack choice: Healthy snack choice is reduced in the post-exercise state. *Nutrients, 10*, E1941.

Hazell, T. J., Townsend, L. K., Hallworth, J. R., Doan, J., & Copeland, J. L. (2017). Sex differences in the response of total PYY and GLP-1 to moderate-intensity continuous and sprint interval cycling exercise. *European Journal of Applied Physiology, 117*, 431–440.

Herman Hansen, B., Børtnes, I., Hildebrand, M., Holme, I., Kolle, E., & Anderssen, S. A. (2014). Validity of the ActiGraph GT1M during walking and cycling. *Journal of Sports Sciences, 32*, 510–516.

Höchsmann, C., Dorling, J. L., Apolzan, J. W., Johannsen, N. M., Hsia, D. S., & Martin, C. K. (2020a). Baseline habitual physical activity predicts weight loss, weight compensation, and energy intake during aerobic exercise. *Obesity, 28*, 882–892.

Höchsmann, C., Knaier, R., Infanger, D., & Schmidt-Trucksäss, A. (2020b). Validity of smartphones and activity trackers to measure steps in a free-living setting over three consecutive days. *Physiological Measurement, 41*, Article 015001.

Hopkins, M., Finlayson, G., Duarte, C., Whybrow, S., Ritz, P., Horgan, G. W., Blundell, J. E., & Stubbs, R. J. (2016). Modelling the associations between fat-free mass, resting metabolic rate and energy intake in the context of total energy balance. *International Journal of Obesity, 40*, 312–318.

Hubner, S., Boron, J. B., & Koehler, K. (2021). The effects of exercise on appetite in older adults: A systematic review and meta-analysis. *Frontiers in Nutrition, 8*, Article 734267.

Jones, T. E., Basilio, J. L., Brophy, P. M., McCammon, M. R., & Hickner, R. C. (2009). Long-term exercise training in overweight adolescents improves plasma peptide YY and resistin. *Obesity, 17*, 1189–1195.

Karlsson, J., Persson, L. O., Sjöström, L., & Sullivan, M. (2000). Psychometric properties and factor structure of the Three-Factor Eating Questionnaire (TFEQ) in obese men and women. Results from the Swedish Obese Subjects (SOS) study. *International Journal of Obesity and Related Metabolic Disorders, 24*, 1715–1725.

King, N. A., Burley, V. J., & Blundell, J. E. (1994). Exercise-induced suppression of appetite: Effects on food intake and implications for energy balance. *European Journal of Clinical Nutrition, 48*, 715–724.

King, J. A., Deighton, K., Broom, D. R., Wasse, L. K., Douglas, J. A., Burns, S. F., Cordery, P. A., Petherick, E. S., Batterham, R. L., Goltz, F. R., Thackray, A. E., Yates, T., & Stensel, D. J. (2017). Individual variation in hunger, energy intake, and ghrelin responses to acute exercise. *Medicine & Science in Sports & Exercise, 49*, 1219–1228.

Koehler, K., Beckford, S. E., Thayer, E., Martin, A. R., Boron, J. B., & Stevens, J. R. (2021). Exercise shifts hypothetical food Choices toward greater amounts and more immediate consumption. *Nutrients, 13*, 347.

Martin, C. K., Johnson, W. D., Myers, C. A., Apolzan, J. W., Earnest, C. P., Thomas, D. M., Rood, J. C., Johannsen, N. M., Tudor-Locke, C., Harris, M., Hsia, D. S., & Church, T. S. (2019). Effect of different doses of supervised exercise on food intake, metabolism, and non-exercise physical activity: The E-MECHANIC randomized controlled trial. *American Journal of Clinical Nutrition, 110*, 583–592.

Martins, C., Morgan, L. M., Bloom, S. R., & Robertson, M. D. (2007). Effects of exercise on gut peptides, energy intake and appetite. *Journal of Endocrinology, 193*, 251–258.

Mayer, J., Roy, P., & Mitra, K. P. (1956). Relation between caloric intake, body weight, and physical work: Studies in an industrial male population in West Bengal. *American Journal of Clinical Nutrition, 4*, 169–175.

Murphy, K. G., & Bloom, S. R. (2006). Gut hormones and the regulation of energy homeostasis. *Nature, 444*, 854–859.

Ranganath, L. R., Beety, J. M., Morgan, L. M., Wright, J. W., Howland, R., & Marks, V. (1996). Attenuated GLP-1 secretion in obesity: Cause or consequence? *Gut, 38*, 916–919.

Stensel, D. (2010). Exercise, appetite and appetite-regulating hormones: Implications for food intake and weight control. *Annals of Nutrition & Metabolism, 57*, 36–42.

Stunkard, A. J., & Messick, S. (1985). The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *Journal of Psychosomatic Research, 29*, 71–83.

Vainik, U., García-García, I., & Dagher, A. (2019). Uncontrolled eating: A unifying heritable trait linked with obesity, overeating, personality and the brain. *European Journal of Neuroscience, 50*, 2430–2445.

Weise, C. M., Hohenadel, M. G., Krakoff, J., & Votruba, S. B. (2014). Body composition and energy expenditure predict ad-libitum food and macronutrient intake in humans. *International Journal of Obesity, 38*, 243–251.

Publication 2

Baseline habitual physical activity predicts weight loss, weight compensation, and energy intake during aerobic exercise.

Christoph Höchsmann [1], James L Dorling [1], John W Apolzan [1], Neil M Johannsen [1,2], Daniel S Hsia [1], Corby K Martin [1]

[1] Pennington Biomedical Research Center, Baton Rouge, LA, USA

[2] School of Kinesiology, Louisiana State University, Baton Rouge, LA, USA

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
Obesity, 28(5): 882–892	2020	10.1002/oby.22766	5.0 (2020)	Q1	Scopus: 10 Google Scholar: 13

Author contribution:

First author; developed the research question, conducted the statistical analyses, drafted the manuscript, and created tables and figures.

Original Article
CLINICAL TRIALS AND INVESTIGATIONS

Obesity

Baseline Habitual Physical Activity Predicts Weight Loss, Weight Compensation, and Energy Intake During Aerobic Exercise

Christoph Höchsmann [1], *James L. Dorling* [1], *John W. Apolzan* [1], *Neil M. Johannsen* [1,2], *Daniel S. Hsia* [1], *and Corby K. Martin* [1]

Objective: This study aimed to determine whether different measures of habitual physical activity (PA) at baseline predict weight change, weight compensation, and changes in energy intake (EI) during a 24-week supervised aerobic exercise intervention.

Methods: Data from 108 participants (78 women; 48.7 [SD: 11.6] years; BMI 31.4 [SD: 4.6] kg/m^2), randomly assigned to either the moderate-dose exercise group (8 kcal/kg of body weight per week) or the high-dose exercise group (20 kcal/kg of body weight per week) of the Examination of Mechanisms of Exercise-induced Weight Compensation (E-MECHANIC) trial, were analyzed. Moderate-to-vigorous PA (MVPA), steps per day, and PA energy expenditure (PAEE) were measured with SenseWear armbands (BodyMedia, Pittsburgh, Pennsylvania), and total activity energy expenditure and EI were estimated with doubly labeled water, all over 2 weeks, before and toward the end of the intervention. Multiple linear regression models, adjusted for sex, exercise group, and baseline value of the outcome, were used.

Results: Baseline habitual MVPA levels predicted weight change ($\beta = -0.275$; $P = 0.020$), weight compensation ($\beta = -0.238$; $P = 0.043$), and change in EI ($\beta = -0.318$; $P = 0.001$). Associations between baseline PAEE and outcomes were comparable, whereas steps per day and, importantly, total activity energy expenditure (via doubly labeled water) did not significantly predict change in weight-related outcomes.

Conclusions: While acknowledging substantial variability in the data, on average, lower baseline habitual MVPA and PAEE levels were associated with less weight loss from exercise, higher compensation, and increased EI.

Obesity (2020) **28**, 882-892.

Study Importance

What is already known?

- Exercise is recommended for weight management.
- Exercise-induced weight loss often is less than expected based on measured energy expenditure (EE).
- This is called *weight compensation* and results primarily from increased energy intake (EI).

What does this study add?

- Moderate-to-vigorous physical activity (MVPA) and physical activity EE (PAEE) (≥3 metabolic equivalents) levels prior to engaging in a moderate- to high-dose aerobic exercise intervention predict weight loss, weight compensation, and changes in EI during the intervention.
- Prior MVPA and PAEE have a superior predictive value compared with steps per day and total activity-related EE, as estimated by doubly labeled water, regarding these outcomes.

How might these results change the direction of research?

- Further research is needed to understand why participants with lower baseline habitual MVPA and PAEE levels lose less weight from structured exercise, show higher weight compensation, and increase their EI more than those who are more active at baseline to develop strategies to mitigate this detrimental effect.

Introduction

The prevalence of overweight and obesity has grown into a worldwide epidemic in recent years (1), and excess body weight substantially increases the risk of adverse health conditions (2). Exercise has been shown to support the prevention and management of obesity (3); however, when used for weight loss, exercise interventions consistent with the physical activity (PA) guidelines for weight loss and weight loss maintenance (>225 min/wk of moderate-intensity PA) frequently produce less weight loss than expected based on energy

[1] Pennington Biomedical Research Center, Baton Rouge, Louisiana, USA. Correspondence: Corby K. Martin (Corby.Martin@pbrc.edu) [2] Louisiana State University, Baton Rouge, Louisiana, USA.

 Received: 15 October 2019; **Accepted:** 22 January 2020; **Published online 6 March 2020. doi:10.1002/oby.22766**

expended in exercise (4-7). This discrepancy is called weight compensation (8), and it results primarily from exercise-induced increases in appetite and energy intake (EI) as opposed to changes in metabolism or activity (6,9).

It is unknown whether factors pertaining to one's lifestyle prior to starting an exercise program affect weight compensation and food intake (FI). An individual's habitual PA level at baseline might be such a determinant of the observed difference between actual weight change and predicted weight loss from the energy balance model. As suggested by Westerterp (10), it is possible that a lower habitual PA level at baseline allows an exercise-induced increase in energy expenditure (EE) without or with less compensatory increase in EI. Conversely, and based on previous research indicating that EI and energy balance are better regulated at higher levels of activity-related EE (11,12), lower habitual PA levels at baseline might be associated with larger compensatory increases in EI in response to an exercise-induced increase in EE (10).

To further elucidate the mechanisms for weight compensation in response to exercise, the aim of this analysis was to determine whether different measures of habitual PA at baseline predict weight change, weight compensation, and changes in EI during a 24-week, supervised, controlled aerobic exercise intervention. Specifically, we aimed to compare the predictive value of (1) minutes spent in moderate-to-vigorous PA (MVPA), (2) steps per day, and (3) EE through PA (PAEE), assessed via two validated methods, with regard to these outcomes. Based on previous work (11,12), we hypothesized that participants with lower PA levels at baseline would show greater weight compensation and larger exercise-induced increases in EI. Although PAEE is directly related to the energy balance model and a significant association with our outcomes might be expected, we aimed to additionally assess the association of MVPA and steps per day with our outcomes, as PA recommendations based on these parameters are commonly communicated to patients, and a predictive value of these parameters would consequently be of interest to many clinical and research settings.

Methods

Design and participants

This report is a secondary analysis of the Examination of Mechanisms of Exercise-induced Weight Compensation (E-MECHANIC) study (ClinicalTrials.gov identifier NCT01264406) that was approved by the institutional review board and conducted between November 2010 and December 2015 at Pennington Biomedical Research Center (Baton Rouge, Louisiana). The complete design, methods, and primary outcomes of the E-MECHANIC study have been previously published (6,13). In brief, this 24-week randomized controlled trial recruited 198 healthy men and women with overweight or obesity (BMI ≥ 25 kg/m^2 to ≤ 45 kg/m^2) and low levels of PA (≤ 20 minutes of structured exercise on ≤ 3 d/wk based on self-report; < 8,000 steps per day (14) assessed during 1 week of accelerometer data [SenseWear armband; BodyMedia, Pittsburgh, Pennsylvania]). Participants were randomly allocated in a 1:1:1 ratio to either a moderate-dose exercise group (8 kcal/kg of body weight per week [KKW]), a high-dose exercise group (20 KKW), or a nonexercise control group (13). The selected exercise doses reflect recommendations for general health (8 KKW) and for weight loss (20 KKW) (15). Exercise intensity during the supervised exercise sessions was self-selected between 65% and 85% peak oxygen uptake, and sessions varied in length to meet each participant's EE goal (16).

Participants were excluded if they were currently participating in a weight loss program (and/or had ≥ 4-kg weight change in the past 6 months), were currently pregnant or had been pregnant within the past 6 months, or were diagnosed with diabetes, cardiovascular disease, or arrhythmia. All participants provided written informed consent prior to inclusion in the study.

The primary aim of the E-MECHANIC study was to identify mechanisms of exercise-induced weight compensation (i.e., less than expected weight loss) by examining the effect of the two different doses of exercise training on EI over the 24-week intervention period. The study found significantly higher weight compensation in the high-dose exercise group compared with the moderate-dose exercise group, which resulted primarily from increased EI and concomitant increases in appetite (6).

In this report, to examine the impact of baseline levels of habitual PA on outcome measures during a supervised exercise intervention, only participants allocated to the two exercise groups ($n = 110$) who completed the trial per protocol were included in the main analyses. Demographics of those exercisers who did not complete the trial ($n = 25$) did not differ significantly from completers (all P values ≥ 0.093).

Outcome variables

Anthropometry and body composition. At baseline and follow-up, body weight was assessed under fasting conditions using a Tanita scale (Arlington Heights, Illinois), and waist circumference was determined using a nonextensible tape measure (Gulick II; Sammons Preston, Chicago, Illinois). Dual-energy x-ray absorptiometry (DXA) (Lunar iDXA and Encore software version 13.60; GE Healthcare, Madison, Wisconsin) was used to assess fat mass.

Weight compensation. Weight compensation is the difference between the amount of weight loss predicted from exercise-associated EE and observed weight loss from baseline to follow-up (actual – predicted weight change). Predicted weight loss was calculated using a validated dynamic energy balance model that overcomes the limitations of the conventional assumption that 1 kg of body weight equals 7,700 kcal/kg (7,17,18), accounting for adaptations that occur when body mass changes, including adaptations to resting metabolic rate (RMR), dietary-induced thermogenesis, and nonexercise activity thermogenesis (19).

EI. EI was estimated with doubly labeled water (DLW) and FI tests at baseline and follow-up. DLW data were collected over a 2-week period at both time points. DLW measures total daily EE (TDEE), which equals total daily EI during weight stability (20,21). The DLW period at baseline occurred before participants began exercising. During the DLW period at follow-up, participants exercised at their prescribed dose. During both DLW periods, participants were weight stable (≤ 0.15-kg change in weight during the 2-week period). Change in EI by DLW was calculated with and without adjusting for change in RMR. For participants who were weight stable or who gained weight during the 6-month trial, follow-up TDEE was subtracted from baseline TDEE to quantify the change in EI, as any changes in RMR from weight gain are reflected in the TDEE value from DLW. For participants who lost weight during the intervention, this calculation fails to consider decreased basal metabolic requirements;

therefore, the difference between RMR from baseline to follow-up was added to the difference in TDEE for these participants to quantify the change in EI during the intervention period.

In addition, at baseline and follow-up, validated laboratory-based FI tests were conducted at lunch and dinner. Following a standard breakfast between 0700 and 0800 consisting of a 190-kcal nutrition bar, participants returned to the center between 1100 and 1200 to complete their test lunch, which consisted of ad libitum sandwiches, potato chips, cookies, water, and choice of an artificially-sweetened soda or tea or sugar-sweetened soda or tea. At 5.5 hours after the start of their lunch, participants returned to the center again to complete their dinner meal, which consisted of a previously described 18-food-item buffet meal (22), presented to the participants all at once within arm's reach. At both test meals, participants were instructed to eat as much or as little of the presented food items as desired and to avoid distractions (e.g., mobile phone use), focusing completely on the meal. FI testing at follow-up occurred at least 24 hours after the last exercise session. We quantified FI at lunch and dinner by covertly weighing food provision and waste and combined EI (kilocalories) from both meals for the analyses presented in this paper.

RMR. We measured RMR with indirect calorimetry over 30 minutes after a 12-hour overnight fast with Max-II metabolic carts (AEI Technologies, Pittsburgh, Pennsylvania) at baseline and follow-up. The change in RMR was calculated as RMR at follow-up minus RMR at baseline. Calculations adjusted for change in body composition (i.e., lean mass measured with DXA), sex, and age did not differ meaningfully from the basic change scores; hence, the basic change scores are reported.

PA. SenseWear armbands measured the minutes per day spent in activities of different intensities, steps per day, and PAEE during a 2-week period at baseline and follow-up. In the MVPA-related analyses presented in this paper, only activities ≥3 metabolic equivalents (MET) are included (3), and, congruent with the most recent Physical Activity Gudelines for Americans (23), all MVPA was considered rather than only that accumulated in bouts of at least 10 minutes as recommended previously. The SenseWear software classifies any activity ≤3 MET as sedentary; hence, PAEE included only activities ≥3 MET (24). Participants were instructed to wear the armbands continuously and to take them off only during activities involving water. The SenseWear armbands detect and record wear time, and only full days of data, defined as a wear time of ≥95% (equating to 22 hours and 48 minutes or 1,368 minutes), were included in the analyses. During the monitoring period at follow-up, participants exercised at their prescribed dose; therefore, PA data collected by the SenseWear armbands during these sessions were removed before analysis. To account for differing wear times between participants caused by varying durations of the exercise sessions and different nonwear times within the time frame of 22 hours and 48 minutes, the total number of minutes of daily activity was divided by the total daily wear time (minutes) and then extrapolated out to a 24-hour day.

In addition to the PAEE estimates by the SenseWear armband, we calculated the gold standard of activity EE (AEE) based on the DLW-estimated TDEE (DLW-AEE = TDEE − [RMR + thermic effect of food]), which captures all PA-related EE. The thermic effect of food was estimated as 10% of TDEE.

Questionnaires. Retrospective visual analog scales assessed average ratings of appetite during the previous week (25) at baseline and follow-up. The Eating Inventory (26) was used to assess eating behavior, specifically restraint, disinhibition, and hunger at baseline and follow-up. Additional questionnaires included the Multifactorial Assessment of Eating Disorders Symptoms (27), Food Preference Questionnaire (28), and Food Craving Inventory (29).

Statistical analyses

The distribution of variables was verified using the Shapiro-Wilk test and by visual inspection of histograms and quantile-quantile plots of the residuals. The influence of outliers was estimated using studentized residuals, and multicollinearity was assessed via the variance inflation factor. Exclusion of outliers (≤2 for all models) did not change the results meaningfully; therefore, the models including outliers are reported. Descriptive data are reported as mean and SD. We used multiple linear regression models to estimate the effect of SenseWear-assessed habitual MVPA levels (minutes per day), steps per day, PAEE, and DLW-estimated AEE at baseline on weight change (kilograms) and weight compensation (kilograms), as well as on changes in waist circumference (centimeters), fat mass (kilograms), EI (by DLW in kilocalories per day), EI during FI testing (kilocalories at a test lunch and test dinner combined), RMR (kilocalories per day), and habitual MVPA levels (minutes per day), steps per day, and PAEE (kilocalories per day). Covariates in the models were sex, exercise group, and baseline value of the respective outcome. Results of analyses that included age, ethnicity, and baseline BMI did not differ meaningfully; therefore, the models without these additional covariates are reported. Similarly, interaction terms for sex and exercise group were nonsignificant; therefore, results are reported without the interaction terms in the models. Pearson product moment correlation analysis was used to assess the association between habitual MVPA levels and questionnaire-assessed eating behaviors at baseline. The analyses were conducted using SPSS Statistics for Windows, version 25 (IBM Corp., Armonk, New York), with the significance level set to 0.05 (two-sided).

Results

Two participants were excluded from the analyses because they did not provide baseline accelerometer data. Baseline characteristics of all included 108 participants are shown in Table 1. Baseline characteristics of the control group (not included in main analyses) are provided in Supporting Information Table S1. At baseline, average wear time of the armbands was 1,415.2 min/d (SD: 9.1 min/d), equating to 98.3% (SD: 0.6%); at follow-up, average wear time (excluding study-related exercise sessions) was 1,393.1 min/d (SD: 31.7 min/d) or 97.8% (SD: 2.1%). Baseline habitual MVPA was 61.2 min/d (SD: 46.9 min/d) on average, with an average intensity of 3.7 MET (SD: 0.2 MET) and 99.2% (SD: 0.2%) of all MVPA below 6 MET. Total habitual PA, measured as steps per day, was 6,300 (SD: 2,301) at baseline. Total duration and intensity of daily habitual MVPA, habitual steps per day, and habitual PAEE (all outside of the structured exercise sessions) did not change significantly from baseline to follow-up (all P values ≥0.094). Average self-chosen exercise intensity during the intervention was 6.9 MET (SD: 1.0 MET), with no significant difference between the 20-KKW group and the 8-KKW group ($P = 0.074$). This average exercise intensity corresponds to vigorous PA (3).

Table 2 and Figure 1 (MVPA), Table 3 (steps per day), Table 4 (SenseWear PAEE), and Table 5 (DLW-estimated AEE) show the results of the multiple linear regression analyses. We found significant negative

TABLE 1 Baseline characteristics of 108 included participants

	All ($N=108$)	8 KKW ($n=57$)	20 KKW ($n=51$)
Female, *n* (%)	78 (72.2)	42 (73.7)	36 (70.6)
Ethnicity, *n* (%)			
Caucasian	74 (68.5)	37 (64.9)	37 (72.5)
African American	32 (29.6)	20 (35.1)	12 (23.5)
Hispanic/other	2 (1.9)	0 (0.0)	2 (4.0)
Age (y)	48.7 (11.6)	48.3 (11.0)	49.1 (12.4)
Height (cm)	167.1 (8.2)	167.2 (8.7)	167.0 (7.6)
Weight (kg)	87.8 (15.5)	89.0 (16.0)	86.5 (15.1)
Waist circumference (cm)	97.8 (12.0)[a]	98.7 (12.1)[b]	97.0 (11.9)
BMI (kg/m^2)	31.4 (4.6)	31.8 (4.6)	30.9 (4.5)
Fat mass (kg)	36.8 (9.8)	37.3 (9.7)	36.2 (9.9)
EI, DLW (kcal/d)	2,497.7 (462.5)	2,530.1 (438.8)	2,461.5 (489.4)
EI, buffet (kcal at lunch and dinner combined)	1,795.5 (550.7)	1,820.1 (489.7)	1,768.1 (615.7)
RMR (kcal/d)	1,529.1 (297.1)[a]	1,525.8 (261.2)	1,532.8 (334.7)
MVPA (min/d)	61.2 (46.9)	63.9 (49.5)	58.2 (44.0)
Average intensity of MVPA (MET)	3.7 (0.2)	3.8 (0.2)	3.7 (0.2)
Steps per day	6,300 (2,301)	6,576 (2,613)	5,992 (1,870)
PAEE, SenseWear (kcal/d)	336.7 (257.8)	349.1 (257.6)	322.9 (259.8)
AEE, DLW (kcal/d)	717.5 (216.6)[a]	749.3 (201.3)[b]	682.6 (229.2)

Data are mean (SD) if not stated otherwise. ANOVA (continuous variables) and χ^2 test (categorical variables) used to test for baseline differences between two groups. The 8-KKW and 20-KKW groups did not differ significantly in any baseline measures presented in table.
[a]Data available in 107 of 108 participants.
[b]Data available in 56 of 57 participants.
AEE, activity energy expenditure; DLW, doubly labeled water; EI, energy intake; KKW, kcal/kg of body weight/wk; MET, metabolic equivalent; MVPA, moderate-to-vigorous physical activity; PAEE, physical AEE; RMR, resting metabolic rate.

associations between baseline habitual MVPA levels and weight change ($P=0.020$; Figure 1A), weight compensation ($P=0.043$; Figure 1B), and change in DLW-estimated EI both with ($P=0.001$; Figure 1C) and without ($P=0.001$; not shown in Figure 1) adjustment for change in RMR. The analyses further showed significant negative associations between baseline habitual MVPA levels and changes in waist circumference ($P=0.030$), fat mass ($P=0.025$), and habitual MVPA levels ($P<0.001$; Figure 1D). Although there is substantial variability in the data (Figure 1), these results suggest that, on average, for every 15-minute decrease in habitual MVPA per day at baseline, participants lost 0.23 kg less weight, compensated 0.20 kg more, and increased DLW-estimated daily EI from baseline to follow-up by 21.5 kcal/d (adjusted for RMR of 23.2 kcal/d).

Compared with women, men lost 1.9 kg less weight, compensated 1.8 kg more, and increased DLW-estimated EI by 182.4 kcal/d (adjusted for RMR of 171.6 kcal/d) (Table 2). Further, compared with participants in the 8-KKW group, participants in the 20-KKW group lost 1.4 kg more weight but showed 1-kg-higher weight compensation (Table 2).

Baseline levels of habitual MVPA were significantly correlated with the disinhibition subscale of the Eating Inventory ($r=-0.229$; $P=0.018$) and with the binge eating subscale of the Multifactorial Assessment of Eating Disorders Symptoms ($r=-0.230$; $P=0.018$). No other correlations between baseline levels of habitual MVPA and eating behavior–related constructs, as assessed by questionnaires, were significant.

Baseline PA levels measured as steps per day significantly predicted change in steps per day ($\beta=-0.382$; $P<0.001$); however, no associations between steps per day at baseline and change in any other of the outcome variables were significant (Table 3). Associations between average intensity (MET) of baseline habitual MVPA and all outcomes were nonsignificant (all $P>0.1$, data not shown).

Associations between baseline habitual PAEE and outcomes were similar to those of baseline habitual MVPA, albeit slightly attenuated, as indicated by the regression coefficients (Table 4). DLW-estimated AEE only significantly predicted change in DLW-estimated EI (Table 5); all other associations were nonsignificant (all P values ≥ 0.709).

As described above, habitual MVPA, steps per day, and PAEE (all outside of structured exercise sessions) did not change significantly from baseline to follow-up on a group level. However, on an individual level, baseline habitual MVPA (Table 2, Figure 1D), steps per day (Table 3), and PAEE (Table 4) were significantly inversely associated with change in the respective measure.

Supporting Information Tables S2-S5 show the results of the multiple linear regression analyses for the control group. For habitual MVPA (Supporting Information Table S2), steps per day (Supporting Information Table S3), and PAEE (Supporting Information Table S4), only change in each PA measure was significantly associated with the respective baseline value. A Fisher *r*-to-*z* transformation revealed that the correlation coefficients for habitual MVPA did not differ between exercisers and the control

TABLE 2 Multiple linear regression analysis for association between baseline habitual MVPA levels and changes in body weight, fat mass, EI, and MVPA levels

	R^2	B	SE	β	*P*
Weight change (kg)	0.124				
Habitual MVPA at baseline (min/d)		−0.015	0.006	−0.275	**0.020**
Weight at baseline (kg)		−0.033	0.019	−0.197	0.095
Sex[a]		1.880	0.716	0.328	**0.010**
Group[b]		−1.442	0.483	−0.280	**0.004**
Waist circumference change (cm)	0.074				
Habitual MVPA at baseline (min/d)		−0.019	0.009	−0.267	**0.030**
Waist circumference at baseline (cm)		−0.034	0.034	−0.122	0.317
Sex[a]		1.357	0.947	0.185	0.155
Group[b]		−1.325	0.638	−0.201	**0.040**
Weight compensation (kg)	0.127				
Habitual MVPA at baseline (min/d)		−0.013	0.006	−0.238	**0.043**
Weight at baseline (kg)		−0.011	0.020	−0.064	0.585
Sex[a]		1.826	0.723	0.315	**0.013**
Group[b]		1.049	0.488	0.201	**0.034**
Fat mass change (kg)	0.153				
Habitual MVPA at baseline (min/d)		−0.014	0.006	−0.257	**0.025**
Fat mass at baseline (kg)		−0.070	0.027	−0.271	**0.011**
Sex[a]		1.173	0.562	0.209	**0.039**
Group[b]		−1.478	0.462	−0.293	**0.002**
Change in EI, DLW (kcal/d)	0.220				
Habitual MVPA at baseline (min/d)		−1.546	0.442	−0.336	**0.001**
EI, DLW at baseline (kcal/d)		−0.206	0.054	−0.443	**<0.001**
Sex[a]		182.400	57.888	0.381	**0.002**
Group[b]		−11.680	37.824	−0.027	0.758
Change in EI, adjusted DLW kcal/d)[c]	0.205				
Habitual MVPA at baseline (min/d)		−1.436	0.438	−0.318	**0.001**
EI, DLW at baseline (kcal/d)		−0.195	0.054	−0.427	**<0.001**
Sex[a]		171.594	57.370	0.365	**0.003**
Group[b]		3.388	37.486	0.008	0.928
Change EI, buffet (kcal at lunch and dinner combined)	0.158				
Habitual MVPA at baseline (min/d)		−0.938	0.951	−0.098	0.326
EI, buffet at baseline (kcal at lunch and dinner combined)		−0.287	0.077	−0.354	**<0.001**
Sex[a]		319.655	102.282	0.322	**0.002**
Group[b]		21.191	80.993	0.024	0.794
Change in RMR, indirect calorimetry (kcal/d)	0.067				
Habitual MVPA at baseline (min/d)		0.210	0.641	0.036	0.744
RMR at baseline (kcal/d)		−0.160	0.109	−0.195	0.146
Sex[a]		−44.702	77.817	−0.081	0.567
Group[b]		41.109	50.933	0.083	0.422
Change in habitual MVPA (min/d)	0.223				
Habitual MVPA at baseline (min/d)		−0.274	0.058	−0.450	**<0.001**
Sex[a]		−2.425	5.996	−0.039	0.687
Group[b]		7.963	4.992	0.141	0.114

Bold font indicates statistical significance ($P<0.05$). Independent variable in all models: habitual MVPA levels (min/d) at baseline.
[a]Female = 0, male = 1.
[b]8 kcal/kg of body weight per week = 0, 20 kcal/kg of body weight per week = 1.
[c]Adjusted for change in RMR.
DLW, doubly labeled water; EI, energy intake; MVPA, moderate-to-vigorous physical activity; RMR, resting metabolic rate.

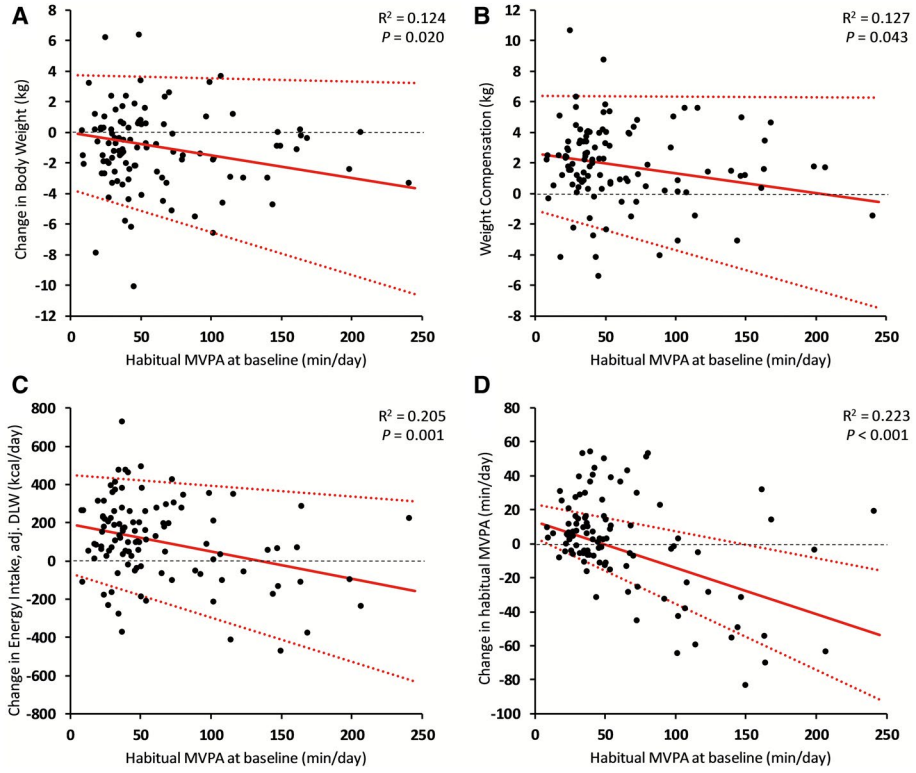


Figure 1 Association between habitual moderate-to-vigorous physical activity (MVPA) at baseline and change in (**A**) body weight, (**B**) weight compensation, (**C**) change in doubly labeled water (DLW)-measured energy intake, adjusted for change in resting metabolic rate, and (**D**) change in MVPA. Regression line (solid line) in each panel represents the relationship for the fully adjusted model with 95% confidence intervals (dotted lines). [Color figure can be viewed at wileyonlinelibrary.com]

group (data not shown). For habitual PAEE, the difference between exercisers and control participants was significant, with a markedly more pronounced association for the control participants.

Discussion

To our knowledge, this is the first study to determine and compare the effect of prior habitual MVPA, steps per day, and PAEE on changes in weight, EI, RMR, and MVPA, steps per day, and PAEE in response to a moderate- to high-dose aerobic exercise intervention. The results show that, on average, lower levels of habitual MVPA and/or PAEE at baseline are related to less weight loss and greater weight compensation during the exercise intervention, supporting our hypothesis. Importantly, lower levels of habitual MVPA and/or PAEE at baseline were also associated with greater increases in EI, which likely contributed to the lower weight compensation in those with higher baseline levels of habitual MVPA and/or PAEE, particularly because changes in RMR were not associated with baseline habitual MVPA and/or PAEE levels. Interestingly, we found substantial heterogeneity in the weight loss/compensation response, which likely influenced the results of the regression analysis. Although many participants across all baseline MVPA levels successfully lost weight during the intervention, some participants with low MVPA at baseline actually gained weight, whereas no one with higher baseline MVPA gained weight.

In line with previous findings (30,31), participants with lower habitual MVPA levels showed higher tendencies for disinhibition and binge eating at baseline, factors that may have influenced the greater increases in EI and subsequent greater weight compensation in response to the exercise intervention. This assumption is supported by previous findings showing that individuals with lower levels of measured MVPA have weaker appetite control and satiety response to food and thus have an impaired regulation of energy balance compared with their more active counterparts (32-34). Consequently, in our study, participants with lower levels of habitual MVPA and/or PAEE at baseline may have had a more impaired regulation of energy balance than those with higher levels of habitual MVPA and/or PAEE, and this became particularly apparent with the onset of the exercise intervention. Although participants were weight stable during the 2-week baseline accelerometer assessment, suggesting an adequately regulated energy balance during that period, the exercise intervention and, subsequently, the substantial increase in daily EE disrupted this balance. This disruption may have revealed the potentially impaired energy balance regulation in participants with lower baseline levels of habitual MVPA and/or PAEE, as the intervention-related increases in MVPA and/or PAEE (i.e., structured exercise sessions) were

TABLE 3 Multiple linear regression analysis for association between habitual steps per day at baseline and changes in body weight, fat mass, EI, and steps per day

	R^2	B	SE	β	*P*
Weight change (kg)	0.091				
Habitual PA at baseline (steps/d)		−0.0001	0.0001	−0.1233	0.211
Weight at baseline (kg)		−0.0156	0.0177	−0.0945	0.379
Sex[a]		1.0394	0.6043	0.1812	0.088
Group[b]		−1.3705	0.4924	−0.2664	**0.006**
Waist circumference change (cm)	0.075				
Habitual PA at baseline (steps/d)		−0.0001	0.0001	−0.2324	0.057
Waist circumference at baseline (cm)		−0.0223	0.0311	−0.0807	0.476
Sex[a]		0.5600	0.7905	0.0763	0.480
Group[b]		−1.3730	0.6409	−0.2082	**0.035**
Weight compensation (kg)	0.098				
Habitual PA at baseline (steps/d)		−0.0001	0.0001	−0.0818	0.404
Weight at baseline (kg)		0.0052	0.0179	0.0310	0.772
Sex[a]		1.0611	0.6094	0.1828	0.085
Group[b]		1.1322	0.4965	0.2174	**0.025**
Fat mass change (kg)	0.124				
Habitual PA at baseline (steps/d)		−0.0001	0.0001	−0.1223	0.212
Fat mass at baseline (kg)		−0.0505	0.0252	−0.1954	**0.048**
Sex[a]		0.6968	0.5256	0.1238	0.188
Group[b]		−1.4435	0.4715	−0.2860	**0.003**
Change in EI, DLW (kcal/d)	0.147				
Habitual PA at baseline (steps/d)		−0.0138	0.0089	−0.1475	0.128
EI, DLW at baseline (kcal/d)		−0.1923	0.0589	−0.4127	**0.001**
Sex[a]		112.1006	58.4187	0.2340	0.058
Group[b]		−7.8955	39.6422	−0.0183	0.843
Change in EI, adjusted DLW kcal/d)[c]	0.140				
Habitual PA at baseline (steps/d)		−0.0128	0.0088	−0.1397	0.150
EI, DLW at baseline (kcal/d)		−0.1823	0.0580	−0.3985	**0.002**
Sex[a]		106.2766	57.5923	0.2260	0.068
Group[b]		6.8974	39.0800	0.0163	0.860
Change EI, buffet (kcal at lunch and dinner combined)	0.173				
Habitual PA at baseline (steps/d)		−0.0298	0.0177	−0.1537	0.095
EI, buffet at baseline (kcal at lunch and dinner combined)		−0.2771	0.0769	−0.3416	**<0.001**
Sex[a]		288.0988	93.6300	0.2901	**0.003**
Group[b]		10.5387	80.6567	0.0118	0.896
Change in RMR, indirect calorimetry (kcal/d)	0.066				
Habitual PA at baseline (steps/d)		−0.0006	0.0109	−0.0057	0.956
RMR at baseline (kcal/d)		−0.1629	0.1086	−0.1984	0.137
Sex[a]		−35.6318	72.7619	−0.0647	0.626
Group[b]		38.1756	51.0429	0.0774	0.456
Change in habitual PA (steps/d)	0.173				
Habitual PA at baseline (steps/d)		−0.3591	0.0870	−0.3823	**<0.001**
Sex[a]		74.8373	438.7446	0.0156	0.865
Group[b]		500.2422	399.3502	0.1159	0.213

Bold font indicates statistical significance ($P<0.05$). Independent variable in all models: habitual PA levels (steps/d) at baseline.
[a]Female = 0, male = 1.
[b]8 kcal/kg of body weight per week = 0, 20 kcal/kg of body weight per week = 1.
[c]Adjusted for change in RMR.
DLW, doubly labeled water; EI, energy intake; PA, physical activity; RMR, resting metabolic rate.

TABLE 4 Multiple linear regression analysis for association between baseline habitual PAEE as assessed by SenseWear armband and changes in body weight, fat mass, EI, and SenseWear-assessed PAEE

	R^2	B	SE	β	*P*
Weight change (kg)	0.121				
Habitual PAEE at baseline (kcal/d)		−0.003	0.001	−0.260	**0.024**
Weight at baseline (kg)		−0.023	0.018	−0.136	0.208
Sex[a]		1.856	0.718	0.324	**0.011**
Group[b]		−1.400	0.481	−0.272	**0.004**
Waist circumference change (cm)	0.059				
Habitual PAEE at baseline (kcal/d)		−0.003	0.002	−0.220	0.074
Waist circumference at baseline (cm)		−0.019	0.032	−0.068	0.553
Sex[a]		1.222	0.980	0.167	0.215
Group[b]		−1.265	0.641	−0.192	0.051
Weight compensation (kg)	0.121				
Habitual PAEE at baseline (kcal/d)		−0.002	0.001	−0.212	0.065
Weight at baseline (kg)		−0.001	0.018	−0.008	0.944
Sex[a]		1.756	0.727	0.303	**0.017**
Group[b]		1.093	0.487	0.210	**0.027**
Fat mass change (kg)	0.156				
Habitual PAEE at baseline (kcal/d)		−0.003	0.001	−0.261	**0.021**
Fat mass at baseline (kg)		−0.059	0.025	−0.230	**0.019**
Sex[a]		1.376	0.597	0.245	**0.023**
Group[b]		−1.462	0.460	−0.290	**0.002**
Change in EI, DLW (kcal/d)	0.228				
Habitual PAEE at baseline (kcal/d)		−0.322	0.088	−0.385	**<0.001**
EI, DLW at baseline (kcal/d)		−0.169	0.055	−0.362	**0.003**
Sex[a]		186.132	57.668	0.389	**0.002**
Group[b]		−8.899	37.567	−0.021	0.813
Change in EI, adjusted DLW kcal/d)[c]	0.212				
Habitual PAEE at baseline (kcal/d)		−0.298	0.087	−0.363	**0.001**
EI, DLW at baseline (kcal/d)		−0.160	0.055	−0.351	**0.004**
Sex[a]		174.869	57.211	0.372	**0.003**
Group[b]		5.993	37.269	0.014	0.873
Change EI, buffet (kcal at lunch and dinner combined)	0.164				
Habitual PAEE at baseline (kcal/d)		−0.246	0.184	−0.142	0.184
EI, buffet at baseline (kcal at lunch and dinner combined)		−0.284	0.077	−0.350	**<0.001**
Sex[a]		351.316	107.794	0.354	**0.002**
Group[b]		19.233	80.663	0.022	0.812
Change in RMR, indirect calorimetry (kcal/d)	0.066				
Habitual PAEE at baseline (kcal/d)		−0.025	0.127	−0.023	0.846
RMR at baseline (kcal/d)		−0.163	0.108	−0.198	0.136
Sex[a]		−29.457	79.345	−0.053	0.711
Group[b]		37.075	51.013	0.075	0.469
Change in habitual PAEE (kcal/d)	0.289				
Habitual PAEE at baseline (kcal/d)		−0.400	0.069	−0.564	**<0.001**
Sex[a]		52.423	38.970	0.131	0.182
Group[b]		43.443	30.461	0.121	0.157

Bold font indicates statistical significance ($P<0.05$). Independent variable in all models: habitual PAEE (kcal/d) at baseline as assessed by SenseWear armband.
[a]Female = 0, male = 1.
[b]8 kcal/kg of body weight per week = 0, 20 kcal/kg of body weight per week = 1.
[c]Adjusted for change in RMR.
DLW, doubly labeled water; EI, energy intake; PAEE, physical activity energy expenditure; RMR, resting metabolic rate.

TABLE 5 Multiple linear regression analysis for association between baseline habitual AEE as assessed by DLW and changes in body weight, fat mass, and EI

	R^2	B	SE	β	*P*
Weight change (kg)	0.076				
Habitual AEE at baseline (kcal/d)		−0.001	0.001	−0.041	0.709
Weight at baseline (kg)		−0.009	0.018	−0.053	0.630
Sex[a]		1.009	0.620	0.177	0.107
Group[b]		−1.250	0.496	−0.244	**0.013**
Waist circumference change (cm)	0.027				
Habitual AEE at baseline (kcal/d)		0.001	0.002	0.035	0.749
Waist circumference at baseline (cm)		0.003	0.030	0.012	0.911
Sex[a]		0.098	0.829	0.014	0.906
Group[b]		−0.999	0.654	−0.153	0.130
Weight compensation (kg)	0.098				
Habitual AEE at baseline (kcal/d)		−0.001	0.001	−0.013	0.906
Weight at baseline (kg)		0.009	0.018	0.053	0.626
Sex[a]		1.031	0.624	0.178	0.101
Group[b]		1.237	0.499	0.237	**0.015**
Fat mass change (kg)	0.109				
Habitual AEE at baseline (kcal/d)		−0.001	0.001	−0.019	0.854
Fat mass at baseline (kg)		−0.040	0.025	−0.155	0.111
Sex[a]		0.749	0.583	0.134	0.202
Group[b]		−1.324	0.479	−0.263	**0.007**
Change in EI, DLW (kcal/d)	0.171				
Habitual AEE at baseline (kcal/d)		−0.317	0.137	−0.317	**0.023**
EI, DLW at baseline (kcal/d)		−0.095	0.078	−0.204	0.223
Sex[a]		99.726	58.661	0.208	0.092
Group[b]		−13.740	39.562	−0.032	0.729
Change in EI, adjusted DLW kcal/d)[c]	0.167				
Habitual AEE at baseline (kcal/d)		−0.320	0.135	−0.326	**0.020**
EI, DLW at baseline (kcal/d)		−0.082	0.076	−0.179	0.287
Sex[a]		92.237	57.688	0.196	0.113
Group[b]		0.181	38.906	0.001	0.996
Change in EI, buffet (kcal at lunch and dinner combined)	0.149				
Habitual AEE at baseline (kcal/d)		0.070	0.210	0.034	0.741
EI, buffet at baseline (kcal at lunch and dinner combined)		−0.295	0.079	−0.366	**<0.001**
Sex[a]		264.367	101.080	0.267	**0.010**
Group[b]		25.578	82.706	0.029	0.758
Change in RMR, indirect calorimetry (kcal/d)	0.067				
Habitual AEE at baseline (kcal/d)		0.043	0.128	0.038	0.736
RMR at baseline (kcal/d)		−0.165	0.108	−0.201	0.132
Sex[a]		−42.818	75.767	−0.078	0.573
Group[b]		42.681	51.771	0.087	0.412

Bold font indicates statistical significance ($P < 0.05$). Independent variable in all models: habitual AEE (kcal/d) at baseline as assessed by DLW (total daily energy expenditure − [RMR + thermic effect of eating]) Thermic effect of food estimated as 10% of total daily energy expenditure.
[a]Female = 0, male = 1.
[b]8 kcal/kg of body weight per week = 0, 20 kcal/kg of body weight per week = 1.
[c]Adjusted for change in RMR.
AEE, activity energy expenditure; DLW, doubly labeled water; EI, energy intake; RMR, resting metabolic rate.

met by increases in EI, leading to the observed weight compensation. Participants with higher MVPA and/or PAEE levels at baseline might have already experienced this compensatory effort before the start of the intervention, explaining, at least partially, the observed results. In addition to being driven by homeostatic mechanisms such as the aforementioned changes in appetite and satiety, the observed

increases in EI may also be related to hedonic processes such as food reward behaviors (35).

It is noteworthy that although habitual MVPA and/or PAEE levels did not change on a group level, on an individual level, these parameters were significantly inversely associated with change in the respective measure, indicating the substitution of habitual PA with prescribed PA (i.e., structured exercise session) in some participants (36,37). As shown by the results of a Fisher *r*-to-*z* transformation, however, the correlation coefficients for habitual MVPA did not differ between exercisers and the control group, suggesting that any decrease in MVPA in the exercisers was likely not caused by the structured exercise sessions but instead ocurred more likely because of regression to the mean. For habitual PAEE, the difference between exercisers and control participants was significant, with a substantially more pronounced association for the control participants, suggesting that the structured exercise sessions actually protected against decreases in habitual PAEE. It is further noteworthy that participants with greater prior habitual MVPA and/or PAEE remained more active compared with those with lower levels ($B = -0.274$ [MVPA] and $B = -0.400$ [PAEE]). Therefore, considering the magnitude of the change in habitual MVPA and/or PAEE levels and, more importantly, the opposite directionality compared with weight change, it is unlikely that the decrease in habitual MVPA and/or PAEE affected participants' weight compensation. Rather, higher absolute levels of MVPA and/or PAEE at follow-up, along with the reduced increase in EI during the intervention, contributed to the lower weight compensation in those who were more active at baseline.

The identification of baseline habitual MVPA and/or PAEE levels as predictors of weight loss, weight compensation, and changes in EI in this study may have important ramifications for future exercise interventions targeting weight loss. Less than expected weight loss from exercise likely leads to frustration and possibly causes discontinuation of the newly started exercise regimen because of the perceived lack of benefit. Assessing prior habitual PA levels may help determine when the exercise prescription should be combined with a concomitant lifestyle, dietary, or possibly pharmacological intervention to counteract weight compensation and increase the weight loss intervention–related health benefits.

Although habitual MVPA and PAEE predicted our outcomes quite comparably, daily step counts at baseline did not have the same predictive value with regard to weight loss, weight compensation, or EI during the intervention as habitual MVPA and/or PAEE. The better predictive value of PAEE compared with steps per day was expected because of the fact that PAEE is directly related to the energy balance model. The better predictive value of habitual MVPA compared with steps per day is likely because of the fact that MVPA includes an intensity component, whereas steps per day does not. Therefore, to identify individuals with a higher risk for exercise-induced weight compensation, baseline levels of habitual MVPA or PAEE should be considered. It should be noted that AEE, as estimated by DLW, did not predict most of our outcomes, with a substantial discrepancy compared with the associations from SenseWear-assessed PAEE. This suggests that the intensity component included in PAEE (and MVPA) made these parameters better predictors with regard to our outcomes. Therefore, total AEE seems to be less important than EE at an intensity ≥3 MET, which is different from Mayer et al.'s original suggestion (11,12). The use of MVPA as a predictor offers the advantages of being accurately assessable via most current accelerometers and of accelerometer data being more straightforward compared with PAEE data such as that of the SenseWear armband, which is based on a complex pattern-recognition algorithm consisting of heat flux, skin temperature, near-body ambient temperature, and galvanic skin response, in addition to the accelerometer-recorded activity counts.

The present study has several strengths. E-MECHANIC was a large randomized controlled trial, in which exercise dose was strictly monitored and supervised. Habitual PA was measured with validated accelerometers that allow an estimation of the intensity and EE of habitual PA. Additionally, EE and EI (via DLW) and RMR (via indirect calorimetry) were measured with the gold-standard methods to comprehensively assess all aspects of energy balance. The assessment of EI via validated laboratory-based FI tests and via DLW over 2 weeks is a particularly major strength, as self-reported EI, which is still commonly used in many trials today, has been found to be fundamentally inaccurate (38,39). A limitation of this analysis is that although PA assessment at follow-up was performed while participants were still exercising at their prescribed dose, we did not measure habitual MVPA, steps per day, and PAEE continuously throughout the intervention period and thus have no record of the effect of the exercise training on these outcomes over the course of the intervention.

In conclusion, taking into account the substantial variability in the data, our results show that habitual MVPA and/or PAEE levels before engaging in a structured exercise intervention predict weight loss, weight compensation, and changes in EI during that intervention. Importantly, habitual MVPA and/or PAEE (≥3 MET) at baseline showed a superior predictive value with regard to these outcome measures compared with steps per day and total AEE, suggesting that time spent and energy expended during MVPA rather than total activity-related EE before an exercise intervention targeting weight loss are protective against weight compensation. In this regard, habitual MVPA may be the preferable parameter compared with PAEE because of its easier, more economical, and (likely) more accurate assessment. Future studies are needed to elucidate the observed heterogeneous relationship between baseline habitual MVPA and/or PAEE levels and weight loss and compensation to develop individualized strategies to mitigate the detrimental compensatory increase in EI in response to an exercise-induced increase in EE in some individuals.O

Acknowledgments

The authors would like to thank participants for their time and commitment to the study. Access to individual deidentified data that underlie the results reported in this article is possible via Pennington Biomedical Research Center's data sharing policy.

Funding agencies: Research reported in this publication was supported by NIH via the National Heart, Lung, and Blood Institute, with the multiple principal investigators being CKM and T. Church (R01 HL102166); Nutrition Obesity Research Center (NORC) Grant P30 DK072476, titled "Nutritional Programming: Environmental and Molecular Interactions" sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK); and the National Institute of General Medical Sciences, which funds the Louisiana Clinical and Translational Science Center (U54 GM104940). CH is funded by an NIH NIDDK National Research Service Award (T32 DK064584).

Disclosure: CKM gives lectures and presentations on the topic to lay and educational groups, including the Academy of Nutrition and Dietetics. The authors report no other competing interests related to this study.

Clinical trial registration: ClinicalTrials.gov identifier NCT01264406.

Supporting information: Additional Supporting Information may be found in the online version of this article.

References

1. NCD Risk Factor Collaboration (NCD-RisC). Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19·2 million participants. *Lancet* 2016;387:1377-1396.
2. Heymsfield SB, Wadden TA. Mechanisms, pathophysiology, and management of obesity. *N Engl J Med* 2017;376:254-266.
3. Garber CE, Blissmer B, Deschenes MR, et al. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Med Sci Sports Exerc* 2011;43:1334-1359.
4. Swift DL, McGee JE, Earnest CP, Carlisle E, Nygard M, Johannsen NM. The effects of exercise and physical activity on weight loss and maintenance. *Prog Cardiovasc Dis* 2018;61:206-213.
5. Church TS, Martin CK, Thompson AM, Earnest CP, Mikus CR, Blair SN. Changes in weight, waist circumference and compensatory responses with different doses of exercise among sedentary, overweight postmenopausal women. *PLoS One* 2009;4:e4515. doi:10.1371/journal.pone.0004515
6. Martin CK, Johnson WD, Myers CA, et al. Effect of different doses of supervised exercise on food intake, metabolism, and non-exercise physical activity: The E-MECHANIC randomized controlled trial. *Am J Clin Nutr* 2019;110:583-592.
7. Thomas DM, Bouchard C, Church T, et al. Why do individuals not lose more weight from an exercise intervention at a defined dose? An energy balance analysis. *Obes Rev* 2012;13:835-847.
8. Blundell JE, Stubbs RJ, Hughes DA, Whybrow S, King NA. Cross talk between physical activity and appetite control: does physical activity stimulate appetite? *Proc Nutr Soc* 2003;62:651-661.
9. King NA, Hopkins M, Caudwell P, Stubbs RJ, Blundell JE. Individual variability following 12 weeks of supervised exercise: identification and characterization of compensation for exercise-induced weight loss. *Int J Obes (Lond)* 2008;32:177-184.
10. Westerterp KR. Exercise for weight loss. *Am J Clin Nutr* 2019;110:540-541.
11. Mayer J, Roy P, Mitra KP. Relation between caloric intake, body weight, and physical work: studies in an industrial male population in West Bengal. *Am J Clin Nutr* 1956;4:169-175.
12. Church T, Martin CK. The obesity epidemic: a consequence of reduced energy expenditure and the uncoupling of energy intake? *Obesity (Silver Spring)* 2018;26:14-16.
13. Myers CA, Johnson WD, Earnest CP, et al. Examination of mechanisms (E-MECHANIC) of exercise-induced weight compensation: study protocol for a randomized controlled trial. *Trials* 2014;15:212. doi:10.1186/1745-6215-15-212
14. Tudor-Locke C, Leonardi C, Johnson WD, Katzmarzyk PT, Church TS. Accelerometer steps/day translation of moderate-to-vigorous activity. *Prev Med* 2011;53:31-33.
15. West Suitor C, Kraak VI; Institute of Medicine. *Adequacy of Evidence for Physical Activity Guidelines Development: Workshop Summary*. Washington, DC: The National Academies Press; 2007.
16. American College of Sports Medicine. *ACSM's Guidelines for Exercise Testing and Prescription*. 9th ed. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins Health; 2014.
17. Thomas DM, Martin CK, Lettieri S, et al. Can a weight loss of one pound a week be achieved with a 3500-kcal deficit? Commentary on a commonly accepted rule. *Int J Obes (Lond)* 2013;37:1611-1613.
18. Heymsfield SB, Thomas D, Nguyen AM, et al. Voluntary weight loss: systematic review of early phase body composition changes. *Obes Rev* 2011;12:e348-e361. doi:10.1111/j.1467-789X.2010.00767.x
19. Thomas DM, Ciesla A, Levine JA, Stevens JG, Martin CK. A mathematical model of weight change with adaptation. *Math Biosci Eng* 2009;6:873-887.
20. Livingstone MBE, Black AE. Markers of the validity of reported energy intake. *J Nutr* 2003;133(suppl 3):895S-920S.
21. Schoeller DA. How accurate is self-reported dietary energy intake? *Nutr Rev* 1990; 48:373-379.
22. Martin CK, Coulon SM, Markward N, Greenway FL, Anton SD. Association between energy intake and viewing television, distractibility, and memory for advertisements. *Am J Clin Nutr* 2009;89:37-44.
23. Piercy KL, Troiano RP, Ballard RM, et al. The Physical Activity Guidelines for Americans. *JAMA* 2018;320:2020-2028.
24. Jones V, Bults R, de Wijk R, Widya I, Batista R, Hermens H. Experience with using the SenseWear BMS sensor system in the context of a health and wellbeing application. *Int J Telemed Appl* 2011:671040. doi:10.1155/2011/671040
25. Womble LG, Wadden TA, Chandler JM, Martin AR. Agreement between weekly vs. daily assessment of appetite. *Appetite* 2003;40:131-135.
26. Stunkard AJ, Messick S. *Eating Inventory Manual (The Psychological Corporation)*. San Antonio, TX: Harcourt Brace; 1988.
27. Anderson DA, Williamson DA, Duchmann EG, Gleaves DH, Barbin JM. Development and validation of a multifactorial treatment outcome measure for eating disorders. *Assessment* 1999;6:7-20.
28. Geiselman PJ, Anderson AM, Dowdy ML, West DB, Redmann SM, Smith SR. Reliability and validity of a macronutrient self-selection paradigm and a food preference questionnaire. *Physiol Behav* 1998;63:919-928.
29. White MA, Whisenhunt BL, Williamson DA, Greenway FL, Netemeyer RG. Development and validation of the food-craving inventory. *Obes Res* 2002;10:107-114.
30. Shook RP, Hand GA, Drenowatz C, et al. Low levels of physical activity are associated with dysregulation of energy intake and fat mass gain over 1 year. *Am J Clin Nutr* 2015;102:1332-1338.
31. Myers A, Gibbons C, Finlayson G, Blundell J. Associations among sedentary and active behaviours, body fat and appetite dysregulation: investigating the myth of physical inactivity and obesity. *Br J Sports Med* 2017;51:1540-1544.
32. Beaulieu K, Hopkins M, Long C, Blundell J, Finlayson G. High habitual physical activity improves acute energy compensation in nonobese adults. *Med Sci Sports Exerc* 2017;49:2268-2275.
33. Beaulieu K, Hopkins M, Blundell J, Finlayson G. Homeostatic and non-homeostatic appetite control along the spectrum of physical activity levels: an updated perspective. *Physiol Behav* 2018;192:23-29.
34. Beaulieu K, Hopkins M, Blundell J, Finlayson G. Does habitual physical activity increase the sensitivity of the appetite control system? A systematic review. *Sports Med* 2016;46:1897-1919.
35. Dalton M, Finlayson G, Esdaile E, King N. Appetite, satiety, and food reward in obese individuals: a behavioral phenotype approach. *Curr Nutr Rep* 2013;2:207-215.
36. Thompson D, Peacock OJ, Betts JA. Substitution and compensation erode the energy deficit from exercise interventions. *Med Sci Sports Exerc* 2014;46:423. doi:10.1249/MSS.0000000000000164
37. Thompson D, Karpe F, Lafontan M, Frayn K. Physical activity and exercise in the regulation of human adipose tissue physiology. *Physiol Rev* 2012;92:157-191.
38. Dhurandhar NV, Schoeller D, Brown AW, et al. Energy balance measurement: when something is not better than nothing. *Int J Obes (Lond)* 2015;39:1109-1113.
39. Dorling J, Broom DR, Burns SF, et al. Acute and chronic effects of exercise on appetite, energy intake, and appetite-related hormones: the modulating effect of adiposity, sex, and habitual physical activity. *Nutrients* 2018;10:E1140. doi:10.3390/nu10091140

Publication 3

Initial weight change and long-term changes in weight and compensation during supervised exercise training.

James L Dorling [1], **Christoph Höchsmann** [1], S Nicole Fearnbach [1], John W Apolzan [1], Daniel S Hsia [1], Neil M Johannsen [1,2], Tim S Church [1], Corby K Martin [1]

[1] Pennington Biomedical Research Center, Baton Rouge, LA, USA

[2] School of Kinesiology, Louisiana State University, Baton Rouge, LA, USA

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
Medicine & Science in Sports & Exercise, 53(8):1675–1684	2021	10.1249/MSS.0000000000002633	6.3 (2021)	Q1	Scopus: 4 Google Scholar: 8

Author contribution:

Second author; co-developed the research question and statistical model, co-drafted the manuscript, and created parts of the figures.

Initial Weight Change and Long-Term Changes in Weight and Compensation during Supervised Exercise Training

JAMES L. DORLING[1], CHRISTOPH HÖCHSMANN[1], S. NICOLE FEARNBACH[1], JOHN W. APOLZAN[1], DANIEL S. HSIA[1], NEIL M. JOHANNSEN[1,2], TIM S. CHURCH[1], and CORBY K. MARTIN[1]

[1]Pennington Biomedical Research Center, Baton Rouge, LA; and [2]Louisiana State University, Baton Rouge, LA

ABSTRACT

DORLING, J. L., C. HÖCHSMANN, S. N. FEARNBACH, J. W. APOLZAN, D. S. HSIA, N. M. JOHANNSEN, T. S. CHURCH, and C. K. MARTIN. Initial Weight Change and Long-Term Changes in Weight and Compensation during Supervised Exercise Training. *Med. Sci. Sports Exerc.*, Vol. 53, No. 8, pp. 1675–1684, 2021. **Introduction:** Our primary aim was to investigate the association between initial weight change and longer-term changes in weight and compensation (predicted weight loss–observed weight loss) during exercise. As secondary aims, we investigated if initial weight change was related to change in cardiometabolic risk markers and energy balance modulators. **Methods:** Two 6-month randomized controlled exercise trials conducted in individuals with overweight or obesity were analyzed (study 1, $n = 312$; study 2, $n = 102$). In both studies, participants in an exercise condition (4 kcal·kg^{-1}·wk^{-1} [KKW], 8 KKW, 12 KKW, or 20 KKW) were split into tertiles based on percent weight change from baseline to week 4. Tertiles 1 and 3 exhibited the least and most initial weight loss, respectively. Changes in end points were compared between tertiles. **Results:** At month 6, weight loss was lower in tertile 1 than tertile 3 (study 1: −3.6%, 95% confidence interval [CI] = −4.6 to −2.6; study 2: −1.8%, 95% CI = −3.1 to −0.4; $P \leq 0.034$). Tertile 1 also showed greater compensation than tertile 3 in study 1 (3.0 kg, 95% CI = 2.2 to 3.9) and study 2 (1.5 kg, 95% CI = 0.3 to 2.6; $P \leq 0.048$). Changes in triglycerides and, in study 1, HDL cholesterol were less favorable in tertile 1 versus tertile 3 ($P \leq 0.043$); however, changes in other cardiometabolic markers were similar ($P \geq 0.209$). In study 2, tertile 1 increased energy intake and exhibited maladaptive changes in eating behaviors relative to tertile 3 ($P < 0.050$). No between-tertile differences in cumulative exercise energy expenditure and physical activity were evident ($P \geq 0.321$). **Conclusions:** Less initial weight loss was associated with longer-term attenuated weight loss and greater compensation during aerobic exercise training. Individuals who display less initial weight loss during exercise may require early interventions to decrease compensation and facilitate weight loss. **Key Words:** CARDIOMETABOLIC HEALTH, ENERGY INTAKE, FOOD PREFERENCES, COMPENSATORY HEALTH BELIEFS, INITIAL WEIGHT LOSS, WEIGHT MANAGEMENT

Excess body weight is associated with increased risk for various physical and psychological conditions (1), and over two-thirds of Americans are either overweight or obese (2). Exercise training is advocated in individuals with overweight or obesity because it reduces the risk for a plethora of chronic diseases (3) and can stimulate clinically significant reductions in weight (4). However, interindividual variability in exercise-induced weight change exists. Within many exercise training studies, some individuals present clinically significant weight loss and others gain weight (5). Although exercise training triggers weight-independent benefits (4), substandard weight loss or weight gain attenuates many improvements of exercise (6,7). Thus, during exercise training, it is crucial to identify characteristics of individuals who experience attenuated weight loss or weight gain and determine the mechanisms underlying substandard weight change.

A reliable predictor of long-term weight loss within dietary interventions is initial weight change (e.g., weight change in first 4 wk) (8). Individuals who exhibit less initial weight loss display poorer changes in weight and cardiometabolic end points at the end of studies (9,10). This early indicator of long-term weight loss success allows interventionists to identify individuals who are likely to experience substandard weight-related outcomes and apply procedures to optimize end points (8). Such strategies look to decrease compensation, which is a discrepancy between weight loss achieved and expected (11), and can include more rigorous modifications of key dietary patterns, such as eating behaviors and attitudes (12).

No study has examined if initial weight change during exercise is related to longer-term changes in weight, compensation,

Address for correspondence: Corby K. Martin, Ph.D., Pennington Biomedical Research Center, 6400 Perkins Rd., Baton Rouge, LA 70808; E-mail: Corby.Martin@pbrc.edu.

Submitted for publication November 2020.
Accepted for publication February 2021.
Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.acsm-msse.org).

0195-9131/21/5308-1675/0
MEDICINE & SCIENCE IN SPORTS & EXERCISE®

DOI: 10.1249/MSS.0000000000002633

and cardiometabolic risk markers. Such investigations are important, given the heterogeneity in exercise-induced weight loss and the unknown indicators of this variability (13). It is also important to examine modulators of energy balance to ascertain the mechanisms underlying differences in weight change and compensation, and to inform approaches that augment the benefits of exercise training.

Our primary aim was to investigate the association between initial weight change from baseline to week 4 and changes in weight and compensation after 6 months of exercise training. We hypothesized that individuals who lose less weight from baseline to week 4 would experience lower weight loss and greater compensation after 6 months of training. As secondary aims, we investigated the associations between initial weight change and 6-month change in cardiometabolic risk markers, components of energy expenditure and energy intake, eating behaviors and attitudes, and compensatory health beliefs.

METHODS

The current analysis uses data from two supervised 6-month exercise intervention studies: the Dose–Response to Exercise in postmenopausal Women (DREW) study and the Examination of Mechanisms of Exercise-Induced Weight Compensation (E-MECHANIC) study. Trial design, randomization methods, trial dates, blinding procedures, and sample size calculations of both studies have been published (14–17). Previous interventions have found robust relationships between weight change from baseline to week 4 and prolonged weight change (18–20). Accordingly, we chose to examine the association between initial weight change at week 4 and weight change and compensation at the end of these studies. Compensation was defined as the difference between predicted weight loss and observed weight loss. Predicted weight loss was estimated in both studies using a dynamic energy balance model, which accounts for metabolic adaptation and body composition changes (fat mass and lean mass) during aerobic exercise training, and which overcomes limitations with conventional estimates that assume a 7700-kcal energy deficit leads to a 1-kg reduction in weight (see Appendix, Supplemental Digital Content, Supplementary Information, http://links.lww.com/MSS/C274) (17,21,22). This model has been validated on previous aerobic training studies (21).

To improve validity of our findings, participants were included if they 1) were randomized to an exercise condition, 2) completed the trial, 3) performed weight measurements at week 4, and 4) achieved >75% compliance (number of exercise sessions/prescribed exercise sessions) and/or adherence (achieved exercise energy expenditure/prescribed energy expenditure) to the exercise regimen (17,23,24).

The DREW Study

The DREW study (Clinical Trials.gov: NCT00011193) was performed at the Cooper Institute and was approved by the institute's Institutional Review Board. All participants provided written informed consent before screening. The study included females with overweight or obesity (body mass index [BMI] = 25.0–43.0 $kg \cdot m^{-2}$) and elevated systolic blood pressure (120.0–159.9 mm Hg). Other exclusion criteria have been detailed (14).

Detailed descriptions of the exercise intervention have been published (14,15). Participants were enrolled into either a no-exercise control group or one of three exercise groups for 6 months. The three exercise groups included a group that aimed to expend 4 $kcal \cdot kg^{-1} \cdot wk^{-1}$ (KKW), one that aimed to expend 8 KKW, and another that aimed to expend 12 KKW. All participants expended 4 KKW during the first week; thereafter, participants enrolled to the 4 KKW group continued at this dose, whereas the 8 KKW and the 12 KKW groups ramped up their exercise dose by 1 KKW every week until their prescribed dose was reached. Body weight measurements were collected before exercise sessions every week in the exercise training facility on an electronic scale (Siemens Medical Solutions, Malvern, PA). Exercise training was performed on semirecumbent cycle ergometers and treadmills, with the intensity set at a heart rate equivalent to 50% of baseline peak oxygen uptake ($\dot{V}O_2$). Using standard American College of Sports Medicine (ACSM) equations (25), the energy expenditure of exercise was calculated in real time based on the participant's weight and either watts (cycle ergometer) or speed and gradient (treadmill). Exercise time was adjusted by dividing the stipulated daily caloric dose by energy expenditure rate. All sessions were monitored to ensure the prescribed exercise dose (energy expenditure) was closely met. The intervention was intended to take place over 24 wk, but two additional weeks were allowed for participants who had not met their exercise doses (14,15).

Outcomes. Outcome measures were assessed at baseline and month 6. Weight change and compensation at month 6 were primary end points, whereas others were secondary or exploratory. During clinical assessment visits, fasting body weight was measured on a calibrated electronic scale (Siemens Medical Solutions). Compensation was calculated as described previously, with predicted weight loss estimated using the dynamic energy balance model (17,21).

Waist circumference was determined (26), and fitness tests were performed as documented to measure peak absolute and relative $\dot{V}O_2$ (14). Further, triglycerides, cholesterol (total cholesterol, LDL cholesterol [LDL-C], HDL cholesterol [HDL-C]), glucose, and insulin concentrations were measured, and blood pressure was determined (14,27).

The E-MECHANIC Study

The E-MECHANIC study (ClinicalTrials.gov: NCT01264406) was conducted at Pennington Biomedical Research Center, with approval of the center's Institutional Review Board. Males and females with overweight or obesity (BMI = 25.0–45.0 $kg \cdot m^{-2}$) were recruited, with all participants providing written informed consent before enrolment. Exclusion criteria have been reported (16).

The details of the intervention have been reported (16,17). Participants were randomized to either a no-exercise control

group or two exercise groups: 8 KKW and 20 KKW (17). Participants assigned to the 8 KKW group performed their prescribed dose from the start; conversely, the 20 KKW group ramped up their exercise prescription from 8 KKW during week 1 to 14 KKW during week 2 and 20 KKW during week 3. A Tanita scale (Tanita Corporation, Arlington Heights, IL) was used weekly before exercise sessions to measure body weight. All exercise was performed on a treadmill at a speed and gradient that kept participants within a heart range equivalent to 65% to 85% of baseline peak $\dot{V}O_2$. Energy expenditure was calculated in real time based on treadmill speed, treadmill gradient, and participant weight using standard ACSM equations (28). Moreover, energy expenditure was measured using a metabolic cart at weeks 2, 4, 6, 8, 12, 16, and 20 to monitor changes in metabolic and/or biomechanical efficiency. The duration of exercise was adjusted to meet participant's energy expenditure targets (16,17). Participants aimed to achieve their total intervention energy expenditure within 24 wk, although an additional 3 wk was allowed if needed.

Outcomes. Outcomes were collected at baseline, month 6, and (for questionnaire data only) week 4. Weight change and compensation at month 6 were primary end points; all other end points were secondary. Body weight measurements at baseline and month 6 were the average of three fasting weights on a calibrated Tanita scale collected over a 14-d period (days 0, 7, and 14 of the 14-d period). Compensation was calculated using methods identical to DREW (see Appendix, Supplemental Digital Content, Supplementary Information, http://links.lww.com/MSS/C274) (17,21).

Waist circumference was measured at baseline and month 6 in clinical assessment visits via a nonextensible tape measurer (Gulick II; Sammons Preston, Chicago, IL), and fat mass and lean mass were determined through dual-energy x-ray absorptiometry (DXA; iDXA, encore software version 13.60; GE Healthcare, Chicago, IL) before the exercise intervention and on completion of the trial. As described previously (16), fitness tests were performed at baseline and follow-up to measure peak and relative $\dot{V}O_2$. The aforementioned cardiometabolic disease risk markers that were assessed in DREW were measured in E-MECHANIC (16,17,29).

Energy intake was estimated over a 2-wk period at baseline and month 6 using doubly labeled water. This method, considered the gold standard of free-living energy requirements (30), assesses total daily energy expenditure while accounting for body composition changes during exercise (31,32). Energy intake can also be measured using unadjusted energy expenditure and energy expenditure adjusted for metabolic rate (17), but our results were not meaningfully affected when using these alternative methods. Daily steps were measured for 2 wk using SenseWear armbands (Body Media, Pittsburgh, PA), with steps from exercise sessions excluded from the month 6 measurement period. After a 12-h fast, resting metabolic rate (RMR) was measured for 30 min with Max-II metabolic carts at baseline and month 6 (AEI Technologies, Pittsburgh, PA).

Several validated questionnaires were administered at baseline, week 4, and month 6 to measure constructs of eating behaviors and attitudes and physical activity. These included appetite ratings on visual analog scales (33) that were administered on two occasions: the laboratory after the consumption of a 190-kcal nutrition bar and retrospectively during the previous week (34). The Activity Temperament Questionnaire (35), the Compensatory Health Belief Scale (CHBS) (36), the Eating Inventory (37), the Food Craving Inventory (38), the Food Preference Questionnaire (FPQ) (39), and the Multifactorial Assessment of Eating Disorder Symptoms (40) were also administered.

Statistical Analysis

The present manuscript is a *post hoc* analysis of the DREW and E-MECHANIC studies; accordingly, the present analysis used the sample size obtained from both studies. In both studies, participants were divided into tertiles based on percent weight change from baseline to week 4, given residuals of changes were nonnormally distributed and skewed. Participants in tertile 1 and tertile 3 had the least and most percent weight loss at week 4, respectively.

Similar analyses were performed on both DREW and E-MECHANIC data. Between-tertile differences in continuous and categorical measures were assessed via a one-way ANOVA and a chi-square test, respectively, for baseline and descriptive data. A one-way ANCOVA was used to compare change in primary and secondary end points from baseline, with age, exercise group, self-reported race, baseline values, and (for E-MECHANIC only) sex used as covariates. If between-tertile differences in cardiometabolic risk markers were evident, we also included percent weight change as a covariate to determine if variations were weight dependent. As an exploratory analysis, for participants with all weekly weight data up to week 24, a two-way mixed (tertile–time) ANCOVA was performed using the same covariates to examine weight change and compensation during the trial. ANCOVA analyses were used irrespective of normality (41). If data were aspherical, a Greenhouse–Geisser correction was applied for epsilon <0.75, whereas the Huynh–Feldt correction was used for less severe asphericity (epsilon >0.75). Where significance occurred, adjusted *post hoc* pairwise comparisons (Holm–Bonferroni) located differences. Analyses were performed using SPSS version 25 (IBM Corp., Armonk, NY), and α was set at 0.05. Unless noted otherwise, baseline and descriptive data are reported as mean ± SD, whereas outcome measures are presented as estimated marginal mean ± 95% confidence interval (CI).

RESULTS

The DREW Study

Participant characteristics. The DREW study enrolled 362 participants into the three exercise groups. In the present study, 312 participants were analyzed for the primary and secondary end points, with 50 participants removed for reasons related to attrition and abidance to the exercise regimen (Supplementary Fig. 1, see Appendix, Supplemental Digital Content,

TABLE 1. Demographic and baseline characteristics of participants in the DREW study across tertiles of initial percent weight change.

	All (n = 312)		Tertile 1 (n = 104), Least Initial Weight Loss		Tertile 2 (n = 104)		Tertile 3 (n = 104), Most Initial Weight Loss		
Variable	**Mean**	**SD/%**	**Mean**	**SD/%**	**Mean**	**SD/%**	**Mean**	**SD/%**	***P***
Ethnicity (n)									0.433
White	195	62.5	71	68.3	63	60.6	61	58.7	
African American	96	30.8	28	26.9	33	31.7	35	33.7	
Hispanic/other	21	6.7	5	4.8	8	7.7	8	7.7	
Exercise group (n)									0.077
4 KKW	134	42.9	43	41.3	46	44.2	45	43.3	
8 KKW	87	27.9	31	29.8	20	19.2	36	34.6	
12 KKW	91	29.2	30	28.8	38	36.5	23	22.1	
Age (yr)	57.3	6.5	58.0	6.5	57.3	6.8	56.6	6.0	0.272
Weight (kg)	83.7	11.8	81.9	11.3	84.6	11.6	84.6	12.5	0.179
BMI ($kg \cdot m^{-2}$)	31.5	3.8	30.8	3.4	31.9	3.9	31.9	4.1	**0.049**
Waist circumference (cm)	100.3	11.7	99.5	11.0	101.8	11.7	99.7	12.5	0.299
Fitness variables									
Peak absolute $\dot{V}O_2$ ($L \cdot min^{-1}$)	1.29	0.24	1.25	0.24	1.31	0.24	1.30	0.25	0.175
Peak relative $\dot{V}O_2$ ($mL \cdot kg^{-1} \cdot min^{-1}$)	15.5	2.9	15.4	2.6	15.6	2.9	15.5	3.0	0.798
Cardiometabolic disease risk markers									
Triglycerides ($mg \cdot dL^{-1}$)	128.0	62.8	126.4	69.2	131.9	60.1	125.8	59.2	0.740
Total cholesterol ($mg \cdot dL^{-1}$)	201.1	29.6	198.5	28.2	201.8	31.9	203.0	28.8	0.536
LDL-C ($mg \cdot dL^{-1}$)	117.7	26.4	113.0	24.8	119.6	27.2	120.4	26.7	0.090
HDL-C ($mg \cdot dL^{-1}$)	58.0	14.6	60.6	14.6	55.9	12.5	57.4	16.1	0.063
Glucose ($mg \cdot dL^{-1}$)	94.5	8.6	94.5	8.5	96.0	7.5*	92.9	9.6	**0.036**
Insulin ($pmol \cdot L^{-1}$)	73.6	41.6	76.4	47.0	73.1	35.9	71.2	41.7	0.693
Systolic blood pressure (mm Hg)	139.0	13.3	140.8	13.8	138.1	13.1	138.1	13.0	0.236
Diastolic blood pressure (mm Hg)	80.8	8.7	81.5	9.1	79.6	8.3	81.3	8.6	0.209

P is derived from ANOVA for continuous variables and contingency chi-square test for categorical variables.
Values are presented as mean and SD for continuous variables and number (n) and percentage for categorical variables.
Bold indicates significant ($P < 0.05$).
*Significant difference in tertile 2 vs tertile 3 ($P = 0.03$).

http://links.lww.com/MSS/C274). The baseline and descriptive characteristics of the tertiles are shown in Table 1. Most variables were similar between tertiles ($P \geq 0.063$), although there was a significant effect of tertile on BMI and glucose ($P \leq 0.049$).

Week 4 percent weight change differed between all tertiles per the design of this analysis ($P < 0.001$), with tertile 1 (least initial weight loss/initial weight gain), tertile 2, and tertile 3 (most initial weight loss) exhibiting 1.6% ± 1.0%, −0.2% ± 0.4%, and −2.9% ± 2.0% mean ± SD weight change, respectively, at week 4. Cumulative energy expended at exercise sessions was not different among tertiles at week 4 (tertile 1, 1547 ± 370 kcal; tertile 2, 1619 ± 325 kcal; tertile 3, 1569 ± 367 kcal; $P = 0.321$) and at the end of the trial (tertile 1, 15,127 ± 7686 kcal; tertile 2, 14,761 ± 6455 kcal; tertile 3, 13,809 ± 6077 kcal; $P = 0.351$).

Change in outcome data at month 6. Weight and BMI change at follow-up were different between tertiles ($P < 0.001$; Table 2). *Post hoc* tests revealed tertile 1 (least initial weight loss/initial weight gain) exhibited less weight loss compared with tertile 2 (−1.0 kg, 95% CI = −1.8 to −0.1; $P = 0.022$) and tertile 3 (most initial weight loss; −3.0 kg, 95% CI = −3.8

TABLE 2. Predicted weight change, compensation, and change in outcome variables after 6 months of exercise in participants from the DREW study.

	Tertile 1 (n = 104), Least Initial Weight Loss		Tertile 2 (n = 104)		Tertile 3 (n = 104), Most Initial Weight Loss		
Variable	**EM Mean**	**95% CI**	**EM Mean**	**95% CI**	**EM Mean**	**95% CI**	***P***
Weight (kg)	−0.2	−0.8 to 0.4	−1.2[c]	−1.8 to −0.6	−3.2[a,b]	−3.8 to −2.6	**<0.001**
Weight (%)	−0.2	−0.9 to 0.5	−1.4[c]	−2.1 to −0.7	−3.8[a,b]	−4.5 to −3.1	**<0.001**
BMI ($kg \cdot m^{-2}$)	−0.1	−0.4 to 0.1	−0.5	−0.7 to −0.2	−1.3[a,b]	−1.5 to −1.0	**<0.001**
Waist circumference (cm)	−3.1	−4.4 to −1.8	−1.9	−3.2 to −0.6	−3.2	−4.6 to −1.9	0.270
Predicted weight change	−2.0	−2.1 to −1.9	−2.0	−2.2 to −1.9	−2.0	−2.1 to −1.9	0.856
Weight compensation (kg)	1.8	1.2 to 2.4	0.8[c]	0.2 to 1.4	−1.2[a,b]	−1.8 to −0.6	**<0.001**
Fitness variables							
Peak absolute $\dot{V}O_2$ ($L \cdot min^{-1}$)	0.05	0.02 to 0.09	0.07	0.04 to 0.10	0.05	0.02 to 0.09	0.756
Peak relative $\dot{V}O_2$ ($mL \cdot kg^{-1} \cdot min^{-1}$)	0.7	0.3 to 1.2	1.0	0.6 to 1.4	1.3	0.9 to 1.7	0.194
Cardiometabolic disease risk markers							
Triglycerides ($mg \cdot dL^{-1}$)	3.6	−4.7 to 12.0	0.5	−7.8 to 8.8	−11.2[a]	−19.5 to −2.9	**0.035**
Total cholesterol ($mg \cdot dL^{-1}$)	−1.4	−5.7 to 2.8	3.8	−0.4 to 8.0	0.5	−3.7 to 4.8	0.229
LDL-C ($mg \cdot dL^{-1}$)	−0.4	−4.7 to 3.9	2.9	−1.4 to 7.1	1.6	−2.7 to 5.8	0.568
HDL-C ($mg \cdot dL^{-1}$)	−3.5	−4.9 to −2.1	−0.1[c]	−1.5 to 1.3	1.1[a]	−0.3 to 2.5	**<0.001**
Glucose ($mg \cdot dL^{-1}$)	−1.8	−3.1 to −0.5	−1.0	−2.3 to 0.3	−1.2	−2.5 to 0.1	0.690
Insulin ($pmol \cdot L^{-1}$)	−2.0	−7.9 to 3.9	−5.6	−11.3 to 0.2	−3.3	−9.1 to 2.4	0.690
Systolic blood pressure (mm Hg)	−0.6	−3.0 to 1.7	−1.5	−3.9 to 0.8	−0.5	−2.8 to 1.9	0.795
Diastolic blood pressure (mm Hg)	0.7	−0.7 to 2.0	−0.5	−1.9 to 0.9	0.7	−0.7 to 2.1	0.395

P is derived from ANCOVA. When significant, *post hoc* comparisons among tertiles were adjusted with Holm–Bonferroni corrections.
Values are estimated marginal means and 95% CI adjusted for age, ethnicity, group, and baseline.
Bold indicates significant ($P < 0.05$).
[a]Significant difference between tertile 1 and tertile 3 ($P < 0.05$).
[b]Significant difference between tertile 2 and tertile 3 ($P < 0.05$).
[c]Significant difference between tertile 1 and tertile 2 ($P < 0.05$).
EM, estimated marginal.

to −2.2; $P < 0.001$), plus tertile 2 showed less weight loss compared with tertile 3 (−2.0 kg, 95% CI = −2.8 to −1.2; $P < 0.001$; Table 2). Tertile 1 and tertile 2 also showed a smaller reduction in BMI compared with tertile 3 ($P < 0.001$), but no differences between tertiles for change in waist circumference were seen ($P = 0.270$).

Predicted weight change did not differ between tertiles ($P = 0.856$), yet one-way ANCOVA showed that initial weight change tertile was linked to compensation ($P < 0.001$). *Post hoc* tests revealed that tertile 1 showed 3.0 kg (95% CI = 2.2 to 3.9) more compensation than tertile 3 ($P < 0.001$). Further, tertile 1 exhibited greater compensation compared with tertile 2 ($P = 0.025$), and tertile 3 displayed lower compensation than tertile 2 ($P < 0.001$; Table 2).

There was a significant effect of tertile on change in HDL-C ($P < 0.001$; Table 2), with tertile 1 having a decrease in HDL-C relative to tertile 2 ($P = 0.002$) and tertile 3 ($P < 0.001$). Moreover, change in triglycerides was different between tertiles ($P = 0.035$); specifically, tertile 1 showed an increase in triglycerides versus tertile 3 ($P = 0.043$; Table 2). The between-tertile variations remained significant for change in HDL-C ($P < 0.001$) but not change in triglycerides ($P = 0.329$) when controlled for percent weight change at month 6. Change in other cardiometabolic risk markers and fitness variables were similar between tertiles (all $P \geq 0.194$; Table 2).

Weekly weight change and compensation. In participants with all weekly weights ($n = 110$), there was a main effect of tertile on weight change ($P < 0.001$), although no time or tertile–time effects were seen ($P \geq 0.175$; Fig. 1). *Post hoc* comparisons for main tertile effect showed that tertile 1 displayed less weight loss versus tertile 2 and tertile 3 ($P \leq 0.027$), whereas weight loss was also lower in tertile 2 versus tertile 3 ($P < 0.001$). Although predicted weight was expected to decrease over time (main effect of time; $P < 0.001$), there were no effects of tertile and no tertile–time interaction for predicted weight change ($P \geq 0.083$; Supplementary Fig. 2, see Appendix, Supplemental Digital Content, http://links.lww.com/MSS/C274). There was a main effect of tertile on compensation ($P < 0.001$), with tertile 1 exhibiting greater compensation than tertile 3 ($P < 0.001$) and tertile 2 showing higher compensation than tertile 3 ($P < 0.001$; Fig. 1). No main effect of time and no tertile–time interaction were revealed, however ($P \geq 0.174$). All weekly weight change and compensation data are shown in Supplementary Figure 3 (see Appendix, Supplemental Digital Content, http://links.lww.com/MSS/C274).

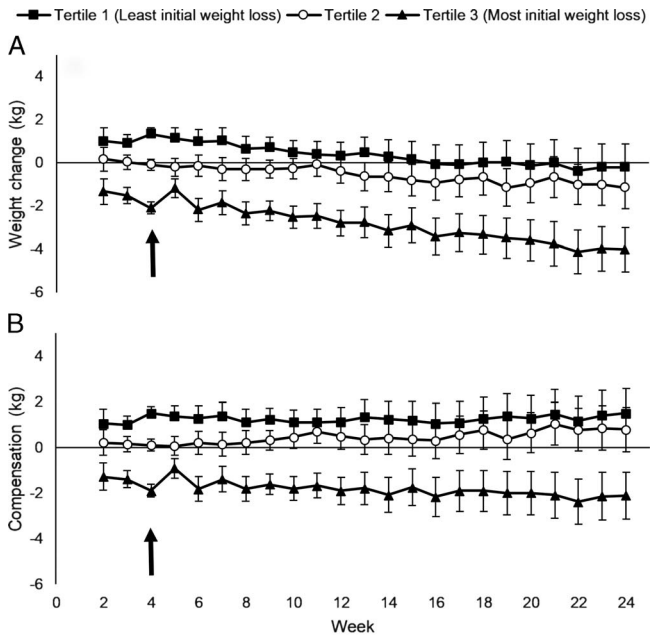


FIGURE 1—A, Weight change data for participants in the DREW study with all weekly weight data up to week 24 ($n = 110$ [tertile 1, $n = 34$; tertile 2, $n = 41$; tertile 3, $n = 35$]). B, Compensation data for participants in the DREW study with all weekly weight data up to week 24 ($n = 110$ [tertile 1, $n = 34$; tertile 2, $n = 41$; tertile 3, $n = 35$]). Data are from weekly weight measurements performed before exercise sessions. Participants in tertile 1 and tertile 3 had the least and most percent weight loss at week 4, respectively. *Black arrows* represent point where tertiles were calculated. Values are estimated marginal means (95% CI) adjusted for age, ethnicity, group, and baseline.

The E-MECHANIC Study

Participant characteristics. In total, 133 participants were enrolled into exercise groups, although 102 were used in the analysis because of attrition, inadequate adherence/compliance, or no weight data at week 4 (Supplementary Fig. 4, see Appendix, Supplemental Digital Content, http://links.lww.com/MSS/C274). There were no differences in most baseline characteristics among tertiles ($P \geq 0.053$), although significant differences were seen for glucose (Table 3), CHBS total score, restraint, and Food Craving Inventory sweet score (all $P \leq 0.034$; Supplementary Table 1, see Appendix, Supplemental Digital Content, http://links.lww.com/MSS/C274).

In accord with analysis design, mean ± SD week 4 percent weight change was different between tertile 1 (least initial weight loss/most initial weight gain; 2.1% ± 0.9%), tertile 2 (0.8% ± 0.3%), and tertile 3 (most initial weight loss; −0.5% ± 0.7%) at week 4 ($P < 0.001$). Cumulative energy expenditure at exercise sessions was not different among tertiles at week 4 (tertile 1, 3915 ± 1425 kcal; tertile 2, 3779 ± 1323 kcal; tertile 3, 3990 ± 1565 kcal; $P = 0.829$) and at the end of the trial (tertile 1, 28,231 ± 11,570 kcal; tertile 2, 26,328 ± 12,142 kcal; tertile 3, 27,810 ± 13,357 kcal; $P = 0.802$).

Change in outcome data at month 6. Initial weight change tertile was related to weight change at month 6 ($P \leq 0.042$), with tertile 1 (least initial weight loss/most initial weight gain) presenting less weight loss compared with tertile 3 (most initial weight loss; −1.5 kg, 95% CI = −2.7 to −0.3; $P \leq 0.042$; Table 4). There was only a tendency for a difference in change in BMI between tertiles ($P = 0.052$), and change in waist circumference did not differ between tertiles ($P = 0.227$). Further, changes in fat mass and lean mass were not significantly different between tertiles ($P \geq 0.189$).

There were no between-tertile differences in predicted weight change ($P = 0.641$). One-way ANCOVA showed a significant effect on compensation ($P = 0.043$). *Post hoc* comparisons demonstrated that tertile 1 had greater compensation relative to tertile 3 (1.5 kg, 95% CI = 0.3 to 2.6; $P = 0.048$).

TABLE 3. Demographic and baseline characteristics of participants in the E-MECHANIC study across tertiles of initial percent weight change.

	All (n = 102)		Tertile 1 (n = 34), Least Initial Weight Loss		Tertile 2 (n = 34)		Tertile 3 (n = 34), Most Initial Weight Loss		
Variable	**Mean**	**SD/%**	**Mean**	**SD/%**	**Mean**	**SD/%**	**Mean**	**SD/%**	***P***
Ethnicity (n)									0.608
White	71	69.6	24	70.6	23	67.6	24	70.6	
African American	29	28.4	10	29.4	11	32.4	8	23.5	
Hispanic/other	2	2.0	0	0.0	0	0.0	2	5.9	
Exercise group (n)									0.624
8 KKW	54	52.9	16	47.1	20	58.9	18	52.9	
20 KKW	48	47.1	18	52.9	14	41.2	16	47.1	
Sex (n)									0.371
Male	30	29.4	8	23.5	13	38.2	9	26.5	
Female	72	70.6	26	76.5	21	61.8	25	73.5	
Age (yr)	48.8	11.9	50.7	11.8	47.3	12.3	48.6	11.6	0.506
Weight (kg)	87.1	15.4	85.7	15.8	87.1	15.1	88.5	15.7	0.753
BMI ($kg \cdot m^{-2}$)	31.1	4.5	30.8	4.9	30.9	4.4	31.6	4.3	0.732
Waist circumference (cm)	97.3	11.9	95.1	10.9	97.4	12.3	99.3	12.3	0.347
Fat mass (kg)	36.2	9.5	36.2	10.5	35.1	9.3	37.3	8.6	0.652
Lean mass (kg)	47.9	10.1	46.7	10.0	49.0	10.5	48.1	9.9	0.654
Fitness variables									
Peak absolute $\dot{V}O_2$ ($L \cdot min^{-1}$)	2.05	0.54	1.93	0.49	2.14	0.66	2.07	0.43	0.273
Peak relative $\dot{V}O_2$ ($mL \cdot kg^{-1} \cdot min^{-1}$)	23.9	5.3	22.8	4.0	24.9	6.7	23.9	4.9	0.253
Cardiometabolic disease risk markers									
Triglycerides ($mg \cdot dL^{-1}$)	109.5	51.8	108.3	62.4	114.6	47.7	110.1	44.4	0.695
Total cholesterol ($mg \cdot dL^{-1}$)	204.9	35.8	197.7	40.3	213.6	33.3	203.3	32.3	0.175
LDL-C ($mg \cdot dL^{-1}$)	123.4	27.3	117.9	30.6	128.2	26.7	124.2	23.9	0.295
HDL-C ($mg \cdot dL^{-1}$)	59.5	16.9	59.0	20.3	62.5	16.3	57.1	13.6	0.414
Glucose ($mg \cdot dL^{-1}$)	92.4	7.6	89.6	7.1	94.1	6.8*	93.3	8.2	**0.034**
Systolic blood pressure (mm Hg)	119.8	9.9	118.2	12.1	120.1	9.0	120.9	8.3	0.512
Diastolic blood pressure (mm Hg)	76.7	7.2	75.6	8.8	76.9	6.1	77.6	6.4	0.512
Energy intake ($kcal \cdot d^{-1}$)	2498	471	2389	479	2513	494	2591	430	0.207
Steps per day	6180	2254	5968	1939	6093	2179	6468	2612	0.645
RMR ($kcal \cdot d^{-1}$)	1519	284	1474	304	1502	276	1582	268	0.271

P is derived from ANOVA for continuous variables and contingency chi-square test for categorical variables.
Values are mean and SD for continuous variables and number (n) and percentage for categorical variables.
Bold indicates significant ($P \leq 0.05$).
*Significant difference in tertile 2 vs tertile 1 ($P \leq 0.045$).

There was an association between initial weight change and change in triglycerides at month 6 ($P = 0.013$). Individuals categorized in tertile 3 had a reduction in triglyceride levels compared with tertile 1 ($P = 0.011$; Table 4). The effect of initial weight change was reduced when percent weight change at month 6 was controlled, with only a tendency for a between-tertile difference in change in triglycerides seen ($P = 0.050$). Change in other cardiometabolic markers and fitness variables were similar among tertiles (all $P \geq 0.115$; Table 4).

Change in energy intake differed between initial weight change tertiles ($P = 0.031$), with change in energy intake higher in tertile 1 compared with tertile 3 ($P = 0.003$). No between-tertile differences in change in steps per day and RMR were seen (all $P \geq 0.567$; Table 4). Tertile 1 had a reduction in the Multifactorial Assessment of Eating Disorder Symptoms avoidance of forbidden foods relative to tertile 3 at month 6 ($P = 0.020$; Table 4), yet no other between-tertile changes were seen (all $P \geq 0.085$; Supplementary Table 2, see Appendix, Supplemental Digital Content, http://links.lww.com/MSS/C274).

Change in outcome data at week 4. There was an effect of tertile on change in CHBS score ($P = 0.039$; Table 4), although there was only a tendency for a difference after pairwise adjustments ($P = 0.050$). Tertile 3 showed a decrease in several FPQ-assessed food preferences compared with tertile 1, including high fat, high fat/high complex carbohydrate, high sugar, and low fat/high sugar foods (all $P \leq 0.041$; Table 4). No other between-tertile differences in questionnaire change scores were seen (all $P \geq 0.081$; Supplementary Table 3, see Appendix, Supplemental Digital Content, http://links.lww.com/MSS/C274).

Weekly weight change and compensation. For participants with all weekly weight data up to week 24 ($n = 63$), there was a main effect of tertile on weight change ($P = 0.001$) but no main effect of time and no tertile–time interaction ($P \geq 0.586$; Fig. 2). The main tertile effect indicated that tertile 1 showed less weight loss than tertile 3 ($P = 0.001$); further, tertile 2 had less weight loss compared with tertile 3 ($P = 0.019$). Two-way ANCOVA for predicted weight change showed that weight was predicted to decrease (main effect of time; $P < 0.001$), yet no main effect of tertile ($P = 0.961$) and no tertile–time interaction ($P = 0.840$) was observed (Supplementary Fig. 5, see Appendix, Supplemental Digital Content, http://links.lww.com/MSS/C274). There was a main effect of tertile on compensation ($P = 0.001$), with lower compensation in tertile 3 relative to tertiles 1 ($P = 0.001$) and 2 ($P = 0.020$; Fig. 2); however, no time or tertile–time effects were shown ($P \geq 0.580$). All weekly weight change and compensation data for participants are depicted in Supplementary Figure 6 (see Appendix, Supplemental Digital Content, http://links.lww.com/MSS/C274).

DISCUSSION

The present analysis demonstrated that during aerobic exercise training, individuals with overweight or obesity who showed less initial weight loss (or most initial weight gain)

TABLE 4. Predicted weight change, compensation, and change in outcome variables after 6 months and 4 wk (questionnaire data only) of exercise in participants from the E-MECHANIC study.

	Tertile 1 (n = 34), Least Initial Weight Loss		Tertile 2 (n = 34)		Tertile 3 (n = 34), Most Initial Weight Loss		
Variable	**EM Mean**	**95% CI**	**EM Mean**	**95% CI**	**EM Mean**	**95% CI**	***P***
Weight (kg)	−0.3	−1.2 to 0.5	−0.8	−1.7 to 0.0	−1.8[a]	−2.7 to −1.0	**0.042**
Weight (%)	−0.5	−1.4 to 0.5	−1.0	−1.9 to −0.0	−2.2[a]	−3.2 to −1.3	**0.033**
BMI (kg·m^{-2})	−0.2	−0.5 to 0.2	−0.3	−0.6 to −0.0	−0.7	−1.0 to −0.4	0.052
Waist circumference (cm)	−0.5	−1.6 to 0.6	−0.8	−1.9 to 0.3	−1.8	−2.9 to −0.7	0.227
Fat mass (kg)	−0.6	−1.3 to 0.1	−0.7	−1.4 to 0.4	−1.5	−2.2 to −0.7	0.189
Lean mass (kg)	−0.1	−0.5 to 0.3	−0.2	−0.6 to 0.1	−0.4	−0.8 to −0.1	0.512
Predicted weight change (kg)	−3.0	−3.1 to −2.9	−3.1	−3.2 to −3.0	−3.0	−3.2 to −2.9	0.641
Weight compensation (kg)	2.7	1.8 to 3.5	2.3	1.4 to 3.1	1.2[a]	0.4 to 2.0	**0.043**
Fitness variables							
Peak absolute $\dot{V}O_2$ (L·min^{-1})	0.18	0.11 to 0.26	0.17	0.10 to 0.24	0.17	0.09 to 0.24	0.933
Peak relative $\dot{V}O_2$ (mL·kg^{-1}·min^{-1})	1.5	0.2 to 2.8	2.0	0.7 to 3.3	2.0	0.7 to 3.3	0.831
Cardiometabolic risk markers							
Triglycerides (mg·dL^{-1})	5.3	−5.8 to 16.3	−9.1	−20.1 to 1.9	−18.2[a]	−29.1 to −7.3	**0.013**
Total cholesterol (mg·dL^{-1})	−4.3	−10.5 to 2.0	−5.0	−11.3 to 1.2	0.3	−5.9 to 6.4	0.428
LDL-C (mg·dL^{-1})	−4.4	−9.8 to 1.0	−2.3	−7.6 to 3.1	2.2	−3.1 to 7.5	0.209
HDL-C (mg·dL^{-1})	−1.0	−3.2 to 1.1	−0.9	−3.1 to 1.3	1.8	−0.3 to 3.9	0.115
Glucose (mg·dL^{-1})	0.0	−2.0 to 2.0	−1.7	−3.7 to 0.3	−0.2	−2.1 to 1.8	0.424
Systolic blood pressure (mm Hg)	−2.5	−4.8 to −0.1	−4.7	−7.0 to −2.3	−4.7	−7.0 to −2.3	0.329
Diastolic blood pressure (mm Hg)	−1.4	−3.2 to 0.3	−1.9	−3.6 to −0.2	−1.1	−2.8 to 0.7	0.789
Energy intake (kcal·d^{-1})	321	155 to 487	76	−90 to 242	−92[a]	−259 to 75	**0.003**
Steps per day	−487	−1218 to 245	−329	−1024 to 366	−509	−1199 to 180	0.926
RMR (kcal·d^{-1})	17	−53 to 88	45	−28 to 117	72	−1 to 145	0.567
Questionnaires, week 4							
CHBS	−2.2	−4.0 to −0.4	1.0	−0.8 to 2.8	0.3	−1.4 to 2.1	**0.039**
FPQ, high fat	0.0	−0.5 to 0.5	−0.6	−1.1 to −0.1	−0.9[a]	−1.3 to −0.4	**0.044**
FPQ, high fat and high complex carbohydrates	0.2	−0.3 to 0.6	−0.4	−0.9 to 0.1	−0.9[a]	−1.3 to −0.4	**0.018**
FPQ, high sugars	0.0	−0.4 to 0.5	−0.5	−0.9 to −0.0	−0.8[a]	−1.2 to −0.4	**0.040**
FPQ, high protein	−0.0	−0.5 to 0.5	−0.7	−1.2 to −0.3	−0.9	−1.4 to −0.4	**0.039**
FPQ, low fat	0.0	−0.4 to 0.5	−0.5	−1.0 to −0.1	−0.7	−1.1 to −0.3	**0.049**
FPQ, low fat and high protein	−0.0	−0.6 to 0.5	−0.8	−1.3 to −0.3	−0.9	−1.4 to −0.4	**0.043**
FPQ, low fat and high sugars	0.2	−0.2 to 0.6	−0.3	−0.7 to 0.1	−0.7[a]	−1.1 to −0.3	**0.016**
Questionnaires, month 6							
MAEDS, avoidance of forbidden foods	−1.5	−3.4 to 0.4	0.1	−1.8 to 1.9	2.3[a]	0.4 to 4.2	**0.024**

P is derived from ANCOVA. When significant, *post hoc* comparisons among tertiles were adjusted with Holm–Bonferroni corrections.
Values are estimated marginal means (95% CI) adjusted for age, sex, ethnicity, group, and baseline.
Bold indicates significant ($P < 0.05$).
[a]Significant difference between tertile 1 and tertile 3 ($P < 0.05$).
EM, estimated marginal; MAEDS, Multifactorial Assessment of Eating Disorder Symptoms.

at week 4 displayed less weight loss, greater compensation, and poorer changes in blood lipids at the end of 6 months compared with those with most initial weight loss. In addition, we showed that individuals with the least initial weight loss showed a greater rise in energy intake and maladaptive alterations in eating attitudes and behaviors relative to those with most initial weight loss, supplying potential mechanisms that drove differences in weight-related end points.

Exercise training is recommended for individuals with overweight or obesity due to the numerous health benefits it stimulates (3), but there is substantial unexplained heterogeneity in weight change, which can alter many improvements seen (5,42). Our findings showed that individuals with less initial weight loss exhibited poorer long-term changes in weight in two exercise training studies. All individuals in tertile 1 showed weight gain at week 4 in both exercise studies. Generally speaking, this is at odds with dietary interventions that have characterized individuals with less initial weight loss as those who attain 0.5%–3.0% weight loss at 1–2 months (8). Although this discrepancy could occur because—at odds with many dietary studies—weight loss was not a primary objective of exercise training in DREW and E-MECHANIC, the lower initial weight loss shown in our study compared with these studies is unsurprising, as it is easier to induce greater energy deficits through diet than exercise (17). It is less clear whether initial weight change affects long-term changes in body composition because we observed nonsignificant between-tertile variations in fat mass and lean mass in E-MECHANIC participants. Still, our findings broadly support work demonstrating that weight change from baseline to week 4 is associated with weight loss after 12 wk of exercise (43) and suggest, akin to dietary interventions (44), that initial weight change at week 4 can be monitored by interventionists to forecast weight loss during long-term exercise interventions. In light of previous trials (10,19), it is tempting to speculate that the associations we observed occur for even longer periods after the onset of exercise training, but additional studies beyond 6 months are needed.

In dietary studies, individuals with less initial weight loss often display inadequate adherence, and this leads to lower energy deficits and in turn poorer changes in weight-associated end points (8,19). Consequently, initial weight change can be used as an indicator of adherence to dietary restriction and can pinpoint individuals who require early support to improve adherence, increase energy deficits, and attain weight loss targets (8). We found that the amount of exercise performed is not implicated in the associations between initial weight change and weight change at month 6, as we saw no between-tertile differences in cumulative exercise energy expenditure. Participants

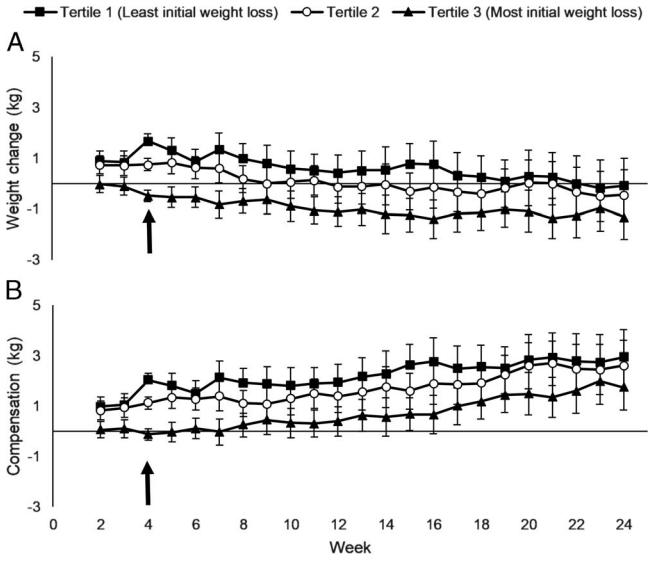


FIGURE 2—A, Weight change data for participants in the E-MECHANIC study with all weekly weight data up to week 24 ($n = 63$ [tertile 1, $n = 17$; tertile 2, $n = 21$; tertile 3, $n = 25$]). B, Compensation data for participants in the E-MECHANIC study with all weekly weight data up to week 24 ($n = 63$ [tertile 1, $n = 17$; tertile 2, $n = 21$; tertile 3, $n = 25$]). Data are from weekly weight measurements performed before exercise sessions. Participants in tertile 1 and tertile 3 had the least and most percent weight loss at week 4, respectively. *Black arrows* represent point where tertiles were calculated. Values are estimated marginal means (95% CI) adjusted for age, ethnicity, group, baseline, and sex.

who completed higher exercise doses in DREW (8 KKW and 12 KKW) and E-MECHANIC (20 KKW) ramped up their exercise prescription during the initial phases of exercise training (14,16), and this would have lessened cumulative energy expenditure differences between exercise groups (doses) at week 4. It is interesting, however, that between-tertile differences in weight change at month 6 also occurred independent of cumulative exercise energy expenditure. Our findings are consistent with the primary weight loss results from our two studies (5,17) and imply that at exercise doses typically recommended for health, interindividual variations in weight occur, with exercise energy expenditure exerting limited influence on initial and long-term weight change variability (21).

Our results indicate that the associations between initial weight change and long-term weight change are due to differences in compensation to exercise-induced energy deficits. Specifically, during exercise training, individuals who present less initial weight loss show greater weight compensation. Consistent with previous studies (17,42), the variations in compensation in response to exercise were likely driven by differences in energy intake. Indeed, in E-MECHANIC, we saw a greater rise in free-living energy intake at month 6 in individuals who lost less weight (or gained most weight) initially compared with those who lost most weight, yet we saw no between-tertile differences in change in physical activity or RMR. We additionally saw that those with less initial weight loss presented a reduction in avoidance of forbidden foods at month 6 and a smaller decline in food preferences, including those foods high in fat and sugar, at week 4 compared with individuals with the most initial weight loss. Shifts in food preferences (45) and an elevation in avoidance of forbidden foods (46) affect food intake. It is thus possible, given the differences in week 4 food preferences, that changes in eating patterns started during the initial phases of exercise. This suggests that interventionists could implement early strategies that attenuate food preferences and increase avoidance of unhealthy foods in individuals who display substandard weight change promptly during exercise. Such strategies have been frequently used in dietary regimens (12) and could comprise early nutritional classes that assist individuals controlling portions sizes and food cravings, with a particular focus on unhealthy foods as defined by the FPQ (cakes, doughnuts, and potato chips) (39). Similarly, as demonstrated by the findings of Unick and colleagues (47), early interventional support focused on goal setting and meal planning may be effective (8). Nevertheless, studies with such interventions during exercise training are needed, particularly as changes in several other eating-related constructs (e.g., restraint and food cravings) were similar between tertiles.

Our findings show that individuals who present less initial weight loss experience poorer changes in blood lipids after exercise training. Intriguingly, in DREW, the relationship between initial weight change and change in HDL-C at month 6 remained after weight change was statistically controlled. Consonant with some postulations (48), it is possible that those with less initial weight loss exhibited an increase in energy intake early in response to exercise, which resulted in decreases in HDL-C that were not mitigated by longer-term weight changes. However, between-tertile differences in triglyceride were attenuated when weight change was controlled, indicating the greater weight loss experienced by those in tertile 3 drove these findings. We also saw similar responses in other cardiometabolic disease risk markers, waist circumference, and fitness between tertiles. This implies that exercise induces metabolic and health improvements irrespective of weight loss, supporting previous work (15,49).

There are two notable strengths of our study: it comprises a large sample of participants with overweight or obesity from two exercise training studies, and exercise sessions were supervised, with stringent monitoring of exercise doses. Our study has limitations, however. We indeed did not perform measurements of DXA at week 4, and although our primary end point measurements were performed in controlled conditions, our measurements of weight before exercise sessions were performed in testing facilities where less standardization procedures were enforced. Furthermore, most of our sample (93%) consisted of females, suggesting that additional studies in males are required. We were also unable to obtain sophisticated measures of energy balance and body composition in the larger-powered DREW study. Nevertheless, despite the smaller sample size, we used advanced measures of energy intake, physical activity, RMR, and weight-related constructs in E-MECHANIC at baseline and month 6.

In conclusion, less initial weight loss from baseline to week 4 was associated with diminished weight loss at 6 months in two supervised aerobic exercise interventions comprising

individuals with overweight or obesity. Individuals who initially lost less weight or gained weight also showed greater compensation at month 6, and this was likely linked to an increase in energy intake and changes in eating behaviors and preferences conducive to poorer dietary patterns. Although exercise training should be universally advocated, individuals with overweight or obesity who show poor weight change initially during exercise training may benefit from early support to improve eating patterns, decrease compensation, and assist weight loss.

The authors acknowledge the late Conrad P. Earnest, PhD. Dr. Earnest was integral to many key aspects of the DREW and E-MECHANIC studies.

This work was supported by the National Institutes of Health (grant nos. HL66262 and HL102166), the NORC (grant no. P30 DK072476, titled "Nutritional Programming: Environmental and Molecular Interactions" sponsored by the NIDDK), and the National Institute of General Medical Sciences (grant no. U54 GM104940). J. L. D. is funded by the American Heart Association (grant: 20POST35210907); C. H. is funded by an NIH NIDDK Award (grant no. T32 DK064584). C. K. M. gives lectures and presentations on the topic to groups, including the Academy of Nutrition and Dietetics. C. K. M. is also a member of advisory boards for EHE and Naturally Slim. T. S. C. is employed at Naturally Slim. All other authors report no conflicts of interest related to this study. The sponsors had no role in the design, data collection, data analysis, interpretation of results, or preparation of the results for publication. The results are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. They also do not constitute endorsement by the ACSM.

REFERENCES

1. GBD 2015 Obesity Collaborators, Afshin A, Forouzanfar MH, Reitsma MB, et al. Health effects of overweight and obesity in 195 countries over 25 years. *N Engl J Med*. 2017;377(1):13–27.
2. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011–2012. *JAMA*. 2014; 311(8):806–14.
3. Booth FW, Roberts CK, Laye MJ. Lack of exercise is a major cause of chronic diseases. *Compr Physiol*. 2012;2:1143–211.
4. Donnelly JE, Blair SN, Jakicic JM, Manore MM, Rankin JW, Smith BK, American College of Sports Medicine. American College of Sports Medicine Position Stand: appropriate physical activity intervention strategies for weight loss and prevention of weight regain for adults. *Med Sci Sports Exerc*. 2009;41: 459–71.
5. Church TS, Martin CK, Thompson AM, Earnest CP, Mikus CR, Blair SN. Changes in weight, waist circumference and compensatory responses with different doses of exercise among sedentary, overweight postmenopausal women. *PLoS One*. 2009;4(2):e4515.
6. Church TS, Earnest CP, Thompson AM, et al. Exercise without weight loss does not reduce C-reactive protein: the INFLAME study. *Med Sci Sports Exerc*. 2010;42(4):708–16.
7. Swift DL, Johannsen NM, Lavie CJ, Earnest CP, Blair SN, Church TS. Effects of clinically significant weight loss with exercise training on insulin resistance and cardiometabolic adaptations. *Obesity*. 2016; 24(4):812–9.
8. Unick JL, Pellegrini CA, Demos KE, Dorfman L. Initial weight loss response as an indicator for providing early rescue efforts to improve long-term treatment outcomes. *Curr Diab Rep*. 2017;17(9):69.
9. Unick JL, O'Leary KC, Dorfman L, Thomas JG, Strohacker K, Wing RR. Consistency in compensatory eating responses following acute exercise in inactive, overweight and obese women. *Br J Nutr*. 2015;113:1170–7.
10. Unick JL, Neiberg RH, Hogan PE, et al. Weight change in the first 2 months of a lifestyle intervention predicts weight changes 8 years later. *Obesity*. 2015;23(7):1353–6.
11. Blundell JE, Stubbs RJ, Hughes DA, Whybrow S, King NA. Cross talk between physical activity and appetite control: does physical activity stimulate appetite? *Proc Nutr Soc*. 2003;62(3): 651–61.
12. Rickman AD, Williamson DA, Martin CK, et al. The CALERIE study: design and methods of an innovative 25% caloric restriction intervention. *Contemp Clin Trials*. 2011;32:874–81.
13. Melanson EL, Keadle SK, Donnelly JE, Braun B, King NA. Resistance to exercise-induced weight loss: compensatory behavioral adaptations. *Med Sci Sports Exerc*. 2013;45(8):1600–9.
14. Morss GM, Jordan AN, Skinner JS, et al. Dose–Response to Exercise in Women age 45–75 yr (DREW): design and rationale. *Med Sci Sports Exerc*. 2004;36(2):336–44.
15. Church TS, Earnest CP, Skinner JS, Blair SN. Effects of different doses of physical activity on cardiorespiratory fitness among sedentary, overweight or obese postmenopausal women with elevated blood pressure. *JAMA*. 2007;297(19):2081–91.
16. Myers CA, Johnson WD, Earnest CP, et al. Examination of mechanisms (E-MECHANIC) of exercise-induced weight compensation: study protocol for a randomized controlled trial. *Trials*. 2014;15:212.
17. Martin CK, Johnson WD, Myers CA, et al. Effect of different doses of supervised exercise on food intake, metabolism, and non-exercise physical activity: the E-MECHANIC randomized controlled trial. *Am J Clin Nutr*. 2019;110(3):583–92.
18. Unick JL, Hogan PE, Neiberg RH, et al. Evaluation of early weight loss thresholds for identifying nonresponders to an intensive lifestyle intervention. *Obesity*. 2014;22:1608–16.
19. Unick JL, Leahey T, Kent K, Wing RR. Examination of whether early weight loss predicts 1-year weight loss among those enrolled in an Internet-based weight loss program. *Int J Obes (Lond)*. 2015;39(10): 1558–60.
20. Unick JL, Ross KM, Wing RR. Factors associated with early non-response within an Internet-based behavioural weight loss program. *Obes Sci Pract*. 2019;5(4):324–32.
21. Thomas DM, Bouchard C, Church T, et al. Why do individuals not lose more weight from an exercise intervention at a defined dose? An energy balance analysis. *Obes Rev*. 2012;13(10): 835–47.
22. Thomas DM, Schoeller DA, Redman LA, Martin CK, Levine JA, Heymsfield SB. A computational model to determine energy intake during weight loss. *Am J Clin Nutr*. 2010;92(6):1326–31.
23. Swift DL, Johannsen NM, Tudor-Locke C, et al. Exercise training and habitual physical activity: a randomized controlled trial. *Am J Prev Med*. 2012;43(6):629–35.
24. Fearnbach SN, Johannsen NM, Myers CA, et al. Adaptations to exercise in compensators and noncompensators in the E-MECHANIC Trial. *J Appl Physiol*. 2020;129(2):317–24.
25. American College of Sports Medicine. *American College of Sports Medicine Guidelines for Exercise Testing and Prescription*. 5th ed. Philadelphia (PA): Williams & Wilkins; 1995.
26. Lohman TJ, Roache AF, Martorell R. *Anthropometric Standardization Reference Manual*. Champaign (IL): Human Kinetics Books; 1988.
27. Arsenault BJ, Côté M, Cartier A, et al. Effect of exercise training on cardiometabolic risk markers among sedentary, but metabolically healthy overweight or obese post-menopausal women with elevated blood pressure. *Atherosclerosis*. 2009;207(2):530–3.
28. American College of Sports Medicine. *American College of Sports Medicine's Guidelines for Exercise Testing and Prescription*. 9th ed. Philadelphia (PA): Lippincott Williams & Wilkins; 2013.
29. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma,

without use of the preparative ultracentrifuge. *Clin Chem*. 1972; 18(6):499–502.
30. Westerterp KR. Doubly labelled water assessment of energy expenditure: principle, practice, and promise. *Eur J Appl Physiol*. 2017;117(7): 1277–85.
31. Tataranni PA, Harper IT, Snitker S, et al. Body weight gain in free-living Pima Indians: effect of energy intake vs expenditure. *Int J Obes Relat Metab Disord*. 2003;27:1578–83.
32. Racette SB, Das SK, Bhapkar M, et al, CALERIE Study Group. Approaches for quantifying energy intake and %calorie restriction during calorie restriction interventions in humans: the multicenter CALERIE study. *Am J Physiol Endocrinol Metab*. 2012;302:E441–8.
33. Flint A, Raben A, Blundell JE, Astrup A. Reproducibility, power and validity of visual analogue scales in assessment of appetite sensations in single test meal studies. *Int J Obes Relat Metab Disord*. 2000; 24(1):38–48.
34. Womble LG, Wadden TA, Chandler JM, Martin AR. Agreement between weekly vs. daily assessment of appetite. *Appetite*. 2003;40(2): 131–5.
35. Anderson SE, Bandini LG, Dietz WH, Must A. Relationship between temperament, nonresting energy expenditure, body composition, and physical activity in girls. *Int J Obes (Lond)*. 2004;28(2):300–6.
36. Knäuper B, Rabiau M, Cohen O, Patriciu N. Compensatory health beliefs: scale development and psychometric properties. *Psychol Health*. 2004;19(5):607–24.
37. Stunkard AJ, Messick S. The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *J Psychosom Res*. 1985;29(1):71–83.
38. White MA, Whisenhunt BL, Williamson DA, Greenway FL, Netemeyer RG. Development and validation of the food-craving inventory. *Obes Res*. 2002;10(2):107–14.
39. Geiselman PJ, Anderson AM, Dowdy ML, West DB, Redmann SM, Smith SR. Reliability and validity of a macronutrient self-selection paradigm and a Food Preference Questionnaire. *Physiol Behav*. 1998;63(5):919–28.
40. Anderson DA, Williamson DA, Duchmann EG, Gleaves DH, Barbin JM. Development and validation of a multifactorial treatment outcome measure for eating disorders. *Assessment*. 1999;6(1):7–20.
41. Maxwell S, Delaney H. *Designing Experiments and Analyzing Data: A Model Comparison Perspective*. 2nd ed. Routledge; 1990. 1990 p.
42. King NA, Hopkins M, Caudwell P, Stubbs RJ, Blundell JE. Individual variability following 12 weeks of supervised exercise: identification and characterization of compensation for exercise-induced weight loss. *Int J Obes (Lond)*. 2008;32(1):177–84.
43. Sawyer BJ, Bhammar DM, Angadi SS, et al. Predictors of fat mass changes in response to aerobic exercise training in women. *J Strength Cond Res*. 2015;29:297–304.
44. Nackers LM, Ross KM, Perri MG. The association between rate of initial weight loss and long-term success in obesity treatment: does slow and steady win the race? *Int J Behav Med*. 2010;17(3):161–7.
45. Martin CK, Rosenbaum D, Han H, et al. Change in food cravings, food preferences, and appetite during a low-carbohydrate and low-fat diet. *Obesity*. 2011;19:1963–70.
46. Williamson DA, Martin CK, Anton SD, et al. Is caloric restriction associated with development of eating-disorder symptoms? Results from the CALERIE trial. *Health Psychol*. 2008;27:S32–42.
47. Unick JL, Dorfman L, Leahey TM, Wing RR. A preliminary investigation into whether early intervention can improve weight loss among those initially non-responsive to an internet-based behavioral program. *J Behav Med*. 2016;39:254–61.
48. Montani JP, Schutz Y, Dulloo AG. Dieting and weight cycling as risk factors for cardiometabolic diseases: who is really at risk? *Obes Rev*. 2015;16:7–18.
49. Phillips SM, Joyner MJ. Out-running "bad" diets: beyond weight loss there is clear evidence of the benefits of physical activity. *Br J Sports Med*. 2019;53(14):854–5.

Publication 4

Why exercise by itself is often ineffective for weight loss but crucial for weight management.

Christoph Höchsmann [1] & Karsten Koehler [1]

[1] Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
Sports & Exercise Medicine Switzerland (SEMS) Journal; 70(4)	2022	10.34045/SEMS/2022/21	N/A	Q4	Scopus: 1 Google Scholar: 1

Author contribution:

First and corresponding author, received invitation to submit review article; drafted the manuscript, and created tables and figures.

SEMS-journal
Sports & exercise medicine switzerland

published online on 01.12.2022
https://doi.org/10.34045/SEMS/2022/21

REVIEW

Why exercise by itself is often ineffective for weight loss but crucial for weight management

EXERCISE IS MEDICINE

Bild: Shutterstock

Höchsmann Christoph, Koehler Karsten
Department of Sport and Health Sciences, Technical University of Munich, Germany

Abstract

The prevalence of obesity is increasing worldwide, and excess body weight is associated with a substantially increased risk of adverse health conditions. Exercise supports the prevention and management of obesity; however, when used for weight loss, exercise (even at high volumes) is usually relatively ineffective, frequently producing less weight loss than expected based on measured energy expenditure. The difference between observed and expected weight loss is called compensation and it is

published online on 01.12.2022
https://doi.org/10.34045/SEMS/2022/21

primarily caused by increases in energy intake in response to exercise (compensatory eating). On the other hand, it has been shown that energy balance and body weight are better regulated in individuals with moderate to high levels of physical activity (i.e., energy intake = energy expenditure) compared to those with a sedentary lifestyle (energy expenditure < energy intake), demonstrating that physical activity and exercise are crucial for long-term maintenance of a healthy weight. Weight loss approaches should combine dietary components (calorie restriction) and physical activity for increased success. Calorie restriction facilitates weight loss while physical activity can support the conservation of fat-free mass to avoid a state of increased hunger, often occurring following calorie restriction interventions due to the associated loss in fat-free mass, which ultimately encourages weight regain.

Zusammenfassung

Die Prävalenz von Adipositas ist in den letzten Jahren weltweit dramatisch angestiegen, und übermässiges Körpergewicht (Körperfett) ist mit einem deutlich erhöhten Risiko für langfristige gesundheitliche Beeinträchtigungen verbunden. Körperliche Bewegung unterstützt die Prävention und Behandlung von Adipositas; als alleiniges Mittel zur Gewichtsabnahme ist Bewegung (selbst bei hohen Trainingsumfängen) in der Regel jedoch relativ unwirksam und führt häufig zu einem geringeren Gewichtsverlust als aufgrund des gemessenen Energieverbrauchs zu erwarten wäre. Die Differenz zwischen dem beobachteten und dem erwarteten Gewichtsverlust wird als Kompensation bezeichnet und ist in erster Linie auf eine erhöhte Energiezufuhr als Reaktion auf das Training zurückzuführen (kompensatorisches Essen). Andererseits hat sich gezeigt, dass der Energiehaushalt und das Körpergewicht bei Personen mit mässiger bis hoher körperlicher Aktivität (d.h. Energieaufnahme = Energieverbrauch) besser reguliert sind als bei Personen mit einem sitzenden Lebensstil (Energieverbrauch < Energieaufnahme), was verdeutlicht, dass körperliche Aktivität und Bewegung für die langfristige Aufrechterhaltung eines gesunden Gewichts entscheidend sind. Ansätze zur Gewichtsabnahme sollten ernährungsbezogene Komponenten (Kalorienrestriktion) und körperliche Aktivität kombinieren, um den Erfolg zu steigern. Eine Kalorienrestriktion erleichtert die Gewichtsabnahme, während körperliche Aktivität die Erhaltung der fettfreien Masse unterstützen kann. Dies kann wiederum helfen, das verstärkte Hungergefühl zu reduzieren, das nach einer Kalorienrestriktion aufgrund des damit verbundenen Verlusts an fettfreier Masse häufig auftritt und letztlich eine erneute Gewichtszunahme begünstigt.

Introduction

Overweight and obesity rates have reached epidemic proportions worldwide. In Europe, approximately 60% of citizens are either overweight (body mass index [BMI] ≥25 kg/m^2) or have obesity (BMI ≥30 kg/m^2) [1]. It has been estimated that the worldwide obesity prevalence will reach 18-21% by 2025, suggesting that over one billion people will be affected by obesity [1]. Obesity is associated with an increased risk for serious health conditions, presenting a substantial public health and economic burden worldwide [2]. Weight loss is one of the most important reasons why individuals with overweight and obesity choose to engage in an exercise or physical activity (PA) program [3], and the fitness industry frequently advertises specific ‘weight loss workouts’ to target these individuals. Regular exercise and PA are associated with a plethora of health benefits, including improvements in psychiatric, neurological, metabolic,

published online on 01.12.2022
https://doi.org/10.34045/SEMS/2022/21

cardiovascular, respiratory, and musculoskeletal conditions and diseases as well as many cancers are well established [4]. However, despite the many important health benefits of exercise and PA, by itself, exercise is often ineffective for weight loss, with most individuals losing less weight than expected based on measured energy expenditure, and some individuals even gaining weight after engaging in an exercise or PA program [5]. Villareal et al. showed that average weight loss during a 1-year randomized controlled trial (N=107) was only around 1% (not significant) following an exercise intervention (3x90min per week) compared to around 10% following a diet intervention (daily energy deficit of 500-750 kcal), independent of whether or not the diet intervention was combined with an exercise program [6].
Various factors contribute to the seeming ineffectiveness of exercise in producing weight loss. Most importantly it is due to an insufficient energy deficit primarily caused by changes (increases) in energy intake in response to the exercise program that (over-) compensate the exercise-induced energy deficit and negate weight loss.

Energy-related aspects of weight loss

Both PA, which is defined as any bodily movement produced by skeletal muscles [7], and exercise, a subset of PA that is planned, structured, and repetitive and aims to improve or maintain physical fitness [8], increase energy expenditure. The amount of energy expenditure during exercise is dependent on the type, intensity, and duration of the activity. One metabolic equivalent (MET) equals an oxygen uptake of 3.5 mL/kg/min or 1 kcal/kg/h, which roughly corresponds to a person's resting metabolic rate. For example, at a running speed of 12 km/h (11.5 MET), a person weighing 70 kg would expend ~400 kcal during a 30-min run, whereas a person weighing 90 kg would expend ~ 520 kcal [9]. Of course, the exercise-related energy expenditure (EEE) only accounts for parts of the total daily energy expenditure (TDEE), with resting metabolic rate (RMR; ~60-70% of TDEE), thermic effect of food (TEF; ~5-10% of TDEE), and non-exercise activity thermogenesis (NEAT = TDEE−[RMR+TEF]), making up the other components. As illustrated in *Figure 1*, regular PA and exercise increase TDEE; however a moderate PA level (PAL=1.8) only increases TDEE by around 20% compared to a sedentary lifestyle and even in very active individuals such as competitive athletes with a training load of 2-3h per day, TDEE is only about 50-60% higher than in physically inactive individuals. This illustrates why calorie restriction diets such as a very low-calorie diet (VLCD, ~600-800 kcal/d) and even more moderate diet approaches (1600-1800 kcal/d) can induce a much greater calorie deficit, which is subsequently reflected in more substantial weight loss.

published online on 01.12.2022
https://doi.org/10.34045/SEMS/2022/21

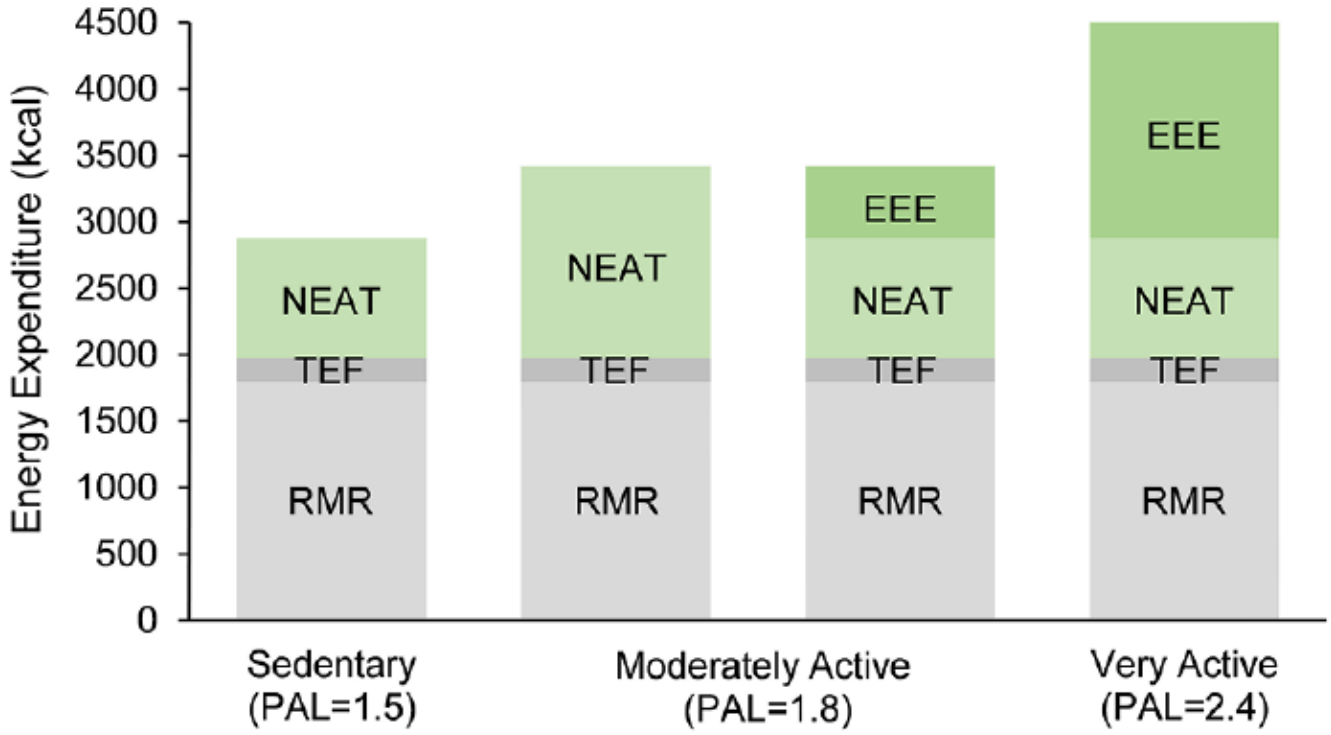


Figure 1: Relative contribution of exercise-related energy expenditure and non-exercise activity thermogenesis to total daily energy expenditure. EEE, energy-related energy expenditure; NEAT, non-exercise activity thermogenesis; PAL, physical activity level; RMR, resting metabolic rate; TEF, thermic effect of food

Additionally, regular exercise alters the energy deficit necessary to lose 1 kg of weight. During a typical calorie restriction intervention, 3/4 of the lost weight consists of fat mass while 1/4 consists of fat-free mass (FFM). In a weight loss intervention that combines calorie restriction and exercise, this ratio shifts to approximately 7/8 (fat mass) and (1/8 FFM), and in an exercise intervention without calorie restriction, almost all FFM is conserved, and the weight loss is almost exclusively due to a reduction in fat mass [10]. While the conservation of FFM is, of course, positive and beneficial for health and longevity [11], it explains the slower rate of weight loss. Because fat mass has a substantially higher energy density than FFM (9400 vs. 1800 kcal/kg) [12], an about 25% greater energy deficit is needed to lose the same amount of weight solely from exercise compared to solely from calorie restriction.

Further, while PA and exercise increase TDEE, most individuals increase energy intake in response to the PA or exercise program [13], which reduces or even negates the exercise-induced energy deficit and consequently the weight lost from exercise (weight compensation). The 6-month E-Mechanic (Examination of Mechanisms of Exercise-Induced Weight Compensation) study examined the mechanisms of weight compensation by comparing two exercise interventions with different weekly energy expenditures: 8 kcal/kg/wk (8KKW) and 20 kcal/kg/wk (20KKW) [5]. For an individual with a body weight of 70 kg, this would equate to a total weekly EEE of 560 kcal (8KKW) and 1400 kcal (20KKW), respectively. The study found that only 58% (8KKW) and 77% (20KKW) of participants lost work at all during the 6-month intervention and that in 76% (8KKW) and 90% (20KKW) of participants weight compensation (less weight loss than expected based on measured EEE) occurred. On average, participants in the 8KKW group lost 0.4 kg (vs. the expected 1.9 kg) and participants in the 20KKW group lost 1.6 kg (vs. the expected 4.3 kg).

published online on 01.12.2022
https://doi.org/10.34045/SEMS/2022/21

Measurement of energy intake via doubly-labeled water further showed that participants increased their daily energy intake by 91 kcal (8KKW) and 124 kcal (20KKW) on average compared to baseline [5]. Various factors contribute to compensatory eating, for example, physiological changes such as an increased release of appetite-stimulating hormones or psychological factors such as food reward after an exhausting workout [14].

Effect of exercise on eating behavior

As early as the 1950s, Jean Mayer found that an increased TDEE leads to an increased energy intake, but that conversely, a decrease in TDEE does not cause a reduction in energy intake. Specifically, in experiments on factory workers in India, Mayer et al. found the greatest energy intake both in workers with the highest work-related energy expenditure and those with the lowest work-related energy expenditure. Those workers conducting light-to-moderate work had the lowest energy intake [15]. This J-shaped association between energy expenditure and energy intake has been confirmed in several studies [16], demonstrating that becoming sedentary does not downregulate energy intake (unregulated zone of energy intake), which consequently leads to higher body weight and BMI over time. In the unregulated zone, non-homeostatic factors such as the availability of food influence food intake [17]. Increasing PA on the other hand improves satiety signaling, and homeostatic factors (i.e., factors to maintain body weight) influence energy intake (regulated zone) [17,18]. Importantly, this association also holds for the intake of sugar and nutrient-dense foods. An analysis of NHANES (National Health and Nutrition Examination Survey) data further showed *(Figure 2)* that intake of sugar and sweetened beverages (energy-dense and low-nutrient foods) was highest in sedentary individuals and those with very high PA levels, while it was the lowest among moderately active individuals [19]. Conversely, intake of nutrient-dense, “healthy” foods (fruit and vegetables, fiber, whole grain, dairy products) increased from sedentary to moderately active but then remained stable despite further increases in energy intake *(Figure 2).* Finally, and consistent with the Mayer curve, a secondary analysis of the E-Mechanic study found that PA behavior before participating in the 6-month exercise intervention predicted weight compensation, with greater amounts of habitual PA (moderate-to-vigorous intensity) being associated with less weight compensation and more weight loss during the intervention [20].

published online on 01.12.2022
https://doi.org/10.34045/SEMS/2022/21

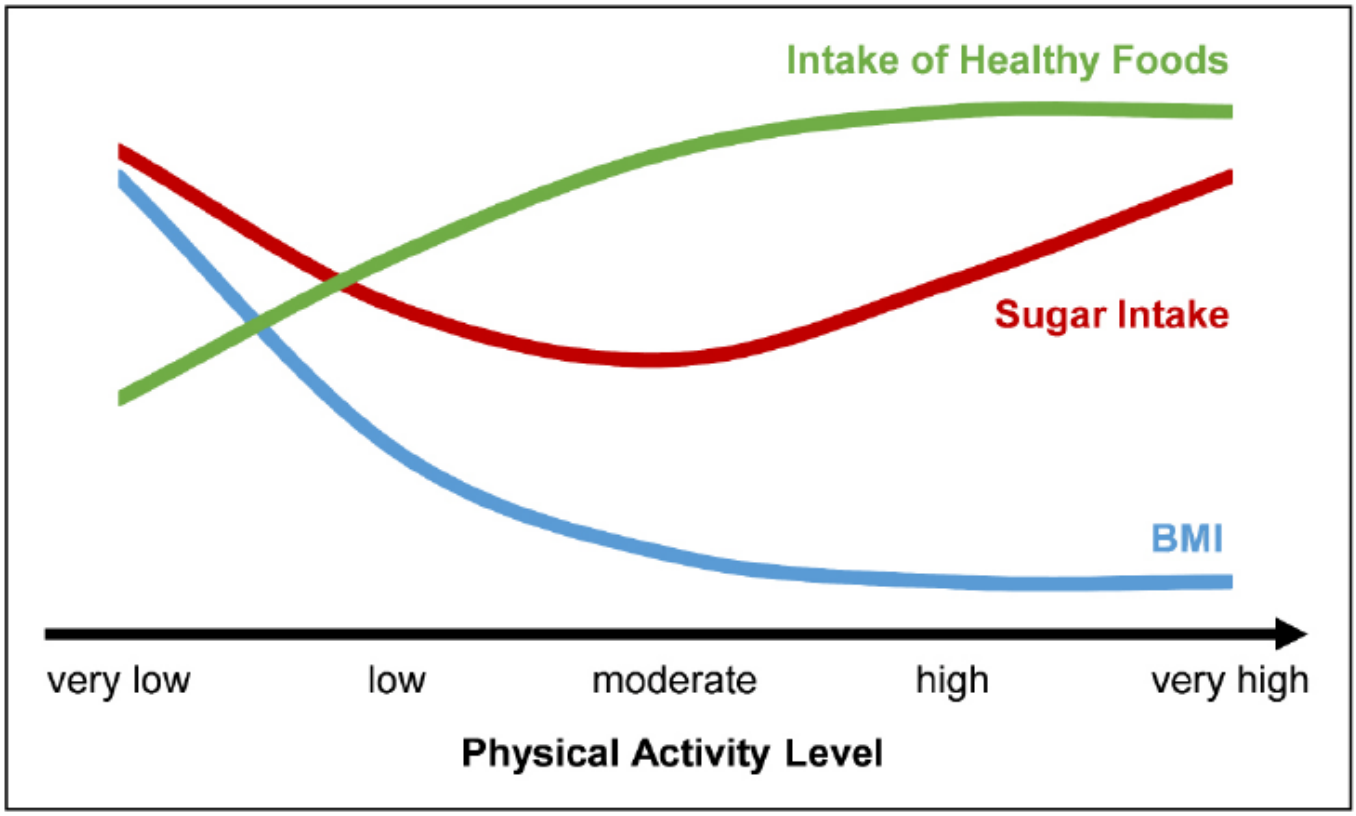


Figure 2: Intake of added sugar and healthy foods (fruit and vegetables, fiber, whole grain, dairy products) as well as BMI at different levels of physical activity. Figure modified from Koehler et al. (2019)[19].

In addition to affecting general long-term eating habits and energy balance, PA and exercise also acutely affect how much and what we want to eat. A recent study examined hypothetical food choices before, immediately after and 30 min after a 45-min exercise session on a bike ergometer (60% VO2peak) compared to a rest condition of identical duration [21]. Specifically, participants viewed a series of food images displaying hypothetical food choices with varying palatability and energy density (sweet/low fat, non-sweet/low fat, sweet/high fat, and non-sweet/high fat, respectively) and rated their food amount preference at these time points after each study condition (bike ergometer vs. rest). The results showed that the selected food amount (kcal) increased after the exercise condition, with increases of 23% (immediately after) and 30% (30 min after), whereas the rest condition did not induce such increases in food amount preference [21]. Another field-based study asked participants to choose between a “healthy” snack (apple) and an “unhealthy” snack (brownie) either before or after a gym visit, with the consumption of the snack taking place after the gym visit in either case [22]. The proportion of participants, who chose the healthy snack option (to be eaten after the gym visit) decreased from 74% (choice before the gym visit) to 55% (choice after the gym visit) and the proportion of the unhealthy snack option increased from 14% to 20% from before to after [22]. These findings are consistent with the behavioral phenomenon of *immediate gratification* that leads to more impulsive food choices in a state of increased hunger, irrespective of longer-term consequences, and has been linked to an increased risk of obesity [23]. Further, these results suggest that food choices are generally better (i.e., less food, healthier choices) before compared to after exercise. This may have implications for clinical practice, as food choices and preparation before exercise

published online on 01.12.2022
https://doi.org/10.34045/SEMS/2022/21

may help reduce compensatory energy intake and thereby support long-term weight loss through exercise.

Exercise and long-term weight management

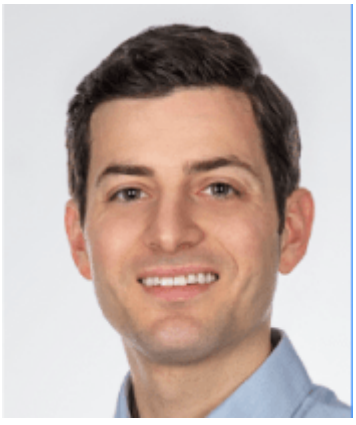
While the aforementioned phenomena are responsible for a slowed initial weight loss through exercise, this explains at the same time, why (the addition of) PA and exercise are crucial for long-term weight management and advantageous compared to mere calorie restriction interventions. Importantly, the conservation of FFM through exercise can help prevent rapid weight regain after the end of the calorie restriction intervention. That is because the loss of FFM during a mere calorie restriction intervention leads to a state of hyperphagia that persists until FFM is fully recovered. The recovery of FFM is inevitably accompanied by fat deposition, causing weight regain [24]. Conserving FFM during a weight loss intervention through moderate-to-high levels of PA will consequently facilitate better maintenance of the new weight, as demonstrated in a follow-up study to a large weight loss intervention [25]. Participants in that study who were highly active (2500 kcal/wk) during the 6-month behavioral weight loss intervention had a lower weight 2 years after the end of the intervention compared to those who were moderately active (1000 kcal/wk) during the intervention. Even more impressive, those participants who sustained the high PA levels during the 2 years after the weight loss intervention maintained 12 kg weight loss, whereas those who ceased to be physically active returned to their baseline weight [25].

Conclusions

Despite the numerous positive effects in the prevention and treatment of various diseases, the effects of PA and exercise on body weight are at least in the initial period of weight loss interventions small to negligible. It is, therefore, crucial to lower (unrealistic) expectations of rapid exercise-induced weight loss in clinical practice and to counteract compensatory eating behaviors, which further reduce the weight lost during an exercise-based weight loss intervention. However, the positive effects of PA and exercise for long-term weight management are undisputed, and lasting benefits for energy balance are particularly achieved through the lifelong implementation of a physically active lifestyle.

Corresponding author

Christoph Höchsmann, PhD
Department of Sport and Health Sciences
Technical University of Munich
Connollystrasse 32, 80809 Munich, Germany



published online on 01.12.2022
https://doi.org/10.34045/SEMS/2022/21

References

1. Boutari C, Mantzoros CS. A 2022 update on the epidemiology of obesity and a call to action: as its twin COVID-19 pandemic appears to be receding, the obesity and dysmetabolism pandemic continues to rage on. Metabolism. 2022 Aug;133:155217.
2. Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. BMC Public Health. 2009 Mar 25;9:88.
3. Statista. Why people work out at a gym in the US 2016 [Internet]. 2016 [cited 2022 Sep 19]. Available from: https://www.statista.com/statistics/639169/reasons-behind-gym-exercise-in-us/
4. Pedersen BK, Saltin B. Exercise as medicine – evidence for prescribing exercise as therapy in 26 different chronic diseases. Scand J Med Sci Sports. 2015 Dec;25 Suppl 3:1-72.
5. Martin CK, Johnson WD, Myers CA, Apolzan JW, Earnest CP, Thomas DM, et al. Effect of different doses of supervised exercise on food intake, metabolism, and non-exercise physical activity: The E-MECHANIC randomized controlled trial. Am J Clin Nutr. 2019 Sep 1;110(3):583-92.
6. Villareal DT, Chode S, Parimi N, Sinacore DR, Hilton T, Armamento-Villareal R, et al. Weight Loss, Exercise, or Both and Physical Function in Obese Older Adults. N Engl J Med. 2011 Mar 31;364(13):1218-29.
7. World Health Organization. WHO | Physical activity [Internet]. 2017 [cited 2017 Aug 29]. Available from: http://www.who.int/mediacentre/factsheets/fs385/en/
8. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep. 1985;100(2):126-31.
9. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc. 2011 Aug;43(8):1575-81.
10. Weinheimer EM, Sands LP, Campbell WW. A systematic review of the separate and combined effects of energy restriction and exercise on fat-free mass in middle-aged and older adults: implications for sarcopenic obesity. Nutr Rev. 2010 Jul;68(7):375-88.
11. Sedlmeier AM, Baumeister SE, Weber A, Fischer B, Thorand B, Ittermann T, et al. Relation of body fat mass and fat-free mass to total mortality: results from 7 prospective cohort studies. Am J Clin Nutr. 2021 Mar 11;113(3):639-46.
12. Hall KD. What is the Required Energy Deficit per unit Weight Loss? Int J Obes 2005. 2008 Mar;32(3):573-6.
13. Blundell JE, Gibbons C, Caudwell P, Finlayson G, Hopkins M. Appetite control and energy balance: impact of exercise. Obes Rev Off J Int Assoc Study Obes. 2015 Feb;16 Suppl 1:67-76.
14. Hall KD, Heymsfield SB, Kemnitz JW, Klein S, Schoeller DA, Speakman JR. Energy balance and its components: implications for body weight regulation123. Am J Clin Nutr. 2012 Apr;95(4):989-94.
15. Mayer J, Roy P, Mitra KP. Relation between caloric intake, body weight, and physical work: studies in an industrial male population in West Bengal. Am J Clin Nutr. 1956 Apr;4(2):169-75.
16. Church T, Martin CK. The Obesity Epidemic: A Consequence of Reduced Energy Expenditure and the Uncoupling of Energy Intake? Obes Silver Spring Md. 2018;26(1):14-6.
17. Beaulieu K, Oustric P, Finlayson G. The Impact of Physical Activity on Food Reward: Review and Conceptual Synthesis of Evidence from Observational, Acute, and Chronic Exercise Training Studies. Curr Obes Rep. 2020;9(2):63-80.

published online on 01.12.2022
https://doi.org/10.34045/SEMS/2022/21

18. Beaulieu K, Hopkins M, Long C, Blundell J, Finlayson G. High Habitual Physical Activity Improves Acute Energy Compensation in Nonobese Adults. Med Sci Sports Exerc. 2017 Nov;49(11):2268-75.
19. Koehler K, Boron JB, Garvin TM, Bice MR, Stevens JR. Differential relationship between physical activity and intake of added sugar and nutrient-dense foods: A cross-sectional analysis. Appetite. 2019 Sep 1;140:91-7.
20. Höchsmann C, Dorling JL, Apolzan JW, Johannsen NM, Hsia DS, Martin CK. Baseline Habitual Physical Activity Predicts Weight Loss, Weight Compensation, and Energy Intake During Aerobic Exercise. Obes Silver Spring Md. 2020 May;28(5):882-92.
21. Koehler K, Beckford SE, Thayer E, Martin AR, Boron JB, Stevens JR. Exercise Shifts Hypothetical Food Choices toward Greater Amounts and More Immediate Consumption. Nutrients. 2021 Jan 24;13(2):347.
22. Gustafson CR, Rakhmatullaeva N, Beckford SE, Ammachathram A, Cristobal A, Koehler K. Exercise and the Timing of Snack Choice: Healthy Snack Choice is Reduced in the Post-Exercise State. Nutrients. 2018 Dec 7;10(12):1941.
23. Schiff S, Amodio P, Testa G, Nardi M, Montagnese S, Caregaro L, et al. Impulsivity toward food reward is related to BMI: Evidence from intertemporal choice in obese and normal-weight individuals. Brain Cogn. 2016 Dec;110:112-9.
24. Dulloo AG, Jacquet J, Montani JP, Schutz Y. How dieting makes the lean fatter: from a perspective of body composition autoregulation through adipostats and proteinstats awaiting discovery. Obes Rev Off J Int Assoc Study Obes. 2015 Feb;16 Suppl 1:25-35.
25. Tate DF, Jeffery RW, Sherwood NE, Wing RR. Long-term weight losses associated with prescription of higher physical activity goals. Are higher levels of physical activity protective against weight regain? Am J Clin Nutr. 2007 Apr;85(4):954-9.

COMPENSATORY EATING ENERGY BALANCE WEIGHT COMPENSATION WEIGHT LOSS WEIGHT MAINTENANCE

Publication 5

Initial weight loss and early intervention adherence predict long-term weight loss during the Promoting Successful Weight Loss in Primary Care in Louisiana lifestyle intervention.

Christoph Höchsmann [1,2], Corby K Martin [2], John W Apolzan [2], James L Dorling [3], Robert L Newton Jr [2], Kara D Denstel [2], Emily F Mire [2], William D Johnson [2], Dachuan Zhang [2], Connie L Arnold [4], Terry C Davis [4], Vivian Fonseca [5], Tina K Thethi [6], Carl J Lavie [7], Benjamin Springgate [8], Peter T Katzmarzyk [2] for the PROPEL Research Group

[1] Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany

[2] Pennington Biomedical Research Center, Baton Rouge, LA, USA

[3] Human Nutrition, School of Medicine, Dentistry and Nursing, University of Glasgow, Glasgow, UK

[4] Louisiana State University Health Sciences Center, Shreveport, LA, USA

[5] Department of Medicine, Tulane University Health Sciences Center, New Orleans, LA, USA

[6] AdventHealth, Translational Research Institute, Orlando, FL, USA

[7] John Ochsner Heart and Vascular Institute, New Orleans, LA, USA

[8] Louisiana State University School of Medicine, New Orleans, LA, USA

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
Obesity, 31(9): 2272–2282	2023	10.1002/oby.23854	6.9 (2022)	Q1	Scopus: 1 Google Scholar: 1

Author contribution:

First author; developed the research question, co-developed the statistical model, drafted the manuscript, and created tables and figures.

Received: 23 August 2022 | Revised: 30 May 2023 | Accepted: 3 June 2023

DOI: 10.1002/oby.23854

Obesity A Research Journal | THE OBESITY SOCIETY | WILEY

ORIGINAL ARTICLE

Clinical Trials and Investigations

Initial weight loss and early intervention adherence predict long-term weight loss during the Promoting Successful Weight Loss in Primary Care in Louisiana lifestyle intervention

Christoph Höchsmann[1,2] | Corby K. Martin[2] | John W. Apolzan[2] | James L. Dorling[3] | Robert L. Newton Jr[2] | Kara D. Denstel[2] | Emily F. Mire[2] | William D. Johnson[2] | Dachuan Zhang[2] | Connie L. Arnold[4] | Terry C. Davis[4] | Vivian Fonseca[5] | Tina K. Thethi[5] | Carl J. Lavie[6] | Benjamin Springgate[7] | Peter T. Katzmarzyk[2] | for the PROPEL Research Group

[1]Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany

[2]Pennington Biomedical Research Center, Baton Rouge, Louisiana, USA

[3]Human Nutrition, School of Medicine, Dentistry and Nursing, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK

[4]Department of Medicine and Feist-Weiller Cancer Center, Louisiana State University Health Sciences Center, Shreveport, Louisiana, USA

[5]AdventHealth, Translational Research Institute, Orlando, Florida, USA

[6]Department of Cardiovascular Diseases, John Ochsner Heart and Vascular Institute, New Orleans, Louisiana, USA

[7]Department of Internal Medicine, Louisiana State University School of Medicine, New Orleans, Louisiana, USA

Correspondence
Peter T. Katzmarzyk, Pennington Biomedical Research Center, 6400 Perkins Rd., Baton Rouge, LA 70808, USA.
Email: peter.katzmarzyk@pbrc.edu

Funding information
National Institute of Diabetes and Digestive and Kidney Diseases, Grant/Award Number: T32DK064584; National Institute of General Medical Sciences, Grant/Award Number: U54 GM104940; Nutrition Obesity Research Centers, Grant/Award Number: P30 DK072476; Patient-Centered Outcomes Research Institute, Grant/Award Number: OB-1402-10977

Abstract

Objective: This study tested whether initial weight change (WC), self-weighing, and adherence to the expected WC trajectory predict longer-term WC in an underserved primary-care population with obesity.

Methods: Data from the intervention group ($n = 452$; 88% women; 74% Black; BMI 37.3 kg/m^2 [SD: 4.6]) of the Promoting Successful Weight Loss in Primary Care in Louisiana trial were analyzed. Initial (2-, 4-, and 8-week) percentage WC was calculated from baseline clinic weights and daily at-home weights. Weights were considered adherent if they were on the expected WC trajectory (10% at 6 months with lower [7.5%] and upper [12.5%] bounds). Linear mixed-effects models tested whether initial WC and the number of daily and adherent weights predicted WC at 6, 12, and 24 months.

Results: Percentage WC during the initial 2, 4, and 8 weeks predicted percentage WC at 6 ($R^2 = 0.15$, $R^2 = 0.28$, and $R^2 = 0.50$), 12 ($R^2 = 0.11$, $R^2 = 0.19$, and $R^2 = 0.32$), and 24 ($R^2 = 0.09$, $R^2 = 0.11$, and $R^2 = 0.16$) months (all $p < 0.01$). Initial daily and adherent weights were significantly associated with WC as individual predictors, but they only marginally improved predictions beyond initial weight loss alone in multivariable models.

Conclusions: These results highlight the importance of initial WC for predicting long-term WC and show that self-weighing and adherence to the expected WC trajectory can improve WC prediction.

INTRODUCTION

Almost half of all US adults are affected by obesity [1], presenting a substantial public health burden with an increased risk for serious health conditions [2]. Particularly alarming is that 9% of adults have severe or Class III obesity with body mass index (BMI) ≥ 40 kg/m^2 [1]. Certain demographic groups, such as Black and Hispanic populations as well as those with a low household income, exhibit higher rates of obesity [3, 4]. To decrease these health disparities, it is important to identify effective weight management methods for individuals with obesity in these populations.

Intensive lifestyle interventions (ILIs) are the current non-medical gold-standard approach to promote weight loss (WL) and effectively treat obesity and manage associated health risks [5]. However, there is considerable variability in the response to lifestyle-based WL treatment, and many participants in ILIs do not achieve clinically significant WL (≥5%), particularly in the medium to long term [6, 7]. Therefore, identifying predictors of medium-to-long-term WL is important to guide treatment and identify patients who may need increased support for WL [8]. The ultimate goal is to predict an individual's WL as early as possible and to intervene with alternative treatment approaches if needed. Patient-specific characteristics such as male sex [9–11], older age [12–14], and White versus Black race [15] have been shown to be predictive of greater WL; however, overall, there is only limited evidence for the association between baseline patient-specific characteristics and WL outcomes in lifestyle-based WL studies [8]. In addition to being modifiable behaviors, dynamic, intervention-specific characteristics such as greater initial WL and adherence to an ILI program are more reliable predictors of medium-to-long-term WL [8].

For example, in a 12-month behavioral WL intervention, initial (1-month) WL made the strongest unique contribution to the prediction of WL at 12 months [10]. Similarly, in a 10-week dietary intervention (−600 kcal/d), halfway (5-week) WL explained 77% of the variance in 10-week WL, and even very early (1-week) WL explained 28% of the variance [11]. Furthermore, initial WL was shown to predict long-term WL, with those losing >4% in the first month and >6% in the first 2 months having 2- and 2.8-times greater odds of achieving ≥5% WL after 8 years in the Action for Health in Diabetes (Look AHEAD) trial [16]. In addition to initial WL, behavioral adherence (i.e., session attendance, self-monitoring of behavior, and self-weighing) is associated with greater WL. In the Preventing Overweight Using Novel Dietary Strategies (POUNDS Lost) Study [17], early (first 6 months) adherence was predictive of greater WL after 6 and 24 months, and in two other clinical behavioral intervention studies in academic health centers, self-weighing adherence predicted WL after 12 weeks [18] and 12 months [19].

To date, no study, to our knowledge, has examined the effects of initial WL or initial adherence to a self-weighing protocol or the expected WL trajectory on medium-to-long-term WL during a pragmatic ILI program conducted in primary care. We aimed to test whether WL, self-weighing adherence, and adherence to the expected WL trajectory during different time points in the early phase of a pragmatic WL intervention (2 weeks, 4 weeks, and 8 weeks) predicted medium-to-long-term (6, 12, and 24 months) WL. Identifying the time at which the prediction is best is important to inform clinical application, as it allows deployment of targeted countermeasures in those who struggle with initial adherence and WL as early as possible to improve the odds of medium-to-long-term WL. We used data from the Promoting Successful Weight Loss in Primary Care in Louisiana (PROPEL) trial for our analyses, which demonstrated significant WL of 4.5% (95% confidence interval[CI]: 3.1%–5.9%) in underserved individuals with obesity following a 24-month ILI compared with usual care [20], along with significant improvements in cardiovascular disease risk factors [21]. We hypothesized that greater initial WL, self-weighing adherence, and adherence to the expected WL trajectory would predict greater WL at 6, 12, and 24 months. We further hypothesized that a longer time frame of initial WL and adherence (i.e., 8 weeks) would yield the strongest prediction of medium-to-long-term WL.

Study Importance

What is already known?

- There is limited evidence for the association between baseline patient-specific characteristics and weight loss outcomes in lifestyle-based weight loss studies.
- Dynamic, intervention-specific characteristics such as greater initial weight loss and adherence to the lifestyle intervention are more reliable predictors of medium-to-long-term weight loss.

What does this study add?

- Percentage weight change during the initial 2, 4, and 8 weeks of an intensive lifestyle intervention delivered to an underserved population with obesity in primary care predicted percentage weight change at 6, 12, and 24 months.
- Initial self-weighing and particularly adherence to the predicted weight loss trajectory further improved the prediction models, but the improvements were small.

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

- The present results highlight the benefits of the weight graph approach in pragmatic weight loss interventions that allows early identification of individuals who may perform poorly and need additional support for long-term weight loss.
- These results further demonstrate the importance for those administering weight loss interventions in primary care of ensuring participants' early intervention adherence (diet/calorie restriction, daily self-weighing) to improve long-term weight loss.

METHODS

Design and participants

This report is a secondary analysis of the 24-month PROPEL trial (ClinicalTrials.gov identifier NCT02561221). PROPEL was a cluster-randomized, two-arm trial conducted in primary-care clinics between April 2016 and September 2019. The complete design, methods, and primary and secondary outcomes of the PROPEL trial have been previously published [20–22]. Briefly, 18 primary-care clinics across Louisiana were randomly allocated in equal numbers to either an ILI group or a usual care group. Participants were recruited from the participating clinics and eligible if they were 20 to 75 years old, had BMI of 30 to 50, and were patients at one of the participating clinics. Participants were excluded if they were participating in a structured WL program or using WL medication at the time of enrollment, had ever undergone bariatric surgery or planned to within the next 2 years, or had lost >10 lb of weight within the last 6 months. The complete list of eligibility criteria is provided in the protocol [22]. All procedures of the PROPEL trial were approved by the Institutional Review Board of the Pennington Biomedical Research Center, and all participants provided written informed consent before inclusion in the study.

The primary aim of the PROPEL trial was to develop and test the effectiveness of a 24-month, patient-centered, and pragmatic obesity treatment program delivered within primary care in an underserved population [22]. The trial found clinically relevant WL over 24 months in the ILI group of 5.0% (95% CI: 4.0%–6.0%; $p < 0.01$) [20].

In these analyses, to test whether WL, self-weighing adherence, and adherence to the expected WL trajectory during the initial 2, 4, and 8 weeks of the intervention predicted medium-to-long-term WL, only participants allocated to the ILI group ($n = 452$) were included. Participants at clinics allocated to the ILI group received a comprehensive high-intensity lifestyle intervention, based on previous successful behavioral lifestyle regimens such as the Diabetes Prevention Program [23], Look AHEAD [24], and Comprehensive Assessment of Long-term Effects of Reducing Intake of Energy (CALERIE) [25]. The pragmatic PROPEL program was consistent with the 2013 American Heart Association (AHA)/American College of Cardiology (ACC)/The Obesity Society (TOS) Guidelines for Managing Overweight and Obesity in Adults [5]. The ILI regimen was administered by trained health coaches embedded in primary-care clinics in weekly sessions (16 in-person sessions and 6 via phone) during the first 6 months and at least monthly sessions for the remaining 18 months. The objective for participants in the ILI group was to lose 10% of their body weight during the first 6 months and to maintain that WL in the remaining 18 months. The health coaches worked with participants to develop and adhere to personalized action plans focusing on changes in eating, diet, and physical activity behavior. An important intervention component of PROPEL's ILI program was daily self-weighing and the incorporation of a personalized weight graph, which automatically plotted participants' wirelessly transmitted weight data (BodyTrace scale, BodyTrace Inc., Palo Alto, California) via a computer tracking system (CTS) in relation to the individualized WL (target) zone. The WL zone reflected the expected WL trajectory (10% WL after 6 months) if the personalized daily energy targets were met, including lower and upper bounds representing ~7.5% and ~12.5% of WL to accommodate biological variation and error in the prediction model [22]. Weights were considered adherent if they were on that WL trajectory. An example of the personalized weight graph for a single patient is provided in the protocol [22]. Daily self-weighing has been shown to promote self-efficacy and self-regulatory behavior in between intervention sessions, and participants and health coaches were able to access the weight graph via a website at any time, allowing them to quickly detect deviations from the intended WL progress [26, 27]. If deviations occurred, adjustments were made to the personalized action plans via a toolbox approach (i.e., specific nutritional, physical activity, and behavioral strategies tailored to the needs of the patient, such as increased contact, modifying recipes to decrease energy density, strategies to avoid impulse eating, or adding variety to physical activity routines), as applied in previous clinical trials [23–25].

Measures

Body weight

Body weight was measured during the clinic visits at baseline and at 6, 12, and 24 months, as well as daily via cellular-connected scales in participants' homes. In the clinic, trained staff measured weight in duplicate to the nearest 0.1 kg using a digital scale (Seca Model 876, Weigh and Measure LLC, Olney, Maryland) with participants in light clothing and without shoes. Daily at-home weights were measured by participants autonomously via a BodyTrace scale. Participants were instructed to weigh themselves after getting up in the morning in a similar state (e.g., before breakfast, after the first void) and with similar clothing (e.g., after disrobing to get in the shower), and health coaches worked with participants to foster consistency. Percentage weight change (WC) during the initial 2, 4, and 8 weeks was calculated by subtracting the CTS-recorded weights at 2 weeks (Day 14), 4 weeks (Day 28), and 8 weeks (Day 56) from the baseline clinic weight.

Self-weighing and adherence to the predicted WL trajectory

Self-weighing adherence (the total number of daily at-home weights) and adherence to the expected WL trajectory (the total number of weights in or below the projected WL zone [i.e., adherent weights]) during the initial 2, 4, and 8 weeks of the intervention were assessed via the CTS data.

Statistical analyses

Descriptive data are reported as means and SDs for continuous variables and as frequencies (percentages) for categorical variables. The

primary end points in the present analysis were percentage WC at 6, 12, and 24 months. We used linear mixed-effects models (LMMs) to evaluate the predictive effect of initial WL, self-weighing, and adherence to the predicted WL trajectory (all at week 2, week 4, and week 8, respectively) on WL at each consecutive follow-up time point (6, 12, and 24 months). In the first step, we ran separate models fitting initial percentage WC, number of recorded weights, and number of adherent weights to predict follow-up percentage WC at each time point. These results provided a preliminary view of relationships. Next, all three predictors (initial percentage WC, number of recorded weights, and number of adherent weights) of a specific initial week were entered in the model as candidate independent variables with percentage WC at a follow-up time point as the dependent variable. The model selection process involved comparing models with different combinations of predictor variables and selecting the model with the lowest Akaike information criterion as the final model. We assessed the assumptions of LMMs, including normality of residuals, constant variance, and linearity of the predictor variables, using diagnostic plots and statistical tests. Coefficient estimates of single-predictor models and final models were provided to quantify the rate at which each predictor forecast the outcome. Because it was reasonable to assume that patients within clinics were more homogeneous than between clinics, we took account of the nine clinics as random cluster variance components that deviate from the main fixed effect. The full model can be expressed as the following:

$$\text{Weight loss}_{ij} = \beta_0 + \beta_1 \text{Initial\%weight loss}_{ij} + \beta_2 \text{Number of weights}_{ij} + \beta_3 \text{Number of adherent weights}_{ij} + \gamma_j + \varepsilon_{ij}$$

where i denoted the *ith* patient and j denoted the *jth* clinic. The clustering effect of clinics was considered as a random intercept term $\gamma_j \sim N(0, \sigma_c^2)$, which consequently made the covariance matrix block diagonal. Additionally, we calculated the generalized coefficients of determination (R^2) for the single-predictor and optimal LMMs to quantify the proportion of variance explained by each model [28]. Furthermore, to determine optimal cutoff values to discriminate between successfully achieving versus not achieving clinically relevant WL (≥3%, ≥5%, and ≥10%) at 6, 12, and 24 months based on initial (2-, 4-, and 8-week) WL, we used receiver operating characteristic curves including area under the curve estimates from logistic regression models. Clinically significant WL is typically defined as a reduction of ≥5% of initial weight [5]. However, because modest WL (~3%) is associated with some health benefits in those with obesity and substandard cardiovascular disease risk factors, and more WL (≥10%) is associated with even greater benefits [29], we also determined optimal cutoff values for these categories of WL. The point which minimizes the Euclidean distance of sensitivity (true positive) and specificity (true negative) from perfection (sensitivity = 1 and specificity = 1) was identified as the respective optimal cutoff. All analyses were conducted with SAS version 9.4 (SAS Institute Inc., Cary, North Carolina) for Windows with the significance level set to 0.05 (two-sided).

TABLE 1 Participant characteristics at baseline and descriptive statistics for dependent and independent variables

	$N = 452$
Race, *n* (%)	
Black	332 (73.5)
White	95 (21.0)
Other	25 (5.5)
Sex, *n* (%)	
Male	54 (11.9)
Female	398 (88.1)
Age (y)	48.8 (12.7)
Weight (kg)	101.6 (16.4)
BMI (kg/m^2)	37.3 (4.6)
Initial weight change	
Percentage weight change in initial 2 weeks[a]	−2.9 (1.8)
Percentage weight change in initial 4 weeks[a]	−3.9 (2.2)
Percentage weight change in initial 8 weeks[a]	−5.4 (3.2)
Self-weighing adherence	
At-home weights in initial 2 weeks (*n*)[a]	12.9 (2.0)
At-home weights in initial 4 weeks (*n*)[a]	24.9 (4.7)
At-home weights in initial 8 weeks (*n*)[a]	46.9 (11.9)
Adherence to predicted weight loss trajectory	
Adherent at-home weights in initial 2 weeks (*n*)[a,b]	12.8 (2.2)
Adherent at-home weights in initial 4 weeks (*n*)[a,b]	24.6 (5.1)
Adherent at-home weights in initial 8 weeks (*n*)[a,b]	45.4 (13.0)
Medium-to-long-term weight change	
Percentage weight change at 6 months[c]	−7.4 (5.6)
Percentage weight change at 12 months[d]	−7.0 (7.1)
Percentage weight change at 24 months[e]	−5.1 (7.7)

Note: Data are mean (SD) unless stated otherwise.
[a]Data available for 449 of 452 participants.
[b]Adherent at-home weight was defined as weight in or below the projected weight loss zone.
[c]Data available for 391 of 452 participants.
[d]Data available for 373 of 452 participants.
[e]Data available for 370 of 452 participants.

RESULTS

Participant characteristics

A total of 452 participants were included in the present analyses. Participants (74% Black, 88% women) were 48.8 years (SD = 12.7) old on average and had an average BMI of 37.3 (SD = 4.6). Baseline characteristics are shown in Table 1. On average, participants

TABLE 2 Associations between initial weight change (%), self-weighing, and adherence to the predicted weight loss trajectory (single predictors) during the first 2, 4, and 8 weeks and medium-to-long-term weight change

	R^2	B	SE	*df*	*t*	*p*
Percentage weight change at 6 months						
Weight change						
Percentage weight change in initial 2 weeks	0.15	1.15	0.16	379	7.38	**<0.01**
Percentage weight change in initial 4 weeks	0.28	1.46	0.11	379	12.81	**<0.01**
Percentage weight change in initial 8 weeks	0.50	1.28	0.06	379	20.11	**<0.01**
Self-weighing adherence						
Weights recorded in initial 2 weeks (*n*)	0.10	−0.66	0.15	379	−4.47	**<0.01**
Weights recorded in initial 4 weeks (*n*)	0.14	−0.47	0.06	379	−6.92	**<0.01**
Weights recorded in initial 8 weeks (*n*)	0.19	−0.24	0.03	379	−8.92	**<0.01**
Adherence to predicted weight loss trajectory						
Adherent weights in initial 2 weeks (*n*)[a]	0.10	−0.68	0.14	379	−4.90	**<0.01**
Adherent weights in initial 4 weeks (*n*)[a]	0.15	−0.46	0.06	379	−7.49	**<0.01**
Adherent weights in initial 8 weeks (*n*)[a]	0.26	−0.26	0.02	379	−11.48	**<0.01**
Percentage weight change at 12 months						
Weight change						
Percentage weight change in initial 2 weeks	0.11	1.11	0.20	361	5.48	**<0.01**
Percentage weight change in initial 4 weeks	0.19	1.46	0.16	361	9.33	**<0.01**
Percentage weight change in initial 8 weeks	0.32	1.27	0.10	361	13.35	**<0.01**
Self-weighing adherence						
Weights recorded in initial 2 weeks (*n*)	0.09	−0.85	0.21	361	−4.07	**<0.01**
Weights recorded in initial 4 weeks (*n*)	0.11	−0.51	0.09	361	−5.67	**<0.01**
Weights recorded in initial 8 weeks (*n*)	0.14	−0.24	0.04	361	−6.86	**<0.01**
Adherence to predicted weight loss trajectory						
Adherent weights in initial 2 weeks (*n*)[a]	0.10	−0.92	0.20	361	−4.72	**<0.01**
Adherent weights in initial 4 weeks (*n*)[a]	0.13	−0.53	0.08	361	−6.65	**<0.01**
Adherent weights in initial 8 weeks (*n*)[a]	0.19	−0.27	0.03	361	−9.08	**<0.01**
Percentage weight change at 24 months						
Weight change						
Percentage weight change in initial 2 weeks	0.09	1.07	0.22	358	4.82	**<0.01**
Percentage weight change in initial 4 weeks	0.11	1.17	0.18	358	6.59	**<0.01**
Percentage weight change in initial 8 weeks	0.16	0.96	0.11	358	8.35	**<0.01**
Self-weighing adherence						
Weights recorded in initial 2 weeks (*n*)	0.06	−0.58	0.23	358	−2.49	**0.01**
Weights recorded in initial 4 weeks (*n*)	0.07	−0.40	0.10	358	−4.02	**<0.01**
Weights recorded in initial 8 weeks (*n*)	0.08	−0.18	0.04	358	−4.55	**<0.01**
Adherence to predicted weight loss trajectory						
Adherent weights in initial 2 weeks (*n*)[a]	0.07	−0.71	0.21	358	−3.31	**<0.01**
Adherent weights in initial 4 weeks (*n*)[a]	0.09	−0.44	0.09	358	−5.07	**<0.01**
Adherent weights in initial 8 weeks (*n*)[a]	0.11	−0.20	0.03	358	−6.15	**<0.01**

Note: Models accounted for the random clustering effects of clinics. Bold font indicates statistical significance ($p < 0.05$).
Abbreviations: B, unstandardized regression coefficient; *df*, degrees of freedom.
[a]An adherent at-home weight was defined as a weight in or below the projected weight loss zone.

recorded 12.9 (SD = 2.0), 24.9 (SD = 4.7), and 46.9 (SD = 11.9) daily weights during the initial 2, 4, and 8 weeks of the intervention, respectively, corresponding to a self-weighing adherence of 92%, 89%, and 84%. Of these weights, 99% (initial 2 weeks), 98% (initial 4 weeks), and 96% (initial 8 weeks) were in or below the projected WL zone (i.e., adherent), on average. Average WL was 2.9%

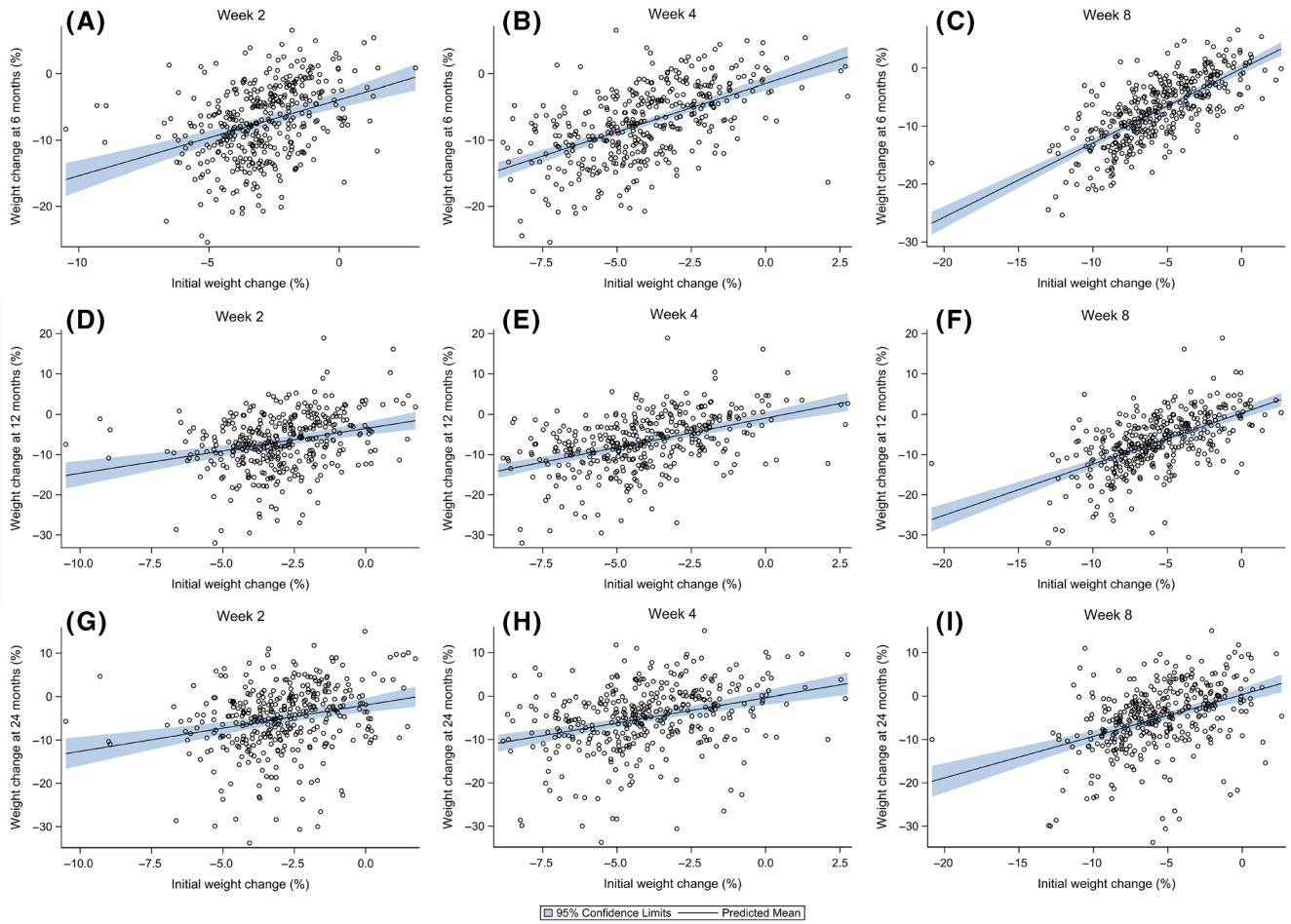


FIGURE 1 Association between initial weight change and medium-to-long-term weight change. Panels A–C show the associations between initial weight change at 2 weeks, 4 weeks, and 8 weeks and 6-month weight change. Panels D–F show the associations between initial weight change at 2 weeks, 4 weeks, and 8 weeks and 12-month weight change. Panels G–I show the associations between initial weight change at 2 weeks, 4 weeks, and 8 weeks and 24-month weight change. [Color figure can be viewed at wileyonlinelibrary.com]

(SD = 1.8%) at 2 weeks, 3.9% (SD = 2.2%) at 4 weeks, and 5.4% (SD = 3.2%) at 8 weeks (Table 1).

Association between initial WL and medium-to-long-term WL

Percentage WC during the initial 2, 4, and 8 weeks predicted percentage WC at 6, 12, and 24 months (all $p < 0.01$; Table 2, Figure 1). At 6 months, initial WL explained 15% (2 weeks), 28% (4 weeks), and 50% (8 weeks) of the variance in WC. At 12 months, these values were 11% (2 weeks), 19% (4 weeks), and 32% (8 weeks) and at 24 months were 9% (2 weeks), 11% (4 weeks), and 16% (8 weeks). Table 3 shows the optimal cutoff values to discriminate between successfully achieving versus not achieving clinically relevant WL (≥3%, ≥5%, and ≥10%) at 6, 12, and 24 months based on initial (2-, 4-, and 8-week) WL. For example, selecting the threshold of 2.7% WL in the initial 2 weeks provided the best compromise between sensitivity (true positive rate) and specificity (true negative rate) of correctly identifying WL of ≥10%/<10% at 6 months. Because the area under the curve estimates were <0.7 in several cases (all receiver operating characteristic curves are provided in Supporting Information Table S1), the presented cut points should be interpreted with caution.

Association between self-weighing and adherence to the predicted WL trajectory and medium-to-long-term WL

Self-weighing and adherence to the predicted WL trajectory were associated with medium-to-long-term WL (Table 2). At 6 months, self-weighing explained 10% (2 weeks), 14% (4 weeks), and 19% (8 weeks) of the variance in WC. At 12 months, these values were 9% (2 weeks), 11% (4 weeks), and 14% (8 weeks) and at 24 months were 6% (2 weeks), 7% (4 weeks), and 8% (8 weeks). At 6 months, adherence to the predicted WL trajectory explained 10% (2 weeks), 15% (4 weeks), and 26% (8 weeks) of the variance in WC. At 12 months, these values were 10% (2 weeks), 13% (4 weeks), and 19% (8 weeks) and at 24 months were 7% (2 weeks), 9% (4 weeks), and 11% (8 weeks).

TABLE 3 Cutoff values to discriminate between successfully achieving versus not achieving clinically significant weight loss (≥3%, ≥5%, and ≥10%) at 6, 12, and 24 months based on initial (2-, 4-, and 8-week) weight loss

	Weight loss		
	≥3%[a]	≥5%[a]	≥10%[a]
At 6 months			
Percentage weight loss in initial 2 weeks	2.5	2.7	2.7
AUC (95% CI)	0.70 (0.64–0.76)	0.68 (0.63–0.73)	0.59 (0.54–0.64)
Sensitivity (95% CI)	0.68 (0.63–0.73)	0.65 (0.60–0.71)	0.65 (0.59–0.72)
Specificity (95% CI)	0.69 (0.59–0.78)	0.69 (0.61–0.76)	0.65 (0.59–0.72)
Percentage weight loss in initial 4 weeks	3.4	4.0	4.0
AUC (95% CI)	0.77 (0.72–0.82)	0.75 (0.71–0.80)	0.65 (0.59–0.70)
Sensitivity (95% CI)	0.71 (0.66–0.76)	0.65 (0.59–0.70)	0.65 (0.58–0.72)
Specificity (95% CI)	0.75 (0.67–0.84)	0.80 (0.74–0.87)	0.61 (0.55–0.67)
Percentage weight loss in initial 8 weeks	4.4	5.3	6.0
AUC (95% CI)	0.81 (0.77–0.86)	0.79 (0.75–0.83)	0.71 (0.66–0.76)
Sensitivity (95% CI)	0.73 (0.68–0.77)	0.67 (0.62–0.73)	0.64 (0.57–0.71)
Specificity (95% CI)	0.80 (0.71–0.88)	0.81 (0.75–0.87)	0.70 (0.65–0.76)
At 12 months			
Percentage weight loss in initial 2 weeks	2.6	2.7	2.7
AUC (95% CI)	0.66 (0.60–0.72)	0.63 (0.57–0.68)	0.54 (0.48 0.59)[b]
Sensitivity (95% CI)	0.62 (0.57–0.68)	0.63 (0.57–0.68)	0.60 (0.53–0.67)
Specificity (95% CI)	0.65 (0.56–0.74)	0.63 (0.55–0.70)	0.50 (0.44–0.56)
Percentage weight loss in initial 4 weeks	3.4	4.0	4.0
AUC (95% CI)	0.72 (0.66–0.77)	0.68 (0.63–0.73)	0.57 (0.52–0.62)
Sensitivity (95% CI)	0.69 (0.64–0.74)	0.61 (0.55–0.66)	0.57 (0.50–0.65)
Specificity (95% CI)	0.67 (0.58–0.76)	0.70 (0.63–0.77)	0.55 (0.49–0.61)
Percentage weight loss in initial 8 weeks	4.8	4.9	6.6
AUC (95% CI)	0.75 (0.70–0.80)	0.73 (0.68–0.78)	0.63 (0.58–0.68)
Sensitivity (95% CI)	0.68 (0.63–0.73)	0.70 (0.65–0.75)	0.52 (0.45–0.59)
Specificity (95% CI)	0.71 (0.63–0.80)	0.67 (0.60–0.75)	0.73 (0.67–0.78)
At 24 months			
Percentage weight loss in initial 2 weeks	2.7	2.7	2.7
AUC (95% CI)	0.64 (0.58–0.69)	0.60 (0.54–0.65)	0.55 (0.49–0.60)[b]
Sensitivity (95% CI)	0.63 (0.57–0.68)	0.63 (0.57–0.69)	0.58 (0.51–0.66)
Specificity (95% CI)	0.63 (0.55–0.71)	0.57 (0.50–0.63)	0.50 (0.44–0.56)
Percentage weight loss in initial 4 weeks	3.4	4.2	4.5
AUC (95% CI)	0.64 (0.59–0.70)	0.62 (0.57–0.67)	0.54 (0.48–0.59)[b]
Sensitivity (95% CI)	0.70 (0.65–0.75)	0.55 (0.49–0.61)	0.45 (0.37–0.53)
Specificity (95% CI)	0.55 (0.47–0.63)	0.64 (0.57–0.70)	0.63 (0.58–0.69)
Percentage weight loss in initial 8 weeks	5.4	5.4	5.3
AUC (95% CI)	0.68 (0.63–0.73)	0.66 (0.61–0.71)	0.56 (0.51–0.62)
Sensitivity (95% CI)	0.61 (0.55–0.66)	0.61 (0.55–0.67)	0.56 (0.49–0.64)
Specificity (95% CI)	0.68 (0.60–0.75)	0.62 (0.55–0.68)	0.51 (0.45–0.57)

Abbreviation: AUC, area under the curve.

[a]Cutoffs determined via receiver operating characteristic (ROC) curves from logistic regression models. The point, which minimizes the Euclidean distance of sensitivity (true positive) and specificity (true negative) from perfection, was identified as the optimal cutoff. AUC estimates including 95% CI are provided for each ROC curve.

[b]Likely nondiscriminant.

TABLE 4 Multivariable associations between initial weight loss, self-weighing, and adherence to the predicted weight loss trajectory and medium-to-long-term weight change

	Predictor time	Predictor	B	SE	*df*	*p*	R^2
Percentage weight change at 6 months	2 weeks	Percentage weight change	1.05	0.16	378	**<0.01**	0.17
	2 weeks	Weights recorded (*n*)	−0.48	0.14	378	**<0.01**	
Percentage weight change at 12 months	2 weeks	Percentage weight change	0.92	0.21	360	**<0.01**	0.13
	2 weeks	Adherent weights (*n*)	−0.68	0.20	360	**<0.01**	
Percentage weight change at 24 months	2 weeks	Percentage weight change	0.93	0.23	357	**<0.01**	0.09
	2 weeks	Adherent weights (*n*)	−0.45	0.22	357	**0.04**	
Percentage weight change at 6 months	4 weeks	Percentage weight change	1.38	0.13	377	**<0.01**	0.32
	4 weeks	Weights recorded (*n*)	−0.53	0.18	377	**<0.01**	
	4 weeks	Adherent weights (*n*)	0.28	0.17	377	0.11	
Percentage weight change at 12 months	4 weeks	Percentage weight change	1.28	0.16	360	**<0.01**	0.22
	4 weeks	Weights recorded (*n*)	−0.31	0.09	360	**<0.01**	
Percentage weight change at 24 months	4 weeks	Percentage weight change	0.93	0.20	357	**<0.01**	0.13
	4 weeks	Adherent weights (*n*)	−0.24	0.10	357	**0.01**	
Percentage weight change at 6 months	8 weeks	Percentage weight change	1.11	0.07	378	**<0.01**	0.53
	8 weeks	Adherent weights (*n*)	−0.09	0.02	378	**<0.01**	
Percentage weight change at 12 months	8 weeks	Percentage weight change	1.06	0.11	360	**<0.01**	0.35
	8 weeks	Adherent weights (*n*)	−0.12	0.03	360	**<0.01**	
Percentage weight change at 24 months	8 weeks	Percentage weight change	0.79	0.14	357	**<0.01**	0.17
	8 weeks	Adherent weights (*n*)	−0.08	0.04	357	**0.02**	

Note: The optimal linear mixed-effects models presented were selected based on the lowest Akaike information criterion (not reported). Bold font indicates statistical significance ($p < 0.05$).
Abbreviations: B, unstandardized regression coefficient; *df*, degrees of freedom.

Multivariable associations between initial WL, self-weighing, and adherence to the predicted WL trajectory and medium-to-long-term WL

The optimal models selected and their coefficients at each time point are reported in Table 4. WL was computed as follow-up minus baseline, and consequently positive values represent weight gain. Positive values in the percentage WC coefficients in the initial weeks suggest that greater initial percentage WC predicts less follow-up percentage weight gain. Negative values in the coefficients of the number of weights recorded/adherent weights in initial weeks suggest that a greater number of weights recorded/adherent weights in initial weeks predicts less follow-up percentage weight gain.

All optimal models showed that more than one predictor is important. Among them, percentage WC during the initial 2, 4, and 8 weeks was a predictor in all models (β ranging from 0.79 to 1.38, all $p < 0.01$; Table 4). The number of weights recorded and the number of adherent weights were mostly not selected in the same model because of correlation between the two. The number of weights recorded in the initial 2 ($\beta = -0.48$, $p < 0.01$) and 4 ($\beta = -0.53$, $p < 0.01$) weeks was a predictor for percentage WC at 6 months. The number of weights recorded in the initial 4 ($\beta = -0.31$, $p < 0.01$) weeks was also selected in the model for 12 months. At all other time points, the number of adherent weights in the initial weeks (rather than the number of weights recorded) was a better predictor. Notably, for WL at 6 months, both the number of weights recorded and the number of adherent weights in the initial 4 weeks were selected.

At 6 months, the best model explained 17% (2 weeks), 32% (4 weeks), and 53% (8 weeks) of the variance in WC. At 12 months, these values were 13% (2 weeks), 22% (4 weeks), and 35% (8 weeks). At 24 months, the best model explained 9% (2 weeks), 13% (4 weeks), and 17% (8 weeks) of the variance in WC. The comparison of Akaike information criterion (not reported) across the best models with the same dependent variable suggests that initial predictors at week 8 were stronger than those at week 2 and 4.

DISCUSSION

The present analysis demonstrated that initial WL, self-weighing adherence, and adherence to the expected WL trajectory were associated with medium-to-long-term WL during a comprehensive ILI delivered to patients in primary care. Specifically, and in line with our hypothesis and previous findings [10, 16, 30, 31], greater initial WL predicted WL at 6, 12, and 24 months. Furthermore, greater initial daily self-weighing adherence and adherence to the predicted WL trajectory were also associated with longer-term WL, highlighting the importance of getting participants off to a good start in ILIs by ensuring early program (diet, physical activity, calorie restriction) and

behavioral (self-weighing) adherence. However, in multivariable models, the best predictor remained initial WL.

As hypothesized, a longer time frame of initial WL and adherence (i.e., 8 weeks) yielded stronger predictions of medium-to-long-term WL compared with shorter time frames (2 weeks and 4 weeks). Of note, the predictive value of the initial 2-, 4-, and 8-week WL was generally highest for the 6-month WL and it decreased for the 12-month and particularly the 24-month WL. The improved prediction with later early time (8 weeks) and more proximal later time (6 months) is likely because predicting the future is difficult and, therefore, the shortest period is likely to be the most accurate.

The significant associations of 2-week WL (albeit attenuated compared with 4- and 8-week WL, in line with previous findings) [32, 33] with WL at 6, 12, and 24 months may be particularly relevant, as they demonstrated that individuals undergoing primary care-based ILIs may be identified very early on as needing additional support to increase longer-term WL. Contrary to previous findings showing that 2-week WL predicted only short-term (6 months) but not later (1 year or 2 years) WL and 4-week WL predicted WL at 6 months and 1 year but not 2 years [32], our study demonstrated that 2-week and 4-week WL was predictive of WL at 6, 12, and 24 months. Consequently, our results indicate that WL may be evaluated as early as 2 weeks to determine whether treatment needs to be altered to improve medium-to-long-term WL. However, because 4-week and particularly 8-week WL was a stronger predictor of medium-to-long-term WL, WL progression should be reevaluated at these time points.

Importantly, our results also show that greater initial daily self-weighing adherence and adherence to the individualized weight graph are associated with medium-to-long-term WL, showing that even after 2 weeks of an ILI program, lower levels of program (calorie restriction) adherence, displayed as poorer adherence to the expected WL trajectory, can already have an impact on longer-term WL. These findings demonstrated the benefits of the weight graph approach as a method to promote early program (calorie restriction) adherence and improve medium-to-long-term WL. The visualization of daily weights on the weight graph can help the interventionist detect deviations from the projected WL trajectory quickly (i.e., within days) and get the patient back on track by deploying targeted countermeasures according to the individual progress and preferences (toolbox approach) [22, 24, 25]. In patients with low initial WL (and poor adherence to the expected WL trajectory), other treatment options ("rescue efforts") such as more digital messages, different diets, intensifying the calorie restriction, meal planning or meal replacements, additional physical activity, or pharmacotherapy as add-ons may be considered [34]. These strategies can help reduce participant frustration (and attrition) that is likely to occur from absent WL and increase interest in continued weight management efforts. Importantly, through the weight graph, patients can detect deviations from the expected WL trajectory and even independently of the interventionists' feedback [26, 27, 35]. This can help facilitate self-efficacy and effective problem solving as the patient can more easily link his/her behavior to the outcome of interest (i.e., WL). When adherent to the calorie restriction goals, the patient receives daily confirmation of the WL progress, as visualized on the weight graph, which can act as a great motivator and facilitate continued intervention adherence [36]. Ultimately, the weight graph approach can also increase the cost-effectiveness of ILIs: less interventionist-patient interaction may be required, especially in the longer term [26], which may be of particular interest in primary care.

In the PROPEL trial, meal replacements proved to be a particularly effective strategy to drive initial adherence in the trial's underserved, low-income population. In fact, as reported by the health coaches in informal conversations, participants asked for continued provision of the meal-replacement shakes that were originally intended only for the first 4 weeks [22], and they were provided them as long as supplies lasted. The popularity of the meal replacements was somewhat unexpected, but it demonstrated that the very limited grocery budget in individuals with low income (and food insecurity), such as the PROPEL sample, played an important role in the lack of access to WL-supporting foods, simply because these foods (e.g., meal-replacement shakes) were expensive enough that they could not be regularly incorporated into the tight grocery budget. Aside from the benefits for WL, the meal replacements may have reduced some of the general food insecurity and led to a less limited (grocery) budget, which likely contributed to their popularity, something that should be considered for clinical practice.

More studies are needed to test the aforementioned rescue efforts for improving long-term WL in those with poor initial WL, especially in racially diverse, low-income populations with obesity. Continued provision of meal replacements may be an effective strategy in these populations that should be tested. Future studies should also determine whether other differences (e.g., genetics, personal health beliefs, cultural factors) exist between those who lose more versus less weight initially and how interventionists can tailor ILIs in a primary-care setting to address those differences.

Strengths of this study include the large and diverse sample of typically underserved individuals from urban and rural regions of Louisiana, who generally are underrepresented in clinical research. Consequently, the present results apply to the large underserved, low-income population in the United States. Further strengths include the objective measures of daily self-weighing and the weight graph approach that illustrate a potential model for WL programs in primary care. A limitation of this study is that the PROPEL sample was mostly women, which limits the generalizability of the present results for both sexes. However, this is not uncommon in lifestyle interventions [37].

CONCLUSION

In this 2-year lifestyle intervention delivered to underserved patients with obesity in primary care, greater WL during the initial 2, 4, and 8 weeks predicted greater WL at 6, 12, and 24 months. Self-weighing and adherence to the expected WL trajectory in the initial weeks marginally improved WL prediction beyond initial WL alone. Those administering WL interventions in primary care should ensure participants' early adherence to daily self-weighing and the predicted WL trajectory, particularly when initial WL is suboptimal.O

ACKNOWLEDGMENTS

The authors acknowledge the patients and stakeholders who are members of the PROPEL Patient Advisory Boards, Community Monitoring Board, and Project Management Committee (Chris Lodge and Ava Zebrick, MSHCM), who significantly affected the trial design and conduct. The authors are also indebted to the PROPEL patients, assessment technicians, and health coaches, without whom this study would not have been possible. The authors gratefully acknowledge the contributions of Willie C. White III, MPH, and the David Raines Community Health Centers; Dr. Gary Wiltz and the Teche Action Clinic sites; Michael G. Griffin, Dr. Robert Post, and the Daughters of Charity Services of New Orleans; and the Ochsner Health System and Access Health Louisiana clinic sites. The authors also thank the Data and Safety Monitoring Board for monitoring patient safety and the overall conduct of the trial: Robert Ross, PhD (chair), John Lefante, PhD, Michael Rolfsen, MD, and Chris Lodge. Study data were collected and managed with REDCap (Research Electronic Data Capture) tools hosted at Pennington Biomedical Research Center. Access to individual de-identified data that underlie the results reported in this article is possible via Pennington Biomedical Research Center's Data Sharing policy.

The PROPEL Research Group includes: Pennington Biomedical Research Center: Peter T. Katzmarzyk, PhD (principal investigator), Robert L. Newton Jr, PhD (outcomes assessment director), Corby K. Martin, PhD (intervention director), John W. Apolzan, PhD (intervention codirector), William Johnson, PhD (biostatistician), Kara D. Denstel, MPH (project manager), Emily F. Mire, MS (data manager), Robert K. Singletary Jr, MHS, Cheryl Lewis, MPH, Phillip Brantley, PhD, Ronald Horswell, PhD, Betty Kennedy, PhD, Dachuan Zhang, MAppStats, Stephanie Authement, RD, LDN, MS, Shiquita Brooks, RDN, LDN, Danielle S. Burrell, MEd, MCHES, Leslie Forest-Everage, MA, Angelle Graham Ullmer, RDN, LDN, MS, Laurie Murphy, RDN, LDN, Cristalyn Reynolds, MA, Kevin Sanders, MS, RDN, LDN, Stephen Bower, MS, Daishaun Gabriel, MHA, Hillary Gahagan, MPH, Tabitha K. Gray, MA, Jill Hancock, MPH, Marsha Herrera, Brittany Molinere, Georgia Morgan, MA, Brittany Neyland, Stephanie Rincones, Deanna Robertson, MA, Ekambi Shelton, MPH, Russell J. Tassin, MS, Kaili Williams; Louisiana State University Health Sciences Center at New Orleans: Benjamin F. Springgate, MD; Louisiana State University Health Sciences Center at Shreveport: Terry C. Davis, PhD, Connie L. Arnold, PhD; Ochsner Health System: Eboni Price-Haywood, MD, Carl J. Lavie, MD, Jewel Harden-Barrios, MEd; Tulane University Medical School: Vivian A. Fonseca, MD, Tina K. Thethi, MD (MEDICAL Monitor), Jonathan Gugel, MD; Xavier University: Kathleen B. Kennedy, PhD, Daniel F. Sarpong, PhD, Amina D. Massey.

FUNDING INFORMATION

The research reported in this article was funded through a Patient-Centered Outcomes Research Institute (PCORI) award (OB-1402-10977). The statements in this article are solely the responsibility of the authors and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute, its Board of Governors, or its Methodology Committee. Additional support was provided by U54 GM104940 from the National Institute of General Medical Sciences of the National Institutes of Health, which funds the Louisiana Clinical and Translational Science Center, and Nutrition Obesity Research Centers grant P30 DK072476 titled "Nutrition and Metabolic Health Through the Lifespan" sponsored by National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). The authors also thank Health and Nutrition Technology (Carmel, California) for providing the HealthOne formula and Nutrisystem (Fort Washington, Pennsylvania) for providing the meal replacements used in the study. Christoph Höchsmann was supported by an NIDDK National Research Service Award (T32DK064584).

CONFLICT OF INTEREST STATEMENT

Louisiana State University, Pennington Biomedical Research Center, and Montclair State University have interest in and patent applications related to the intellectual property surrounding the weight graph that was used in the intervention, and Corby K. Martin, among others, is an inventor of the technology. Licensing of that technology results in financial benefits to Louisiana State University, Pennington Biomedical Research Center, Montclair State University, and the inventors. Corby K. Martin further reports grants from WW International, Inc., the Academy of Nutrition and Dietetics, ABGIL, Eli Lilly and Co., and Evidation, as well as personal fees from Wondr Health, the Academy of Nutrition and Dietetics, and ABGIL. John W. Apolzan reports grants from WW International, Inc., and non-financial support from Health and Nutrition Technology and Nutrisystem. Connie L. Arnold reports a subcontract with Pennington Biomedical Research Center during the conduct of the study. Vivian Fonseca reports grants from Bayer, Boehringer Ingelheim, and Gilead; personal fees from Takeda, Novo Nordisk A/S, Sanofi-Aventis, Eli Lilly and Co., Abbott, AstraZeneca, and Intarcia; and stock options from Microbiome Technologies, Insulin Algorithms, BRAVO4Health, and Amgen. Tina K. Thethi reports personal fees from Novo Nordisk A/S. Carl J. Lavie reports personal fees from Amgen, AstraZeneca, Esperion, Sanofi, Regeneron, and PAI Health. Peter T. Katzmarzyk reports grants from Patient-Centered Outcomes Research Institute and non-financial support from Health and Nutrition Technology and Nutrisystem. All other authors report no conflicts of interest. The work related to this article occurred before Dr. Price-Haywood's appointment to the Patient-Centered Outcomes Research Institute Board of Governors.

CLINICAL TRIAL REGISTRATION

ClinicalTrials.gov identifier NCT02561221.

ORCID

Christoph Höchsmann https://orcid.org/0000-0003-2007-3007
Corby K. Martin https://orcid.org/0000-0002-8125-4015
John W. Apolzan https://orcid.org/0000-0001-8492-7820
James L. Dorling https://orcid.org/0000-0002-2586-3983
Robert L. Newton Jr https://orcid.org/0000-0002-6863-913X
Peter T. Katzmarzyk https://orcid.org/0000-0002-9280-6022

REFERENCES

1. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity and severe obesity among adults: United States, 2017–2018. *NCHS Data Brief*, no. 360. National Center for Health Statistics; 2020.

2. Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. *BMC Public Health*. 2009;9:88. doi:10.1186/1471-2458-9-88
3. Ogden CL, Fakhouri TH, Carroll MD, et al. Prevalence of obesity among adults, by household income and education – United States, 2011–2014. *MMWR Morb Mortal Wkly Rep*. 2017;66:1369-1373.
4. Hruby A, Hu FB. The epidemiology of obesity: a big picture. *Pharmacoeconomics*. 2015;33:673-689.
5. Jensen MD, Ryan DH, Apovian CM, et al. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Obesity Society. *Circulation*. 2014;129(25 suppl 2):S102-S138.
6. Wadden TA, West DS, Neiberg R, et al; Look AHEAD Research Group. One-year weight losses in the Look AHEAD study: factors associated with success. *Obesity*. 2009;17:713-722.
7. Wadden TA, Tronieri JS, Butryn ML. Lifestyle modification approaches for the treatment of obesity in adults. *Am Psychol*. 2020;75:235-251.
8. Chopra S, Malhotra A, Ranjan P, et al. Predictors of successful weight loss outcomes amongst individuals with obesity undergoing lifestyle interventions: a systematic review. *Obes Rev*. 2021;22:e13148. doi:10.1111/obr.13148
9. Kerrigan SG, Clark M, Convertino A, Forman EM, Butryn ML. The association between previous success with weight loss through dietary change and success in a lifestyle modification program. *J Behav Med*. 2018;41:152-159.
10. Ortner Hadžiabdić M, Mucalo I, Hrabač P, Matić T, Rahelić D, Božikov V. Factors predictive of drop-out and weight loss success in weight management of obese patients. *J Hum Nutr Diet*. 2015;28:24-32.
11. Handjieva-Darlenska T, Holst C, Grau K, et al. Clinical correlates of weight loss and attrition during a 10-week dietary intervention study: results from the NUGENOB project. *Obes Facts*. 2012;5:928-936.
12. Batterham M, Tapsell LC, Charlton KE. Baseline characteristics associated with different BMI trajectories in weight loss trials: a case for better targeting of interventions. *Eur J Clin Nutr*. 2016;70:207-211.
13. Karlsen T-I, Søhagen M, Hjelmesæth J. Predictors of weight loss after an intensive lifestyle intervention program in obese patients: a 1-year prospective cohort study. *Health Qual Life Outcomes*. 2013;11:165. doi:10.1186/1477-7525-11-165
14. Delahanty LM, Peyrot M, Shrader PJ, et al; DPP Research Group. Pretreatment, psychological, and behavioral predictors of weight outcomes among lifestyle intervention participants in the Diabetes Prevention Program (DPP). *Diabetes Care*. 2013;36:34-40.
15. Beavers KM, Neiberg RH, Kritchevsky SB, et al. Association of sex or race with the effect of weight loss on physical function: a secondary analysis of 8 randomized clinical trials. *JAMA Netw Open*. 2020;3:e2014631. doi:10.1001/jamanetworkopen.2020.14631
16. Unick JL, Neiberg RH, Hogan PE, et al; Look AHEAD Research Group. Weight change in the first 2 months of a lifestyle intervention predicts weight changes 8 years later. *Obesity*. 2015;23:1353-1356.
17. Williamson DA, Anton SD, Han H, et al. Early behavioral adherence predicts short and long-term weight loss in the POUNDS LOST study. *J Behav Med*. 2010;33:305-314.
18. Tanenbaum ML, Ross KM, Wing RR. Overeat today, skip the scale tomorrow: an examination of caloric intake predicting non-adherence to daily self-weighing. *Obesity*. 2016;24:2341-2343.
19. Zheng Y, Burke LE, Danford CA, Ewing LJ, Terry MA, Sereika SM. Patterns of self-weighing behavior and weight change in a weight loss trial. *Int J Obes*. 2016;40:1392-1396.
20. Katzmarzyk PT, Martin CK, Newton RL, et al. Weight loss in underserved patients – a cluster-randomized trial. *N Engl J Med*. 2020;383:909-918.
21. Höchsmann C, Dorling JL, Martin CK, et al. Effects of a 2-year primary care lifestyle intervention on cardiometabolic risk factors: a cluster-randomized trial. *Circulation*. 2021;143:1202-1214.
22. Katzmarzyk PT, Martin CK, Newton RL, et al. Promoting Successful Weight Loss in Primary Care in Louisiana (PROPEL): rationale, design and baseline characteristics. *Contemp Clin Trials*. 2018;67:1-10.
23. Diabetes Prevention Program (DPP) Research Group. The Diabetes Prevention Program (DPP): description of lifestyle intervention. *Diabetes Care*. 2002;25:2165-2171.
24. Look AHEAD Research Group; Wadden TA, West DS, Delahanty L, et al. The Look AHEAD study: a description of the lifestyle intervention and the evidence supporting it. *Obesity*. 2006;14:737-752.
25. Rickman AD, Williamson DA, Martin CK, et al. The CALERIE study: design and methods of an innovative 25% caloric restriction intervention. *Contemp Clin Trials*. 2011;32:874-881.
26. Martin CK, Miller AC, Thomas DM, Champagne CM, Han H, Church T. Efficacy of SmartLoss^{SM}, a smartphone-based weight loss intervention: results from a randomized controlled trial. *Obesity*. 2015;23:935-942.
27. Martin CK, Gilmore LA, Apolzan JW, Myers CA, Thomas DM, Redman LM. Smartloss: a personalized mobile health intervention for weight management and health promotion. *JMIR Mhealth Uhealth*. 2016;4:e18. doi:10.2196/mhealth.5027
28. Nakagawa S, Schielzeth H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol Evol*. 2013;4:133-142.
29. Ryan DH, Yockey SR. Weight loss and improvement in comorbidity: differences at 5%, 10%, 15%, and over. *Curr Obes Rep*. 2017;6:187-194.
30. Fabricatore AN, Wadden TA, Moore RH, Butryn ML, Heymsfield SB, Nguyen AM. Predictors of attrition and weight loss success: results from a randomized controlled trial. *Behav Res Ther*. 2009;47:685-691.
31. Elfhag K, Rössner S. Initial weight loss is the best predictor for success in obesity treatment and sociodemographic liabilities increase risk for drop-out. *Patient Educ Couns*. 2010;79:361-366.
32. Garvin JT, Hardy D, Xu H. Initial response to program, program participation, and weight reduction among 375 MOVE! participants, Augusta, Georgia, 2008–2010. *Prev Chronic Dis*. 2016;13:E55. doi:10.5888/pcd13.150598
33. Miller CK, Nagaraja HN, Weinhold KR. Early weight-loss success identifies nonresponders after a lifestyle intervention in a worksite diabetes prevention trial. *J Acad Nutr Diet*. 2015;115:1464-1471.
34. Unick JL, Pellegrini CA, Demos KE, Dorfman L. Initial weight loss response as an indicator for providing early rescue efforts to improve long-term treatment outcomes. *Curr Diab Rep*. 2017;17:69. doi:10.1007/s11892-017-0904-1
35. Wing RR, Tate D, LaRose JG, et al. Frequent self-weighing as part of a constellation of healthy weight control practices in young adults. *Obesity*. 2015;23:943-949.
36. Tronieri JS, Wadden TA, Walsh O, Berkowitz RI, Alamuddin N, Chao AM. Measures of adherence as predictors of early and total weight loss with intensive behavioral therapy for obesity combined with liraglutide 3.0 mg. *Behav Res Ther*. 2020;131:103639. doi:10.1016/j.brat.2020.103639
37. Pagoto SL, Schneider KL, Oleski JL, Luciani JM, Bodenlos JS, Whited MC. Male inclusion in randomized controlled trials of lifestyle weight loss interventions. *Obesity*. 2012;20:1234-1239.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Höchsmann C, Martin CK, Apolzan JW, et al. Initial weight loss and early intervention adherence predict long-term weight loss during the Promoting Successful Weight Loss in Primary Care in Louisiana lifestyle intervention. *Obesity (Silver Spring)*. 2023;31(9):2272-2282. doi:10.1002/oby.23854

Publication 6

Mediators of weight change in underserved patients with obesity: exploratory analyses from the PROPEL cluster-randomized trial.

James L Dorling [1,2], Corby K Martin [2], Qingzhao Yu [3], Wentao Cao [3], **Christoph Höchsmann** [2,4], John W Apolzan [2], Robert L Newton Jr [2], Kara D Denstel [2], Emily F Mire [2], Peter T Katzmarzyk [2] for the PROPEL Research Group

[1] Human Nutrition, School of Medicine, Dentistry and Nursing, University of Glasgow, Glasgow, UK

[2] Pennington Biomedical Research Center, Baton Rouge, LA, USA

[3] School of Public Health, Louisiana State University Health Sciences Center, New Orleans, LA, USA

[4] Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
American Journal of Clinical Nutrition, 116(4): 1112–1122	2022	10.1093/ajcn/nqac179	7.1 (2022)	Q1	Scopus: 0 Google Scholar: 2

Author contribution:

Co-author; co-developed the research question and statistical model, contributed to interpreting the results, and provided critical revisions for important intellectual content.

Mediators of weight change in underserved patients with obesity: exploratory analyses from the Promoting Successful Weight Loss in Primary Care in Louisiana (PROPEL) cluster-randomized trial

James L Dorling,[1,2] Corby K Martin,[2] Qingzhao Yu,[3] Wentao Cao,[3] Christoph Höchsmann,[2,4] John W Apolzan,[2] Robert L Newton, Jr,[2] Kara D Denstel,[2] Emily F Mire,[2] and Peter T Katzmarzyk[2]

[1]Human Nutrition, School of Medicine, Dentistry and Nursing, College of Medical, Veterinary & Life Sciences, University of Glasgow, Glasgow, United Kingdom; [2]Pennington Biomedical Research Center, Baton Rouge, LA, USA; [3]School of Public Health, Louisiana State University Health Sciences Center, New Orleans, LA, USA; and [4]Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany

ABSTRACT

Background: Intensive lifestyle interventions (ILIs) stimulate weight loss in underserved patients with obesity, but the mediators of weight change are unknown.

Objectives: We aimed to identify the mediators of weight change during an ILI compared with usual care (UC) in underserved patients with obesity.

Methods: The PROPEL (Promoting Successful Weight Loss in Primary Care in Louisiana) trial randomly assigned 18 clinics ($n = 803$) to either an ILI or UC for 24 mo. The ILI group received an intensive lifestyle program; the UC group had routine care. Body weight was measured; further, eating behaviors (restraint, disinhibition), dietary intake (percentage fat intake, fruit and vegetable intake), physical activity, and weight- and health-related quality of life constructs were measured through questionnaires. Mediation analyses assessed whether questionnaire variables explained between-group variations in weight change during 2 periods: baseline to month 12 ($n = 779$) and month 12 to month 24 ($n = 767$).

Results: The ILI induced greater weight loss at month 12 compared with UC (between-group difference: −7.19 kg; 95% CI: −8.43, −6.07 kg). Improvements in disinhibition (−0.33 kg; 95% CI: −0.55, −0.10 kg), percentage fat intake (−0.25 kg; 95% CI: −0.50, −0.01 kg), physical activity (−0.26 kg; 95% CI: −0.41, −0.09 kg), and subjective fatigue (−0.28 kg; 95% CI: −0.46, −0.10 kg) at month 6 during the ILI partially explained this between-group difference. Greater weight loss occurred in the ILI at month 24, yet the ILI group gained 2.24 kg (95% CI: 1.32, 3.26 kg) compared with UC from month 12 to month 24. Change in fruit and vegetable intake (0.13 kg; 95% CI: 0.05, 0.21 kg) partially explained this response, and no variables attenuated the weight regain of the ILI group.

Conclusions: In an underserved sample, weight change induced by an ILI compared with UC was mediated by several psychological and behavioral variables. These findings could help refine weight management regimens in underserved patients with obesity. This trial was registered at clinicaltrials.gov as NCT02561221. *Am J Clin Nutr* 2022;116:1112–1122.

Keywords: comprehensive lifestyle intervention, diet, eating attitudes, health disparities, minority groups, primary health care, weight loss, weight regain

Supported by Patient-Centered Outcomes Research Institute (PCORI) award OB-1402-10977 (to PTK). The statements in this article are wholly the responsibility of the authors and do not necessarily represent the views of the PCORI, its Board of Governors, or its Methodology Committee. Supported in addition by grant U54 GM104940 from the National Institute of General Medical Sciences of the National Institutes of Health (NIH), which funds the Louisiana Clinical and Translational Science Center, and by grant P30DK072476 ("Nutrition and Metabolic Health through the Lifespan") from the Nutrition and Obesity Research Center, sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases. JLD was supported by American Heart Association grant 20POST35210907. CH was supported by NIH National Research Service award T32 DK064584.

The intellectual property surrounding the mathematical code that creates the weight graph used in this study is owned by Louisiana State University/Pennington Biomedical and Montclair State University. CKM is an inventor of the intellectual property, which is included in a US and European patent application. The code has also been licensed and Louisiana State University/Pennington Biomedical, Montclair State University, and CKM have received royalties.

Supplemental Tables 1 and 2 are available from the "Supplementary data" link in the online posting of the article and from the same link in the online table of contents at https://academic.oup.com/ajcn/.

Address correspondence to PTK (e-mail: peter.katzmarzyk@pbrc.edu).

Abbreviations used: EI, Eating Inventory; ES, effect size; ILI, intensive lifestyle intervention; IPAQ, International Physical Activity Questionnaire; IWQOL, Impact of Weight on Quality of Life-Lite; MET, metabolic equivalent of task; NCI, National Cancer Institute; PROMIS-29, Patient-Reported Outcomes Measurement Information System-29; PROPEL, Promoting Successful Weight Loss in Primary Care in Louisiana; UC, usual care.

Received March 15, 2022. Accepted for publication June 21, 2022.

First published online June 28, 2022; doi: https://doi.org/10.1093/ajcn/nqac179.

Introduction

Obesity is a public health disease that increases the risk of type 2 diabetes, cardiovascular disease, cancer, and premature death (1, 2). Overall, obesity affects ∼40% of adults in the United States (3), and health disparities are present. Obesity is more prevalent in certain demographic groups with a low annual income (4). Moreover, compared with non-Hispanic white adults, black and Hispanic populations exhibit higher rates of obesity (5). It is thus important to identify effective weight-management methods for individuals with obesity in these populations to attain national health targets and decrease health disparities.

Usual care (UC) for weight loss and weight management within primary care typically involves behavioral counseling and therapy to improve dietary habits and physical activity, yet such regimens often yield substandard weight loss because of time constraints and a lack of training among practitioners (6). Intensive lifestyle interventions (ILIs) are recommended as alternative programs for weight loss in individuals with obesity in primary care (7). These aim to stimulate energy deficits and weight loss through reduced-calorie diets, improvements in physical activity, and behavioral therapy in an on-site and intensive (≥14 sessions in the first 6 mo) regimen delivered by trained interventionists (7, 8).

In the PROPEL (Promoting Successful Weight Loss in Primary Care in Louisiana) trial, we demonstrated that underserved patients with obesity lose more weight and improve cardiometabolic risk markers during an ILI compared with UC over 24 mo (9, 10). However, it is unclear what factors drove the increased weight loss produced by the ILI relative to UC. It is in addition not known if the factors associated with midterm (6–12 mo) weight loss during the ILI were effective at attenuating weight regain, which is common and can decrease the health benefits associated with lifestyle interventions (11). These factors could include those that have been associated with weight loss and were linked to behaviors and strategies covered in counseling sessions of the ILI, such as increased dietary restraint (i.e., the intent and ability to restrict food intake), reduced dietary disinhibition (i.e., the tendency to overeat) (12), increased intake of healthy foods with low fat (13), increased physical activity (14), and improved quality of life (15). Identifying the factors that mediated weight loss and weight-loss maintenance during the PROPEL trial is important because strategies and behaviors can be targeted and tested in future interventions, enhancing the efficacy of weight-management programs that are delivered to underserved individuals with obesity in primary care.

The aim of this exploratory investigation was to use mediation analyses to identify the mediators of weight change during an ILI compared with UC in underserved patients with obesity. We hypothesized that improvements in eating behaviors (increased dietary restraint and reduced dietary disinhibition), dietary intake (reductions in dietary fat and increases in fruit and vegetable intake), physical activity, and quality of life shown in the ILI compared with UC would mediate improved weight change during the ILI.

Methods

Patients

Primary inclusion criteria for PROPEL (NCT02561221) included an age of 20–75 y, a BMI (in kg/m^2) of 30.0–50.0, and being a patient at a participating primary care clinic. Patients were excluded if they used weight-loss medication, were presently partaking in a structured weight-loss program, previously had bariatric surgery or planned to have bariatric surgery within 2 y, or had lost >10 lbs (4.5 kg) in the last 6 mo. A full list of inclusion and exclusion criteria has been previously published (9, 16), and all these criteria applied to these analyses.

Study design

The PROPEL study was a cluster-randomized trial consisting of 18 primary care clinics from 5 health systems across Louisiana. Details of the trial's design, randomization and recruitment methods, and protocol have been published (9, 16). The Pennington Biomedical Research Center Institutional Review Board approved the study. All procedures followed the ethical standards set by this Institutional Review Board, and all patients provided written informed consent. A self-report demographic questionnaire was used to obtain information about sex, race, and income.

Clinics were randomly assigned in a 1:1 allocation ratio to provide patients with an ILI or UC for 24 mo. Randomization was stratified by health system, with the random allocation method generated by a study statistician. Patients were not blinded to their group assignment because randomization occurred at the clinic level and the interventions are distinct. Efforts were nonetheless made to blind staff involved in data collection to the clinic randomization, and intervention staff were blinded to the patient's official study measures. The PROPEL trial data were collected and managed via the use of Research Electronic Data Capture (REDCap) resources hosted by the Pennington Biomedical Research Center (9, 17). The trial was conducted between April 2016 and September 2019, finishing when recruited patients who completed the trial had their month 24 assessments (9, 16).

Patients in the ILI received a pragmatic, intensive lifestyle program, which was based on previous lifestyle regimens (18–20) and consistent with the 2013 recommendations for the management of overweight and obesity set out by the American Heart Association, American College of Cardiology, and The Obesity Society (8). The ILI regimen was administered by appropriately trained health coaches embedded within primary care clinics and comprised weekly sessions in the first 6 mo (16 face-to-face and 6 delivered via telephone), followed by sessions that were held at least monthly. The objective for patients in the ILI was to lose 10% of their body weight through numerous strategies which aimed to change eating behaviors and physical activity. Strategies incorporated in the ILI included the provision of suitable prepackaged foods and meal replacements, coaching on appropriate portion sizes, and information on how to purchase and prepare healthy foods. It also included encouragement to increase physical activity to 175 min/wk in line with the physical activity goal of the Look AHEAD (Action for Health in Diabetes) trial (19). In addition to these strategies, a weight-loss calculator was used to formulate personalized energy intake targets and then display predicted weight loss to patients and health coaches (21).

Patients assigned to UC received the care routinely delivered by their clinic for the duration of the trial. They were also provided 6 newsletters that covered numerous topics such as sitting and health, goal setting, memory health, self-care, sleep hygiene, and smoking cessation. Primary care providers in the

UC clinics received information at baseline and annually on the Centers for Medicare and Medicaid Services approach to behavioral therapy for obesity (22).

Measures

Body weight.

Body weight was measured using a digital scale (Seca Model 876) at assessment visits conducted at baseline and at months 6, 12, 18, and 24. Patients were instructed to wear light clothes and no shoes while measurements were conducted. Anthropometric measurements were made in duplicate, although a third measurement was taken if weight differed by 0.5 kg. The mean of the 2 closest measurements was recorded.

Questionnaires.

All questionnaires used in the present analysis were administered at baseline, month 6, month 12, and month 24.

The Eating Inventory (EI) is a 51-item tool that assesses dietary restraint, dietary disinhibition, and hunger (23). However, only restraint and disinhibition were assessed and thus a shortened 37-item EI was provided to PROPEL patients, with items assessing hunger removed (9). Dietary restraint is defined as the intent and ability to restrict food intake; a higher score is generally positive for weight control when disinhibition is low (24). Dietary disinhibition is defined as the tendency to overeat, and a higher value is associated with eating disorder symptoms and poor weight control (25). Greater scores for restraint and disinhibition were indicative of higher levels of the eating behavior assessed.

A customized questionnaire was administered to measure aspects of dietary intake. The questionnaire utilized scales from several sources to measure 3 outcomes: a National Cancer Institute (NCI) fat screener assessed percentage fat intake (26); a 7-item screener devised by the NCI and National 5 A Day Program examined fruit and vegetable consumption (27, 28); and 3 questions from the Brief Questionnaire to Assess Habitual Beverage Intake (BEVQ-15) assessed the frequency of alcohol intake (29).

Weight-related quality of life was measured through the 31-item Impact of Weight on Quality of Life-Lite (IWQOL) questionnaire (30, 31). This measures obesity-related aspects of quality of life, with a total quality of life score and separate scores for physical function, self-esteem, sexual life, public distress, and work or daily activities yielded. Scores are transformed to a 0–100 scale; a score of 100 represents the highest quality of life. The questionnaire asks patients to reflect on quality of life constructs because of their weight (30). Hence, in line with other analyses (15), only the total IWQOL score was utilized in the current analysis to limit the inclusion of variables that may be causally affected by weight change.

The Patient-Reported Outcomes Measurement Information System-29 (PROMIS-29) questionnaire was also administered to measure health-related quality of life (32). This 29-item questionnaire assesses health-related domains related to physical function, anxiety, fatigue, depression, sleep disturbance, ability to partake in social roles and activities, pain interference, and pain intensity. All constructs were used except for pain intensity owing to its relation with pain interference.

The International Physical Activity Questionnaire (IPAQ) short form was used to assess physical activity levels (33). The questionnaire, which asks questions related to physical activity over the previous 7 d, provides physical activity scores in median metabolic equivalent of task (MET)-minutes per week. Four constructs of physical activity were assessed in MET-minutes per week: vigorous, moderate, walking, and total. In the PROPEL trial, numerous patients had missing data for particular activity types (vigorous, moderate, and walking), meaning total MET-minutes per week scores could not be calculated for these patients per standardized scoring methods (33). Thus, in the present analysis, we only included vigorous, moderate, and walking MET-minutes per week variables.

Statistical analysis.

The present article is an exploratory analysis; accordingly, the sample size acquired in the trial was studied. As summary statistics, between-group differences in change scores for questionnaire variables were determined using unadjusted independent-samples *t* tests. Absolute Cohen's *d* effect size (ES) values were also assessed for change scores (34). The magnitude of an ES value was considered trivial (<0.20), small (0.20–0.49), medium (0.50–0.79), or large (≥ 0.80) (34).

Our objective was to identify the mediators of weight change in the ILI compared with UC; in other words, we aimed to test the extent to which a set of variables (mediators) explained weight differences between the ILI group and the UC group. Multilevel mediation analysis was used to measure the effects conveyed by intervening variables (mediators) to the observed relation between an exposure variable and an outcome variable (35, 36). In this analysis, the mediator (change in questionnaire variables) and outcome (weight change) variables were continuous, whereas the exposure variable was binary (ILI or UC group). We built random intercept models to account for the correlation among subjects within the clinic. As part of the analysis, the total effect was estimated at the individual level; that is, the average difference in weight change (outcome variable) caused by the ILI compared with UC (exposure variable). The analysis further separated the total effect of the ILI (compared with UC) on weight change into 2 components: the indirect effects from mediators and the direct effect. The indirect effect is the effect of the ILI (compared with UC) on weight change that is driven by each proposed mediator; the direct effect is the effect of the ILI (compared with UC) on weight change that is not explained by the change in the proposed mediators.

In accord with the aims of the article, 2 conceptual models were used to guide the analyses (**Figure 1**). The first model aimed to determine the mediators of weight change induced by the ILI relative to UC during the first 12 mo of the trial. This was chosen to highlight mediators of midterm weight loss (6 to $\leq$12 mo) (37). The exposure variable was the trial group (ILI or UC), the proposed mediators were change in questionnaire variables from baseline to month 6, and the outcome variable was weight change from baseline to month 12. In the second model, the aim was to assess the mediators of weight change during the second 12 mo of the trial. This was broadly chosen to identify mediators of weight change during periods of weight-loss maintenance. This model had the same exposure variable as model 1, although change in questionnaire variables from baseline to month 12

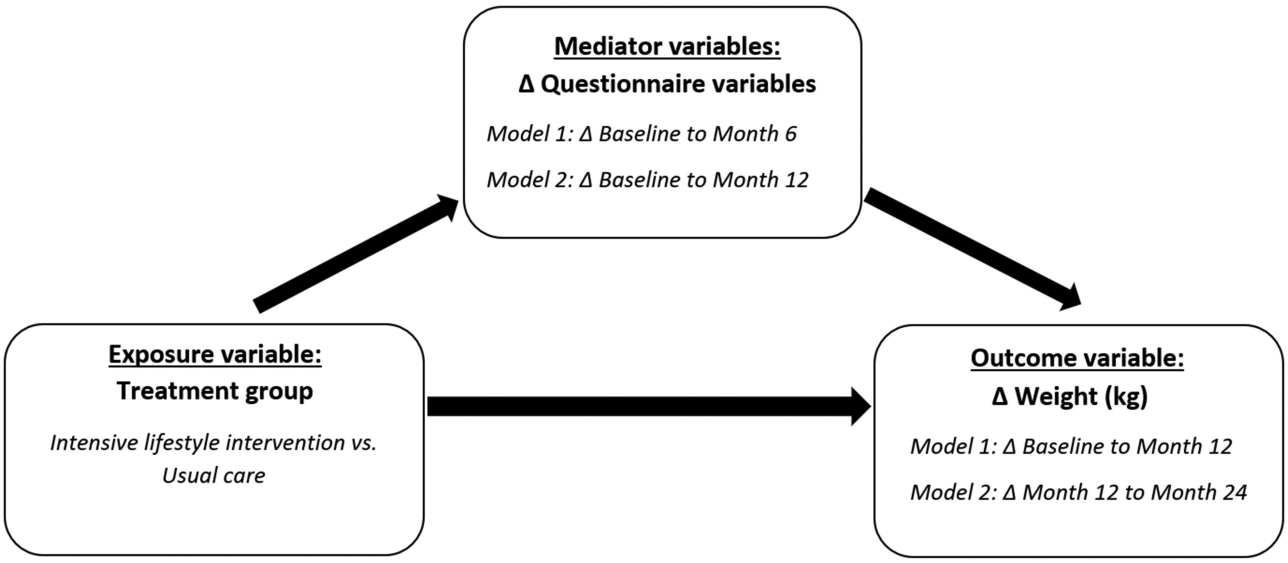


FIGURE 1 Hypothetical mediation models.

were the proposed mediators and change in weight from month 12 to month 24 was the outcome variable. In both models, the proposed mediators preceded the outcome variable, with a time difference (6 mo in model 1 and 12 mo in model 2) between the final measurement of the proposed mediators and the outcome variable. This was to ensure temporal ordering of our exposure variable, proposed mediators, and outcome variable, limiting the confounding influence of reverse causality. In addition, we removed patients with censored weights from the mediation models. Weight measurements were censored if a patient became pregnant, developed a medical condition, or died.

We conducted our analyses using the multilevel mediation analysis method of Yu and colleagues (36, 38), which is implemented in the *mlma* package in the software R. Briefly, potential mediators from our proposed mediators were informally selected if 2 conditions were satisfied. First, the proposed mediator distributed differently with or without the study's intervention (ILI compared with UC). In this regard, we used the ANOVA method to test if the mean of the variable differed between the ILI and UC. Second, the variable was significantly related to the outcome (weight change) while adjusting for all other related factors. This condition was tested through mixed-effect generalized linear models, with linear regression models used for linear outcomes or mediators. If only the second condition was satisfied, the variable was included as a covariate; yet the variable was excluded if the second condition was not satisfied (39). Further to the tests of 2 conditions, the package allows related variables to be forced into the model as mediators or covariates and it can assess joint effects of groups of mediators. Because the PROMIS-29 is used to determine overall health-related quality of life and no total score is obtained in the measure, we forced all PROMIS-29 constructs into the model as potential mediators and their joint effect was estimated. We likewise forced vigorous, moderate, and walking MET scores into the model as potential mediators and estimated the joint effect of these variables. Age, sex, race, baseline values for selected mediators, and weight (baseline weight for model 1; month 12 weight for model 2) were added as covariates. We estimated absolute total, direct, and indirect effects, as well as relative direct and indirect effects that provide the magnitude of these effects as a proportion of the total effect. For both the absolute and relative effect estimates, the SE and asymmetric 95% CIs around estimates were calculated, with inferences made using the bootstrap method. Unless noted otherwise, within the text, data are displayed as mean ± SD and 95% CI.

Results

Patient characteristics

A total of 803 patients with obesity (BMI: 37.2 ± 4.7) and a mean ± SD age of 49.4 ± 13.1 y were enrolled in the trial from 18 clinics: 452 patients from 9 clinics enrolled into the ILI and 351 patients from 9 clinics enrolled into UC (**Figure 2**). Details of the sample and the numbers who missed visits and withdrew are reported in the primary outcome article (9). The majority of patients were female (n = 678; 84.4%), were black (n = 540; 67.2%), and had a total household income <$40,000 ($n$ = 515; 64.1%) (**Table 1**). Moreover, 247 patients (30.8%) were food insecure.

During the trial, 24 patients had their month 12 weight censored, whereas a further 12 had their weight censored at month 24 (Figure 2). Therefore, the first mediation analysis and related summary comparisons (change scores in mediators from baseline to month 6) included 779 (439 ILI; 340 UC) patients, whereas the second mediation analysis and related summary comparisons (change scores for mediators from baseline to month 12) included 767 patients (433 ILI; 334 UC). **Supplemental Tables 1** and **2** show baseline characteristics of these analytical samples.

Change scores

The analysis in the primary outcome article showed that weight loss in the ILI group was greater than in the UC group at month

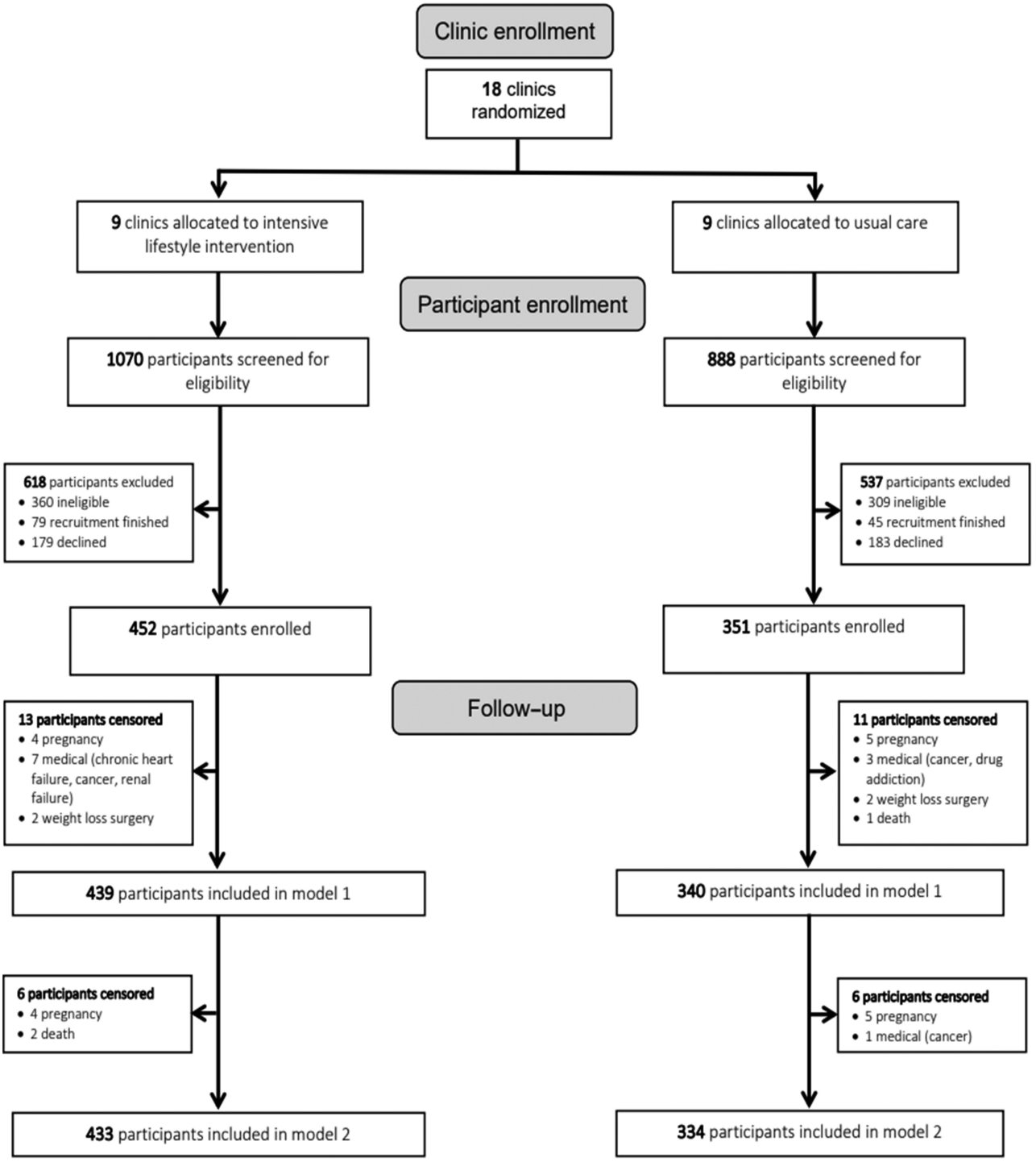


FIGURE 2 Participant flowchart for the analyses.

12 (ILI: −7.22 kg; 95% CI: −8.25, −6.19 kg; UC: −0.99 kg; 95% CI: −2.08, 0.09 kg) and month 24 (ILI: −5.43 kg; 95% CI: −6.52, −4.34 kg; UC: −0.91 kg; 95% CI: −2.07, 0.24 kg) (9).

Unadjusted independent-sample *t* tests suggested that the ILI group displayed a significant and large increase in restraint compared with the UC group at month 6 and month 12 ($P < 0.001$; ES $\geq$ 1.16), whilst a 0.9-point reduction in disinhibition was shown in the ILI relative to UC at month 6 ($P < 0.001$; ES = 0.33) (**Table 2**). At month 6 and month 12, compared with the UC group, the ILI group showed a small reduction in percentage fat intake and an increase in fruit and vegetable intake ($P \leq 0.010$; ES $\geq$ 0.20); yet both groups reported a similar change in alcohol intake at months 6 and 12 ($P \geq 0.662$; ES $\leq$ 0.03) (Table 2). The ILI group reported an increase in all physical activity constructs at month 6 ($P \leq 0.035$; ES $\geq$ 0.17), although only change in vigorous physical activity

TABLE 1 Baseline characteristics and measures of the PROPEL (Promoting Successful Weight Loss in Primary Care in Louisiana) trial cohort[1]

	All ($n = 803$)	ILI ($n = 452$)	UC ($n = 351$)
Age, y	49.4 ± 13.1	48.8 ± 12.7	50.1 ± 13.6
Sex			
Male	125 (15.6)	54 (11.9)	71 (20.2)
Female	678 (84.4)	398 (88.1)	280 (79.8)
Race			
White	208 (25.9)	95 (21.0)	113 (32.2)
Black	540 (67.2)	332 (73.5)	208 (59.3)
Other	55 (6.8)	25 (5.5)	30 (8.5)
Total annual household income, $			
<10,000	156 (19.4)	86 (19.0)	70 (19.9)
10,000–19,999	168 (20.9)	95 (21.0)	73 (20.8)
20,000–39,999	191 (23.8)	112 (24.8)	79 (22.5)
40,000–59,999	117 (14.6)	69 (15.3)	48 (13.7)
>60,000	154 (19.2)	83 (18.4)	71 (20.2)
Missing	17 (2.1)	7 (1.5)	10 (2.8)
Household food security status			
Food insecure	247 (30.8)	129 (28.5)	118 (33.6)
Food secure	556 (69.2)	323 (71.5)	233 (66.4)
Weight, kg	102.1 ± 16.7	101.6 ± 16.4	102.7 ± 17.0
BMI, kg/m^2	37.2 ± 4.7	37.3 ± 4.6	37.2 ± 4.8
EI			
EI, restraint	9.6 ± 4.5	9.6 ± 4.5	9.5 ± 4.5
EI, disinhibition	6.9 ± 3.7	7.0 ± 3.6	6.7 ± 3.7
Dietary intake questionnaire			
NCI, percentage fat intake	35.3 ± 6.4	35.9 ± 6.7	34.6 ± 5.9
NCI, fruit and vegetable intake	2.2 ± 1.7	2.2 ± 1.6	2.3 ± 1.8
BEVQ-15, alcohol intake	0.2 ± 0.4	0.2 ± 0.4	0.2 ± 0.4
Physical activity, MET-min/wk			
IPAQ, vigorous	561.4 ± 956.1	504.7 ± 891.1	634.3 ± 1030.3
IPAQ, moderate	475.2 ± 839.4	435.9 ± 803.3	525.2 ± 881.9
IPAQ, walking	808.9 ± 1011.2	780.8 ± 1027.3	844.0 ± 991.2
Weight-related quality of life			
IWQOL, total score	73.9 ± 19.0	72.8 ± 19.5	75.3 ± 18.3
Health-related quality of life			
PROMIS-29, sadness	47.5 ± 8.6	47.0 ± 8.5	48.1 ± 8.7
PROMIS-29, pain interference	51.9 ± 9.6	51.5 ± 9.7	52.5 ± 9.4
PROMIS-29, physical function	48.6 ± 8.0	48.9 ± 7.9	48.1 ± 8.1
PROMIS-29, social functioning	54.8 ± 9.0	55.2 ± 8.9	54.3 ± 9.1
PROMIS-29, fatigue	50.1 ± 10.1	49.4 ± 9.8	50.9 ± 10.4
PROMIS-29, anxiety	51.9 ± 9.9	51.7 ± 9.7	52.2 ± 10.1
PROMIS-29, sleep disturbance	50.7 ± 9.4	50.2 ± 9.2	51.5 ± 9.5

[1]Values are mean ± SD for continuous data and n (%) for categoric variables. BEVQ-15, Brief Questionnaire to Assess Habitual Beverage Intake; EI, Eating Inventory; ILI, intensive lifestyle intervention; IPAQ, International Physical Activity Questionnaire; IWQOL, Impact of Weight on Quality of Life-Lite; MET, metabolic equivalent of task; NCI, National Cancer Institute; PROMIS-29, Patient-Reported Outcomes Measurement Information System-29; UC, usual care.

was greater in the ILI group than in the UC group at month 12 (Table 2).

There was an increase in weight-related quality of life in the ILI group relative to the UC group at months 6 and 12 ($P < 0.001$; ES ≥ 0.62) (Table 2). Apart from sadness ($P = 0.177$; ES = 0.10), all health-related quality of life constructs of the PROMIS-29 were significantly improved in the ILI at month 6 compared with UC, with trivial-to-small ESs observed ($P \leq 0.008$; ES ≥ 0.20) (Table 2). At month 12, however, statistically significant improvements were only observed for pain interference, physical function, social functioning, and fatigue in the ILI group relative to the UC group ($P \leq 0.041$; ES ≥ 0.16) (Table 2).

Mediation analysis

Table 3 summarizes results from the mediation analyses. In model 1 (baseline to month 12 weight change), restraint, disinhibition, percentage fat intake, and weight-related quality of life total score met the 2 criteria and were selected as potential mediators alongside the IPAQ and PROMIS-29 variables and their composite scores. Similar to the primary outcome article (9), the total effect of the ILI (compared with UC) on weight change at month 12 was −7.19 kg (95% CI: −8.43, −6.07 kg). The direct effect [i.e., effect of ILI (compared with UC) on 12-mo weight change independent of change in mediators] was −5.36 kg (95% CI: −6.90, −3.94 kg), with a relative effect estimate showing that 75% of the between-group weight change

TABLE 2 Change scores in questionnaire variables at month 6 and month 12 during the PROPEL (Promoting Successful Weight Loss in Primary Care in Louisiana) trial[1]

	Baseline to month 6				Baseline to month 12			
	ILI ($n = 439$)	UC ($n = 340$)	P	Cohen's d	ILI ($n = 433$)	UC ($n = 334$)	P	Cohen's d
EI								
EI, restraint	6.3 ± 4.5	0.7 ± 3.6	<0.001	1.37	5.6 ± 4.4	0.8 ± 3.9	<0.001	1.16
EI, disinhibition	−1.8 ± 3.1	−0.9 ± 2.5	<0.001	0.33	−1.3 ± 3.1	−1.0 ± 2.7	0.109	0.12
Dietary intake questionnaire								
NCI, percentage fat intake	−3.7 ± 6.0	−1.2 ± 5.5	<0.001	0.43	−2.9 ± 6.0	−1.0 ± 5.1	<0.001	0.34
NCI, fruit and vegetable intake	0.2 ± 1.6	−0.2 ± 1.7	0.003	0.23	0.2 ± 1.7	−0.1 ± 1.9	0.010	0.20
BEVQ-15, alcohol intake	0.0 ± 0.3	0.0 ± 0.3	0.662	0.03	0.0 ± 0.3	0.0 ± 0.3	0.881	0.01
Physical activity, MET-min/wk								
IPAQ, vigorous	300.5 ± 1038.2	43.4 ± 1172.5	0.003	0.23	272.8 ± 1124.7	74.0 ± 1173.4	0.030	0.17
IPAQ, moderate	189.0 ± 1049.8	17.3 ± 989.6	0.035	0.17	238.0 ± 1047.3	138.8 ± 964.0	0.228	0.10
IPAQ, walking	211.1 ± 1203.9	9.5 ± 1047.3	0.034	0.18	125.4 ± 1283.0	−4.4 ± 1117.0	0.215	0.11
Weight-related quality of life								
IWQOL, total score	10.9 ± 14.0	3.2 ± 10.8	<0.001	0.62	12.1 ± 14.4	3.4 ± 11.5	<0.001	0.67
Health-related quality of life								
PROMIS-29, sadness	−0.1 ± 7.1	0.7 ± 7.9	0.177	0.10	0.2 ± 7.7	0.8 ± 7.4	0.281	0.08
PROMIS-29, pain interference	−1.3 ± 8.1	0.3 ± 8.3	0.008	0.20	−0.9 ± 9.2	0.9 ± 8.3	0.009	0.20
PROMIS-29, physical function	2.3 ± 6.4	0.1 ± 6.7	<0.001	0.34	1.8 ± 6.9	−0.1 ± 6.6	<0.001	0.28
PROMIS-29, social functioning	2.0 ± 7.5	0.1 ± 7.4	<0.001	0.26	2.0 ± 8.2	0.4 ± 7.8	0.007	0.21
PROMIS-29, fatigue	−3.0 ± 9.2	−0.6 ± 8.2	<0.001	0.28	−2.2 ± 9.2	−0.8 ± 8.9	0.041	0.16
PROMIS-29, anxiety	−1.4 ± 9.2	0.6 ± 8.9	0.003	0.22	−1.0 ± 8.9	0.0 ± 8.5	0.153	0.11
PROMIS-29, sleep disturbance	−1.8 ± 8.0	0.3 ± 7.7	<0.001	0.28	−0.8 ± 8.7	0.1 ± 8.5	0.174	0.10

[1] Values are mean ± SD unless indicated otherwise. Independent-sample t tests compared change scores between groups at month 6 and month 12. Absolute Cohen's d effect size values were used to compare change scores between groups at month 6 and month 12. BEVQ-15, Brief Questionnaire to Assess Habitual Beverage Intake; EI, Eating Inventory; ILI, intensive lifestyle intervention; IPAQ, International Physical Activity Questionnaire; IWQOL, Impact of Weight on Quality of Life-Lite; MET, metabolic equivalent of task; NCI, National Cancer Institute; PROMIS-29, Patient-Reported Outcomes Measurement Information System-29; UC, usual care.

was not caused by mediators. Of the selected potential mediators, disinhibition, percentage fat intake, moderate physical activity, walking, and fatigue change from baseline to month 6 were significant mediators of the improved weight loss displayed by the ILI group compared with the UC group at month 12. Specifically, month 6 change in disinhibition, percentage fat intake, moderate physical activity, walking, and fatigue explained −0.33 kg (95% CI: −0.55, −0.10 kg), −0.25 kg (95% CI: −0.50, −0.01 kg), −0.13 kg (95% CI: −0.23, −0.03 kg), −0.11 kg (95% CI: −0.21, −0.02 kg), and −0.28 kg (95% CI: −0.46, −0.10 kg), respectively, of the 12-mo weight change caused by the ILI (compared with UC). The joint indirect effect of physical activity (composite score for vigorous physical activity, moderate physical activity, and walking) was also significant and explained −0.26 kg (95% CI: −0.41, −0.09 kg) of the 12-mo weight change caused by the ILI (compared with UC). The relative effect estimates indicated that disinhibition, percentage fat intake, physical activity (joint effect), and fatigue explained 5%, 4%, 4%, and 4%, respectively, of the improved weight change seen in the ILI group compared with the UC group at month 12. Restraint was not a statistically significant mediator (−0.70 kg; 95% CI: −1.44, 0.03 kg). Similarly, the individual and joint effects of other PROMIS-29 variables and the change in weight-related quality of life did not significantly mediate month 12 weight change induced by the ILI (compared with UC) (Table 3).

In model 2 (month 12 to month 24 weight change), only fruit and vegetable intake met the 2 criteria and was selected as a potential mediator with the IPAQ and PROMIS variables. The ILI group displayed a significant 2.24-kg (95% CI: 1.32, 3.26 kg) increase in weight from month 12 to month 24 compared with the UC group (Table 3). The direct effect in this model was 2.00 kg (95% CI: 1.09, 3.02 kg), with the relative effect estimate suggesting 89% of the increase in weight exhibited by the ILI group (compared with UC) was not explained by the selected mediators. The change in fruit and vegetable intake from baseline to month 12 was a significant mediator of the increase in weight shown by the ILI group relative to the UC group from month 12 to month 24 (0.13 kg; 95% CI: 0.05, 0.21 kg); the relative effect estimate suggested that this explained 6% of the weight gain shown by the ILI (compared with UC). None of the other indirect effects of the selected mediators were significant (Table 3), suggesting 12-mo change in these selected mediators from baseline did not explain or attenuate (i.e., inconsistent mediation) the increase in weight seen by the ILI group compared with the UC group from month 12 to month 24.

TABLE 3 Total, direct, and indirect effects of the PROPEL (Promoting Successful Weight Loss in Primary Care in Louisiana) ILI (compared with UC) on weight change, with questionnaire variables as mediators[1]

	Absolute effect		Relative effect	
	Estimate	95% CI	Estimate	95% CI
Model 1 (baseline to month 12 weight change)[2]				
EI				
EI, restraint	−0.70 ± 0.40	−1.44, 0.03	0.10 ± 0.05	−0.01, 0.20
EI, disinhibition	−0.33 ± 0.11	−0.55, −0.10	0.05 ± 0.02	0.02, 0.07
Dietary intake questionnaire				
NCI, percentage fat intake	−0.25 ± 0.12	−0.50, −0.01	0.04 ± 0.02	0.00, 0.07
Physical activity, MET-min/wk				
IPAQ, joint effect of constructs[3]	−0.26 ± 0.08	−0.41, −0.09	0.04 ± 0.01	0.01, 0.06
IPAQ, vigorous	0.02 ± 0.06	−0.14, 0.11	0.00 ± 0.01	−0.02, 0.02
IPAQ, moderate	−0.13 ± 0.05	−0.23, −0.03	0.02 ± 0.01	0.00, 0.03
IPAQ, walking	−0.11 ± 0.05	−0.21, −0.02	0.02 ± 0.01	0.00, 0.03
Weight-related quality of life				
IWQOL, total score	−0.35 ± 0.21	−0.76, 0.04	0.05 ± 0.03	−0.01, 0.11
Health-related quality of life				
PROMIS-29, joint effect of constructs[4]	0.06 ± 0.13	−0.19, 0.33	−0.01 ± 0.02	−0.05, 0.03
PROMIS-29, sadness	0.01 ± 0.03	−0.05, 0.06	−0.00 ± 0.00	−0.01, 0.01
PROMIS-29, pain interference	−0.02 ± 0.06	−0.13, 0.09	0.00 ± 0.01	−0.01, 0.02
PROMIS-29, physical function	0.14 ± 0.09	−0.03, 0.31	−0.02 ± 0.01	−0.04, 0.01
PROMIS-29, social functioning	0.03 ± 0.08	−0.12, 0.20	−0.00 ± 0.01	−0.03, 0.02
PROMIS-29, fatigue	−0.28 ± 0.09	−0.46, −0.10	0.04 ± 0.01	0.01, 0.06
PROMIS-29, anxiety	0.12 ± 0.08	−0.03, 0.27	−0.02 ± 0.01	−0.04, 0.00
PROMIS-29, sleep disturbance	0.06 ± 0.09	−0.12, 0.25	−0.01 ± 0.01	−0.03, 0.02
Direct effect	−5.36 ± 0.76	−6.90, −3.94	0.75 ± 0.06	0.62, 0.87
Total effect	−7.19 ± 0.60	−8.43, −6.07	—	—
Model 2 (month 12 to month 24 weight change)[5]				
Dietary intake questionnaire				
NCI, fruit and vegetable intake	0.13 ± 0.04	0.05, 0.21	0.06 ± 0.02	0.01, 0.11
Physical activity, MET-min/wk				
IPAQ, joint effect of constructs[3]	0.03 ± 0.04	−0.06, 0.10	0.01 ± 0.02	−0.03, 0.05
IPAQ, vigorous	0.02 ± 0.04	−0.06, 0.09	0.01 ± 0.02	−0.03, 0.04
IPAQ, moderate	0.01 ± 0.02	−0.03, 0.05	0.00 ± 0.01	−0.02, 0.02
IPAQ, walking	0.00 ± 0.03	−0.05, 0.06	0.00 ± 0.01	−0.02, 0.03
Health-related quality of life				
PROMIS-29, joint effect of constructs[4]	0.09 ± 0.07	−0.06, 0.23	0.04 ± 0.03	−0.03, 0.10
PROMIS-29, sadness	0.02 ± 0.02	−0.02, 0.06	0.01 ± 0.01	−0.01, 0.03
PROMIS-29, pain interference	0.06 ± 0.06	−0.06, 0.17	0.02 ± 0.03	−0.03, 0.08
PROMIS-29, physical function	0.09 ± 0.07	−0.05, 0.23	0.04 ± 0.03	−0.03, 0.11
PROMIS-29, social functioning	−0.03 ± 0.05	−0.13, 0.07	−0.02 ± 0.03	−0.07, 0.03
PROMIS-29, fatigue	−0.01 ± 0.04	−0.08, 0.06	0.00 ± 0.02	−0.04, 0.03
PROMIS-29, anxiety	−0.05 ± 0.03	−0.10, 0.01	−0.02 ± 0.01	−0.05, 0.01
PROMIS-29, sleep disturbance	0.01 ± 0.02	−0.04, 0.05	0.00 ± 0.01	−0.02, 0.03
Direct effect	2.00 ± 0.49	1.09, 3.02	0.89 ± 0.04	0.81, 0.98
Total effect	2.24 ± 0.49	1.32, 3.26	—	—

[1]Values and 95% CIs were calculated with the *mlma* package of Yu and colleagues (36, 38). Absolute effects are estimated means ± SEs, whereas relative direct and indirect effects are the corresponding direct or indirect effect divided by the total effect ± SE. EI, Eating Inventory; ILI, intensive lifestyle intervention; IPAQ, International Physical Activity Questionnaire; IWQOL, Impact of Weight on Quality of Life-Lite; MET, metabolic equivalent of task; NCI, National Cancer Institute; PROMIS-29, Patient-Reported Outcomes Measurement Information System-29; UC, usual care.

[2]The exposure variable was group (ILI compared with UC), the proposed mediators were change in questionnaire variables from baseline to month 6, and the outcome variable was weight change from baseline to month 12. Adjusted for age, sex, race, baseline questionnaire variables for selected mediators, and baseline weight; $n = 779$ (439 ILI; 340 UC).

[3]Indirect effect is a composite score of the joint effect of all constructs: vigorous MET-mins/wk, moderate MET-mins/wk, and walking MET-mins/wk.

[4]Indirect effect is a composite score of the joint effect of all constructs: sadness, pain interference, physical function, social functioning, fatigue, anxiety, and sleep disturbance.

[5]The exposure variable was group (ILI compared with UC), the proposed mediators were change in questionnaire variables from baseline to month 12, and the outcome variable was weight change from month 12 to month 24. Adjusted for age, sex, race, baseline questionnaire variables for selected mediators, and month 12 weight; $n = 767$ (433 ILI; 334 UC).

Discussion

Over 24 mo, an ILI induced weight loss relative to UC in a sample of underserved patients with obesity. These analyses showed that month 6 change in disinhibition, percentage fat intake, physical activity, and subjective fatigue partially mediated the weight change seen in the ILI group relative to the UC group at 12 mo. The ILI group lost more weight than the UC group at month 24, but weight gain of 2.24 kg was observed in the ILI compared with UC from month 12 to month 24, with fruit and vegetable intake identified as a mediator. Analyses showed that the change in questionnaire constructs explained a small amount of the between-group weight change. More specifically, each mediator explained ≤10% of the between-group weight change, and relative direct effect values indicated that ≥75% of the between-group weight change was not explained by assessed constructs. Nonetheless, although other unmeasured factors could drive between-group weight variations, these results could help improve weight regimens by highlighting critical constructs and behaviors for weight loss.

Behavioral lifestyle interventions typically offer counseling sessions that aim to improve the eating behaviors of individuals with obesity via an increase in dietary restraint and a reduction in dietary disinhibition (8). In this analysis, we observed that a decrease in dietary disinhibition was a significant mediator of 12-mo weight loss seen in the ILI group compared with the UC group. This is consonant with studies reporting that a reduction in dietary disinhibition (12) is associated with weight loss in individuals with obesity during lifestyle interventions, and it suggests that regimens provided to underserved cohorts should place particular focus on behavioral strategies linked to disinhibition. Such strategies could consist of those utilized during the PROPEL ILI behavioral sessions, including controlled eating of foods, eating habits in response to stress and negative emotion, and healthy eating during special events. In contrast to the decrease in disinhibition, the increase in dietary restraint was not a significant mediator in our analyses. This supports research showing no association between restraint and weight loss (40), although in contrast to some work (12), we may have been underpowered to detect a positive influence of dietary restraint on weight loss.

A core strategy recommended for weight loss in individuals with obesity is the adoption of healthy dietary patterns. This includes limiting fat and alcohol intake and incorporating fruits, vegetables, and grains into a calorie-deficit diet (8, 37). Our results suggest that a decrease in percentage fat intake was a mediator of the between-group difference in month 12 weight loss, supporting previous analyses (13, 41) and suggesting that a reduction in fat intake is a key practice that assists the development of a calorie deficit and weight loss in underserved individuals with obesity. However, consistent with previous evidence (42), change in alcohol consumption did not mediate weight loss seen in the ILI compared with UC. In addition, although it did not influence weight loss in the first year of the trial, the increase in fruit and vegetable consumption seen in the ILI relative to UC did mediate the relative weight gain from month 12 to month 24. It is possible that fruit and vegetable consumption increased energy intake during a period of relapse in the ILI group, but it should be noted that research examining the influence of fruit and vegetable intake per se on long-term weight maintenance is mixed (14, 43). Therefore, further research is needed to elucidate the role of fruit and vegetable intake during weight-management interventions in underserved individuals with obesity.

Studies show that physical activity combined with dietary modifications stimulate greater weight loss over periods of ≥12 mo than do dietary modifications alone (44). We found, in line with these findings, that increased physical activity at month 6 mediated the greater 12-mo weight loss in the ILI, particularly the increase in moderate physical activity and walking. This suggests future weight-loss regimens in similar patient populations should seek to increase physical activity to improve weight loss. It could also imply that interventions should set more ambitious activity goals that have been recommended for weight loss, such as ≥200 min/wk of walking or moderate physical activity (45). However, because physical activity did not influence between-group weight change from month 12 to month 24, future research should elucidate the long-term role of physical activity during ILIs in underserved populations, especially because physical activity may be important in preventing weight regain (46). These studies should identify methods to sustain elevations in physical activity, given there were no differences in moderate physical activity and walking between the groups at month 12.

In addition to physical activity, model 1 revealed decreased fatigue as a mediator of improved 12-mo weight change during the ILI. Speculatively, although concurrent changes in weight and fatigue may reciprocally affect each other (15), the behavioral strategies of the ILI may have decreased subjective fatigue and led to better adherence to the weight-loss regimen compared with UC. Our analysis nonetheless indicated that other health- and weight-related quality of life constructs did not drive the greater weight loss seen in the ILI group during the first 12 mo, and quality of life changes at month 12 did not influence between-group weight change during the trial from month 12 to month 24.

A strength of the analyses is that they comprise data from a cluster-randomized trial performed in patients with obesity who are underserved and understudied within clinical studies. As a result, broadly, our findings have implications for socioeconomically disadvantaged individuals who are disproportionally affected by obesity and obesity-related conditions, and who face significant barriers for treatment. Another strength is that we collected mediator and weight measurements at multiple points during the 24-mo trial, enabling us to investigate mediators of weight change during periods which, despite variations in definitions (47), can be generally considered midterm weight loss (37) and weight-loss maintenance. This means our results can be utilized to develop enhanced ILIs which target constructs that are important for long-term weight management in similar at-risk populations. A final noteworthy strength is that trained health coaches of the trial were embedded within a team in primary care. This may explain why our analyses revealed many findings that are similar to those derived from more controlled trials, and it could supply a model for weight-management regimens in primary care (9).

The current article has limitations. First, the trial consisted mostly of females, restricting our ability to generalize our results to underserved males with obesity. Second, because there were

no sessions offered in the UC group, we could not incorporate number of sessions attended into our analyses. It is possible that session attendance was a significant driver of weight change, as indicated by some studies (8). Third, these analyses are exploratory, so although our analyses comprised a relatively large number of participants from an understudied population, we included several variables and may be underpowered to detect some effects. Fourth, unmeasured mediator variables and mediator–outcome confounders may be more causally linked to between-group weight differences, limiting our ability to make causal inferences. Finally, measurement errors likely explain, at least in part, why our mediators explained a small proportion of between-group weight differences (48). Indeed, we used self-report assessments of diet and physical activity, which, in contrast to objective measures (e.g., waist devices for physical activity), are prone to systematic and random errors, primarily because they rely on recall and can be influenced by demand characteristics (49). The PROPEL trial was a pragmatic trial performed in a low-literate population in primary care; hence, a large battery of sophisticated assessments was unfeasible, and we decreased the burden of some questionnaires (e.g., not administering the EI hunger subscale). Yet additional research is needed during lifestyle interventions in underserved populations to elucidate the causal drivers of weight change. Such studies could examine further potential mediators of weight change like calorie intake, consistency of eating (50), hunger, energy density (51), and sugar-sweetened beverage consumption (52). Further, where possible, they should utilize reliable and objective assessment methods, particularly for diet (e.g., emerging technologies like food photography) and physical activity (e.g., pedometers or accelerometers) (49).

In conclusion, our analyses indicated that 12-mo weight loss during an ILI compared with UC was explained by improvements in disinhibition, percentage fat intake, physical activity, and fatigue in underserved patients with obesity. These variables did not, however, attenuate the weight gain shown during the ILI compared with UC in the final 12 mo of the trial, and fruit and vegetable intake may partially explain this response. Although additional work is needed using precise assessment methods to elucidate causal drivers of weight change during ILIs, these findings highlight psychological and behavioral constructs that could be targeted to refine interventions and facilitate weight management in underserved patients with obesity.

The PROPEL Research Group includes: Pennington Biomedical Research Center: Peter T. Katzmarzyk, PhD, Robert L. Newton Jr, PhD, Corby K. Martin, PhD, John W. Apolzan, PhD, William Johnson, PhD, Kara D. Denstel, MPH, Emily F. Mire, MS, Robert K. Singletary Jr, MHS, Cheryl Lewis, MPH, Phillip Brantley, PhD, Ronald Horswell, PhD, Betty Kennedy, PhD, Dachuan Zhang, MAppStats, Stephanie Authement, RD, LDN, MS, Shiquita Brooks, RDN, LDN, Danielle S. Burrell, MEd, MCHES, Leslie Forest-Everage, MA, Angelle Graham Ullmer, RDN, LDN, MS, Laurie Murphy, RDN, LDN, Cristalyn Reynolds, MA, Kevin Sanders, MS, RDN, LDN, Stephen Bower, MS, Daishaun Gabriel, MHA, Hillary Gahagan, MPH, Tabitha K. Gray, MA, Jill Hancock, MPH, Marsha Herrera, Brittany Molinere, Georgia Morgan, MA, Brittany Neyland, Stephanie Rincones, Deanna Robertson, MA, Ekambi Shelton, MPH, Russell J. Tassin, MS, Kaili Williams; Louisiana State University Health Sciences Center at New Orleans: Benjamin F. Springgate, MD; Louisiana State University Health Sciences Center at Shreveport: Terry C. Davis, PhD, Connie L. Arnold, PhD; Ochsner Health System: Eboni Price-Haywood,MD, Carl J. Lavie, MD, Jewel Harden-Barrios, MEd; Tulane University Medical School: Vivian A. Fonseca, MD, Tina K. Thethi, MD (MEDICAL Monitor), Jonathan Gugel, MD; Xavier University: Kathleen B. Kennedy, PhD, Daniel F. Sarpong, PhD, Amina D. Massey. Study data were collected and managed with REDCap (Research Electronic Data Capture) electronic data capture tools hosted at Pennington Biomedical. REDCap is a secure, web-based application designed to support data capture for research studies, providing an intuitive interface for validated data entry, audit trails for tracking data manipulation and export procedures, automated export procedures for seamless data downloads to common statistical packages, and procedures for importing data from external sources. We thank Health One (Carmel, CA) and Nutrisystem for providing portion-controlled meals during the study.

The authors' responsibilities were as follows—JLD, CKM, CH, and PTK: conceived the project; CKM, JWA, RLN, and PTK: conceived the study, developed the overall research plan, and performed hands-on conduct of the experiments and data collection; JLD, QY, and WC: oversaw the analysis; JLD, CKM, and PTK: wrote the paper; CKM, JWA, RLN, KDD, EFM, and PTK: oversaw the study; QY and WC: performed the statistical analysis; KDD and EFM: managed the data; PTK: had primary responsibility for the final content; and all authors: read and approved the final manuscript. All other authors report no conflicts of interest.

Data Availability

Complete de-identified data described in the article, code book, and analytic code are available from the corresponding author (PTK) by reasonable request pending the approval of the PROPEL publication committee.

References

1 Visscher TLS, Seidell JC. The public health impact of obesity. Annu Rev Public Health 2001;22:355–75.
2 Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. BMC Public Health 2009;9(1):88.
3 Hales CM, Fryar CD, Carroll MD, Freedman DS, Ogden CL. Trends in obesity and severe obesity prevalence in US youth and adults by sex and age, 2007–2008 to 2015–2016. JAMA 2018;319(16):1723–5.
4 Ogden CL, Fakhouri TH, Carroll MD, Hales CM, Fryar CD, Li X, et al. Prevalence of obesity among adults, by household income and education—United States, 2011–2014. MMWR Morb Mortal Wkly Rep 2017;66(50):1369–73.
5 Hruby A, Hu FB. The epidemiology of obesity: a big picture. Pharmacoeconomics 2015;33(7):673–89.
6 Carvajal R, Wadden TA, Tsai AG, Peck K, Moran CH. Managing obesity in primary care practice: a narrative review. Ann N Y Acad Sci 2013;1281(1):191–206.
7 Jensen MD, Ryan DH. New obesity guidelines: promise and potential. JAMA 2014;311(1):23–4.
8 Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, et al. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. J Am Coll Cardiol 2014;63(25):2985–3023.
9 Katzmarzyk PT, Martin CK, Newton RL, Apolzan JW, Arnold CL, Davis TC, et al. Weight loss in underserved patients—a cluster-randomized trial. N Engl J Med 2020;383(10):909–18.
10 Höchsmann C, Dorling JL, Martin CK, Newton RL, Apolzan JW, Myers CA, et al. Effects of a two-year primary care lifestyle intervention on cardiometabolic risk factors—a cluster-randomized trial. Circulation 2021;143(12):1202–14.
11 Aronne LJ, Hall KD, Jakicic JM, Leibel RL, Lowe MR, Rosenbaum M, et al. Describing the weight-reduced state: physiology, behavior, and interventions. Obesity (Silver Spring) 2021;29(Suppl 1):S9–24.
12 Batra P, Das SK, Salinardi T, Robinson L, Saltzman E, Scott T, et al. Eating behaviors as predictors of weight loss in a 6 month weight loss intervention. Obesity (Silver Spring) 2013;21(11):2256–63.

13 Zheng Y, Sereika SM, Danford CA, Imes CC, Goode RW, Mancino J, et al. Trajectories of weight change and predictors over 18-month weight loss treatment. J Nurs Scholarsh 2017;49(2):177–84.
14 Coughlin JW, Gullion CM, Brantley PJ, Stevens VJ, Bauck A, Champagne CM, et al. Behavioral mediators of treatment effects in the Weight Loss Maintenance Trial. Ann Behav Med 2013;46(3):369–81.
15 Palmeira AL, Markland DA, Silva MN, Branco TL, Martins SC, Minderico CS, et al. Reciprocal effects among changes in weight, body image, and other psychological factors during behavioral obesity treatment: a mediation analysis. Int J Behav Nutr Phys Act 2009;6(1):9.
16 Katzmarzyk PT, Martin CK, Newton RL, Apolzan JW, Arnold CL, Davis TC, et al. Promoting Successful Weight Loss in Primary Care in Louisiana (PROPEL): rationale, design and baseline characteristics. Contemp Clin Trials 2018;67:1–10.
17 Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377–81.
18 Diabetes Prevention Program (DPP) Research Group. The Diabetes Prevention Program (DPP): description of lifestyle intervention. Diabetes Care 2002;25(12):2165–71.
19 ; The Look AHEAD Research Group, Wadden TA, Smith West D, Delahanty L, Jakicic J, Rejeski J, et al. The Look AHEAD study: a description of the lifestyle intervention and the evidence supporting it. Obesity (Silver Spring) 2006;14(5):737–52.
20 Rickman AD, Williamson DA, Martin CK, Gilhooly CH, Stein RI, Bales CW, et al. The CALERIE study: design and methods of an innovative 25% caloric restriction intervention. Contemp Clin Trials 2011;32(6):874–81.
21 Thomas DM, Martin CK, Heymsfield S, Redman LM, Schoeller DA, Levine JA. A simple model predicting individual weight change in humans. J Biol Dyn 2011;5(6):579–99.
22 Jacques L, Jensen TS, Schafer J, McClain S, Chin J. Decision memo for intensive behavioral therapy for obesity (CAG-00423 N) [Internet]. Baltimore, MD: Centers for Medicare & Medicaid Services; 2011 [cited 26 March, 2019]. Available from: https://www.cms.gov/medicare-coverage-database/details/nca-decision-memo.aspx?&NcaName=Intensive Behavioral Therapy for Obesity&bc=ACAAAAAAIAAA&NCAId=253&.
23 Stunkard AJ, Messick S. The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. J Psychosom Res 1985;29(1):71–83.
24 Stewart TM, Martin CK, Williamson DA. The complicated relationship between dieting, dietary restraint, caloric restriction, and eating disorders: is a shift in public health messaging warranted? Int J Environ Res Public Health 2022;19(1):491.
25 Bryant EJ, King NA, Blundell JE. Disinhibition: its effects on appetite and weight regulation. Obes Rev 2008;9(5):409–19.
26 Thompson FE, Midthune D, Subar AF, Kipnis V, Kahle LL, Schatzkin A. Development and evaluation of a short instrument to estimate usual dietary intake of percentage energy from fat. J Am Diet Assoc 2007;107(5):760–7.
27 Havas S, Heimendinger J, Damron D, Nicklas TA, Cowan A, Beresford SAA, et al. 5 a day for better health—nine community research projects to increase fruit and vegetable consumption. Public Health Rep 1995;110(1):68–79.
28 Thompson FE, Kipnis V, Subar AF, Krebs-Smith SM, Kahle LL, Midthune D, et al. Evaluation of 2 brief instruments and a food-frequency questionnaire to estimate daily number of servings of fruit and vegetables. Am J Clin Nutr 2000;71(6):1503–10.
29 Hedrick VE, Savla J, Comber DL, Flack KD, Estabrooks PA, Nsiah-Kumi PA, et al. Development of a brief questionnaire to assess habitual beverage intake (BEVQ-15): sugar-sweetened beverages and total beverage energy intake. J Acad Nutr Diet 2012;112(6):840–9.
30 Kolotkin RL, Crosby RD. Psychometric evaluation of the impact of weight on quality of life-lite questionnaire (IWQOL-Lite) in a community sample. Qual Life Res 2002;11(2):157–71.
31 Kolotkin RL, Crosby RD, Kosloski KD, Williams GR. Development of a brief measure to assess quality of life in obesity. Obes Res 2001;9(2):102–11.
32 Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. J Clin Epidemiol 2010;63(11):1179–94.
33 Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International Physical Activity Questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35(8):1381–95.
34 Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
35 MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. Annu Rev Psychol 2007;58:593–614.
36 Yu Q, Li B. Third-variable effect analysis with multilevel additive models. PLoS One 2020;15(10):e0241072.
37 Chopra S, Malhotra A, Ranjan P, Vikram NK, Sarkar S, Siddhu A, et al. Predictors of successful weight loss outcomes amongst individuals with obesity undergoing lifestyle interventions: a systematic review. Obes Rev 2021;22(3):e13148.
38 Yu Q, Yu M, Zou J, Wu X, Gomez SL, Li B. Multilevel mediation analysis on time-to-event outcomes: exploring racial/ethnic disparities in breast cancer survival in California. Res Methods Med Health Sci 2021;2(4):157–67.
39 Yu Q, Li B. mma: an R package for mediation analysis with multiple mediators. J Open Res Softw 2017;5(1):11.
40 Dorling JL, Bhapkar M, Das SK, Racette SB, Apolzan JW, Fearnbach SN, et al. Change in self-efficacy, eating behaviors and food cravings during two years of calorie restriction in humans without obesity. Appetite 2019;143:104397.
41 Delahanty LM, Meigs JB, Hayden D, Williamson DA, Nathan DM. Psychological and behavioral correlates of baseline BMI in the Diabetes Prevention Program (DPP). Diabetes Care 2002;25(11):1992–8.
42 Morgan PJ, Hollis JL, Young MD, Collins CE, Teixeira PJ. Workday sitting time and marital status: novel pretreatment predictors of weight loss in overweight and obese men. Am J Mens Health 2018;12(5):1431–8.
43 Kaiser KA, Brown AW, Brown MMB, Shikany JM, Mattes RD, Allison DB. Increased fruit and vegetable intake has no discernible effect on weight loss: a systematic review and meta-analysis. Am J Clin Nutr 2014;100(2):567–76.
44 Donnelly JE, Herrmann SD, Lambourne K, Szabo AN, Honas JJ, Washburn RA. Does increased exercise or physical activity alter ad-libitum daily energy intake or macronutrient composition in healthy adults? A systematic review. PLoS One 2014;9(1):e83498.
45 Donnelly JE, Blair SN, Jakicic JM, Manore MM, Rankin JW, Smith BK, et al. American College of Sports Medicine Position Stand. Appropriate physical activity intervention strategies for weight loss and prevention of weight regain for adults. Med Sci Sports Exerc 2009;41(2):459–71.
46 Jakicic JM, Rogers RJ, Davis KK, Collins KA. Role of physical activity and exercise in treating patients with overweight and obesity. Clin Chem 2018;64(1):99–107.
47 Berger SE, Huggins GS, McCaffery JM, Lichtenstein AH. Comparison among criteria to define successful weight-loss maintainers and regainers in the Action for Health in Diabetes (Look AHEAD) and Diabetes Prevention Program trials. Am J Clin Nutr 2017;106(6):1337–46.
48 Pieters R. Meaningful mediation analysis: plausible causal inference and informative communication. J Consum Res 2017;44(3):692–716.
49 Dhurandhar NV, Schoeller D, Brown AW, Heymsfield SB, Thomas D, Sørensen TIA, et al. Energy balance measurement: when something is not better than nothing. Int J Obes 2015;39(7):1109–13.
50 Wing R, Phelan S. Long-term weight loss maintenance. Am J Clin Nutr 2005;82(1):222S–5S.
51 Ledikwe JH, Blanck HM, Khan LK, Serdula MK, Seymour JD, Tohill BC, et al. Dietary energy density is associated with energy intake and weight status in US adults. Am J Clin Nutr 2006;83(6):1362–8.
52 Chen L, Appel LJ, Loria C, Lin P-H, Champagne CM, Elmer PJ, et al. Reduction in consumption of sugar-sweetened beverages is associated with weight loss: the PREMIER trial. Am J Clin Nutr 2009;89(5):1299–306.

Publication 7

The Personalized Nutrition Study (POINTS): Evaluation of a genetically informed weight loss approach. A Randomized Clinical Trial.

Christoph Höchsmann [1,2], Shengping Yang [2], José M Ordovás [3], James L Dorling [4], Catherine M Champagne [2], John W Apolzan [2], Frank L Greenway [2], Michelle I Cardel [5,6], Gary D Foster [5,7], Corby K Martin [2]

[1] TUM School of Medicine and Health, Technical University of Munich, Munich, Germany

[2] Pennington Biomedical Research Center, Baton Rouge, LA, USA

[3] Tufts University, Boston, MA, USA

[4] Human Nutrition, School of Medicine, Dentistry and Nursing, University of Glasgow, Glasgow, UK

[5] WW International, Inc., New York, NY, USA

[6] University of Florida School of Medicine, Gainesville, FL, USA

[7] Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
Nature Communications, 14(1): 6321	2023	10.1038/s41467-023-41969-1	16.6 (2022)	Q1	Scopus: 2 Google Scholar: 8

Author contribution:

First and corresponding author; co-designed the study, oversaw data acquisition, analyzed and interpreted the data, drafted the manuscript, and created tables and figures.

nature communications

Article

https://doi.org/10.1038/s41467-023-41969-1

The Personalized Nutrition Study (POINTS): evaluation of a genetically informed weight loss approach, a Randomized Clinical Trial

Received: 26 April 2023

Accepted: 26 September 2023

Published online: 09 October 2023

Check for updates

Christoph Höchsmann[1,2] ✉, Shengping Yang[2], José M. Ordovás[3], James L. Dorling[4], Catherine M. Champagne[2], John W. Apolzan[2], Frank L. Greenway[2], Michelle I. Cardel[5,6], Gary D. Foster[5,7] & Corby K. Martin[2]

Weight loss (WL) differences between isocaloric high-carbohydrate and high-fat diets are generally small; however, individual WL varies within diet groups. Genotype patterns may modify diet effects, with carbohydrate-responsive genotypes losing more weight on high-carbohydrate diets (and vice versa for fat-responsive genotypes). We investigated whether 12-week WL (kg, primary outcome) differs between genotype-concordant and genotype-discordant diets. In this 12-week single-center WL trial, 145 participants with overweight/obesity were identified a priori as fat-responders or carbohydrate-responders based on their combined genotypes at ten genetic variants and randomized to a high-fat (n = 73) or high-carbohydrate diet (n = 72), yielding 4 groups: (1) fat-responders receiving high-fat diet, (2) fat-responders receiving high-carbohydrate diet, (3) carbohydrate-responders receiving high-fat diet, (4) carbohydrate-responders receiving high-carbohydrate diet. Dietitians delivered the WL intervention via 12 weekly diet-specific small group sessions. Outcome assessors were blind to diet assignment and genotype patterns. We included 122 participants (54.4 [SD:13.2] years, BMI 34.9 [SD:5.1] kg/m^2, 84% women) in the analyses. Twelve-week WL did not differ between the genotype-concordant (−5.3 kg [SD:1.0]) and genotype-discordant diets (−4.8 kg [SD:1.1]; adjusted difference: −0.6 kg [95% CI: −2.1,0.9], $p = 0.50$). With the current ability to genotype participants as fat- or carbohydrate-responders, evidence does not support greater WL on genotype-concordant diets. ClinicalTrials identifier: NCT04145466.

The 2017–2018 National Health and Nutrition Examination Survey (NHANES) showed that almost 43% of US adults aged 20 and over have obesity, including 9.0% with severe obesity, and another 31% are overweight[1]. Excess body fat increases the risk of numerous medical conditions and premature mortality[2], presenting public health and economic challenges[3,4].

Many weight loss (WL) strategies emphasize either high-carbohydrate (and low-fat) or high-fat (low-carbohydrate) diets[5,6]. WL

[1]Department of Health and Sport Sciences, TUM School of Medicine and Health, Technical University of Munich, Munich, Germany. [2]Pennington Biomedical Research Center, Baton Rouge, LA, USA. [3]Tufts University, Boston, MA, USA. [4]Human Nutrition, School of Medicine, Dentistry and Nursing, College of Medical, Veterinary and Life of Sciences, University of Glasgow, Glasgow, UK. [5]WW International, Inc., New York, NY, USA. [6]Department of Health Outcomes and Biomedical Informatics, University of Florida College of Medicine, Gainesville, FL, USA. [7]Center for Weight and Eating Disorders, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ✉e-mail: christoph.hoechsmann@tum.de

differences between isocaloric high-carbohydrate and high-fat diets are generally small or negligible[7]; however, individual WL varies substantially within diet groups[6], suggesting that individuals react differently to high-carbohydrate or high-fat diets. Retrospective data suggest that participants with carbohydrate-responsive polymorphisms lose more weight on high-carbohydrate vs. high-fat diets and vice versa for those with fat-responsive polymorphisms[8]. However, these results have not been confirmed in randomized controlled trials (RCT), and the approach of determining low-fat- and low-carbohydrate-responsive genotypes based on single-nucleotide polymorphisms (SNPs) from three genes (*PPARG*, *ADRB2*, and *FABP2*)[8,9] has been criticized[10]. Overall, reports show that most genotype × diet interactions are not significant, and replication is rare[11]. A more comprehensive and informative risk score (determined a priori), comprised of a greater number of SNPs with demonstrated and validated effects on the responses to high-fat/high-carbohydrate diets, may better define fat- and carbohydrate-responsive genetic predisposition scores.

The present RCT tested the hypothesis that participants assigned to a diet corresponding to their a priori-determined (fat-responsive or carbohydrate-responsive) genotype would lose more weight over 12 weeks than those assigned to a diet discordant with their genotype. Further, we aimed to analyze those with a fat-responsive genotype (subsequently "fat-responders") and carbohydrate-responsive genotype (subsequently "carbohydrate-responders") separately. We hypothesized that (1) fat-responders would lose more weight on the high-fat vs. high-carbohydrate diet and conversely (2) carbohydrate-responders would lose more weight on the high-carbohydrate vs. high-fat diet. A secondary objective of the present RCT was to test the newly-developed genetic risk score to determine fat- and carbohydrate-responsive genotypes that was based on the current

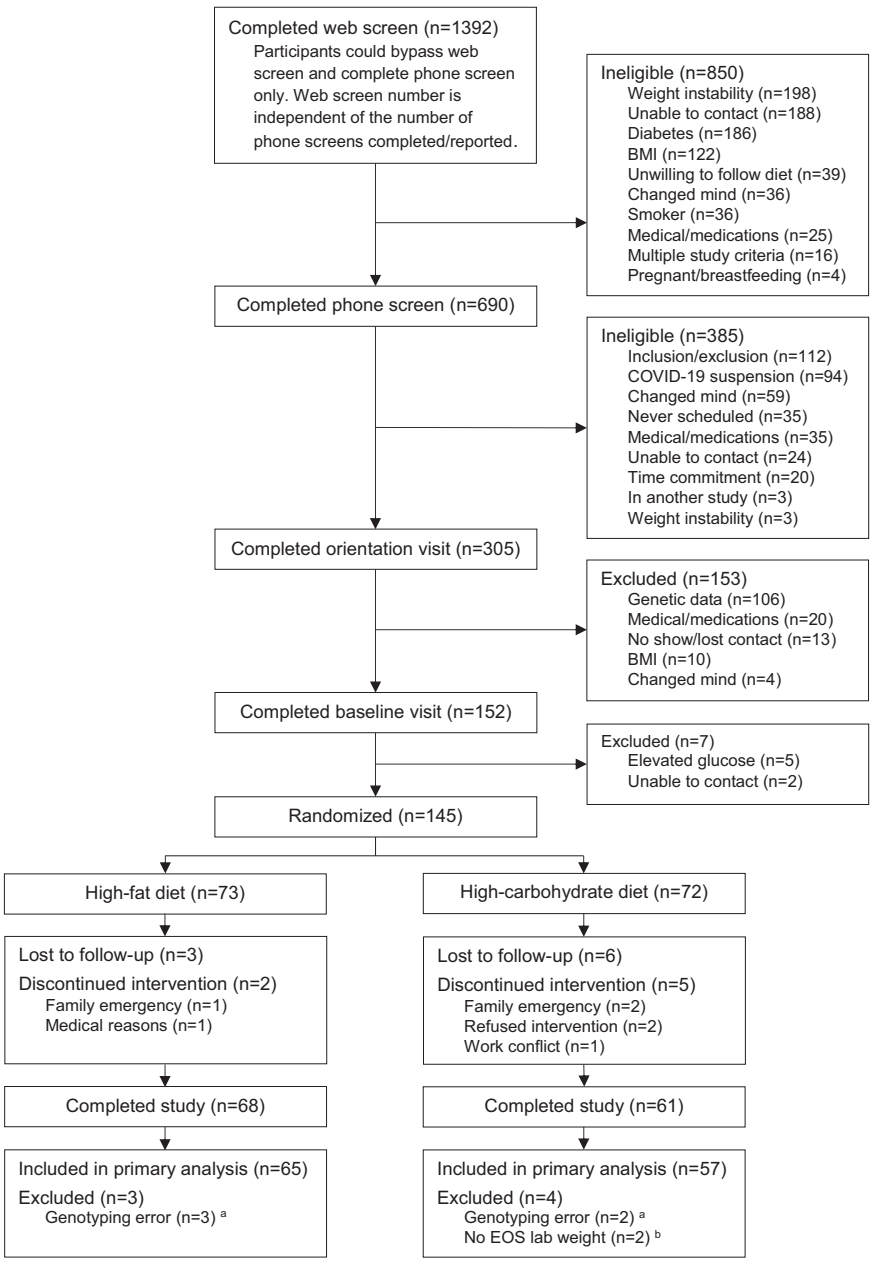


Fig. 1 | CONSORT diagram illustrating the flow of participants through the POINTS trial. [a]An error in the algorithm to determine carbohydrate- and fat-responsive genotypes led to the incorrect classification of these participants. These participants were erroneously enrolled as they did not meet the eligibility criteria. This was reported to the IRB, and, as part of the resolution, their data were removed from the dataset. [b]These participants were unable to attend the W12 visit in person and only completed surveys and questionnaires remotely.

state-of-the-art in nutrigenomics. We also aimed to determine associations between baseline insulin levels and homeostatic model assessment for insulin resistance (HOMA-IR) and differential WL between the diets. These analyses were pursued as previous results were mixed with some studies finding that insulin resistance[12,13] and glucose-stimulated insulin secretion[14] influenced differential weight loss between low-fat and low-carbohydrate diets. In contrast, others found no interaction between glucose-stimulated insulin secretion and diet type on 12-month weight loss[9]. Finally, we examined the diet effects on eating attitudes and behaviors to help elucidate the mechanisms by which any observed differences in WL occurred. As program adherence diminishes over time[15], we chose a 12-week intervention period, which generally has lower attrition (~19%) than 6- (~35%) and 12-month (~54%) programs[16], and short-term WL is associated with long-term results[17,18].

Results

Figure 1 shows the flow of participants through the study. Of the 2082 participants who screened for the study, 305 were eligible following the web/phone screen and were invited to the orientation visit. After eligibility verification based on medical history, medication inventory, and physical measures, 275 remained and completed a genealogy test. Of these 275 individuals, 106 (~39%) were excluded because they had a genotype that was classified as responsive to neither a high-fat nor a high-carbohydrate diet or as responsive to both diets. Of the remaining 169 individuals, 112 (~41%) were fat-responders, and 57 (~20%) were carbohydrate-responders. Before the baseline visit (completed by 152 participants), 17 participants were excluded because we either lost contact between the orientation visit and the baseline visit ($n = 13$) or because participants changed their minds about willingness to participate ($n = 4$). Following the baseline visit, 7 additional participants were excluded due to elevated glucose levels ($n = 5$) or lost contact ($n = 2$). Of the 145 participants randomized, 16 were lost to follow-up (W12), and 129 completed the trial. Seven participants were excluded from the analyses because they were incorrectly genotyped and erroneously enrolled ($n = 5$; removal from dataset suggested by IRB) or failed to provide weight data at W12 ($n = 2$). Baseline characteristics of all 122 included participants (54.4 [SD: 13.2] years, BMI 34.9 [SD: 5.1] kg/m^2, 84% women, 68% White) are provided in Table 1. A comparison of baseline characteristics between non-completers ($n = 16$) and completers ($n = 122$) is provided in Supplementary Table 3.

Change in the primary outcome

Weight change did not differ between genotype-concordant (−5.3 kg [SD: 1.0]) and genotype-discordant diets (−4.8 kg [SD: 1.1]; adjusted difference: −0.6 kg [95% CI: −2.1, 0.9, $p = 0.50$]; Table 2, Fig. 2). Among fat-responders, weight change did not differ between the high-fat (−5.5 kg [SD: 1.2]) and the high-carbohydrate diet (−5.3 kg [SD: 1.3]; adjusted difference: −0.2 kg [95% CI: −2.1, 1.6, $p = 0.78$]; Table 2). Similarly, among carbohydrate-responders, weight change did not differ between the high-carbohydrate (−5.1 kg [SD: 1.6]) and high-fat diet (−4.1 kg [SD: 1.7]; adjusted difference: −1.3 kg [95% CI: −3.9, 1.3, $p = 0.49$]; Table 2). Raw differences are presented in Supplementary Table 5.

Percent weight change and change in body fat and body composition

Similar to absolute weight change, percent weight change (adjusted difference: −0.6% [95% CI: −2.1, 0.9, $p = 0.61$]) and change in body fat (adjusted difference: −0.5% [95% CI: −2.4, 1.4]) did not differ between genotype-concordant and genotype-discordant diets (Table 2, Fig. 2). Among fat-responders, percent weight change (adjusted difference: −0.2% [95% CI: −2.1, 1.7, $p = 0.83$]) and change in body fat (adjusted difference: 0.9% [95% CI: −1.3, 3.0]) did not differ between the high-fat and the high-carbohydrate diet (Table 2). Similarly, among carbohydrate-responders, percent weight change (adjusted difference: −1.2% [95% CI: −4.2, 1.7, $p = 0.57$]) and change in body fat (adjusted difference: −3.4% [95% CI: −7.5, 0.8]) did not differ between the high-carbohydrate and high-fat diet (Table 2). Changes in waist circumference (adjusted difference: −0.5 cm [95% CI: −2.3, 1.3]), hip circumference (adjusted difference: −1.0 cm [95% CI: −3.6, 1.6]), and waist-hip ratio (adjusted difference: 0.00 [95% CI: −0.02, 0.03]) did not differ between genotype-concordant and genotype-discordant diets (Table 2). Raw differences are presented in Supplementary Table 5.

Table 1 | Participant characteristics

		Fat-responders ($n = 85$)		Carbohydrate-responders ($n = 37$)	
	All participants ($N = 122$)	High-fat diet ($n = 44$)	High-carbohydrate diet ($n = 41$)	High-fat diet ($n = 21$)	High-carbohydrate diet ($n = 16$)
Race, *n* (%)					
White	83 (68.0)	30 (68.2)	31 (75.6)	12 (57.1)	10 (62.5)
Black/ African American	36 (29.5)	12 (27.3)	10 (24.4)	9 (42.9)	5 (31.2)
Other	3 (2.5)	2 (4.5)	0 (0.0)	0 (0.0)	1 (6.2)
Sex, *n* (%)					
Female	102 (83.6)	37 (84.1)	35 (85.4)	17 (81.0)	13 (81.2)
Male	20 (16.4)	7 (15.9)	6 (14.6)	4 (19.0)	3 (18.8)
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**
Age, years	54.4 (13.2)	57.4 (11.5)	54.4 (14.2)	49.8 (14.1)	52.4 (13.0)
Weight, kg	94.3 (15.2)	94.2 (14.0)	93.5 (14.4)	95.2 (17.6)	95.6 (18.1)
BMI, kg/m^2	34.9 (5.1)	35.1 (5.0)	34.3 (4.8)	35.8 (5.8)	34.8 (5.3)
Body fat, %	45.1 (9.3)	45.0 (9.4)	45.2 (8.5)	43.8 (11.6)	46.1 (8.4)
Waist circumference, cm	109.0 (12.2)	109.3 (11.8)	108.5 (12.3)	109.3 (11.8)	109.2 (14.8)
Hip circumference, cm	118.9 (12.2)	117.5 (10.7)	118.3 (12.1)	120.1 (12.5)	122.8 (15.8)
Waist-hip ratio	0.92 (0.08)	0.94 (0.09)	0.92 (0.08)	0.91 (0.06)	0.89 (0.10)
SBP, mmHg	121.7 (11.9)	120.5 (11.6)	124.1 (12.9)	121.7 (11.5)	119.4 (10.5)
DBP, mmHg	74.7 (7.4)	75.1 (7.0)	74.5 (7.9)	75.0 (6.1)	73.9 (9.3)

Data are mean (SD) for continuous and n (%) for categorical variables.
BMI body mass index, *DBP* diastolic blood pressure, *SBP* systolic blood pressure, *SD* standard deviation.

https://doi.org/10.1038/s41467-023-41969-1

Table 2 | Change in weight (kg and %), percent body fat, body composition, and blood pressure during the 12-week intervention in those assigned to a diet concordant vs. discordant with the genotype

All participants	Genotype-concordant diet (*n* = 60) Mean (SD)	Genotype-discordant diet (*n* = 62) Mean (SD)	Adjusted difference[a] (95% CI)	*p*-value
Weight change, kg	−5.3 (1.0)	−4.8 (1.1)	−0.6 (−2.1, 0.9)	0.501
Weight change, %	−5.8 (1.0)	−5.4 (1.1)	−0.6 (−2.1, 1.0)	0.605
Change in body fat, %[b]	−1.3 (1.2)	−0.8 (1.3)	−0.5 (−2.4, 1.4)	
Waist circumference, cm	−4.8 (1.1)	−4.3 (1.2)	−0.5 (−2.3, 1.3)	
Hip circumference, cm	−4.6 (1.7)	−3.7 (1.8)	−1.0 (−3.6, 1.6)	
Waist-hip ratio	0.01 (0.00)	0.01 (0.00)	0.00 (−0.02, 0.03)	
SBP, mmHg	1.2 (2.7)	−2.9 (2.9)	4.7 (0.5, 8.8)	
DBP, mmHg	0.4 (1.7)	1.0 (1.9)	−0.1 (−2.8, 2.5)	
Fat-responders	**High-fat diet (*n* = 44) Mean (SD)**	**High-carbohydrate diet (*n* = 41) Mean (SD)**	**Adjusted difference[a] (95% CI)**	***p*-value**
Weight change, kg	−5.5 (1.2)	−5.3 (1.3)	−0.2 (−2.1, 1.6)	0.779
Weight change, %	−5.9 (1.3)	−5.7 (1.4)	−0.2 (−2.1, 1.7)	0.831
Change in body fat, %[c]	−1.1 (1.4)	−1.9 (1.6)	0.9 (−1.3, 3.0)	
Waist circumference, cm	−5.0 (1.4)	−4.4 (1.5)	−0.6 (−2.7, 1.5)	
Hip circumference, cm	−3.9 (1.5)	−4.0 (1.7)	0.2 (−2.1, 2.6)	
Waist-hip ratio	0.00 (0.00)	0.00 (0.00)	−0.01 (−0.03, 0.02)	
SBP, mmHg	4.5 (3.2)	−1.2 (3.5)	6.9 (2.0, 11.8)	
DBP, mmHg	1.7 (2.2)	2.9 (2.4)	−0.5 (−3.8, 2.9)	
Carbohydrate-responders	**High-carbohydrate diet (*n* = 16) Mean (SD)**	**High-fat diet (*n* = 21) Mean (SD)**	**Adjusted difference[a] (95% CI)**	***p*-value**
Weight change, kg	−5.1 (1.6)	−4.1 (1.7)	−1.3 (−3.9, 1.3)	0.487
Weight change, %	−5.7 (1.8)	−4.8 (1.9)	−1.2 (−4.2, 1.7)	0.565
Change in body fat, %[d]	−1.9 (2.5)	1.4 (2.7)	−3.4 (−7.5, 0.8)	
Waist circumference, cm	−4.4 (2.1)	−4.2 (2.3)	−0.3 (−3.9, 3.3)	
Hip circumference, cm	−6.4 (4.3)	−2.7 (4.7)	−3.9 (−11.1, 3.3)	
Waist-hip ratio	0.00 (0.00)	0.00 (0.00)	0.03 (−0.02, 0.08)	
SBP, mmHg	−5.8 (4.6)	−7.2 (5.0)	0.3 (−7.4, 8.0)	
DBP, mmHg	−2.0 (2.8)	−3.0 (3.0)	0.9 (−3.7, 5.5)	

CI confidence interval, *DBP* diastolic blood pressure, *SBP* systolic blood pressure, *SD* standard deviation.
[a]Mixed-effect model, adjusted for sex, race, and baseline value of the outcome for all data.
[b]Data available for 58 of 60 participants (genotype-concordant diet) and 60 of 62 participants (genotype-discordant diet).
[c]Data available for 42 of 44 participants (high-fat diet) and 40 of 41 participants (high-carbohydrate diet).
[d]Data available for 16 of 16 participants (high-carbohydrate diet) and 20 of 21 participants (high-fat diet).

Change in blood pressure

Changes in resting systolic blood pressure (SBP) and DBP did not differ between genotype-concordant and genotype-discordant diets (SBP adjusted difference: 4.7 mmHg [95% CI: 0.5, 8.8]; DBP adjusted difference: −0.1 mmHg [95% CI: −2.8, 2.5]; Table 2, Fig. 3). Similarly, changes in SBP and DBP did not differ between the high-fat and the high-carbohydrate diet among fat-responders (SBP difference: 6.9 mmHg [95% CI: 2.0, 11.8]; DBP difference: −0.5 mmHg [95% CI: −3.8, 2.9]) or between the high-carbohydrate and high-fat diet among carbohydrate responders (SBP difference: 0.3 mmHg [95% CI: −7.4, 8.0]; DBP difference: 0.9 mmHg [95% CI: −3.7, 5.5]; Table 2). Raw differences are presented in Supplementary Table 5.

Association between insulin levels and HOMA-IR and weight loss

Baseline insulin levels ($\beta = -0.036$ [95% CI: −0.125, 0.053, $p = 0.43$]) and HOMA-IR ($\beta = -0.165$ [95% CI: −0.505, 0.175, $p = 0.34$]) were not associated with weight change (Supplementary Figure 1). There was no diet × baseline HOMA-IR interaction on weight change ($p = 0.37$). Similarly, there was no significant diet × baseline HOMA-IR interaction among carbohydrate-responders ($p = 0.62$) or fat-responders ($p = 0.23$; Supplementary Fig. 2).

Change in food cravings, appetitive traits, and food preferences

Changes in food cravings did not differ between the genotype-concordant and genotype-discordant diets (Table 3). Among carbohydrate-responders, those on a high-fat diet decreased cravings for carbohydrates/starches relative to those on the high-carbohydrate diet with an adjusted difference of −0.7 (95% CI: −1.1, −0.4, $p = 0.006$, without Holm-Bonferroni adjustment $p = 0.001$). Changes in all other food cravings did not differ between diets among carbohydrate-responders (Table 3). Among fat-responders, changes in food cravings did not differ between diets (Table 3). Raw differences are presented in Supplementary Table 6. Changes in restraint, disinhibition, and hunger (via EI), and food preferences (FPQ) did not differ between genotype-concordant and genotype-discordant diets (Table 4). Raw differences are presented in Supplementary Table 7 and baseline scores in these instruments are reported in Supplementary Table 4.

Diet personalization and intervention satisfaction

Diet preference (via Diet Personalization Survey, Table 5) and intervention satisfaction (Table 6) did not differ between the genotype-concordant and genotype-discordant diets. Raw differences are presented in Supplementary Table 8.

https://doi.org/10.1038/s41467-023-41969-1

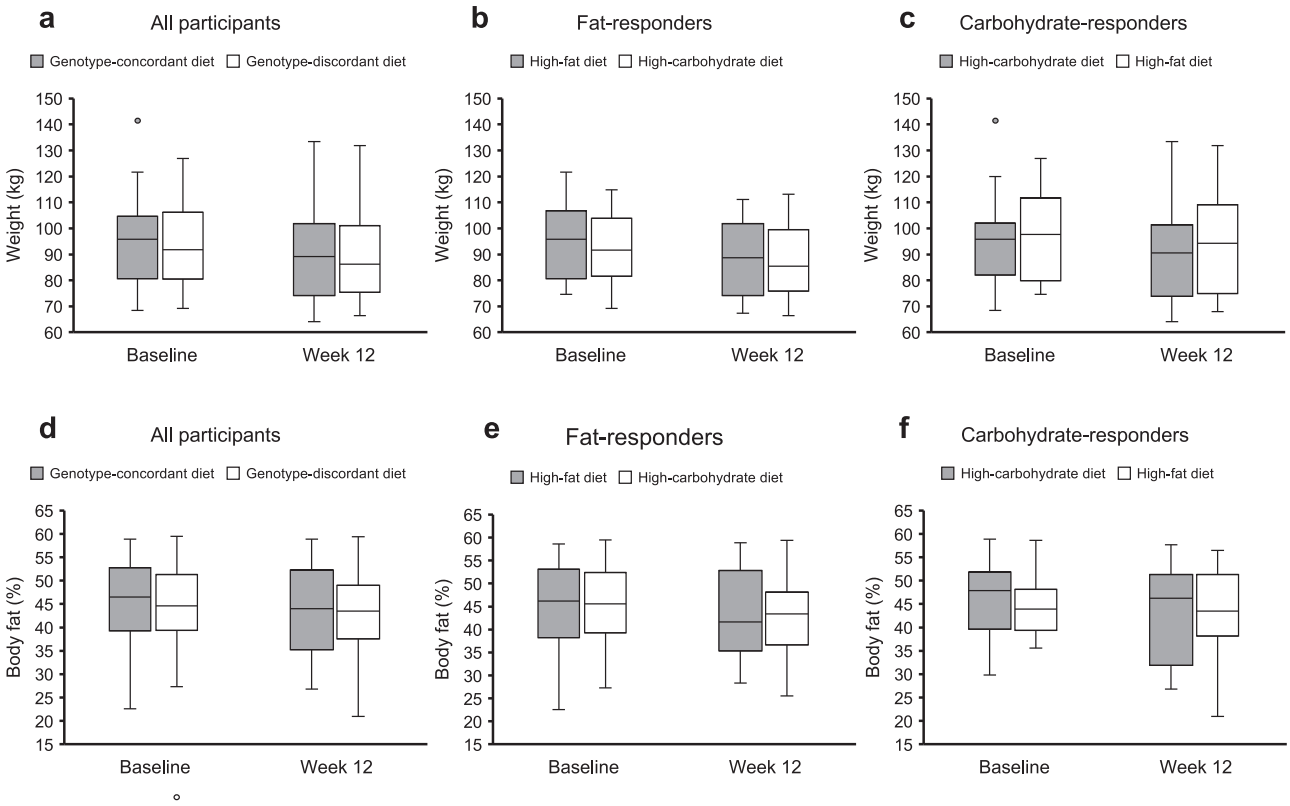


Fig. 2 | Change in weight and percent body fat during the 12-week intervention. Results are presented as boxplots for all participants (**a**, **d**), as well as for fat-responders (**b**, **e**) and carbohydrate responders (**c**, **f**) separately. **a** Genotype-concordant group ($n = 60$, genotype-discordant group ($n = 62$); **b** high-fat diet ($n = 44$), high-carbohydrate diet ($n = 41$); **c** high-carbohydrate diet ($n = 16$), high-fat diet ($n = 21$); **d** genotype-concordant group ($n = 58$), genotype-discordant group ($n = 60$); **e** high-fat diet ($n = 42$), high-carbohydrate diet ($n = 40$); **f** high-carbohydrate diet ($n = 16$), high-fat diet ($n = 20$). In the boxplots, the center line denotes the median value (50th percentile), the bounds of the box represent the 25th and 75th percentiles of the dataset, and the whiskers mark the 5th and 95th percentiles.

Diet adherence

Adherence to the assigned diets is shown in Fig. 4. We encountered difficulties in obtaining the adherence data from participants due, in part, to the pandemic and needing to move to remote intervention delivery. Consequently, these adherence data are only available for 22 of 57 participants (39%) on the high-carbohydrate diet and for 43 of 65 participants (66%) on the high-fat diet (the discrepancy in the percent complete/missing is noted, though we have no reason to believe that it was systematic). On average, participants on the high-carbohydrate diet reported consuming 63.4% (SD: 2.3) of their energy from carbohydrates (target 65%), 20.9% (SD: 2.4) from fat (target 20%), and 16.0% (SD: 1.0) from protein (target 15%) in week 4, 63.3% (SD: 2.8) from carbohydrates, 20.5% (SD: 1.7) from fat, and 15.9% (SD: 1.0) from protein in week 8, and 62.7% (SD: 4.0) from carbohydrates, 20.5% (SD: 2.5) from fat, and 15.7% (SD: 1.8) from protein in week 12. Participants on the high-fat diet reported consuming on average 45.4% (SD: 2.2) of their energy from carbohydrates (target 45%), 39.4% (SD: 2.0) from fat (target 40%), and 15.8% (SD: 1.2) from protein (target 15%) in week 4, 44.7% (SD: 2.2) from carbohydrates, 40.5% (SD: 2.1) from fat, and 15.7% (SD: 2.3) from protein in week 8, and 44.5% (SD: 3.4) from carbohydrates, 39.9% (SD: 2.5) from fat, and 16.1% (SD: 3.3) from protein in week 12.

Session attendance and adverse events

Weekly attendance was similar across the four genotype-diet groups (Supplementary Table 9), with weekly session attendance ranging from 85% to 100%. There were 4 adverse or serious adverse events in total. Two adverse events occurred among fat-responders on a high-carbohydrate diet (unrelated to the study), and there were 2 serious adverse events (1 among fat-responders on a high-carbohydrate diet, 1 among fat-responders on a high-fat diet) that required hospitalization (unrelated to study).

Discussion

The present RCT determined the participant's (fat-responsive or carbohydrate-responsive) genotype a priori via a comprehensive genetic risk score based on published and validated effects and tested the effects of a genotype-concordant diet on WL over 12 weeks. We found no difference in WL between individuals on the genotype-concordant vs. genotype-discordant diet. Further, insulin levels or HOMA-IR were not associated with WL. Food cravings tended to decrease among carbohydrate-responders on a high-fat diet compared to those on a high-carbohydrate diet. Finally, fat-responders on a high-carbohydrate diet tended to decrease resting SBP.

The lack of significant and clinically meaningful differences in WL (-0.6 kg) between genotype-concordant and genotype-discordant diets aligns with the literature[9,11]. In contrast to the well-conducted Gardner et al. study (non-significant difference in WL of 0.7 kg over 12 months)[9], who defined fat vs. carbohydrate-responsive genotypes based on 3 SNPs that were predictive in a preliminary retrospective analysis[8], we determined fat- or carbohydrate-responsive genotypes based on an algorithm involving 10 SNPs. Supported by a recent-meta-analysis (8 trials with 91 SNPs and 63 genetic loci)[11], our results suggest that with the current ability to genotype individuals as fat or carbohydrate-responders, there is no evidence that genotype-concordant diets result in greater WL.

Our sample consisted of substantially fewer carbohydrate-responders ($n = 37$) than fat-responders ($n = 85$). We did not limit recruitment to achieve equal numbers of participants in each genotype-diet group, and this distribution reflects the prevalence in our

https://doi.org/10.1038/s41467-023-41969-1

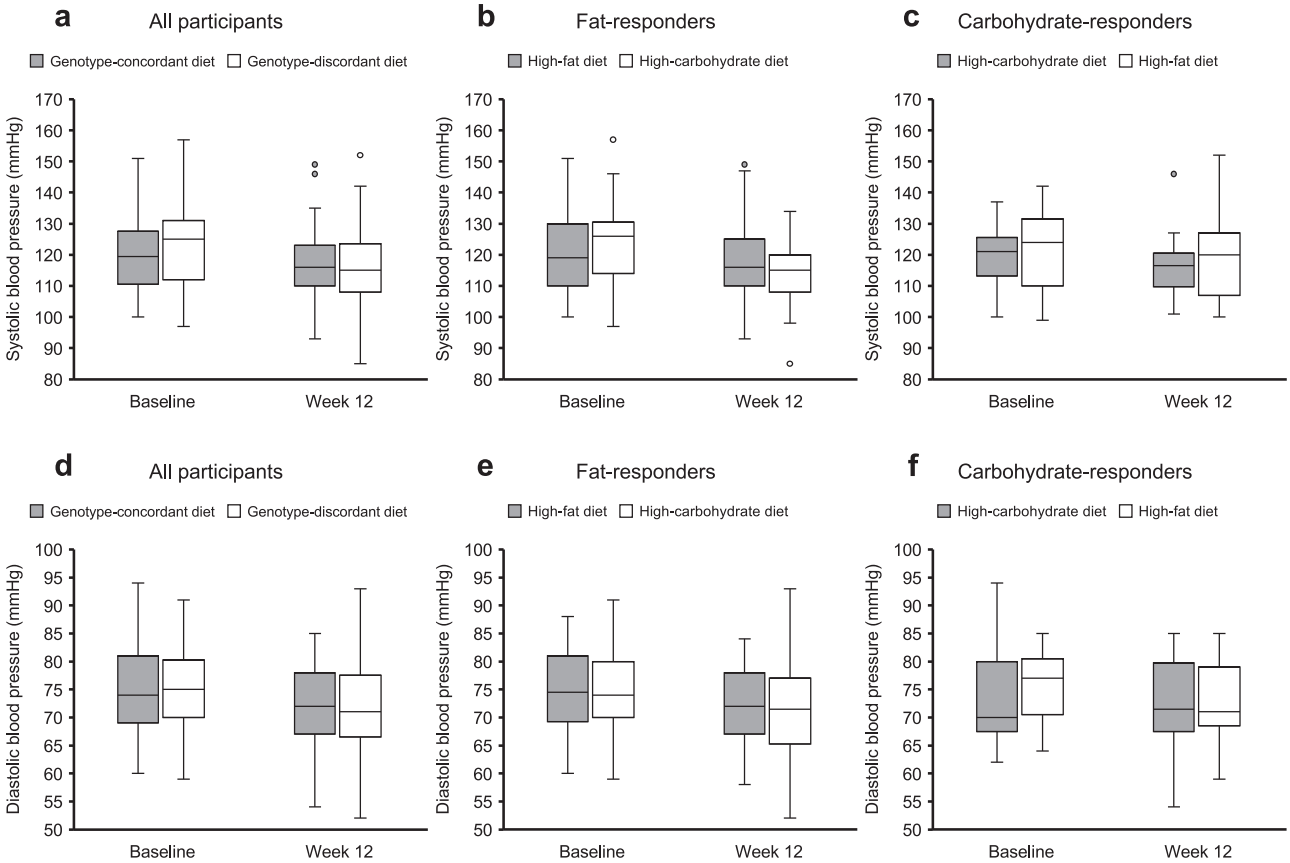


Fig. 3 | Change in systolic and diastolic blood pressure during the 12-week intervention. Results are presented as boxplots for all participants (**a**, **d**; genotype-concordant group, $n = 60$, genotype-discordant group, $n = 62$), as well as for fat-responders (**b**, **e**; high-fat diet, $n = 44$, high-carbohydrate diet, $n = 41$) and carbohydrate responders (**c**, **f**; high-carbohydrate diet, $n = 16$, high-fat diet, $n = 21$) separately. In the boxplots, the center line denotes the median value (50th percentile), the bounds of the box represent the 25th and 75th percentiles of the dataset, and the whiskers mark the 5th and 95th percentiles.

population. As reported in the Results section, 275 individuals completed a genealogy test, of which ~39% had a genotype classified as responsive to both or neither of the two diets, ~41% were fat-responders and ~20% were carbohydrate-responders. Notably, these numbers are somewhat different from what we had estimated during the study's planning phase, as we expected 1/3 of people to be fat-responders, 1/3 carbohydrate-responders, and 1/3 to respond to neither or both of the specified diets. Future studies with larger samples should verify if this uneven distribution between carbohydrate-responders and fat-responders is representative of the general population and further investigate the potential effect on WL among carbohydrate-responders.

Future studies could also consider assigning participants to genotype-concordant diets without specific energy intake targets and examine the diet effects not only on WL but also on cardiovascular risk factors. Previously, a low-carbohydrate diet without energy intake target resulted in greater improvements in body composition, blood lipids, and estimated 10-year coronary heart disease risk compared to a low-fat diet[19]. It would be insightful to investigate whether genotype plays a role in cardiovascular risk reduction following a low-carbohydrate vs. low-fat diet without calorie restriction.

Fasting insulin levels and HOMA-IR did not predict WL. Previous studies reporting a diet × fasting insulin interaction for WL found lower carbohydrate diets to be superior for individuals with greater insulin resistance[13] and high baseline insulin secretion (30 min after a 75 g oral glucose tolerance test)[20], presumably due to a reduced burden on insulin-mediated glucose disposal. However, these studies involved relatively small sample sizes, and findings of the influence of insulin sensitivity[21] and insulin secretion[9,14] on WL via a low-fat vs. a low-carbohydrate diet are inconsistent.

WL can reduce food cravings, particularly for foods restricted on specific diets[22], contributing to the hypothesis that food cravings are a conditioned expression of hunger due to stimuli paired with eating certain foods[23]. Consequently, cravings can be reduced by eliminating or restricting the intake of craved foods. This hypothesis is partially supported by our results as, among carbohydrate-responders, cravings tended to decrease for high-carbohydrate foods on the high-fat diet. Nonetheless, cravings also decreased modestly for high-fat foods, which is to be expected as the amount of all foods was restricted, and cravings for specific foods correlate with each other[24].

Among fat-responders, a high-carbohydrate diet tended to decrease resting SBP. Nonetheless, these individuals had the highest mean SBP of the 4 genotype-diet groups at baseline. Thus, this effect could be explained, in whole or partially, by regression to the mean. Also, all 4 genotype-diet groups had relatively well-controlled blood pressure, leaving little room for improvement through dietary changes, making the non-significant improvements potentially more meaningful.

This trial has some limitations. First, the genetic algorithm to classify individuals as fat- or carbohydrate-responders was created based on published literature[25–38]. However, these (mostly retrospective) studies generally had modest sample sizes, and some of the genotype × diet interactions, which may be false positives, have not been independently replicated. Further, WL is determined by multiple modifiable and non-modifiable (e.g., genetic) factors, and current knowledge accounts for a small percentage of the variability. Further genotypes may have influenced participants' WL responses in directions different from those predicted from the measured genotypes. More comprehensive knowledge of the role of

Table 3 | Changes in food cravings (via the Food Craving Inventory) during the 12-week intervention in those assigned to a diet concordant vs. discordant with the genotype

All participants	Genotype-concordant diet (*n* = 60) Mean (SD)	Genotype-discordant diet (*n* = 62) Mean (SD)	Adjusted difference[a] (95% CI)
High fats[b]	−0.3 (0.1)	−0.4 (0.2)	0.1 (−0.1, 0.4)
Sweets[c]	−0.3 (0.2)	−0.5 (0.2)	0.2 (−0.1, 0.4)
Carbohydrates/Starches[d]	−0.1 (0.2)	−0.4 (0.2)	0.3 (0.0, 0.5)
Fast-food fats[e]	−0.3 (0.2)	−0.4 (0.2)	0.1 (−0.2, 0.4)
Fruits and vegetables[f]	−0.1 (0.2)	−0.4 (0.2)	0.2 (−0.1, 0.5)
Total cravings[g]	−0.2 (0.1)	−0.4 (0.1)	0.2 (−0.1, 0.4)
Fat-responders	**High-fat diet (*n* = 44) Mean (SD)**	**High-carbohydrate diet (*n* = 41) Mean (SD)**	**Adjusted difference[a] (95% CI)**
High fats[b]	−0.4 (0.2)	−0.3 (0.2)	−0.1 (−0.3, 0.3)
Sweets[c]	−0.4 (0.2)	−0.6 (0.2)	0.2 (−0.1, 0.5)
Carbohydrates/Starches[d]	−0.2 (0.2)	−0.3 (0.2)	0.1 (−0.3, 0.4)
Fast-food fats[e]	−0.4 (0.2)	−0.3 (0.3)	−0.1 (−0.4, 0.3)
Fruits and vegetables[f]	−0.3 (0.2)	−0.4 (0.3)	0.1 (−0.3, 0.5)
Total cravings[g]	−0.3 (0.2)	−0.3 (0.2)	0.0 (−0.3, 0.3)
Carbohydrate-responders	**High-carbohydrate diet (*n* = 16) Mean (SD)**	**High-fat diet (*n* = 21) Mean (SD)**	**Adjusted difference[a] (95% CI)**
High fats[b]	−0.2 (0.2)	−0.7 (0.2)	0.5 (0.1, 0.9)
Sweets[c]	−0.1 (0.2)	−0.3 (0.3)	0.2 (−0.2, 0.6)
Carbohydrates/Starches[d]	0.1 (0.2)	−0.7 (0.2)	0.7 (0.4, 1.1)
Fast-food fats[e]	−0.1 (0.3)	−0.5 (0.3)	0.5 (0.0, 1.0)
Fruits and vegetables[f]	0.3 (0.3)	−0.3 (0.3)	0.6 (0.1, 1.1)
Total cravings[g]	0.0 (0.2)	−0.5 (0.2)	0.5 (0.2, 0.9)

CI confidence interval, *SD* standard deviation.

[a]Adjusted for sex, race, and baseline value of the outcome.

[b]Genotype-concordant diet: 55/60 participants; genotype-discordant diet: 60/62 participants. Fat-responders: 41/44 participants (high-fat diet) and 40/41 participants (high-carbohydrate diet). Carbohydrate-responders: 14/16 participants (high-carbohydrate diet) and 20/21 participants (high-fat diet).

[c]Genotype-concordant diet: 59/60 participants; genotype-discordant diet: 60/62 participants. Fat-responders: 43/44 participants (high-fat diet) and 40/41 participants (high-carbohydrate diet). Carbohydrate-responders: 16/16 participants (high-carbohydrate diet) and 20/21 participants (high-fat diet).

[d]Genotype-concordant diet: 59/60 participants; genotype-discordant diet: 61/62 participants. Fat-responders: 44/44 participants (high-fat diet) and 40/41 participants (high-carbohydrate diet). Carbohydrate-responders: 15/16 participants (high-carbohydrate diet) and 21/21 participants (high-fat diet).

[e]Genotype-concordant diet: 58/60 participants; genotype-discordant diet: 61/62 participants. Fat-responders: 43/44 participants (high-fat diet) and 40/41 participants (high-carbohydrate diet). Carbohydrate-responders: 15/16 participants (high-carbohydrate diet) and 20/21 participants (high-fat diet).

[f]Genotype-concordant diet: 58/60 participants; genotype-discordant diet: 60/62 participants. Fat-responders: 43/44 participants (high-fat diet) and 40/41 participants (high-carbohydrate diet). Carbohydrate-responders: 15/16 participants (high-carbohydrate diet) and 20/21 participants (high-fat diet).

[g] Genotype-concordant diet: 54/60 participants; genotype-discordant diet: 67/62 participants. Fat-responders: 41/44 participants (high-fat diet) and 38/41 participants (high-carbohydrate diet). Carbohydrate-responders: 13/16 participants (high-carbohydrate diet) and 19/21 participants (high-fat diet).

genetics in WL is needed and should be obtained from genome-wide association studies; however, the sample size and experimental design required to generate that essential information are beyond reach at this time. Additional limitations of the present study include the relatively small sample size, single-center design, and short time frame. A longer timeframe (6–12-month follow-up) may have increased the amount and differential weight loss between diets. A larger sample size might have also allowed for detecting differences in clinically important secondary outcomes such as changes in body fat and SBP. Further, we did not provide meals in this study, which may have affected dietary adherence (high-fat vs. high-carbohydrate). However, this choice was made by design, as our study was designed as a (pragmatic) effectiveness trial with real-world conditions rather than an efficacy trial. Additionally, the adherence data (albeit limited) suggests that diet adherence was overall satisfactory. In addition to assessing diet adherence continuously throughout the study, future studies should also assess the macronutrient composition of participants' habitual diets to see any differences in the magnitude of the shifts from baseline to the high-fat or high-carbohydrate diet. Further, when assessing a potential effect modification by insulin resistance status, using an oral glucose tolerance test (AUC or INS-30) rather than HOMA-IR to quantify insulin resistance might have been a better option, as HOMA-IR has limited sensitivity due to its reliance on fasting insulin and glucose levels and it does not reflect differences between tissues (e.g., adipose, muscle) or postprandial physiology. Non-fasting methods yield greater variability of the glucose/insulin dynamics and may have been more suitable. Additionally, the assessment of percent body fat via BIA is a limitation as BIA does not provide information on body fat distribution. Finally, participation in "nutri-genomics" studies generally induces improved diet adherence[39–42], independent of the specific recommendations. Therefore, in our study, participants may have responded better to their assigned diets regardless of their genotype matching, obscuring the specific nutri-genomics effects.

In conclusion, in this 12-week RCT, there was no difference in WL between individuals with an a priori determined fat- or carbohydrate-responsive genotype on a high-carbohydrate vs. high-fat diet with specific energy targets and the same level of energy restriction across diets.

Methods

Design and participants

The Personalized Nutrition Study (POINTS, ClinicalTrials.gov identifier: NCT04145466) was a 12-week, single-site, parallel-arm WL trial that was approved by the institutional review board (IRB FWA 00006218) of the

Table 4 | Change in restraint, disinhibition, and hunger and in food preferences during the 12-week intervention in those assigned to a diet concordant vs. discordant with the genotype

All participants	Genotype-concordant diet (*n* = 60) Mean (SD)	Genotype-discordant diet (*n* = 62) Mean (SD)	Adjusted difference[a] (95% CI)
Restraint (EI)[b]	3.6 (0.9)	3.3 (1.0)	0.4 (−1.1, 1.9)
Disinhibition (EI)[c]	−0.1 (0.6)	0.1 (0.7)	0.0 (−1.0, 0.9)
Hunger (EI)[d]	−0.4 (0.5)	−0.9 (0.6)	0.5 (−0.4, 1.4)
HF/HS (FPQ)	−0.1 (0.3)	0.0 (0.4)	0.0 (−0.5, 0.5)
LF/HS (FPQ)	0.1 (0.3)	0.1 (0.3)	0.1 (−0.4, 0.5)
HF/HCCHO (FPQ)	−0.3 (0.3)	−0.3 (0.3)	0.0 (−0.4, 0.5)
LF/HCCHO (FPQ)	−0.1 (0.3)	0.0 (0.3)	−0.1 (−0.5, 0.4)
HF/LCHO/HP (FPQ)	−0.4 (0.3)	−0.4 (0.3)	0.0 (−0.5, 0.4)
LF/LCHO/HP (FPQ)	0.1 (0.3)	0.1 (0.3)	0.0 (−0.4, 0.4)
Fat-responders	**High-fat diet (*n* = 44) Mean (SD)**	**High-carbohydrate diet (*n* = 41) Mean (SD)**	**Adjusted difference[a] (95% CI)**
Restraint (EI)[b]	3.5 (1.2)	2.7 (1.4)	0.8 (−1.3, 2.9)
Disinhibition (EI)[c]	−0.3 (0.8)	0.2 (0.9)	−0.4 (−1.6, 0.9)
Hunger (EI)[d]	−0.9 (0.7)	−1.3 (0.8)	0.4 (−0.8, 1.5)
HF/HS (FPQ)	0.0 (0.4)	0.0 (0.5)	0.1 (−0.6, 0.7)
LF/HS (FPQ)	0.2 (0.4)	0.3 (0.4)	0.0 (−0.6, 0.5)
HF/HCCHO (FPQ)	−0.2 (0.4)	−0.1 (0.4)	0.0 (−0.6, 0.5)
LF/HCCHO (FPQ)	0.0 (0.4)	0.2 (0.4)	−0.2 (−0.7, 0.4)
HF/LCHO/HP (FPQ)	−0.5 (0.4)	−0.6 (0.4)	0.1 (−0.5, 0.7)
LF/LCHO/HP (FPQ)	0.2 (0.4)	0.1 (0.4)	0.1 (−0.4, 0.7)
Carbohydrate-responders	**High-carbohydrate diet (*n* = 16) Mean (SD)**	**High-fat diet (*n* = 21) Mean (SD)**	**Adjusted difference[a] (95% CI)**
Restraint (EI)[b]	3.4 (1.1)	4.6 (1.1)	−0.7 (−2.6, 1.2)
Disinhibition (EI)[c]	0.7 (0.9)	0.0 (0.9)	0.8 (−0.8, 2.5)
Hunger (EI)[d]	0.8 (0.8)	−0.1 (0.9)	1.0 (−0.4, 2.5)
HF/HS (FPQ)	−0.2 (0.5)	0.0 (0.5)	−0.1 (−1.0, 0.7)
LF/HS (FPQ)	−0.1 (0.4)	−0.4 (0.4)	0.3 (−0.3, 0.9)
HF/HCCHO (FPQ)	−0.4 (0.5)	−0.6 (0.5)	0.2 (−0.6, 1.0)
LF/HCCHO (FPQ)	−0.2 (0.4)	−0.5 (0.4)	0.3 (−0.3, 0.9)
HF/LCHO/HP (FPQ)	−0.3 (0.4)	0.1 (0.5)	−0.2 (−1.0, 0.5)
LF/LCHO/HP (FPQ)	−0.1 (0.4)	0.1 (0.4)	−0.2 (−0.8, 0.4)

CI confidence interval, *EI* Eating Inventory, *FPQ* Food Preference Questionnaire, *HF/HS* high fat/high simple sugar, *LF/HS* low fat/high simple sugar, *HF/HCCHO* high fat/high complex carbohydrate, *LF/HCCHO* low fat/high complex carbohydrate, *HF/LCHO/HP* high fat/low carbohydrate/high protein, *LF/LCHO/HP* low fat/low carbohydrate/high protein, *SD* standard deviation.
[a]Adjusted for sex, race, and baseline value of the outcome.
[b]Genotype-concordant diet: 46/60 participants; genotype-discordant diet: 47/62 participants. Fat-responders: 34/44 participants (high-fat diet) and 29/41 participants (high-carbohydrate diet). Carbohydrate-responders: 12/16 participants (high-carbohydrate diet) and 18/21 participants (high-fat diet).
[c]Genotype-concordant diet: 49/60 participants; genotype-discordant diet: 49/62 participants. Fat-responders: 37/44 participants (high-fat diet) and 31/41 participants (high-carbohydrate diet). Carbohydrate-responders: 12/16 participants (high-carbohydrate diet) and 18/21 participants (high-fat diet).
[d]Genotype-concordant diet: 51/60 participants; genotype-discordant diet: 51/62 participants. Fat-responders: 37/44 participants (high-fat diet) and 33/41 participants (high-carbohydrate diet). Carbohydrate-responders: 14/16 participants (high-carbohydrate diet) and 18/21 participants (high-fat diet).

Pennington Biomedical Research Center (PBRC, Baton Rouge, LA). Participants were enrolled between October 7, 2020 and September 8, 2021. Participants were identified a priori as carbohydrate-responders and fat-responders based on their combined genotypes at 10 genetic variant loci and randomized to either a high-carbohydrate or high-fat diet, yielding the following groups: (1) fat-responders receiving a high-fat diet, (2) fat-responders receiving a high-carbohydrate diet, (3) carbohydrate-responders receiving a high-fat diet, and (4) carbohydrate-responders receiving a high-carbohydrate diet.

Participants were recruited from the community. Eligible participants were 18–75 years old, had a BMI of 27.0–47.5 kg/m^2, and had completed or were willing to complete a genealogy test (e.g., Ancestry, 23andMe) and to share the raw data with the investigators. Finally, a genetic profile indicating a predisposition to respond favorably to a high-carbohydrate or high-fat WL diet based on specific SNPs (see below) was required. Exclusion criteria included smoking, weight change ≥10 lbs. in the last 3 months, being pregnant or breastfeeding, conditions, diseases, or medications that affect body weight or metabolism or could affect risk or study completion, and a genotype indicating a predisposition to respond favorably to neither or both of the specified diets. We estimated that approximately 1/3 of people would be fat-responders, 1/3 carbohydrate-responders, and 1/3 would respond favorably to neither or both of the specified diets.

The study included 1 orientation visit, 2 clinic visits (one before and one after the intervention), and 12-weekly intervention sessions. All participants provided written informed consent, and participants who completed the study received a minor compensation of $150.

Genotype determination. Carbohydrate- and fat-responders were identified a priori based on their combined genotypes at the following genetic variants: (1) *FGF21rs838147*[25], (2) *TCF7L2rs12255372*[26,43], (3) *IRS1rs2943641*[28], (4) *APOA5rs662799*[30,31,44], (5) *PLIN1rs894160*[27,32], (6) *APOA2rs5082*[29,33], (7) *FTOrs9939609*[34,35], (8) *PPARGrs1801282*[36], (9) *GIPRrs10423928*[37], and (10) *GYS2rs1478290*[38]. The genetic information

Table 5 | Change in items of the Diet Personalization Survey during the 12-week intervention in those assigned to a diet concordant vs. discordant with the genotype

All participants	Genotype-concordant diet (*n* = 60) Mean[a] (SD)	Genotype-discordant diet (*n* = 62) Mean[a] (SD)	Adj. difference[b] (95% CI)
The assigned diet…			
… fits my typical eating habits	0.9 (0.5)	1.3 (0.6)	−0.3 (−1.2, 0.6)
… fits my lifestyle	0.4 (0.6)	0.2 (0.6)	0.2 (−0.7, 1.1)
… makes it easier to lose weight	0.6 (0.5)	0.7 (0.6)	0.1 (−0.8, 0.8)
I am confident that I can…			
… successfully lose weight on the assigned diet	0.4 (0.5)	0.6 (0.5)	−0.1 (−0.8, 0.7)
… follow the assigned diet	−0.7 (0.4)	−0.4 (0.5)	−0.3 (−0.9, 0.4)
Fat-responders	**High-fat diet (*n* = 44) Mean[a] (SD)**	**High-carbohydrate diet (*n* = 41) Mean[a] (SD)**	**Adj. difference[b] (95% CI)**
The assigned diet…			
… fits my typical eating habits	0.6 (0.6)	1.3 (0.7)	−0.6 (−1.6, 0.4)
… fits my lifestyle	0.0 (0.7)	0.2 (0.8)	−0.1 (−1.2, 1.0)
… makes it easier to lose weight	0.2 (0.7)	0.4 (0.7)	−0.1 (−1.1, 0.9)
I am confident that I can…			
… successfully lose weight on the assigned diet	0.1 (0.6)	0.4 (0.7)	−0.1 (−1.1, 0.8)
… follow the assigned diet	−0.9 (0.6)	−0.5 (0.6)	−0.4 (−1.2, 0.5)
Carbohydrate-responders	**High-carbohydrate diet (*n* = 16) Mean[a] (SD)**	**High-fat diet (*n* = 21) Mean[a] (SD)**	**Adj. difference[b] (95% CI)**
The assigned diet…			
… fits my typical eating habits	1.6 (1.1)	1.3 (1.2)	0.4 (−1.4, 2.2)
… fits my lifestyle	1.1 (1.0)	0.3 (1.1)	1.0 (−0.7, 2.7)
… makes it easier to lose weight	1.3 (0.7)	1.3 (0.8)	0.4 (−0.8, 1.6)
I am confident that I can…			
… successfully lose weight on the assigned diet	0.9 (0.8)	1.1 (0.9)	0.2 (−1.1, 1.5)
… follow the assigned diet	−0.3 (0.7)	−0.2 (0.7)	0.0 (−1.1, 1.1)
… the degree to which the diet helped manage hunger	6.8 (0.9)	6.6 (1.0)	0.3 (−1.3, 1.8)

[a]Mean change during the 12-week intervention.
[b]Adjusted for sex and race.

was accessed via the raw data from the genealogy tests. Initially, only 6 SNPs were included and pilot tested, and the scoring criteria were then modified as few participants were deemed carbohydrate- or fat-responders. The original and updated scoring criteria, including a specific example for 1 SNP, are provided in the Supplementary Methods, including Supplementary Tables 1 and 2. The final risk score comprised 10 SNPs with demonstrated and validated effects on the responses to high-fat/high-carbohydrate diets[25–38,43,44], and validation of this comprehensive and informative risk score was an objective of this study.

Intervention

After enrollment (Week [W] 0 visit), participants were randomized to either a high-carbohydrate diet (rich in whole-grain foods) or a high-fat diet (rich in unsaturated fats/oils). The high-carbohydrate diet consisted of ~20% of energy from fat and ~65% from carbohydrates, whereas the high-fat diet consisted of ~40% energy from fat and ~45% from carbohydrates. Both diets provided 15% of energy from protein. All participants were assigned an energy intake target that would result in a daily deficit of ~750 kcal and provided with a diet-specific meal plan in 200 kcal increments from 1400 to 2800 kcal/day to self-prepare meals during the intervention period. To facilitate meal plan adherence when preparing or selecting meals, the meal plans included a list of ingredients (and their amounts) for all meals of each day (breakfast, lunch, dinner, and 1 daily snack) and instructions for meal preparation and participants were provided a food scale. Baseline energy requirements were calculated with Mifflin-St. Jeor's formulas[45].

The PBRC biostatistics department created the randomization sequence using SAS 9.4 statistical software for Windows (SAS Institute, Cary, NC) and uploaded it to REDCap (Research Electronic Data Capture). REDCap used strata for the inaction of genotype and gender. To ensure a relatively equal baseline BMI between the 4 genotype-diet groups, a 1:1 randomization scheme was devised that adjusted for BMI, gender, and genotype. Gender and genotype were used as strata, while BMI was used in an a-priori-created randomization equation. Within each stratum, this equation used block sizes of 6 (for females) and 4 (for males) at the start of the study and ended with block sizes of 4 and 2, respectively, to ensure relative balance of group assignments. Block sizes were assigned during the study by the biostatistician with access only to information about the enrolment progress (percent enrolled).

Outcome assessors were blind to diet assignment and genotype patterns. Interventionists administering intervention sessions were blind to genotype patterns but not diet type. Participants were only informed of their genotype (carbohydrate- or fat-responder) once they completed the study.

The 12 weekly intervention (group) sessions were diet-specific and had a different focus each week (Supplementary Material). Participants were provided a body weight scale and encouraged to weigh daily throughout the intervention and to send pictures of their weights to their interventionist before each intervention session. With very few exceptions, the first intervention session was conducted in person. Due to the COVID-19 pandemic, almost all subsequent sessions were conducted virtually via webinar (Microsoft Teams).

Table 6 | Change in intervention satisfaction (post-intervention) in those assigned to a diet concordant vs. discordant with the genotype

All participants	Genotype-concordant diet (*n* = 60) Mean[a] (SD)	Genotype-discordant diet (*n* = 62) Mean[a] (SD)	Adj. difference[b] (95% CI)
I am satisfied with...			
... the group format	6.9 (0.4)	7.4 (0.4)	−0.5 (−1.0, 0.1)
... the support from interventionists	7.5 (0.3)	7.5 (0.3)	0.1 (−0.4, 0.5)
... the intervention materials	7.0 (0.3)	7.2 (0.3)	−0.1 (−0.6, 0.4)
... the support from other participants	6.4 (0.4)	6.5 (0.4)	−0.1 (−0.7, 0.6)
... the amount of food in my meal plan	6.5 (0.5)	6.4 (0.5)	0.1 (− 0.6, 0.8)
... the macronutrient content in my meal plan	6.1 (0.4)	5.8 (0.5)	0.3 (−0.4, 1.0)
... my progress toward weight management	6.4 (0.5)	6.3 (0.5)	0.3 (−0.5, 1.0)
... the degree to which the diet helped manage hunger	6.5 (0.5)	6.1 (0.5)	0.5 (−0.3, 1.2)
Fat-responders	**High-fat diet (*n* = 44) Mean[a] (SD)**	**High-carbohydrate diet (*n* = 41) Mean[a] (SD)**	**Adj. difference[b] (95% CI)**
I am satisfied with...			
... the group format	6.8 (0.4)	7.3 (0.5)	−0.6 (−1.2, 0.1)
... the support from interventionists	7.6 (0.4)	7.8 (0.4)	−0.1 (−0.7, 0.4)
... the intervention materials	7.1 (0.4)	7.4 (0.4)	−0.3 (−0.9, 0.3)
... the support from other participants	6.1 (0.5)	6.5 (0.6)	−0.4 (−1.2, 0.4)
... the amount of food in my meal plan	6.3 (0.6)	6.3 (0.6)	0.0 (−0.9, 0.9)
... the macronutrient content in my meal plan	6.0 (0.5)	5.6 (0.6)	0.3 (−0.5, 1.2)
... my progress toward weight management	6.5 (0.6)	6.2 (0.6)	0.4 (−0.5, 1.2)
... the degree to which the diet helped manage hunger	6.4 (0.6)	5.8 (0.6)	0.6 (−0.3, 1.4)
Carbohydrate-responders	**High-carbohydrate diet (*n* = 16) Mean[a] (SD)**	**High-fat diet (n = 21) Mean[a] (SD)**	**Adj. difference[b] (95% CI)**
I am satisfied with...			
... the group format	7.2 (0.7)	7.4 (0.7)	−0.2 (−1.3, 0.9)
... the support from interventionists	7.5 (0.6)	6.9 (0.6)	0.5 (−0.4, 1.4)
... the intervention materials	7.1 (0.5)	6.6 (0.5)	0.4 (−03, 1.1)
... the support from other participants	7.1 (0.6)	6.4 (0.7)	0.8 (−0.3, 1.8)
... the amount of food in my meal plan	7.0 (0.7)	6.6 (0.8)	0.4 (−0.8, 1.6)
... the macronutrient content in my meal plan	6.3 (0.8)	5.9 (0.9)	0.4 (−1.0, 1.7)
... my progress toward weight management	6.2 (0.9)	6.4 (0.9)	−0.1 (−1.5, 1.4)
... the degree to which the diet helped manage hunger	6.8 (0.9)	6.6 (1.0)	0.3 (−1.3, 1.8)

[a] Mean post-intervention value. The Intervention Satisfaction Survey was only assessed at Week 12.
[b]Adjusted for sex and race.

Outcome measures

Anthropometric data. At W0 and W12, fasting body weight and waist and hip circumference were measured in the PBRC outpatient clinic. Clinic weights were also measured at all intervention visits (though not fasting weights). Further, body fat (%, via bioelectrical impedance analysis [BIA]; X-contact 365, Jawon Medical Co., Ltd, Seoul, South Korea) and blood pressure (after 5 min of seated rest) were measured at W0 and W12.

Fasting serum glucose and insulin. Fasting serum glucose and insulin were measured at W0, and HOMA-IR was used to quantify insulin resistance.

Appetitive traits, food cravings, and food preferences. Appetitive traits were measured with the Eating Inventory (EI)[46], food cravings were measured with the Food Craving Inventory (FCI)[24], and hedonic food preferences were measured with the Food Preference Questionnaire (FPQ)[47] at W0 and W12 (see Supplementary Methods for details on outcome materials). Data for these questionnaires were collected and managed using REDCap tools

Diet personalization and intervention satisfaction. The Diet Personalization Survey (Supplementary Methods) was completed at W0 and W12, as well as during the intervention session at W6, and the Intervention Satisfaction Survey (Supplementary Methods) was conducted at W12. Data for these surveys were collected and managed using REDCap tools.

Diet adherence. As stated above, participants were provided with a kitchen scale and could precisely weigh all ingredients specified in the meal plans for the foods consumed at home. Additional foods that were consumed were weighed and added as well. Adherence to the macronutrient content of the assigned diets was assessed for three 7-day periods throughout the intervention (W4, W8, W12).

Statistical analyses

The distribution of variables was evaluated by visual examination and the Shapiro-Wilk test. The primary outcome was weight change (kg) at 12 weeks. All other measures were secondary endpoints. Changes in outcomes are presented as mean and 95% confidence intervals (CI). We used linear mixed models to determine if changes in outcome variables differed among diets. Covariates in the models included baseline value of the outcome, sex, and race. The mixed-effect model accounted for the correlation of the subject over time, and least-square means based on the estimate from the mixed-effect model were used to test for differences in weight change between diets. To evaluate whether

https://doi.org/10.1038/s41467-023-41969-1

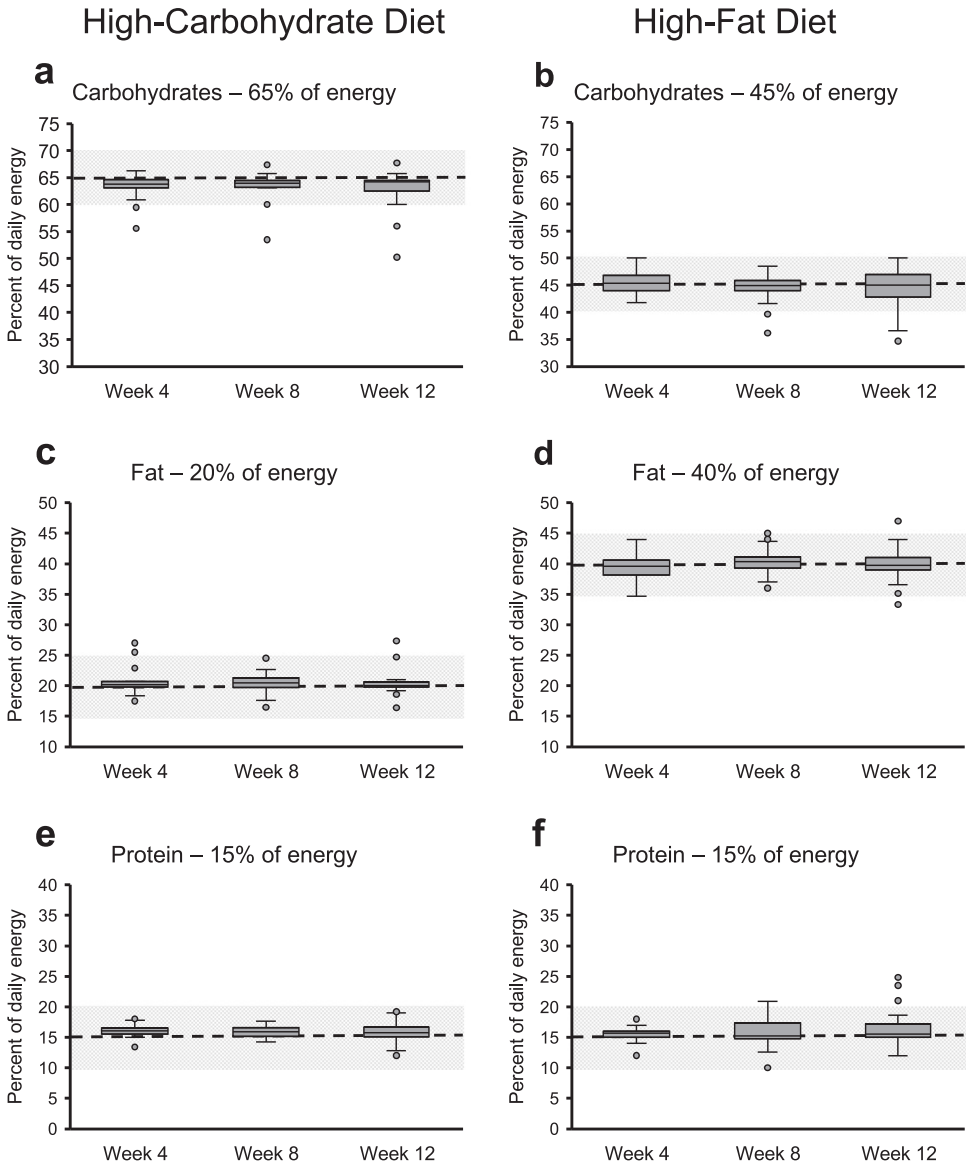


Fig. 4 | Adherence to the macronutrient compositions of the respective diet at week 4, week 8, and week 12. Boxplots showing adherence data for the high-carbohydrate diet (**a**, **c**, **e**) and the high-fat diet (**b**, **d**, **f**). For the high-carbohydrate diet (n = 22 at week 4 and 8 and n = 21 at week 12), target intakes were 65% carbohydrates (**a**), 20% fat (**c**), and 15% protein (**e**) and for the high-fat diet (n = 40 at week 4, n = 38 at week 8, and n = 37 at week 12), they were 45% carbohydrates (**b**), 40% fat (**d**), and 15% protein (**f**). The dashed line shows the target intake with the shaded area representing ±5%. In the boxplots, the center line denotes the median value (50th percentile), the bounds of the box represent the 25th and 75th percentiles of the dataset, and the whiskers mark the 5th and 95th percentiles.

baseline insulin levels and HOMA-IR needed to be included as covariates, their effects on WL were tested using a linear mixed model, adjusted for diet group and other known covariates. Neither baseline insulin levels nor HOMA-IR was significantly associated with WL; hence these variables were not included as covariates. The significance level was set to 0.05 (2-sided). Multiple testing adjustment was performed for secondary outcomes using the Holm-Bonferroni method[48]. All analyses were conducted using SAS (Windows version 9.4; SAS Institute, Cary, NC) and the statistical program R version 4.0.2 (https://cran.r-project.org/).

Power calculations. The present study planned to obtain data from up to 154 participants in total, and we aimed to complete 32 participants per genotype-diet group (128 participants in total) though we did not limit recruitment to achieve equal numbers of participants in each group. We hypothesized that participants on a genotype-concordant diet would lose more weight than those on a genotype-discordant diet. Based on previous studies[49,50], we assumed a standard deviation for between-group differences in weight change of 2.8 kg. To detect a 2.0 kg difference in weight change between group 1 (fat-responders on a high-fat diet) and group 2 (fat-responders on a high-carbohydrate diet) or between group 3 (carbohydrate-responders on a high-fat diet) and group 4 (carbohydrate-responders on a high-carbohydrate diet), with the intended sample size and an alpha level of 0.05, the present study would have 80% power. Further, based on the same assumptions, the present study would have >95% power to test if WL differs between participants on a genotype-concordant diet (groups 1 and 4 combined) and those on a genotype-discordant diet (groups 2 and 3).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All of the data needed to recapitulate the analysis found within this study can be found in the manuscript, figures and supplementary information. Source data are provided with this paper. Due to privacy reasons, de-identified data from the study cannot be shared publicly but will be available from the corresponding author (christoph.hoechsmann@tum.de) immediately following the publication of the paper upon reasonable request. The study protocol and statistical analysis plan will also be available. Source data are provided with this paper.

References

1. Fryar, C. D., Carroll, M. D. & Afful, J. Prevalence of overweight, obesity, and severe obesity among adults aged 20 and over: United States, 1960–1962 through 2017–2018. (2020).
2. Kopelman, P. Health risks associated with overweight and obesity. *Obes. Rev.* **8**, 13–17 (2007).
3. Tremmel, M., Gerdtham, U.-G., Nilsson, P. M. & Saha, S. Economic burden of obesity: a systematic literature review. *Int. J. Environ. Res. Public. Health.* **14**, 435 (2017).
4. Hruby, A. & Hu, F. B. The epidemiology of obesity: a big picture. *Pharmacoeconomics* **33**, 673–689 (2015).
5. Shai, I. et al. Weight loss with a low-carbohydrate, Mediterranean, or low-fat diet. *N. Engl. J. Med.* **359**, 229–241 (2008).
6. Sacks, F. M. et al. Comparison of weight-loss diets with different compositions of fat, protein, and carbohydrates. *N. Engl. J. Med.* **360**, 859–873 (2009).
7. Johnston, B. C. et al. Comparison of weight loss among named diet programs in overweight and obese adults: a meta-analysis. *JAMA.* **312**, 923–933 (2014).
8. Dopler Nelson, M., Prabakar, P., Kondragunta, V., Kornman, K. & Gardner, C. Genetic phenotypes predict weight loss success: the right diet does matter: Paper presented at: joint conference of the 50th Cardiovascular Disease Epidemiology and Prevention and Nutrition, Physical Activity, and Metabolism; March 2–3, 2010, San Francisco, CA.
9. Gardner, C. D. et al. Effect of low-fat vs low-carbohydrate diet on 12-month weight loss in overweight adults and the association with genotype pattern or insulin secretion: The DIETFITS Randomized Clinical Trial. *JAMA.* **319**, 667–679 (2018).
10. Qi, L., Bray, G. A. & Sacks, F. M. Low-fat vs low-carbohydrate diets and weight loss. *JAMA.* **320**, 202–203 (2018).
11. Bayer, S., Winkler, V., Hauner, H. & Holzapfel, C. Associations between genotype–diet interactions and weight loss—a systematic review. *Nutrients.* **12**, 2891 (2020).
12. Cornier, M.-A. et al. Insulin sensitivity determines the effectiveness of dietary macronutrient composition on weight loss in obese women. *Obes. Res.* **13**, 703–709 (2005).
13. McClain, A. D., Otten, J. J., Hekler, E. B. & Gardner, C. D. Adherence to a low-fat vs. low-carbohydrate diet differs by insulin resistance status. *Diabetes Obes. Metab.* **15**, 87–90 (2013).
14. Pittas, A. G. et al. A low-glycemic load diet facilitates greater weight loss in overweight adults with high insulin secretion but not in overweight adults with low insulin secretion in the CALERIE Trial. *Diabetes Care* **28**, 2939–2941 (2005).
15. Martin, C. K. et al. Challenges in defining successful adherence to calorie restriction goals in humans: Results from CALERIE™ 2. *Exp. Gerontol.* **162**, 111757 (2022).
16. Ponzo, V. et al. Predictors of attrition from a weight loss program. A study of adult patients with obesity in a community setting. *Eat. Weight Disord.* **26**, 1729–1736 (2021).
17. Unick, J. L. et al. Weight change in the first 2 months of a lifestyle intervention predicts weight changes 8 years later. *Obesity.* **23**, 1353–1356 (2015).
18. Unick, J. L., Pellegrini, C. A., Demos, K. E. & Dorfman, L. Initial weight loss response as an indicator for providing early rescue efforts to improve long-term treatment outcomes. *Curr. Diab. Rep.* **17**, 69 (2017).
19. Bazzano, L. A. et al. Effects of low-carbohydrate and low-fat diets. *Ann. Intern. Med.* **161**, 309–318 (2014).
20. Ebbeling, C. B., Leidig, M. M., Feldman, H. A., Lovesky, M. M. & Ludwig, D. S. Effects of a low–glycemic load vs low-fat diet in obese young adults: a randomized trial. *JAMA* **297**, 2092–2102 (2007).
21. Gardner, C. D., Offringa, L. C., Hartle, J. C., Kapphahn, K. & Cherin, R. Weight loss on low-fat vs. low-carbohydrate diets by insulin resistance status among overweight adults and adults with obesity: a randomized pilot trial. *Obesity* **24**, 79–86 (2016).
22. Martin, C. K. et al. Change in food cravings, food preferences, and appetite during a low-carbohydrate and low-fat diet. *Obesity* **19**, 1963–1970 (2011).
23. Myers, C. A., Martin, C. K. & Apolzan, J. W. Food cravings and body weight: a conditioning response. *Curr. Opin. Endocrinol. Diabetes Obes.* **25**, 298–302 (2018).
24. White, M. A., Whisenhunt, B. L., Williamson, D. A., Greenway, F. L. & Netemeyer, R. G. Development and validation of the food-craving inventory. *Obes. Res.* **10**, 107–114 (2002).
25. Heianza, Y. et al. Macronutrient intake-associated FGF21 genotype modifies effects of weight-loss diets on 2-year changes of central adiposity and body composition: The POUNDS Lost Trial. *Diabetes Care* **39**, 1909–1914 (2016).
26. Grau, K. et al. TCF7L2 rs7903146-macronutrient interaction in obese individuals' responses to a 10-wk randomized hypoenergetic diet. *Am. J. Clin. Nutr.* **91**, 472–479 (2010).
27. Smith, C. E. et al. Perilipin polymorphism interacts with dietary carbohydrates to modulate anthropometric traits in hispanics of Caribbean origin. *J. Nutr.* **138**, 1852–1858 (2008).
28. Qi, Q. et al. Insulin receptor substrate 1 gene variation modifies insulin resistance response to weight-loss diets in a 2-year randomized trial: the Preventing Overweight Using Novel Dietary Strategies (POUNDS LOST) trial. *Circulation* **124**, 563–571 (2011).
29. Corella, D. et al. APOA2, dietary fat, and body mass index: replication of a gene-diet interaction in 3 independent populations. *Arch. Intern. Med.* **169**, 1897–1906 (2009).
30. Sánchez-Moreno, C. et al. APOA5 gene variation interacts with dietary fat intake to modulate obesity and circulating triglycerides in a Mediterranean population. *J. Nutr.* **141**, 380–385 (2011).
31. Domínguez-Reyes, T. et al. Interaction of dietary fat intake with APOA2, APOA5 and LEPR polymorphisms and its relationship with obesity and dyslipidemia in young subjects. *Lipids Health Dis.* **14**, 106 (2015).
32. Holzbach, L. C., Silveira, A. G. Z., Franco, L. P., Horst, M. A. & Cominetti, C. Polymorphism PLIN1 11482 G>A interacts with dietary intake to modulate anthropometric measures and lipid profile in adults with normal-weight obesity syndrome. *Br. J. Nutr.* **128**, 1004–1012 (2022).
33. Corella, D. et al. Association between the APOA2 promoter polymorphism and body-weight in Mediterranean and Asian populations. Replication of a gene-saturated fat interaction. *Int. J. Obes.* **35**, 666–675 (2011).
34. de Luis, D. A. et al. The rs9939609 gene variant in FTO modified the metabolic response of weight loss after a 3-month intervention with a hypocaloric diet. *J. Investig. Med.* **61**, 22–26 (2013).
35. Xiang, L. et al. FTO genotype and weight loss in diet and lifestyle interventions: a systematic review and meta-analysis. *Am. J. Clin. Nutr.* **103**, 1162–1170 (2016).
36. Garaulet, M., Smith, C. E., Hernández-González, T., Lee, Y.-C. & Ordovás, J. M. PPARγ Pro12Ala interacts with fat intake for obesity

and weight loss in a behavioural treatment based on the Mediterranean diet. *Mol. Nutr. Food Res.* **55**, 1771–1779 (2011).
37. Qi, Q., Bray, G. A., Hu, F. B., Sacks, F. M. & Qi, L. Weight-loss diets modify glucose-dependent insulinotropic polypeptide receptor rs2287019 genotype effects on changes in body weight, fasting glucose, and insulin resistance: the Preventing Overweight Using Novel Dietary Strategies trial. *Am. J. Clin. Nutr.* **95**, 506–513 (2012).
38. Seip, R. L. et al. Physiogenomic comparison of human fat loss in response to diets restrictive of carbohydrate or fat. *Nutr. Metab.* **5**, 4 (2008).
39. Nielsen, D. E. & El-Sohemy, A. Disclosure of genetic information and change in dietary intake: a randomized controlled trial. *PLoS ONE* **9**, e112665 (2014).
40. Celis-Morales, C., Lara, J. & Mathers, J. C. Personalising nutritional guidance for more effective behaviour change. *Proc. Nutr. Soc.* **74**, 130–138 (2015).
41. O'Donovan, C. B., Walsh, M. C., Gibney, M. J., Brennan, L. & Gibney, E. R. Knowing your genes: does this impact behaviour change? *Proc. Nutr. Soc.* **76**, 182–191 (2017).
42. Arkadianos, I. et al. Improved weight management using genetic information to personalize a calorie controlled diet. *Nutr. J.* **6**, 29 (2007).
43. Mattei, J., Qi, Q., Hu, F. B., Sacks, F. M. & Qi, L. TCF7L2 genetic variants modulate the effect of dietary fat intake on changes in body composition during a weight-loss intervention. *Am. J. Clin. Nutr.* **96**, 1129–1136 (2012).
44. Corella, D. et al. APOA5 gene variation modulates the effects of dietary fat intake on body mass index and obesity risk in the Framingham Heart Study. *J. Mol. Med.* **85**, 119–128 (2007).
45. Mifflin, M. D. et al. A new predictive equation for resting energy expenditure in healthy individuals. *Am. J. Clin. Nutr.* **51**, 241–247 (1990).
46. Stunkard, A. J. & Messick, S. The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *J. Psychosom. Res.* **29**, 71–83 (1985).
47. Geiselman, P. J. et al. Reliability and validity of a macronutrient self-selection paradigm and a food preference questionnaire. *Physiol. Behav.* **63**, 919–928 (1998).
48. Eichstaedt, K. E., Kovatch, K. & Maroof, D. A. A less conservative method to adjust for familywise error rate in neuropsychological research: the Holm's sequential Bonferroni procedure. *NeuroRehabilitation* **32**, 693–696 (2013).
49. Martin, C. K. et al. Efficacy of SmartLoss^{SM}, a smartphone-based weight loss intervention: results from a randomized controlled trial. *Obesity* **23**, 935–942 (2015).
50. Heilbronn, L. K. et al. Effect of 6-month calorie restriction on biomarkers of longevity, metabolic adaptation, and oxidative stress in overweight individuals: a randomized controlled trial. *JAMA* **295**, 1539–1548 (2006).

Acknowledgements

The center where the research was conducted is supported in part by U54 GM104940 from the National Institute of General Medical Sciences of the National Institutes of Health, which funds the Louisiana Clinical and Translational Science Center and by the Nutrition Obesity Research Centers grant P30DK072476 titled "Nutrition and Metabolic Health Through the Lifespan" sponsored by National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). C.H. was supported by an NIDDK National Research Service Award (T32DK064584), and J.L.D. was funded by the American Heart Association (Grant # 20POST35210907). The present study was funded by WW International, Inc. (New York, NY, USA). The concept and design of the study were led by the study PI (C.K.M.) and developed in collaboration with five other authors, one of whom is an employee of the sponsor. The sponsor had no role in the execution of the study or statistical analysis. Two sponsor employees are co-authors of the paper and provided editorial comments to the manuscript.

Author contributions

C.K.M. obtained funding for the study. C.K.M., C.H., J.W.A., J.L.D., J.M.O., C.M.C., M.I.C. and G.D.F. designed the study. C.K.M. and C.H. oversaw data acquisition, and S.Y., C.H. and C.K.M. analyzed and interpreted the data. C.H. and C.K.M. drafted the manuscript; all authors provided critical revisions for important intellectual content. C.K.M. was responsible for the overall study supervision, and F.L.G. was responsible for the medical supervision.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

G.D.F. and M.I.C. are shareholders and employees at WW International, Inc. (New York, NY, USA). C.K.M. has previously consulted for WW on a fee-for-service basis, with the latest consultation occurring in 2018. All other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-023-41969-1.

Correspondence and requests for materials should be addressed to Christoph Höchsmann.

Peer review information *Nature Communications* thanks Jennie Brand-Miller, Christopher Gardner and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

PART II

Technology-Assisted Food Intake Assessment

Publication 8

Review of the validity and feasibility of image-assisted methods for dietary assessment.

Christoph Höchsmann [1] & Corby K Martin [1]

[1] Pennington Biomedical Research Center, Baton Rouge, LA, USA

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
International Journal of Obesity, 44(12): 2358–2371	2020	10.1038/s41366-020-00693-2	5.1 (2020)	Q1	Scopus: 40 Google Scholar: 64

Author contribution:

First author; conducted the literature search, drafted the manuscript, and created tables and figures.

International Journal of Obesity (2020) 44:2358–2371
https://doi.org/10.1038/s41366-020-00693-2

REVIEW ARTICLE

Techniques and Methods

Review of the validity and feasibility of image-assisted methods for dietary assessment

Christoph Höchsmann[1] · Corby K. Martin[1]

Received: 28 February 2020 / Revised: 2 September 2020 / Accepted: 26 September 2020 / Published online: 8 October 2020

Abstract

Accurately quantifying dietary intake is essential to understanding the effect of diet on health and evaluating the efficacy of dietary interventions. Self-report methods (e.g., food records) are frequently utilized despite evident inaccuracy of these methods at assessing energy and nutrient intake. Methods that assess food intake via images of foods have overcome many of the limitations of traditional self-report. In cafeteria settings, digital photography has proven to be unobtrusive and accurate and is the method of choice for assessing food provision, plate waste, and food intake. In free-living conditions, image capture of food selection and plate waste via the user's smartphone, is promising and can produce accurate energy intake estimates, though accuracy is not guaranteed. These methods foster (near) real-time transfer of data and eliminate the need for portion size estimation by the user since the food images are analyzed by trained raters. A limitation that remains, similar to self-report methods where participants must truthfully record all consumed foods, is intentional and/or unintentional underreporting of foods due to social desirability or forgetfulness. Methods that rely on passive image capture via wearable cameras are promising and aim to reduce user burden; however, only pilot data with limited validity are currently available and these methods remain obtrusive and cumbersome. To reduce analysis-related staff burden and to allow real-time feedback to the user, recent approaches have aimed to automate the analysis of food images. The technology to support automatic food recognition and portion size estimation is, however, still in its infancy and fully automated food intake assessment with acceptable precision not yet a reality. This review further evaluates the benefits and challenges of current image-assisted methods of food intake assessment and concludes that less burdensome methods are less accurate and that no current method is adequate in all settings.

Introduction

Accurately quantifying food intake (FI) is crucial for investigating the relationship between diet and health in observational studies, understanding the effects of dietary changes on the treatment and management of obesity and obesity-related diseases, and informing public health policies based on empirical data [1]. To date, self-report methods such as food records, food recalls, and food frequency questionnaires are the mainstay of nutritional epidemiology research [2] and commonly used to assess FI in clinical and research settings [3, 4]. While self-report methods have helped to identify associations between consumption of different foods or diet quality and eating behaviors and diseases [1], evidence indicates that these methods frequently inaccurately assess energy and nutrient consumption [5], and their continued use in scientific settings has consequently been questioned and criticized [5, 6]. Limitations of self-report methods and sources of their error include: (1) unintentional underreporting of foods (forgetfulness), (2) intentional underreporting of foods with negative health images (high-fat/high-sugar foods), (3) intentional overreporting of foods that are perceived as healthy (fruits, vegetables), and (4) portion size estimation errors [7, 8]. Further, reactivity due to awareness of being measured can cause changes in eating behavior, resulting in inaccurate reporting and the failure to capture habitual FI data [9]. People also have been found to undereat and lose weight when recording their FI [10]. The last limitation, however, highlights a strength of using self-report methods, as people become more aware of their FI and eating patterns when attempting to manage their body weight, even though

✉ Corby K. Martin
Corby.Martin@pbrc.edu

[1] Pennington Biomedical Research Center, Baton Rouge, LA, USA

the FI data are not necessarily accurate. Thus, self-reported FI remains a frequently used tool in the clinical delivery of weight management services, with problems primarily occurring when these data are used to quantify intake.

Image-assisted methods, which rely on images of foods to estimate FI, are a promising approach to quantify FI that can overcome many of the limitations of self-report. For example, these methods can reduce user burden and eliminate the need for the user to estimate portion size. Additionally, many of these methods transmit food image data to researchers or clinicians in real time or near real time, providing a platform to adapt Ecological Momentary Assessment (EMA) [11] and other methods to detect and minimize missing data [12]. Over the past two decades, several image-assisted methods have been developed that include active or passive image capture and automated or semi-automated analysis of food images. As a result, some methods are better suited for certain conditions or populations vs. other methods. This review presents the strengths and weaknesses of currently available image-assisted methodologies for FI assessment and evaluates their validity in different settings and populations.

Methods

We conducted a literature search through the PubMed electronic database for human studies from inception to February 2020. We included articles published in English, reporting image-assisted methods for FI assessment, assessing their feasibility, and validating them against weighed intake and doubly labeled water (DLW). The following search terms were used individually and in combination: diet*/food/energy intake, digital photography, valid*, reliab*, food record, image-assisted, image-based, portion size, wearable, food recognition. The references of articles were also screened for potentially relevant studies. For this review, methods were categorized as either primarily relying on human raters to estimate FI based on food images vs. methods that claim to be automated or semi-automated. As detailed herein, the term automated or semi-automated is somewhat of a misnomer, however, and those methods still require considerable effort from a human. Further, the reader should be cognizant that the methods used to capture food images can be distinct from the methods used to analyze the images.

Results

The literature search identified 278 articles. Forty-seven articles, reporting 12 distinct methods of image-assisted FI assessment met the inclusion criteria. Table 1 provides an overview of the included methods and their validation in various settings. Figure 1 illustrates the strengths and limitations of the different methodologies regarding their accuracy, feasibility, and ability to detect food waste.

Analysis of food images by human raters

The Digital Photography of Foods Method (DPFM)

The Digital Photography of Foods Method (DPFM) was developed to allow unobtrusive estimation of FI in cafeteria or similar settings [13, 14], and this method or very similar methods have been developed and utilized by many groups [13–23]. These methods use digital video cameras or other devices (e.g., smartphones) to quickly capture images of participants' food selection and plate waste and of precisely weighed standard portions of the foods served on the day of data collection. The images of the weighed standard portions serve as reference images during the analysis of participants' food images, which can occur after data/image collection. The foods in the reference images are linked to foods in the United States Department of Agriculture's (USDA) Food and Nutrient Database for Dietary Studies (FNDDS) [24], an alternative nutrition database, manufacturer's information, or a custom recipe. This allows estimation of energy and nutrient intake. Trained raters analyze the images via computer software that simultaneously displays images of (1) the participant's food selection, (2) plate waste, and (3) the weighed standard portion for each food consumed. The rater then estimates the number of portions of the standard portion of food that was selected and discarded. The software then calculates the amount of food selected, plate waste, and FI, which is the difference between food selection and plate waste.

Portion size estimates from this method have been shown to strongly correlate with weighed portion sizes ($r = 0.92$) [13] and mean overestimation of image-based estimates compared to weighed foods is small, i.e., 5.2 g (standard error [SE] 0.95) or 4.7% of the weighed value. The mean deviation of individual food items such as entrées (17.5 g [SE 4.3]; 6.9%), starches (−1.2 g [SE 1.1]; −1.7%), fruits/vegetables (4.8 g [SE 1.8]; 5.9%), desserts (4.2 g [SE 2.6]; 5.4%), and beverages (7.6 g [SE 3.1]; 4.3%) were likewise small for image-based estimates of total intake compared to weighed estimates; however, condiment intake tended to be overestimated by 4.9 g (SE 4.6; 17%) [13]. This limitation is not unique to this method, and condiments typically do not account for a large proportion of daily FI. In school children ($N = 239$), the mean difference between image-based and weighed estimates of total intake (g) was likewise very small, i.e., 3 g (standard deviation [SD] 20) or 1% [23] and in preschool children ($N = 22$) digital diet estimates were 4% lower than the actual weights [18]. Importantly,

Table 1 Overview of image-assisted methods to measure food intake and studies validating these methods.

Method	Methodology	Review/analysis	Study setting	Sample size	Outcome	Reference method	Reliability/validity
Digital Photography of Foods Method (DPFM) [13, 16, 18, 23]	Images of food selection and plate waste are captured with digital (video) cameras	Human raters compare food images to images of weighed standard portions	Laboratory [13]	60 test meals of 10 different portion sizes	Portion size	Weighed foods	Significant correlation with weighed foods of 0.92. Mean error in portion size was +5.2 g (SE 0.95) or 4.7% relative to weighed foods
			School cafeteria; 5 consecutive days of school lunches [16]	43 school children	EI	Weighed foods	ICC for total EI was 0.93. Convergent validity was supported by significant correlation between food intake and adiposity ($r = 0.45$) and discriminant validity was supported by non-significant correlation between food intake and depressed mood ($r = 0.1$)
			One laboratory-based test meal [18]	22 preschool children	EI	Weighed foods	Significant correlation of DPFM with weighed foods of 0.96 and mean error in total intake of −4% compared to weighed foods
			School cafeteria; 7 days of school lunches and dinners [23]	239 school children	EI	Weighed foods	Mean error in total intake of DPFM of 3 g (SD 20) or 1% compared to weighed foods
Digital Photography + Recall (DP + R) [28]	Images of food selection and plate waste of cafeteria meals including notes to identify ambiguous foods and measuring cups/ spoons to guide portion size estimation. Dietary recall to document any foods or beverages consumed outside the cafeteria	Human raters compare food images to images of weighed standard portions and perform multi-pass dietary recall	Cafeteria and free-living conditions over 7 days	91 adults with overweight/obesity	EI	DLW	The mean EI estimated by DP + R was not significantly different from DLW, overestimating DLW by 264 kJ (SD 3138; 63 kcal [SD 750]) or 6.8% (SD 28) per day. No proportional bias variation as a function of the level of EI ($r = -0.13$, $p = 0.21$)
Remote Food Photography Method (RFPM) [12, 25, 30–33, 39]	Images of food selection and plate waste (including a reference card) are captured via smartphone app and sent to laboratory for analysis	Human raters compare food images to images of weighed standard portions	Free-living conditions (6 days) and 2 laboratory-based buffet meals [12]	50 adults	EI	DLW (free-living) and weighed foods (laboratory)	In free-living conditions, RFPM underestimated total EI by 636 kJ (SD 2904; 152 kcal [SD 694]) or 3.7% (SD 28.7) per day ($p = 0.16$); ICC for daily EI was 0.74 In the laboratory, underestimation for total EI was 17 kJ (SD 305; 4 kcal [SD 73]) or 1.2% (SD 62.8) and error for macronutrients was not significantly different from weighed foods
			Prepacked lunch (consumed in laboratory) and dinner meals (consumed in laboratory or at home) over 3 days [25]	52 adults	EI	Weighed foods	RFPM underestimated EI by 4.7–5.5% (laboratory) and by 6.6% in free-living conditions. ICCs for EI were significant for laboratory ($r = 0.62$; $p < 0.01$) and free-living conditions ($r = 0.68$, $p < 0.01$)

Table 1 (continued)

Method	Methodology	Review/analysis	Study setting	Sample size	Outcome	Reference method	Reliability/validity
			Laboratory; 12-h period [30]	54 preschool children	EI	Weighed foods	RFPM significantly overestimated total EI by 314 kJ (SD 452; 75 kcal [SD 108]) or 7.5% (SD 10.0). The MPE for the macronutrient intakes ranged from 2.9% (fat) to 11.7% (protein), with high variability around the mean
			Free-living conditions over 7 days [31]	39 preschool children	EI	DLW	RFPM underestimated total daily EI by a mean 929 kJ (SD 1146; 222 kcal [SD 274]) or 15.6%
			Laboratory; 2 visits 5–10 days apart [32]	53 adults	EI	Weighed foods	RFPM underestimated EI of 2, 4, and 6 fl oz servings of infant formula by 6.7 kJ (SD 1.7; 1.6 kcal [SD 0.4]), 20.1 kJ (SD 2.5; 4.8 kcal [SD 0.6]), and 25.9 kJ (SD 4.2; 6.2 kcal [SD 1.0]), and overestimated intake by 0.4 kJ (SD 5.0; 0.1 kcal [SD 1.2]) kcals in 8 fl oz servings, but was equivalent to weighed intake within 7.5% for all servings
			Laboratory [33]	7 bottles for each serving size (1, 2, 3, and 4-scoop) containing 5, 10, and 15% more and less formula than recommended	Serving size	Weighed foods	RFPM underestimated servings (1–4 scoops) of powdered instant formula by a mean 0.05 g (90% CI −0.49, 0.40) compared to directly weighed servings, with the MPE ranging from 0.32 to 1.58%. Estimates for all serving sizes were within 5% equivalence bounds
			Free-living conditions over 6 days at 2 time points (early vs. late pregnancy) [39]	23 pregnant women with obesity	EI	DLW	RFPM captured 64.4% (early pregnancy) and 62.2% (late pregnancy) of DLW-measured total daily EI and was not equivalent to DLW within 20% equivalence bounds. The underestimation was significantly associated with low reporting of snacks ($R^2 = 0.4$)
Food Record App (FRapp) [43]	Images of food selection and plate waste including fiducial marker, captured with smartphone app. Additional options to capture food intake are speech-to-text conversions, capturing food label/ nutrition facts images, selecting from recently recorded foods	Human raters analyze recordings (images of food, labels or text recordings) of eating events	Free-living conditions over 3 days	18 adolescents	N/A[a]	N/A[a]	N/A[a]

Table 1 (continued)

Method	Methodology	Review/analysis	Study setting	Sample size	Outcome	Reference method	Reliability/validity
Nutricam Dietary Assessment Method (NuDAM) [44]	Images of food selection (with fiducial marker) combined with a voice recording describing the foods, leftovers, location, and meal occasion, and a brief follow-up phone call the next day	Trained professionals analyze food images, voice recording, and phone calls	Free-living conditions over 3 days	10 adults, diagnosed with T2DM	EI	DLW	NuDAM underestimated total daily EI by 24% compared to DLW
Multiple-pass 24-h dietary recall + SenseCam (MP24 + SC) [45]	SC (worn around the neck on a lanyard) captures images of eating events every 20 s, triggered to turn on by its sensors. Images of eating events are combined with MP24	Review of food images and MP24 with participant to allow modification of self-report; estimation of EI by trained dietitian	Free-living conditions over 3 non-consecutive days	40 adults (20 men, 20 women)	EI	DLW	MP24 + SC underestimated EI by 9% in men and by 7% in women compared to DLW. The addition of SC reduced the error in EI by ~50% compared to MP24 alone
Micro-camera [47]	Micro-camera is worn on the ear and captures audiovisual recordings during meal times. Recordings are combined with food diary entries	Human raters analyze food images and food diaries	Free-living conditions over 2 days	6 adults	EI	DLW	Compared to DLW, daily EI was underestimated by 3912 kJ (SE 1996; 935 kcal [SE 477]) or 34% by food diaries alone and by 3507 kJ (SE 2170; 838 kcal [SE 519]) or 30% when combined with the micro-camera. The difference between the two methods was significant ($p = 0.02$)
mobile Food Record (mFR) [50–52]	Food images are captured with the mFR app and sent to a server for analysis. After review by the user, volume and nutrient content are estimated by the app	Automatic portion size estimation based on statistical pattern recognition techniques of the image	Free-living conditions over a 24-h period [50, 51]	15 adolescents	Portion size	Weighed foods	Mean error in automated weight estimates using mFR compared to known weights ranged from a 38% underestimation to a 26% overestimation, with 75% of all analyzed food being within 7% of the true value
			Free-living conditions over 7.5 days [52]	45 adults	EI	DLW	mFR EI correlated significantly ($r = 0.58$) with DLW-measured daily EI and underestimated EI by 12% (SD 11) for men and 10% (SD 10) for women compared to DLW, with no systematic bias with increasing EI
GoCARB [58]	Food images are captured with a smartphone from two different angles including a reference card	Automatic segmentation and recognition of food items and reconstruction of their 3D shape	Cafeteria	19 adults, 114 test meals	Carbohydrate content, food recognition	Weighed foods	The mean absolute estimation error of GoCARB compared to precisely weighed carbohydrate content was 26.9% (SD 18.9). Automatic food recognition was correct for 85.1% or all food items
FoodCam [59]	The user captures a picture of the food and draws boxes around it to initiate the analysis process. The system populates possible food items and the user selects the best fit	Automatic food recognition and portion size estimation	Laboratory	N/A[b]	N/A[b]	N/A[b]	N/A[b]

Table 1 (continued)

Method	Methodology	Review/analysis	Study setting	Sample size	Outcome	Reference method	Reliability/validity
Snap-n-Eat [60]	The user captures a picture of the food and the system automatically estimates energy and nutrient content	Automatic portion size estimation by image segmentation	Laboratory	2000 food images for 15 food categories	Food classification	N/A	85% accuracy when classifying 2000 images of food items of 15 different categories
eButton [61]	Food images are captured automatically by a chest-worn camera every 2–4 s. Human rater selects 3D models from software's library, overlaying the food, and volume of food is then estimated by the software	Semi-automatic analysis of food images	Laboratory	7 adults capturing 100 pictures of foods	Portion size	Seed displacement method	The mean relative error across all food samples was −2.8% (SD 20.4) and the error for 85 out of 100 foods was between −30 and 30% compared to seed displacement

CHO carbohydrate, *CI* confidence interval, *DLW* doubly labeled water, *EI* energy intake, *ICC* intraclass correlation coefficient, *kJ* kilojoule, *MPE* mean percent error, *SD* standard deviation, *SE* standard error, *T2DM* type 2 diabetes mellitus.

[a]Feasibility study only, to date no validation of the method.

[b]Usability study only, to date no validation of the method.

agreement among raters has been shown to be high (intra-class correlation coefficients of 0.84 [14] and even 0.92 [16] and 0.93 [25], and Cohen's κ of 0.78 [23]).

The DPFM and similar methods have proven to be adaptable and provide a comprehensive assessment of FI-related behaviors, and the accurate quantification of plate waste is a unique strength of this and other image-assisted methods, particularly considering the goal of cutting food waste by 50% in the United States by 2030 [26]. Further, food selection/provision and waste data can be used to determine if efforts to improve diet quality result in higher plate waste due to people not eating the healthier foods, or if food provision and waste systematically differ such that dietary intake is more or less healthy [27]. Examples of the feasibility and utility of digital photography include its ability to estimate FI of large and diverse populations in various settings, including soldiers ($N = 139$) during basic combat training [15], elementary school children ($N = 670$) during school lunches over 2 years [15], and >2000 children from 38 schools over a 3-year period where intake was quantified for 3 days at three different time points [17]. Further, digital photography methods have been used to (1) characterize lunch meals served to preschoolers ($N = 796$) enrolled in Head Start centers [20], (2) estimate FI at family dinners of 231 minority preschool children [19], (3) compare elementary school students' food selection in the school cafeteria to the Institute of Medicine's recommendations across 33 elementary schools, and (4) evaluate the effectiveness of a 28-month school-based obesity prevention intervention (LA Health) at reducing children's selection and consumption of added sugars and sodium during school lunches [22, 27]. Finally, digital photography has been used to assess changes in energy and macronutrient intake during a 16-month exercise trial (Midwest Exercise Trial-2 [21]) in 91 participants over four 7-day periods of ad libitum eating in a university cafeteria.

In summary, the validity and utility of the DPFM and similar methods indicate that they have become the method of choice for quantifying food selection, waste, and intake in cafeteria settings.

Digital Photography + Recall (DP + R)

The Digital Photography + Recall (DP + R) method estimates total daily energy intake (EI) by combining digital photography (pre–post-meal images of food) for assessing EI in a cafeteria setting with dietary recall to record foods consumed outside of the cafeteria setting [28]. The DP + R method includes placing notes on the cafeteria tray to describe any difficult-to-identify food/drink items. Additionally, typical measuring cups and spoons are included in the images to facilitate the estimation of portion size. Multiple-pass dietary recalls are performed at each cafeteria

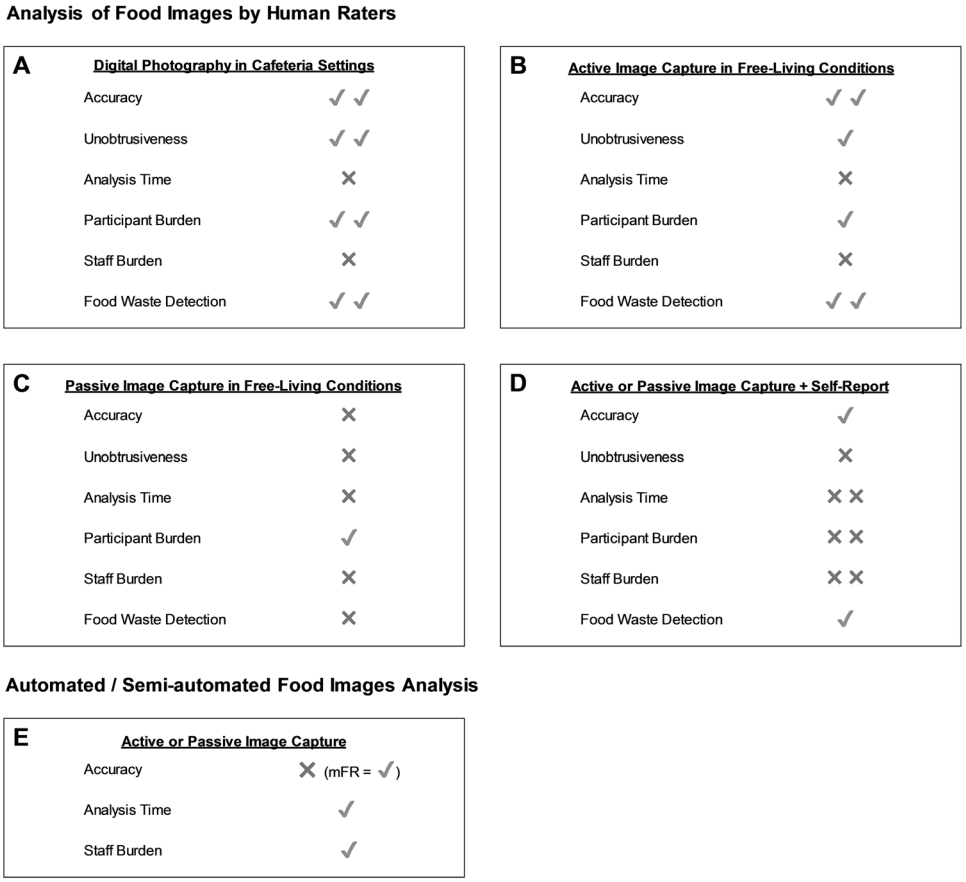


Fig. 1 Overview of different dietary assessment methods concerning accuracy, unobtrusiveness, analysis time, participant burden, staff burden, and food waste detection. Methods that rely on human rater-based analysis where images are captured in cafeteria settings (**a**), actively captured by users in free-living conditions (**b**), passively captured in free-living conditions (**c**), and passively or actively captured and combined with self-report methods (**d**). **e** illustrates systems that automatically or semi-automatically analyze images that are captured actively or passively. It is recognized that these methods differ widely and that many of these systems have not been validated, limiting the information available to perform the ratings displayed in the figure. It is noted, however, that the mFR is among the most studied and validated methods in this category. Each category was rated based on a four-point scale with ✖✖ = poor; ✖ = fair; ✓ = good; ✓✓ = excellent.

meal to document any foods or drinks consumed outside of the cafeteria setting that day [28]. The DP + R method is valid in estimating total EI (required minimum of two cafeteria meals per day) in 91 young adults with overweight or obesity over 7 days. The mean overestimation of EI was 264 kJ (SD 3138; 63 kcal [SD 750]) per day or 6.8% (SD 28) compared to DLW whereby 28.8% of the total estimated daily EI occurred from foods consumed outside of the cafeteria [28]. The implementation of smartphone-captured images of foods consumed outside of the cafeteria may further improve the accuracy of the DP + R method and at the same time reduce participant burden.

Remote Food Photography Method© (RFPM)

The Remote Food Photography Method© (RFPM) resulted from the adaptation of DPFM methods for free-living conditions [12, 25, 29]. When using the RFPM, participants place a reference card next to their food and capture an image of their food selection and plate waste using the SmartIntake® app on a smartphone or other camera-enabled device. For foods that cannot be identified by wrappers or containers, participants briefly annotate the images (e.g., "chicken nuggets"). The annotated images are sent wirelessly to the laboratory via the app. Image information data (date, time, geolocation) are recorded and stored for all food images. In the laboratory, the images are analyzed to estimate FI using methods similar to the DPFM [13, 14]: the foods in the images are linked to a nutrient database via computer software and compared to images of foods with known portion size. The result is detailed data on food selection, plate waste, and FI by difference.

A weakness of the RFPM is that it depends on participants' ability to remember or not neglect to capture images of all consumed foods and calorie-containing beverages. To help address these concerns and ultimately improve data quality and completeness, EMA methods [11] have been incorporated. EMA methods prompt participants to capture images by sending reminders (text messages, push notifications) around participants' typical meal times [25, 29]. A web-based computer system tracks the delivery of prompts as well as participants responses to the prompts, allowing study personnel to more easily detect missing data in near real time. To capture FI data in the case of missing images or phone/app malfunction, participants are asked to additionally use a back-up method.

The reliability and validity of the RFPM have been tested in several different settings and populations [12, 25, 30–35]. First, the RFPM was validated against weighed lunch and dinner meals over 3 days, which participants ($N = 52$)

consumed either in the laboratory or in free-living conditions [25]. The RFPM underestimated daily EI by only 151 kJ (SE 81; 36 kcal [SE 19]; 5.5%) in the laboratory and by 406 kJ (SE 159; 97 kcal [SE 38]; 6.6%) in free-living conditions [25]. Further, the mean difference in estimating EI was stable over different levels of EI and did not differ by body weight or age [25]. Second, the RFPM was validated in adults ($N = 50$) over 6 days in free-living conditions against DLW [12], which is considered accurate for quantifying EI over time in free-living individuals [36]. Total daily EI estimates from the RFPM did not differ significantly from DLW with a mean daily underestimation of 636 kJ (SD 2904; 152 kcal [SD 694]) ($p = 0.16$) or 3.7% (SD 28.7) and a consistent error over different levels of EI [12]. Further, the RFPM's accuracy in estimating nutrient intake was confirmed in two laboratory-based test meals, in which intake of macronutrients and most micronutrients (Calcium, Sodium, Iron, Fiber, Vitamin C) was not significantly different from weighed values [12]. Assessing FI with the RFPM also was not associated with reactivity or changes in EI [12], and, similar to the DPFM, the RFPM has proven feasible and effective at quantifying the plate waste of adults in free-living conditions [37].

The RFPM and SmartIntake® app have proven accurate at measuring infant formula in baby bottles at different stages of preparation (dry powdered formula, prepared formula, liquid waste). The RFPM was equivalent to all weighed servings of formula within 7.5% equivalence bounds and it underestimated EI by ~3% compared to direct weighing [32, 33]. With preschool children who eat solid foods, the RFPM's validity is less consistent, however. Specifically, in preschool children ($N = 54$) who lived in a research unit for 1 day, the RFPM overestimated total intake in grams and kJ by 2.9% (SD 6.6) and 7.5% (SD 10.0), respectively, compared to weighed intake, and bias increased with higher levels of intake [30]. In free-living conditions over 7 days, however, the method underestimated total daily EI by 929 kJ (SD 1146; 222 kcal [SD 274]; 15.6%) when compared to DLW [31]. Although this level of error is in the adequate reporting range identified by Burrows et al. in their review of FI assessment methods in children [38], the results demonstrated that, when the RFPM and SmartIntake® app are used by children's caregivers, the method and app require refinement to obtain the desired level of validity in young children. The authors noted that the biggest challenge in this target group was providing sufficient training to all caregivers (some were not disclosed by the families) and ensuring that images of all meals, snacks, and beverages were captured and sent to the laboratory [30, 31]. In pregnant women with obesity, the RFPM similarly was not able to accurately estimate EI, capturing only around 64% (SE 2.3) of DLW-measured total daily EI [39], which appeared to be due, at least in part, to participants failing to capture images of snacks [39].

The lackluster validity data from the pediatric and pregnancy studies highlighted challenges with the EMA prompts that were used in the older version of SmartIntake®. Specifically, the prompts were previously sent via e-mail, while subsequent versions of the app utilize both push notifications (pop-up messages that are received on one's smartphone, even if the app is not currently in use) and text messages to deliver EMA reminders, improving their effectiveness. Nonetheless, the data indicate that when images are captured, an accurate estimate is typically obtained. The RFPM and SmartIntake® app also have proven feasible and to produce clinically relevant data in demanding conditions, including assessing meal timing, location, level of preparation, and quality of dinner meals among rural, low-income families ($N = 31$) over 1 week (153 dinner meals) [34, 35]. Finally, the RFPM was a feasible and acceptable method for parents of young children ($N = 9$) with type 1 diabetes mellitus to assess breakfast nutrition over 3 days [40].

In summary, the RFPM and SmartIntake® app have many of the same benefits as the DPFM and similar methods, including adaptability to various populations and settings. Additionally, the reference card that is used with the RFPM can facilitate portion size estimates but is not entirely necessary. It does, however, provide a platform for computer imaging algorithms to (1) standardize the images for distance, angle, and color, and (2) attempt to identify and estimate the portion sizes of the foods [41, 42].

Food Record App (FRapp)

The Food Record App (FRapp) uses a methodology similar to the RFPM asking participants to capture and annotate images of all foods and beverages before and after consumption in free-living conditions and to include a fiducial marker (reference card) in each image [43]. FRapp integrates text entry, prompts predefined for eating occasions, and real-time communication between the user and clinician/researcher. The app also allows dietary intake recording via methods other than food images, including speech-to-text conversions with food item extraction, capturing food label/nutrition facts/barcode photos, and selecting from recently consumed food sets [43]. FRapp was an accepted method for dietary intake assessment in community-dwelling adolescents ($N = 18$) in a free-living environment over 3 days; however, only 60% of all eating events with images included the fiducial marker and only 40% included both a pre- and post-meal image, indicating the need for further refinement of the method to improve data completeness in this population [43]. The FRapp has not yet been validated regarding its accuracy in estimating

EI in either a laboratory or free-living setting. Validation of the FRapp will be important to evaluate whether the various options for dietary input, which could affect rater/analysis and user burden, yield any additional benefit to the accuracy of the method compared to methods that rely solely on food images.

The Nutricam Dietary Assessment Method (NuDAM)

The NuDAM combines a phone-captured image of food selection (with reference card) with a voice memo describing the food selection and waste as well as location and type of meal. In addition, on the following day, a brief follow-up phone call is used to probe for commonly underreported foods, and adjustments to the voice memos are made accordingly [44]. The image and accompanying voice recording are analyzed by trained professionals. In a pilot study ($N = 10$) NuDAM was compared to DLW regarding its accuracy in assessing total daily EI over 3 days. NuDAM (8.8 MJ [SD 2.0]; 2102 kcal [SD 478]) underestimated total daily EI compared to DLW (11.8 MJ [SD 2.3]; 2819 kcal [SD 549]) by around 24%, likely due to under- or non-reporting of consumed foods or sugary beverages [44]. The accuracy of NuDAM has only been assessed in a pilot study and further studies with larger sample sizes are needed. However, it is noteworthy that the average underestimation of 24% compared to DLW is rather large compared to that of similar methods that are less burdensome and do not require a follow-up phone call (e.g., the RFPM).

24 h Multiple-pass dietary recall + SenseCam (MP24 + SC)

The MP24 + SC method combines multiple-pass 24 h dietary recall with SenseCam images taken throughout the day on the day before the recall [45]. SenseCam is a wearable camera with a wide-angle lens and built-in accelerometer, heat sensor, and light sensor. It is worn around the neck on a lanyard and captures images approximately every 20 s, as triggered by the sensors [46]. Participants wear the SenseCam continuously; however, they have the option to remove it whenever they are in a location or situation in which they deem photography inappropriate. On the following day, after completion of the dietary recall by trained dietitians, participants may review all SenseCam images in private and delete any images they prefer not to share. Following this, the researcher reviews the SenseCam images with the participant, asking the participant to confirm or modify the self-reported foods without giving any suggestions. Gemming et al. [45] assessed EI with the MP24 + SC method over three non-consecutive 24 h periods in free-living conditions ($N = 40$) and found that on average, total daily EI as assessed by MP24 + SC (13,196 kJ [SD 2529]; 3154 kcal [SD 604]) was underestimated by 9% compared to DLW (14,485 kJ [SD 2632]; 3462 kcal [SD 2632]) in men ($n = 20$) and by 7% (10,091 kJ [SD 1672]; 2412 kcal [SD 400] vs. 10841 kJ [SD 1639]; 2591 kcal [SD 392]) in women ($n = 20$). Compared to MP24 alone, which underestimated average daily EI by 17% (men) and 13% (women) compared to DLW, the addition of the SenseCam reduced the error in daily EI estimation by almost 50%, as previously unreported foods (often snacks) were identified [45]. These data are impressive, though the method has considerable participant and staff burden related to the participant identifying situations/locations in which photography is inappropriate and turning off the SenseCam, the need for the participant to screen all images, and the participant reviewing the images with a staff member.

Micro-camera

This method combines a lightweight, wearable micro-camera, worn on the ear similar to a wireless earpiece for cell phones, with a food diary [47]. The micro-camera has a wide-angle lens (170°) and a microphone for audio recordings during meal times. In a pilot study ($N = 6$), total daily EI estimates from food diary entries over 2 days were analyzed with and without the additional audiovisual micro-camera recordings and compared to EI measured via DLW [47]. The addition of the micro-camera improved the accuracy in estimating total daily EI only slightly from a 34% underestimation (−3912 kJ [SE 1996]; 935 kcal [SE 477]) to a 30% underestimation (−3507 kJ [SE 2170]; 838 kcal [SE 519]) compared to DLW. Much of the underestimation was likely due to underreported foods/snacks and the fact that participants forgot or chose not to turn on the camera during meal times. Interpretation error in estimating intake by the assessors likely further contributed to the large underestimation [47]. Substantial refinement of the method and studies with larger sample sizes are necessary to justify the additional burden of wearing the micro-camera, which in its current state, did not lead to clinically meaningful improvements in EI estimation compared to the food diaries alone.

Automated and semi-automated analysis of food images

Mobile Food Record (mFR)

The Mobile Food Record (mFR) method has been extensively studied and consists of a smartphone app-based food record and a backend secure cloud-like image analysis system [48, 49]. When using mFR, the user captures an image of the food (including a fiducial marker in the image),

which is then transmitted to a server for automatic analysis. The analysis process is based on statistical pattern recognition techniques, identifying food and drink items in the image by comparing the image with those in the database. Next, the labeled image is returned to the participant for review, who then confirms or corrects the automatic labels before sending the image back to the server for final identification and automatic volume estimation via 3D reconstruction of the food items from the images [50]. Finally, identified foods are matched to the USDA FNDDS for nutrient analysis [48, 49].

In a first validation study in adolescents ($N = 15$), the mean error in mFR-estimated weights of individual food items compared to known weights ranged from a 38% underestimation to a 26% overestimation, with 75% of all analyzed foods being within 7% of the true value [50, 51]. In 45 community-dwelling adults, mFR-reported daily EI over 7.5 days correlated significantly ($r = 0.58$) with DLW-measured daily EI and underestimation of total EI was only 12% (SD 11) for men and 10% (SD 10) for women with no systematic bias with increasing EI [52]. Most participants rated the usability of mFR as easy and indicated willingness to use the method for an extended period [52]. Further, the general feasibility and acceptability of the mFR method have been confirmed in 62 young children (3–10 years) [53] and 41 adolescents (11–15 years) [54]; however, variations according to sex and eating occasions in adolescents (higher underreporting in boys and frequently unreported snacks) highlight the need for increased training in the target group to ensure complete data [54]. The mFR method has further been used to characterize adolescents' ($N = 93$) plate waste over 3 days [55] and to assess if 6-month tailored dietary feedback was effective in improving dietary intake of young adults ($N = 143$) [56]. Recently, the automatic portion size estimation of the mFR method was further refined, being now able to estimate portion size and food energy without the need to fit geometric models onto the food but rather by using a complex algorithm that relies on learned energy distribution images [57]. This method's accuracy needs improvement, however, as mean error in estimated EI was 874 kJ (209 kcal; 38%) compared to pre-weighed foods for the 347 analyzed eating occasions. Although further refinement is needed to improve accuracy and include various eating styles and patterns, this development may broaden the applicability of the mFR method to diverse foods and populations.

GoCARB

GoCARB is a smartphone-based food recognition system designed to support patients with type 1 diabetes mellitus in carbohydrate counting [58]. When using GoCARB, the user places a reference card next to their food and uses a smartphone to capture two images of the food from two different angles. The plate is detected via a series of computer vision operations, which automatically segment and recognize the different food items and reconstruct their 3D shape. After food recognition, the carbohydrate content is calculated by linking each food item's volume to the nutritional information provided by the USDA FNDDS [24]. In a pilot study with 19 adults with type 1 diabetes and 114 test meals (one extreme outlier was removed), the mean absolute estimation error of GoCARB compared to precisely weighed carbohydrate content was 26.9% (SD 18.9) [58]. This was a significantly smaller error (−22%; $p = 0.01$) compared to self-report, which had a mean absolute estimation error of 34.3% (SD 24.3) relative to the precisely weighed carbohydrate content. Food recognition was correct for 85.1% or all food items and 90% of participants found GoCARB easy to use and would like to continue to use it in their daily life. GoCARB has to date not been validated in free-living conditions.

FoodCam

FoodCam is a semi-automatic mobile food recognition system. When using FoodCam, the user points a smartphone camera at the food plate and draws bounding boxes around the plates on the smartphone screen to start the food recognition and portion size estimation process. Next, the system's database populates a list of possible food items for the highlighted foods by comparing the captured food items with images stored in the database via a complex algorithm, and the participant selects the best fit. The system does not automatically recognize food volumes and it requires the user to estimated food volumes by touching a slider on the phone screen to adjust the bounding boxes around the food. Finally, calorie and nutrition estimates of each of the recognized food items are calculated based on the image and the food selection from the database [59]. To date, the validity of the FoodCam system has not been tested in laboratory or free-living conditions.

Snap-n-Eat

Snap-n-Eat is designed to recognize foods and estimate the energy and nutrient content of foods automatically [60]. The analytical system recognizes the salient region (food item) in the food image taken by the user and uses hierarchical segmentation to segment the image into regions. Next, the system classifies these regions into different food categories using a linear support vector machine classifier. To estimate portion sizes of the foods, the system counts the number of pixels in each food segment, which then allows the estimation of the energy and nutritional values of the foods. In a feasibility study, the system achieved over 85% accuracy

when classifying 2000 images of food items of 15 different categories [60]. To be a feasible tool for dietary assessment, however, the system needs to be significantly up-scaled to include far more than the 15 different food categories and validity in free-living conditions needs to be established. Additionally, it is unclear if a user can correctly identify misclassified foods, as incorrectly identified foods necessarily result in inaccurate FI estimates.

eButton

The eButton is a small, chest-worn camera, which automatically captures images of consumed foods every 2–4 s. The recorded images are analyzed by computer software to estimate the food's portion size semi-automatically. Specifically, during analysis, food items are identified by the rater and a particular 3D shape model is selected from the software's library and adjusted in location and size to cover the food item in the image as closely as possible. The volume of the food item is then estimated by the software using the volume of the fitted model [61]. In a small pilot study ($N = 7$), eButton was used to capture images of 100 food samples of Asian and Western foods (no liquids) and the software was then used to estimate portion size [61]. The mean relative error across all food samples was −2.8% (SD 20.4) and the error for 85 out of 100 foods was between −30 and 30% compared to the reference method of seed displacement, which is a commonly used method to objectively quantify food volume [62]. The eButton has to date not been validated in free-living conditions.

Discussion

The studies included in this review present image-assisted methodologies to improve the assessment of FI in different settings and populations. Many methods can reduce underreporting observed with traditional self-report methods, though some methods, particularly those relying on automated image analysis, inaccurately estimate FI. In cafeteria settings, the DPFM and similar methods have proven feasible, effective, and highly accurate at estimating FI in large samples of diverse participants [13–23] and can today be considered the method of choice. In free-living conditions, smartphone apps can be used to capture food images and to transfer the images and associated data to a reading center in real time. These methods can produce accurate estimates of energy and nutrient intake [12, 25, 30–33], though accuracy relies on sound methods, such as EMAs, to facilitate data quality and completeness.

A noted weakness of many of the reviewed methods is their limited reliability and validity. For example, many have only been tested in proof-of-concept and pilot studies and laboratory settings and are lacking validation against DLW in free-living conditions. Further, larger sample sizes are needed to make results more generalizable and identify the best method for specific settings and target groups. In general, more accurate methods tend to be less burdensome for the participant but can be more burdensome for the image-analyzing staff. This limits the deployability of these methods on a large scale.

Many of the reviewed methods, particularly those used in free-living conditions, rely on smartphone-captured images. These images are then sent to a reading center for analysis by human raters (RFPM [12, 25, 30–33, 39], FRapp [43], NuDAM [44]) or analyzed semi-automatically via software and additional input by the user (mFR [48, 49], GoCARB [58], FoodCam [59], or Snap-n-Eat [60]). Smartphones are a logical choice for image-assisted dietary assessment since ~3.2 billion people use smartphones daily [63] and most smartphone users carry their phones with them throughout the day [64]. Smartphone apps can reduce missing or incomplete data in free-living conditions by incorporating EMAs and thereby accurately estimate the EI of adults [11]. Failure to capture images of foods consumed due to forgetfulness and/or due to intentional misreporting (e.g., social desirability bias) is a limitation of image-based methods that remains a challenge. Although this limitation applies to any FI assessment method requiring participants to truthfully record all consumed foods, it is still an important limitation that should be considered when using methods with active image capture by the participant. For this reason, passive/automated image capture via wearable devices such as the eButton [61], SenseCam [45], or Micro-camera [47] offer significant advantages since missing food images should occur less frequently and additional contextual information about the eating event can be recorded and annotated at a later date. Currently, however, passive image capture also has limitations which might limit the ability to disseminate these methods widely in the immediate term. For example, the battery life and data storage capacity of the wearable device needs to be sufficient to capture high-quality images throughout the day. The large amount of passively captured images throughout the day further requires a time-consuming review by the participant before images are transmitted to the laboratory for analysis as some pictures may include other people and objects in the participant's environment that the participant does not wish to share due to privacy concerns. While this review process is inevitable and participants would likely have reasonable concerns using systems without the option to censor images, the censorship of certain (food) images could affect the accuracy of these methods. Technological advances promise to dramatically improve these methods in the future.

Many of the more accurate methods rely on the participant or researcher to manage images, verify which images to send, identify foods or verify automatic food identification, and/or estimate or verify portion size. Thus, while some approaches of automated analysis have promise for the future, to date, completely automated food image analysis, including identification of foods, matching of foods to a nutrient database, and estimation of portion size and food waste with sufficient accuracy is not yet a reality. Even with much more advanced recognition technology in the future, the automatic image-based identification and distinction between very similar looking foods will likely remain a significant challenge and may never be possible without at least some degree of user verification. Additionally, the technology to support automatic portion size estimation is still in its infancy and not possible with acceptable precision without at least some form of user feedback.

Because of the limitations of automated food image recognition, many systems and studies in free-living conditions (RFPM [12, 25, 30–33, 39], SenseCam [45], NuDAM [44]) continue to rely on analysis by trained raters who estimate portion sizes and calculate energy and macro-/micronutrient content by matching the foods in the images to a nutrient database. Currently, analysis by human raters is more accurate and less variable than semi-automated image analysis. Importantly, rater-based analysis can rely on existing nutrient databases (USDA, etc.), whereas having to create comprehensive databases for automated food recognition systems can be burdensome and limits feasibility, at least without further technological advances.

Regardless of the method by which portion size is estimated, it is important to recognize the limits of the estimation. For example, portion size estimation of foods with amorphous shapes or higher energy densities tends to be challenging [65]. Further, the correct identification of certain ambiguous foods (e.g., diet soda vs. regular soda), preparation method (e.g., fried vs. baked), and the type and amount of hidden ingredients in a dish (e.g., butter in mashed potatoes) frequently require some form of image annotation by the participant. The precise annotation of images by the participant, of course, relies on self-report with its known flaws, and participants will not always know enough about the ingredients and preparation methods of their food to precisely account for added fat, etc. This problem is not unique to image-based methods, however, and even when directly weighing FI, the recipe and precise amount of ingredients used need to be carefully quantified and recorded. Nevertheless, despite these issues that are likely random [66], image-based methods that use trained raters for image analysis are still far less problematic than the systematic bias observed when food type and portion size are entirely self-reported [1, 67, 68].

In conclusion, image-assisted methods to assess FI will likely remain a provocative force in the literature. Despite technological advances, the more accurate methods still rely on human raters to estimate FI from food images, though significant advances in passive image capture and automated/semi-automated image analysis have opened a new frontier of development. As technology advances, the field can move forward, but only with thorough and critical evaluation of the strengths and weaknesses of the methods. It is unlikely that a single method will be a panacea and applicable to all data collection scenarios, populations, and sample sizes. While the less accurate methods are not suitable to measure FI as an outcome variable, they may still serve as important monitoring tools in behavioral interventions as they may mediate behavior change. In the future, pairing image-based methods with other sensors such as continuous glucose monitoring and using mathematical modeling to integrate the multi-sensor data may increase accuracy of the single methods and improve FI assessment.

Funding This work was supported by National Institutes of Health grants T32 DK064584 (NIDDK National Research Service Award), P30 DK072476 (Nutrition Obesity Research Center), and U54 GM104940 (Louisiana Clinical and Translational Science Center).

Compliance with ethical standards

Conflict of interest The intellectual property surrounding the Remote Food Photography Method© and SmartIntake® application are owned by Louisiana State University/Pennington Biomedical Research Center and CKM is an inventor. There are no other competing interests related to this study to declare.

References

1. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. Addressing current criticism regarding the value of self-report dietary data. J Nutr. 2015;145:2639–45.
2. Walter W. Nutritional epidemiology. 3rd ed. Oxford, New York: Oxford University Press; 2012.
3. Lytle LA, Nicastro HL, Roberts SB, Evans M, Jakicic JM, Laposky AD, et al. Accumulating data to optimally predict obesity treatment (ADOPT) core measures: behavioral domain. Obes Silver Spring Md. 2018;26(Suppl 2):S16–24.
4. Johnson RK. Dietary intake—how do we measure what people are really eating? Obes Res. 2002;10:63S–8S.
5. Archer E, Pavela G, Lavie CJ. The inadmissibility of 'What We Eat In America' (WWEIA) and NHANES dietary data in nutrition & obesity research and the scientific formulation of national dietary guidelines. Mayo Clin Proc. 2015;90:911–26.
6. Dhurandhar NV, Schoeller D, Brown AW, Heymsfield SB, Thomas D, Sørensen TIA, et al. Energy balance measurement: when something is not better than nothing. Int J Obes. 2015;39:1109–13.

7. Basiotis PP, Welsh SO, Cronin FJ, Kelsay JL, Mertz W. Number of days of food intake records required to estimate individual and group nutrient intakes with defined confidence. J Nutr. 1987;117:1638–41.
8. Macdiarmid J, Blundell J. Assessing dietary intake: who, what and why of under-reporting. Nutr Res Rev. 1998;11:231–53.
9. Rebro SM, Patterson RE, Kristal AR, Cheney CL. The effect of keeping food records on eating patterns. J Am Diet Assoc. 1998;98:1163–5.
10. Goris AH, Westerterp-Plantenga MS, Westerterp KR. Undereating and underrecording of habitual food intake in obese men: selective underreporting of fat intake. Am J Clin Nutr. 2000;71:130–4.
11. Stone AA, Shiffman S. Ecological momentary assessment (Ema) in behavioral medicine. Ann Behav Med. 1994;16:199–202.
12. Martin CK, Correa JB, Han H, Allen HR, Rood J, Champagne CM, et al. Validity of the remote food photography method (RFPM) for estimating energy and nutrient intake in near real-time. Obesity (Silver Spring). 2012;20:891–9.
13. Williamson DA, Allen HR, Martin PD, Alfonso AJ, Gerald B, Hunt A. Comparison of digital photography to weighed and visual estimation of portion sizes. J Am Diet Assoc. 2003;103:1139–45.
14. Williamson DA, Allen HR, Martin PD, Alfonso A, Gerald B, Hunt A. Digital photography: a new method for estimating food intake in cafeteria settings. Eat Weight Disord. 2004;9:24–8.
15. Williamson DA, Martin PD, Allen HR, Most MM, Alfonso AJ, Thomas V, et al. Changes in food intake and body weight associated with basic combat training. Mil Med. 2002;167:248–53.
16. Martin CK, Newton RL, Anton SD, Allen HR, Alfonso A, Han H, et al. Measurement of children's food intake with digital photography and the effects of second servings upon food intake. Eat Behav. 2007;8:148–56.
17. Williamson DA, Champagne CM, Harsha D, Han H, Martin CK, Newton R, et al. Louisiana (LA) health: design and methods for a childhood obesity prevention program in rural schools. Contemp Clin Trials. 2008;29:783–95.
18. Nicklas TA, O'Neil CE, Stuff J, Goodell LS, Liu Y, Martin CK. Validity and feasibility of a digital diet estimation method for use with preschool children: a pilot study. J Nutr Educ Behav. 2012;44:618–23.
19. Nicklas TA, O'Neil CE, Stuff JE, Hughes SO, Liu Y. Characterizing dinner meals served and consumed by low-income preschool children. Child Obes Print. 2012;8:561–71.
20. Nicklas TA, Liu Y, Stuff JE, Fisher JO, Mendoza JA, O'Neil CE. Characterizing lunch meals served and consumed by preschool children in head start. Public Health Nutr. 2013;16:2169–77.
21. Washburn RA, Honas JJ, Ptomey LT, Mayo MS, Lee J, Sullivan DK, et al. Energy and macronutrient intake in the Midwest Exercise Trial-2 (MET-2). Med Sci Sports Exerc. 2015;47:1941–9.
22. Hawkins KR, Burton JH, Apolzan JW, Thomson JL, Williamson DA, Martin CK. Efficacy of a school-based obesity prevention intervention at reducing added sugar and sodium in children's school lunches: the LA Health randomized controlled trial. Int J Obes. 2018;42:1845–52.
23. Marcano-Olivier M, Erjavec M, Horne PJ, Viktor S, Pearson R. Measuring lunchtime consumption in school cafeterias: a validation study of the use of digital photography. Public Health Nutr. 2019;22:1745–54.
24. U.S. Department of Agriculture, Agricultural Research Service. USDA Food and Nutrient Database for Dietary Studies 2015-2016. Beltsville, MD: U.S. Department of Agriculture, Agricultural Research Service; 2018.
25. Martin CK, Han H, Coulon SM, Allen HR, Champagne CM, Anton SD. A novel method to remotely measure food intake of free-living individuals in real time: the remote food photography method. Br J Nutr. 2009;101:446–56.
26. U.S. Environmental Protection Agency, U.S. Food and Drug Administration, U.S. Department of Agriculture. Formal agreement between EPA, USDA, and FDA relative to cooperation and coordination on food loss and waste. Washington, DC: US Food Drug Administration; 2020.
27. Martin CK, Thomson JL, LeBlanc MM, Stewart TM, Newton RL, Han H, et al. Children in school cafeterias select foods containing more saturated fat and energy than the Institute of Medicine recommendations. J Nutr. 2010;140:1653–60.
28. Ptomey LT, Willis EA, Honas JJ, Mayo MS, Washburn RA, Herrmann SD, et al. Validity of energy intake estimated by digital photography + recall in overweight and obese young adults. J Acad Nutr Diet. 2015;115:1392–9.
29. Martin CK, Nicklas T, Gunturk B, Correa JB, Allen HR, Champagne C. Measuring food intake with digital photography. J Hum Nutr Diet Off J Br Diet Assoc. 2014;27(Suppl 1):72–81.
30. Nicklas T, Islam NG, Saab R, Schulin R, Liu Y, Butte NF, et al. Validity of a digital diet estimation method for use with preschool children. J Acad Nutr Diet. 2018;118:252–60.
31. Nicklas T, Saab R, Islam NG, Wong W, Butte N, Schulin R, et al. Validity of the remote food photography method against doubly labeled water among minority preschoolers. Obesity (Silver Spring). 2017;25:1633–8.
32. Altazan AD, Gilmore LA, Burton JH, Ragusa SA, Apolzan JW, Martin CK, et al. Development and application of the remote food photography method to measure food intake in exclusively milk fed infants: a laboratory-based study. PLoS ONE. 2016;11: e0163833.
33. Duhé AF, Gilmore LA, Burton JH, Martin CK, Redman LM. The remote food photography method accurately estimates dry powdered foods-the source of calories for many infants. J Acad Nutr Diet. 2016;116:1172–7.
34. Bekelman TA, Bellows LL, McCloskey ML, Martin CK, Johnson SL. Assessing dinner meals offered at home among preschoolers from low-income families with the remote food photography method. Pediatr Obes. 2019;14:e12558.
35. McCloskey ML, Johnson SL, Bekelman TA, Martin CK, Bellows LL. Beyond nutrient intake: use of digital food photography methodology to examine family dinnertime. J Nutr Educ Behav. 2019;51:547–55.e1.
36. Buchowski MS. Doubly labeled water is a validated and verified reference standard in nutrition research. J Nutr. 2014;144:573–4.
37. Roe BE, Apolzan JW, Qi D, Allen HR, Martin CK. Plate waste of adults in the United States measured in free-living conditions. PLoS ONE. 2018;13:e0191813.
38. Burrows TL, Martin RJ, Collins CE. A systematic review of the validity of dietary assessment methods in children when compared with the method of doubly labeled water. J Am Diet Assoc. 2010;110:1501–10.
39. Most J, Vallo PM, Altazan AD, Gilmore LA, Sutton EF, Cain LE, et al. Food photography is not an accurate measure of energy intake in obese, pregnant women. J Nutr. 2018;148:658–63.
40. Rose MH, Streisand R, Aronow L, Tully C, Martin CK, Mackey E. Preliminary feasibility and acceptability of the remote food photography method for assessing nutrition in young children with type 1 diabetes. Clin Pract Pediatr Psychol. 2018;6:270–7.
41. Martin CK, Kaya S, Gunturk BK. Quantification of food intake using food image analysis. Annu Int Conf IEEE Eng Med Biol Soc. 2009;2009:6869–72.
42. Dibiano R, Gunturk BK, Martin CK. Food image analysis for measuring food intake in free living conditions. Med. Imaging 2013 Image Proces. 2013;8669:86693N.
43. Casperson SL, Sieling J, Moon J, Johnson L, Roemmich JN, Whigham L. A mobile phone food record app to digitally capture dietary intake for adolescents in a free-living environment: usability study. JMIR MHealth UHealth. 2015;3:e30.

44. Rollo ME, Ash S, Lyons-Wall P, Russell AW. Evaluation of a mobile phone image-based dietary assessment method in adults with type 2 diabetes. Nutrients. 2015;7:4897–910.
45. Gemming L, Rush E, Maddison R, Doherty A, Gant N, Utter J, et al. Wearable cameras can reduce dietary under-reporting: doubly labelled water validation of a camera-assisted 24 h recall. Br J Nutr. 2015;113:284–91.
46. Hodges S, Williams L, Berry E, Izadi S, Srinivasan J, Butler A, et al. SenseCam: a retrospective memory aid. Proceedings of the 8th International Conference of Ubiquitous Computing (UbiComp 2006), Springer-Verlag; 2006. p. 177–93.
47. Pettitt C, Liu J, Kwasnicki RM, Yang G-Z, Preston T, Frost G. A pilot study to determine whether using a lightweight, wearable micro-camera improves dietary assessment accuracy and offers information on macronutrients and eating rate. Br J Nutr. 2016;115:160–7.
48. Ahmad Z, Khanna N, Kerr DA, Boushey CJ, Delp EJ. A mobile phone user interface for image-based dietary assessment. Proc SPIE Int Soc Opt Eng. 2014;9030:9030.
49. Ahmad Z, Kerr DA, Bosch M, Boushey CJ, Delp EJ, Khanna N, et al. A mobile food record for integrated dietary assessment. MADiMa16. 2016;2016:53–62.
50. Fang S, Liu C, Zhu F, Delp EJ, Boushey CJ. Single-view food portion estimation based on geometric models. 2015 IEEE International Symposium on Multimedia (ISM). Miami, FL: IEEE; 2015, p. 385–90.
51. Lee CD, Chae J, Schap TE, Kerr DA, Delp EJ, Ebert DS, et al. Comparison of known food weights with image-based portion-size automated estimation and adolescents' self-reported portion size. J Diabetes Sci Technol. 2012;6:428–34.
52. Boushey CJ, Spoden M, Delp EJ, Zhu F, Bosch M, Ahmad Z. et al. Reported energy intake accuracy compared to doubly labeled water and usability of the mobile food record among community dwelling adults. Nutrients. 2017;9:312.
53. Aflague TF, Boushey CJ, Guerrero RTL, Ahmad Z, Kerr DA, Delp EJ. Feasibility and use of the mobile food record for capturing eating occasions among children ages 3-10 years in Guam. Nutrients. 2015;7:4403–15.
54. Boushey CJ, Harray AJ, Kerr DA, Schap TE, Paterson S, Aflague T, et al. How willing are adolescents to record their dietary intake? The mobile food record. JMIR MHealth UHealth. 2015;3:e47.
55. Panizza CE, Boushey CJ, Delp EJ, Kerr DA, Lim E, Gandhi K. et al. Characterizing early adolescent plate waste using the mobile food record. Nutrients. 2017;9:93.
56. Shoneye CL, Dhaliwal SS, Pollard CM, Boushey CJ, Delp EJ, Harray AJ. et al. Image-based dietary assessment and tailored feedback using mobile technology: mediating behavior change in young adults. Nutrients. 2019;11:435.
57. Fang S, Shao Z, Kerr DA, Boushey CJ, Zhu F. An end-to-end image-based automatic food energy estimation technique based on learned energy distribution images: protocol and methodology. Nutrients. 2019;11:877.
58. Rhyner D, Loher H, Dehais J, Anthimopoulos M, Shevchik S, Botwey RH. et al. Carbohydrate estimation by a mobile phone-based system versus self-estimations of individuals with type 1 diabetes mellitus: a comparative study. J Med Internet Res. 2016;18:e101.
59. Kawano Y, Yanai K. FoodCam: a real-time food recognition system on a smartphone. Multimed Tools Appl. 2015;74:5263–87.
60. Zhang W, Yu Q, Siddiquie B, Divakaran A, Sawhney H. Snap-n-Eat. J Diabetes Sci Technol. 2015;9:525–33.
61. Jia W, Chen H-C, Yue Y, Li Z, Fernstrom J, Bai Y, et al. Accuracy of food portion size estimation from digital pictures acquired by a chest-worn camera. Public Health Nutr. 2014;17:1671–81.
62. Cauvain SP, Young LS. Technology of breadmaking. New York, NY: Springer Science & Business Media; 2013.
63. Holst A. Number of smartphone users worldwide 2014-2020. Statista. 2019. https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/. Accessed 26 Dec 2019.
64. International Data Corporation. Always connected—how smartphones and social keep us engaged. Framingham, MA: International Data Corporation; 2013.
65. Howes E, Boushey CJ, Kerr DA, Tomayko EJ, Cluskey M. Image-based dietary assessment ability of dietetics students and interns. Nutrients. 2017;9:114.
66. Gibson RS, Charrondiere UR, Bell W. Measurement errors in dietary assessment using self-reported 24-hour recalls in low-income countries and strategies for their prevention. Adv Nutr. 2017;8:980–91.
67. Maurer J, Taren DL, Teixeira PJ, Thomson CA, Lohman TG, Going SB, et al. The psychosocial and behavioral characteristics related to energy misreporting. Nutr Rev. 2006;64:53–66.
68. Subar AF, Crafts J, Zimmerman TP, Wilson M, Mittl B, Islam NG, et al. Assessment of the accuracy of portion size reports using computer-based food photographs aids in the development of an automated self-administered 24-hour recall. J Am Diet Assoc. 2010;110:55–64.

Publication 9

Preference, expected burden, and willingness to use digital and traditional methods to assess food and alcohol intake.

Christoph Höchsmann [1], Nicole Fearnbach [1], James L Dorling [1], Tera L Fazzino [2,3], Candice A Myers [1], John W Apolzan [1], Corby K Martin [1]

[1] Pennington Biomedical Research Center, Baton Rouge, LA, USA

[2] Cofrin Logan Center for Addiction Research and Treatment, University of Kansas, Lawrence, KS, USA

[3] Department of Psychology, University of Kansas, Lawrence, KS, USA

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
Nutrients, 13(10): 3340	2021	10.3390/nu13103340	6.7 (2021)	Q1	Scopus: 6 Google Scholar: 8

Author contribution:

First author; obtained funding, designed the study, performed statistical analyses, drafted the manuscript, and created tables and figures.

nutrients

MDPI

Article

Preference, Expected Burden, and Willingness to Use Digital and Traditional Methods to Assess Food and Alcohol Intake

Christoph Höchsmann [1], Nicole Fearnbach [1], James L. Dorling [1], Tera L. Fazzino [2,3], Candice A. Myers [1], John W. Apolzan [1] and Corby K. Martin [1,*]

[1] Pennington Biomedical Research Center, Baton Rouge, LA 70808, USA; hoechsmann.christoph@gmail.com (C.H.); nicole.fearnbach@pbrc.edu (N.F.); james.dorling@glasgow.ac.uk (J.L.D.); candice.myers@pbrc.edu (C.A.M.); john.apolzan@pbrc.edu (J.W.A.)
[2] Cofrin Logan Center for Addiction Research and Treatment, University of Kansas, Lawrence, KS 66045, USA; tfazzino@ku.edu
[3] Department of Psychology, University of Kansas, Lawrence, KS 66045, USA
* Correspondence: corby.martin@pbrc.edu; Tel.: +1-(225)-763-2585

Abstract: We conducted an online survey to examine the preference, expected burden, and willingness of people to use four different methods of assessing food and alcohol intake such as food/drink record, 24-h recall, Remote Food Photography Method© (RFPM, via SmartIntake® app), and a novel app (PortionSize®) that allows the in-app portion size estimation of foods/drinks by the user. For food (N = 1959) and alcohol (N = 466) intake assessment, 67.3% and 63.3%, respectively, preferred the RFPM/SmartIntake®, 51.9% and 53.4% preferred PortionSize®, 48.0% and 49.3% the food records, and 32.9% and 33.9% the 24-h recalls (difference in preference across all methods was $p < 0.001$ for food and alcohol intake). Ratings of burden and preference of methods were virtually superimposable, and we found strong correlations between high preference and low expected burden for all methods (all $\rho \geq 0.82$; all $p < 0.001$). Willingness (mean (SD)) to use the RFPM/SmartIntake® (food: 6.6 (2.0); alcohol: 6.4 (2.4)) was greater than PortionSize® (food: 6.0 (2.2); alcohol: 6.0 (2.4); all $p < 0.001$) and 24-h recalls (food: 6.1 (2.2); alcohol: 5.7 (2.7); $p < 0.001$), but not different from food records (food: 6.6 (2.0); alcohol: 6.5 (2.3); all $p \geq 0.33$). Our results can be used in conjunction with existing data on the reliability and validity of these methods in order to inform the selection of methods for the assessment of food and alcohol intake.

Keywords: food intake; food records; RFPM; PortionSize; diet recall; alcohol

Citation: Höchsmann, C.; Fearnbach, N.; Dorling, J.L.; Fazzino, T.L.; Myers, C.A.; Apolzan, J.W.; Martin, C.K. Preference, Expected Burden, and Willingness to Use Digital and Traditional Methods to Assess Food and Alcohol Intake. *Nutrients* **2021**, *13*, 3340. https://doi.org/10.3390/nu13103340

Academic Editor: Laura Di Renzo

Received: 18 August 2021
Accepted: 20 September 2021
Published: 24 September 2021

1. Introduction

Several factors affect the suitability of different methods of ingestive behavior assessment in specific study designs, clinical settings, and populations. Validity, reproducibility, usability, and feasibility of the selected method in the target population and data collection setting are indispensable for the collection of high-quality data on ingestive behavior. However, participants' preference as well as the acceptability and perceived burden of a specific method can also play an important role in collecting high-quality data.

It has been reported that adherence to more burdensome self-report methods of assessing dietary intake can be low and typically decreases over time, thereby influencing data quality [1–3]. Furthermore, pen-and-paper food checklists to track food and beverage intake have been reported to be preferred (46% of participants) and perceived as less burdensome compared to both 24-h recalls (29%) and pen-and-paper food records with the additional requirement of weighing all foods consumed (21%) [4]. In a cross-over study, 78% of participants preferred using the online version of a food record compared to 13% who preferred the pen-and-paper record (9% had no preference) after having used each method for 7 days. The online method was perceived as being quicker, more convenient, and overall less burdensome than the paper version [5]. Collecting data online or via an

app offers some additional advantages compared to pen-and-paper methods, including the ability for data to be transferred in real time.

Some ingestive behavior assessment methods attempt to reduce participant burden and improve accuracy by asking participants to capture images of food and drinks with camera-enabled devices such as smartphones, and these images are then analyzed by a trained rater, not the participant, to estimate intake. One such method is called the Remote Food Photography Method© (RFPM), which involves participants capturing images of food and drinks and annotating the images with descriptors in order to identify products that are not readily identifiable by wrappers or logos in the image [6,7]. These images are then sent to researchers or clinicians for analysis in near-real time. The collection of RFPM data is streamlined by a custom-built smartphone app called SmartIntake®, which participants use to capture images of their food/drink selection and plate/drink waste with a smartphone or tablet; the app automatically sends the food images and related data wirelessly to the laboratory for analysis [6,7]. This reduces errors in portion size estimation [6], the largest source of error in self-reported food intake [8]; more importantly, because the burden of estimating portion size is moved from the participant to the researcher or clinician, this may help explain the large difference in preference for this method compared to traditional self-report methods. For example, the RFPM/SmartIntake® has been found to be preferred by 93.6% of participants for assessing food intake compared to pen-and-paper records [7], and 93.3% of participants preferred the RFPM/SmartIntake® compared to online diet recalls for assessing alcohol consumption [9] after using each method for 3 consecutive days in free-living conditions. However, the preference for app-based methods that rely on users to estimate portion size from food images has not been evaluated.

To date, a direct comparison of the perceived burden and preference across these methods (food/drink record, 24-h recall, RFPM) that are commonly used to assess ingestive behavior is lacking. The assumption that app- and image-based methods are perceived as less burdensome and consequently preferred compared to traditional self-report methods (food/drink record, 24-h recall) has not been thoroughly examined. To address this, we conducted an online survey to assess participants' preference, expected burden, and willingness to use four different methods of food/drink intake assessment: (1) a food/drink record; (2) a 24-h recall; (3) the RFPM/SmartIntake® app; and (4) a novel smartphone app (PortionSize®) that allows the estimation of the portion size of foods and drinks by the user directly in the app, which results in immediate food and drink intake feedback without the need for external analysis by a trained clinician or researcher. We hypothesized that for both food intake and alcohol consumption, the RFPM/SmartIntake® and PortionSize® would be rated as more preferred than the traditional methods (food record and 24-h recall) and that the RFPM/SmartIntake® would additionally be perceived as less burdensome and more preferred compared to PortionSize®, which requires participants to estimate and report their portion size.

2. Materials and Methods

2.1. Design and Participants

The Pennington Habits Survey was approved by the Institutional Review Board at Pennington Biomedical Research Center (PBRC, 2019-052-PBRC) and registered at ClinicalTrials.gov (NCT04150510) before the start of recruitment. The survey included a questionnaire assessing demographic and socioeconomic characteristics as well as: (1) the Food Intake Assessment Preference Questionnaire, (2) the Alcohol Consumption Questionnaire, (3) Alcohol Consumption Assessment Preference Questionnaire, (4) the Smoking Questionnaire, (5) the Smoking Assessment Preference Questionnaire, (6) the Vaping Questionnaire, and (7) the Vaping Assessment Preference Questionnaire. A link to the anonymous survey was distributed by paid advertisements on social media platforms, PBRC's webpage, email listservs, and word-of-mouth between February and November 2020. Adults between 18 and 85 years of age, residing in the United States, and with access to the internet were eligible to participate in the Pennington Habits Survey. However, to complete the Assessment

Preference Questionnaires for alcohol consumption, smoking, and vaping, participants had to indicate that they were practicing the respective activities at the time of the survey. Upon opening the survey link, interested individuals received instructions that detailed the purpose of the study. Participants verified that they were adults and provided consent to participate before proceeding with the survey. Data were collected using Research Electronic Data Capture (REDCap) [10]. Participation in the survey was voluntary and the participants had the option not to submit answers or to skip items if they did not wish to complete them. Upon completion of the survey, participants had the option to enter a lottery to win 1 of 10 checks worth USD 50. In this report, only the results from the Food Intake Assessment Preference Questionnaire and the Alcohol Consumption Assessment Preference Questionnaire as well as associated demographic data are presented.

2.2. Demographic and Socioeconomic Characteristics

Before continuing with the Food Intake and Alcohol Consumption Assessment Preference Questionnaires, participants completed a questionnaire assessing demographic and socioeconomic characteristics. The questionnaire captured data such as age, sex, race, education level, and household income. In addition, household food security (assessed with the 6-item Short Form of the United States Household Food Security Survey Module) [11] and subjective social status (assessed with the MacArthur Scale of Subjective Social Status, in which individuals place an "X" on the rung (1–10) of the "social ladder" on which they feel they stand compared to other people in the United States) [12] were assessed. Finally, participants self-reported body weight and height, and past or present diagnosis of one or more of the following diseases: heart disease, type 2 diabetes, hypertension, or dyslipidemia.

2.3. Food Intake and Alcohol Consumption Assessment Preference Questionnaires

Participants were provided with a description of the following 4 methods of food intake and alcohol consumption assessment: (1) a food/drink record; (2) a 24 h recall; (3) the RFPM and SmartIntake® app [7], which has also been used to assess alcohol consumption [9]; and (4) a new smartphone app called PortionSize®. PortionSize® provides immediate food and beverage intake feedback to the participant and researchers since the participant estimates portion size in the app based on their food/drink images without the need for external analysis by a trained clinician or researcher. The verbatim descriptions and images, if applicable, for the 4 methods of food intake assessment that were provided to the participants in the survey are outlined below and also provided as Supplementary Methods. The descriptions for the 4 methods of alcohol consumption assessment were very similar.

2.3.1. Food Record

"Food records are a way to record all of the foods and beverages that you consume. You are usually asked to keep these records for 3–7 days. During this period, you would need to carry the record with you and record all foods and beverages that you consume right when you eat or drink them. The food record can be a paper form that you complete by hand. Other ways to keep these records include using a smartphone to complete the record electronically. To increase the accuracy of the record, you need to carefully estimate or weigh how much food you eat, and how many beverages you drink and record those amounts. Additionally, you need to record details about the food or beverage. Those details include things like what cut of meat you are eating, what condiments you added, and how much condiments you added. Finally, you need to record how the food was cooked or prepared. For example, if the food was fried, baked, sautéed, etc."

2.3.2. The 24-h Recall

"A 24 h recall is another method to track your food and beverage intake. This method is like an interview that is usually conducted via phone or in person. Each interview

takes 20 to 30 min. You would be asked to recall all of the foods and beverages that you consumed over the previous 24 h. You also would need to recall and report how much of each food and beverage you consumed. Finally, you would need to recall and report how the foods were prepared (fried, baked, etc.). Because our food intake varies from day to day, you would typically be asked to complete about 3 of these interviews."

2.3.3. RFPM/SmartIntake®

"Smartphone-based methods can record food intake based on pictures of food that you capture with a smartphone app. Specifically, you would use an app to take pictures of your meals before and after you eat. If it is not clear what you are eating or drinking, you would type in a brief description of those foods. The app then automatically sends the pictures and information you entered to nutrition professionals. Those nutrition professionals can then estimate how much you ate and drank based on the pictures. The app also automatically reminds you to capture images of your meals. Those reminders are customized based on your schedule and eating habits."

2.3.4. PortionSize®

"More recently, smartphone apps have been developed that do not require the analysis of the food images by a nutrition professional. Rather, you would estimate the portion size of the foods directly in the app. You would take a picture of your meal before you ate. You would then identify the foods and beverages in the meal via a drop-down list or with a search function. To estimate the portion size of the foods in the picture, you would do one of two things. First, you can enter in the amount of all food(s) consumed or the size of the food if it is known (e.g., 4 Famous Amos cookies, one 12-ounce regular Coke). Second, you can use templates that appear in the app. You can adjust the size of these templates and move them within the picture of the foods. Hence, to estimate portion size, you would change the size of the template and move it so it covers the food. Figure 1A illustrates an example, in which a template that looks like a deck of cards was placed over scrambled eggs. The app then uses this information to automatically and immediately estimate how much food is on your plate. After the meal, you enter if you ate everything, left a certain amount of the food on your plate, or you can use the templates again to estimate large portions of leftovers. This allows the app to provide a more accurate estimate of how much you ate. The app gives you real-time feedback about how many calories you ate and the nutrient composition of your meal and overall diet." An example of that feedback is provided in Figure 1B.

2.4. Measures

2.4.1. Preference of Methods

Participants ranked the 4 methods for food intake and alcohol consumption from most to least preferred, assuming they would use each method to record their food intake/alcohol consumption for 3 days as part of a clinical or study setting. The timeframe of 3 days is commonly used in clinical studies that aim to assess food and (alcoholic) drink intake [9,13–15].

2.4.2. Expected Burden of Methods

Similar to the preference of methods, participants further ranked the 4 methods for food intake and alcohol consumption from least to most burdensome according to their expected burden, assuming they would use each method to record food intake/alcohol consumption for 3 days as part of a clinical or study setting.

2.4.3. Willingness to Use Methods

Furthermore, participants rated their willingness to use a food/drink record, the RFPM/SmartIntake®, and PortionSize® to record their food intake and alcohol consumption over a period of 3 days on an 8-point Likert scale from 'not at all' (rated as 1) to 'very

much' (rated as 8). For the 24 h recall, participants rated their willingness to complete 3 separate recall interviews on the same Likert scale.

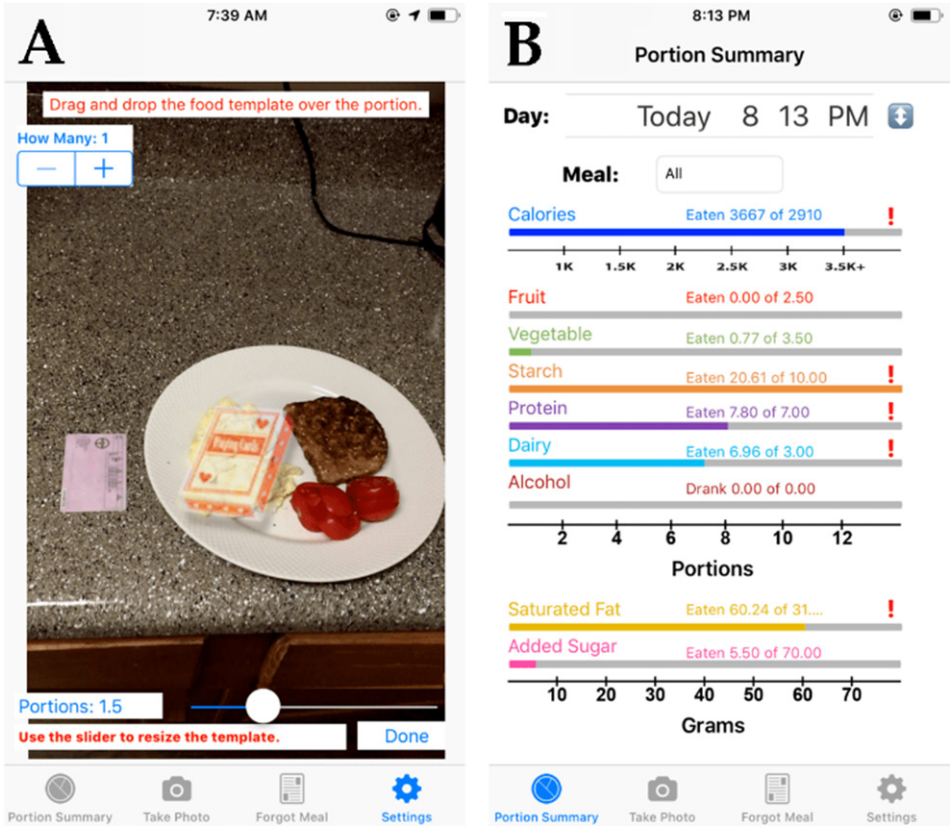


Figure 1. Screenshots from the PortionSize® app illustrating an example in which a template that looks like a deck of cards was placed over scrambled eggs to estimate their portion size (**A**) as well as an example of the real-time feedback (calories and nut and nutrient composition of the meal and the overall diet) provided by the PortionSize® app after the automatic portion size estimation (**B**).

2.5. Statistical Analyses

Data were cross-sectional and analyzed descriptively. We present categorical variables as frequency (%) and continuous data, including Likert scale items, as mean (standard deviation (SD)). For the main analysis, to simplify the presentation of results and allow for a natural binary comparison, we categorized the 'most preferred' (1st choice) and 'second-most preferred' (2nd choice) as 'preferred', and similarly, the 'second-least preferred' (3rd choice) and 'least preferred' (4th choice) as 'not preferred', in addition to the individual ranks (1st through 4th choice). We examined differences between 'preferred' and 'not preferred' for each method as well as differences in ratings across methods with chi-square statistics. In additional exploratory subgroup analyses, we assessed the effect of age (categories: <25 years, 25–34 years, 35–44 years, 45–54 years, 55–64 years, and ≥65 years), sex (male vs. female), race (categories: White, Black, Native American, Asian or Pacific Islander, and other), education (categories: less than high school, high school or equivalent, bachelor's degree, master's degree, doctorate, and other), household income (categories: <USD 10,000, USD 10,000–50,000, USD 50,000–100,000, USD 100,000–150,000, and >USD 150,000), household food security (categories: high food security, low food security, and very low food security), subjective social status, BMI (categories: <25 kg/m^2, 25.0–29.9 kg/m^2, and ≥30.0 kg/m^2, calculated from self-reported height and weight), and cardiometabolic diseases (positive past or present diagnosis vs. no diagnosis) on the preference of methods. For the expected burden of methods, similar to the preference of methods, we categorized the 'least burdensome' (1st choice) and 'second-least burdensome' (2nd choice) as 'low expected burden', and the 'second-most burdensome' (3rd choice) and

'most burdensome' (4th choice) as 'high expected burden', in addition to the individual ranks (1st through 4th choice). We examined differences between 'low expected burden' and 'high expected burden' for each method as well as differences in ratings across methods with chi-square analyses and ran the same subgroup analyses as for the preference of methods. Correlations between preference (1st through 4th choice) and expected burden (1st through 4th choice) of methods for food intake as well as alcohol consumption were analyzed using Spearman's rank correlation coefficient. Differences in willingness to use each method to monitor food intake and alcohol consumption over 3 days were assessed by analysis of covariance (ANCOVA), and we used a Tukey adjustment for post hoc pairwise comparisons. Differences in willingness as examined by the Kruskal–Wallis test did not differ meaningfully (not reported). All analyses were conducted in SPSS version 25. Due to multiple comparisons, the significance level was set to 0.001.

3. Results

3.1. Participant Characteristics

A total of 3245 adults participated in the online Pennington Habits Survey, and 1959 participants completed the Food Intake Assessment Preference Survey and are included in the main analyses. Participant characteristics are provided in Table 1. On average, participants (78.5% women, 78.2% White) were 45.9 (SD: 16.4) years old and had a BMI of 30.8 (SD: 8.8), with the majority of participants having either overweight (26.6%) or obesity (45.0%). Most participants indicated a high school diploma or equivalent (29.9%) or bachelor's degree (34.7%) as their highest level of education, and two-thirds reported a household income between USD 10,000 and USD 100,000. Three-quarters of the participants reported high household food security, and 64.8% of the participants saw themselves on rung 5, 6, or 7 of the MacArthur Scale of Subjective Social Status.

Table 1. Characteristics of participants who completed the Food Intake Assessment Preference Survey (N = 1959).

Sex, *n* (%)		
Male	418	(21.3)
Female	1537	(78.5)
Other	4	(0.2)
Age Category, *n* (%)		
<25 years	214	(10.9)
25–34 years	366	(18.7)
35–44 years	394	(20.1)
45–54 years	309	(15.8)
55–64 years	380	(19.4)
≥65 years	296	(15.1)
Age (years), mean (SD)	45.9	(16.4)
Race		
White	1532	(78.2)
Black	334	(17.0)
Native American	14	(0.7)
Asian or Pacific Islander	41	(2.1)
Other	38	(1.9)
Education, *n* (%)		
Less than High School	15	(0.8)
High School or Equivalent	586	(29.9)
Bachelor's Degree	680	(34.7)
Master's Degree	396	(20.2)
Doctorate	108	(5.5)
Other	174	(8.9)
Household Income, *n* (%)		
<USD 10,000	120	(6.1)
USD 10,000–50,000	633	(32.3)
USD 50,000–100,000	666	(34.0)

Table 1. *Cont.*

USD 100,000–150,000	359	(18.3)
>USD 150,000	181	(9.2)
Household Food Security, *n* (%) [a]		
High Food Security	1450	(74.0)
Low Food Security	227	(11.6)
Very Low Food Security	282	(14.4)
Subjective Social Status (1 = lowest, 10 = highest), *n* (%) [b]		
10	27	(1.4)
9	55	(2.8)
8	175	(8.9)
7	407	(20.8)
6	455	(23.2)
5	408	(20.8)
4	237	(12.1)
3	131	(6.7)
2	41	(2.1)
1	23	(1.2)
BMI Category, *n* (%) [c]		
<25 kg/m^2	537	(27.4)
25.0–29.9 kg/m^2	522	(26.6)
≥30.0 kg/m^2	881	(45.0)
BMI (kg/m^2), mean (SD) [c]	30.8	(8.8)
Cardiometabolic Diseases [d]		
≥1 Disease	830	(42.4)
No Diseases	1129	(57.6)

[a] Assessed with the 6-item Short Form of the U.S. Household Food Security Survey Module. [b] Assessed with the MacArthur Scale of Subjective Social Status. The scale presets a 'social ladder' and asks individuals to place an 'X' on the rung (1–10) on which they feel they stand compared to other people in the United States. [c] Data available for 1940 of 1959 participants. BMI was calculated from self-reported height and weight. [d] Heart disease, type 2 diabetes, hypertension, or dyslipidemia.

A subsample of 466 participants reported consuming alcohol at the time of the survey and completed the Alcohol Consumption Assessment Preference Questionnaire. Characteristics of those participants are provided in Table 2. On average, participants had been consuming alcohol regularly for 26.5 (SD: 15.9) years. Three-quarters of the participants reported drinking no more than 1–2 times per week (34.3%) or only on special occasions (40.3%), and 86.5% of the participants reported drinking ≤ 3 drinks per typical drinking occasion.

Table 2. Characteristics of participants who completed the Alcohol Consumption Assessment Preference Survey (N = 466).

Sex, *n* (%)		
Male	116	(24.9)
Female	350	(75.1)
Age Category, *n* (%)		
<25 years	33	(7.1)
25–34 years	81	(17.4)
35–44 years	103	(22.1)
45–54 years	79	(17.0)
55–64 years	101	(21.7)
≥65 years	69	(14.8)
Age (years), mean (SD)	45.7	(15.6)
Race		
White	387	(83.0)
Black	64	(13.7)
Native American	2	(0.4)
Asian or Pacific Islander	6	(1.3)
Other	7	(1.5)

Table 2. *Cont.*

Education, *n* (%)		
Less than High School	2	(0.4)
High School or Equivalent	101	(21.7)
Bachelor's Degree	188	(40.3)
Master's Degree	110	(23.6)
Doctorate	30	(6.4)
Other	35	(7.5)
Household Income, *n* (%)		
<USD 10,000	18	(3.9)
USD 10,000–50,000	120	(25.8)
USD 50,000–100,000	156	(33.5)
USD 100,000–150,000	108	(23.2)
>USD 150,000	64	(13.7)
Food Security, *n* (%) [a]		
High Food Security	361	(77.5)
Low Food Security	47	(10.1)
Very Low Food Security	58	(12.4)
Subjective Social Status (1 = lowest, 10 = highest), *n* (%) [b]		
10	4	(0.9)
9	15	(3.2)
8	48	(10.3)
7	109	(23.4)
6	137	(29.4)
5	86	(18.5)
4	46	(9.9)
3	15	(3.2)
2	5	(1.1)
1	1	(0.2)
BMI Category, *n* (%) [c]		
<25 kg/m^2	113	24.2
25.0–29.9 kg/m^2	119	25.5
≥30.0 kg/m^2	232	49.8
BMI (kg/m^2), mean (SD) [c]	31.2	(8.5)
Cardiometabolic Diseases [d]		
≥1 disease	203	(43.6)
No diseases	263	(56.4)
Alcohol consumption history (years), mean (SD)	26.5	(15.9)
Average alcohol consumption frequency, *n* (%)		
Everyday	24	(5.2)
3–5 times per week	94	(20.2)
1–2 times per week	160	(34.3)
Only on special occasions	188	(40.3)
Number of drinks on typical drinking days, *n* (%) [e]		
1 drink	173	(37.1)
2–3 drinks	230	(49.4)
3–5 drinks	51	(10.9)
>5 drinks	12	(2.6)

[a] Assessed with the 6-item Short Form of the Food Security Survey Module. [b] Assessed with the MacArthur Scale of Subjective Social Status. The scale presets a 'social ladder' and asks individuals to place an 'X' on the rung (1–10) on which they feel they stand compared to other people in the United States. [c] Data available for 464 of 466 participants. BMI was calculated from self-reported height and weight. [d] Heart disease, type 2 diabetes, hypertension, or dyslipidemia. [e] The following examples for a standard drink were given: 12 oz. regular beer, 5 oz. regular wine, 1.5 oz. distilled spirits.

3.2. Methods of Food Intake Assessment

3.2.1. Preference of Methods

Figure 2A shows the percentage of participants who rated each method as their preferred method of food intake assessment. The RFPM/SmartIntake® was rated as the preferred method by the largest percentage of participants (67.3%), while the 24 h recall was rated as the preferred method by the smallest percentage of participants (32.9%). The

food record and PortionSize® were rated as the preferred method by 48.0% and 51.9%, respectively, with an overall difference in preference across all methods ($p < 0.001$). Pairwise comparisons showed differences in preference ratings between all methods ($p < 0.001$), except between the food record and PortionSize® ($p = 0.06$). Table 3 displays the frequencies (%) for the individual ranks (1st through 4th choice) for the different methods. There was an overall difference in preference ratings across all methods ($p < 0.001$), and pairwise comparisons showed differences between all methods ($p < 0.001$), except between the food record and PortionSize® ($p = 0.35$).

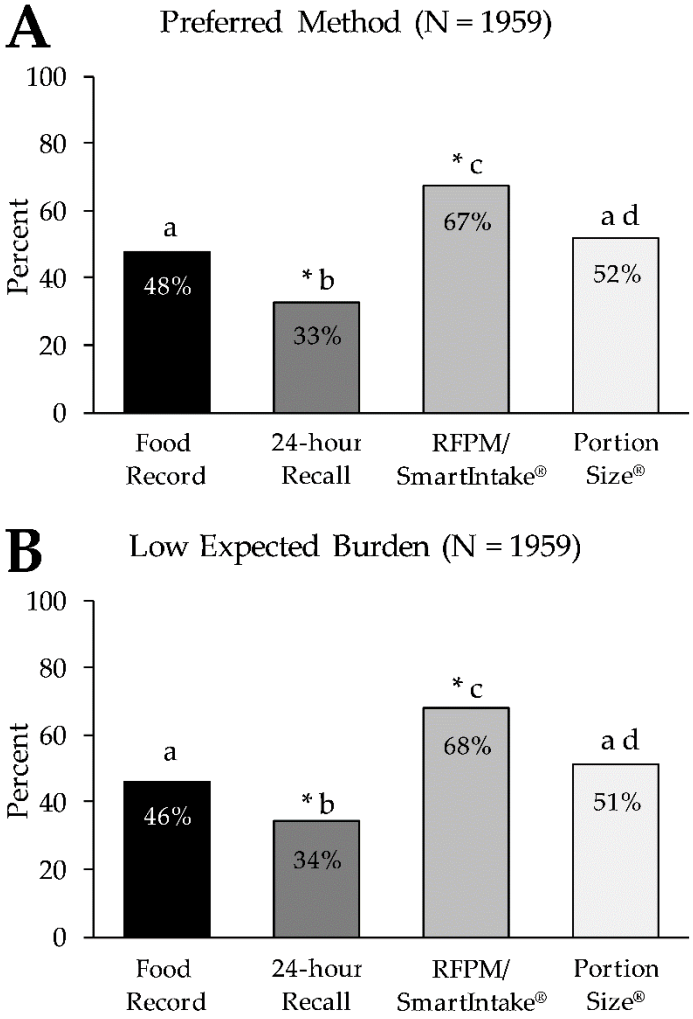


Figure 2. Percentage of participants who rated each method as their preferred method of food intake assessment (Panel **A**) and who rated the expected burden of each method as low (Panel **B**). * Denotes a significant difference in 'not preferred' (Panel **A**) and in 'high expected burden' (Panel **B**) for the respective method ($p < 0.001$) in the same panel. Letters (a–d) that differ from each other indicate differences between methods ($p < 0.001$) in the same panel. RFPM, Remote Food Photography Method.

Table 3. Preference, expected burden, and willingness to use the four methods of food intake assessment (N = 1959).

Preference of Methods, *n* (%) [a]		
Food Record		
Most preferred (first choice)	639	(32.6)
Second-most preferred (second choice)	301	(15.4)
Second-least preferred (third choice)	464	(23.7)
Least preferred (fourth choice)	555	(28.3)

Table 3. *Cont.*

24-h Recall		
Most preferred (first choice)	198	(10.1)
Second-most preferred (second choice)	446	(22.8)
Second-least preferred (third choice)	592	(30.2)
Least preferred (fourth choice)	723	(36.9)
Remote Food Photography Method via SmartIntake® App		
Most preferred (first choice)	631	(32.2)
Second-most preferred (second choice)	687	(35.1)
Second-least preferred (third choice)	530	(27.0)
Least preferred (fourth choice)	111	(5.7)
PortionSize®		
Most preferred (first choice)	491	(25.1)
Second-most preferred (second choice)	525	(26.8)
Second-least preferred (third choice)	373	(19.0)
Least preferred (fourth choice)	570	(29.1)
Expected burden of methods, *n* (%) [b]		
Food Record		
Least burdensome (first choice)	580	(29.6)
Second-least burdensome (second choice)	326	(16.6)
Second-most burdensome (third choice)	477	(24.4)
Most burdensome (fourth choice)	576	(29.4)
24-h Recall		
Least burdensome (first choice)	229	(11.7)
Second-least burdensome (second choice)	442	(22.6)
Second-most burdensome (third choice)	578	(29.5)
Most burdensome (fourth choice)	710	(36.2)
Remote Food Photography Method via SmartIntake® App		
Least burdensome (first choice)	686	(35.0)
Second-least burdensome (second choice)	648	(33.1)
Second-most burdensome (third choice)	518	(26.4)
Most burdensome (fourth choice)	107	(5.5)
PortionSize®		
Least burdensome (first choice)	464	(23.7)
Second-least burdensome (second choice)	543	(27.7)
Second-most burdensome (third choice)	386	(19.7)
Most burdensome (fourth choice)	566	(28.9)
Willingness to use method over 3 days, mean (SD) [c]		
Food Record	6.6	(2.0)
24-h Recall	6.1	(2.2)
Remote Food Photography Method via SmartIntake® App	6.6	(2.0)
PortionSize®	6.0	(2.2)

[a] Participants were asked to rank the four methods from the most to the least preferred. [b] Participants were asked to rank the four methods from the least to the most burdensome. [c] Likert Scale: 1 = not at all willing, 8 = very much willing.

We found a significant age effect on the preference ratings of the four methods ($p < 0.001$). This effect was primarily driven by the difference between those <65 years and those ≥65 years of age (Figure 3). A greater percentage of participants ≥65 years rated the food record (61.4%) and 24 h recall (45.3%) as a preferred method compared to participants <65 years (food record: 45.5%, 24 h recall: 30.6%, all $p < 0.001$). Conversely, a greater percentage of participants <65 years rated the RFPM/SmartIntake® (70.1%) and PortionSize® (53.7%) as their preferred method compared to those ≥65 years (RFPM/SmartIntake®: 51.7%, PortionSize®: 41.6%, all $p < 0.001$). Supplementary Figure S1 shows the preference ratings of the four methods for all age categories. Differences in ratings of the four methods between two age categories were only significant when compared with those ≥65 years, with the single exception of PortionSize®, which had a significantly higher preference rating in those aged 35–44 (61.2%) compared to those aged 25–34 (46.7%, $p < 0.001$). All other comparisons of methods across age categories were not significant (all $p \geq 0.05$).

We also found a significant disease effect on the preference ratings of the four methods ($p = 0.001$). However, pairwise comparisons of the preference ratings of the four methods between those with a past or present diagnosis of cardiometabolic disease and those with no such diagnosis were all not significant (all $p \geq 0.06$). We did not find any effects on the preference ratings for the four methods by sex, race, education level, household income, household food security, subjective social status, or BMI (all $p \geq 0.07$; data not shown).

3.2.2. Expected Burden of Methods

Figure 2B shows the percentage of participants who rated the expected burden of each method of food intake assessment as low. The largest percentage of participants (68.1%) rated the expected burden of the RFPM/SmartIntake® as low, followed by PortionSize® (51.4%), the food record (46.2%), and the 24 h recall (34.3%), with an overall difference in expected burden across all methods ($p < 0.001$). Pairwise comparisons showed differences in preference ratings between all methods ($p < 0.001$), except between the food record and PortionSize® ($p = 0.06$). Table 3 displays the frequencies (%) for the individual ranks (1st through 4th choice) for the different methods. There was an overall difference in preference ratings across all methods ($p < 0.001$), and pairwise comparisons showed differences between all methods ($p < 0.001$), except between the food record and PortionSize® ($p = 0.96$).

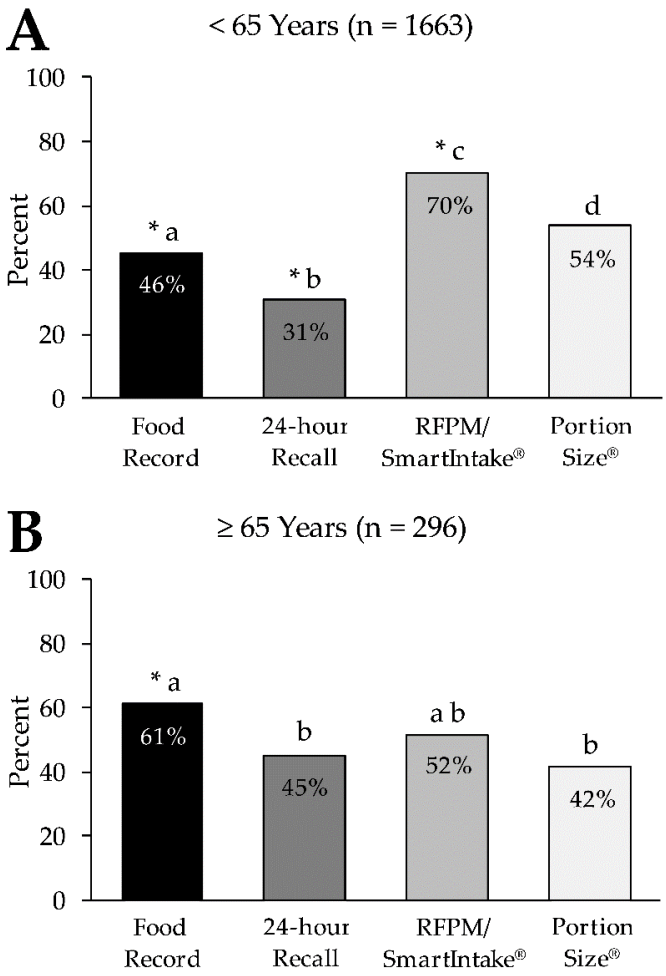


Figure 3. Percentage of participants who rated each method as their preferred method of food intake assessment in those <65 years (Panel **A**) and those ≥65 years of age (Panel **B**). * Denotes a significant difference in 'not preferred' for the respective method ($p < 0.001$) in the same panel. Letters (a–d) that differ from each other indicate differences between methods ($p < 0.001$) in the same panel. RFPM, Remote Food Photography Method.

Similar to the preference of methods, we found a significant age effect on ratings of the expected burden of the four methods ($p < 0.001$), and this effect appeared to be driven by the difference between those <65 years and those ≥65 years of age. A greater percentage of participants ≥65 years rated the expected burden of the food record (61.1%) and 24 h recall (46.6%) as low compared to participants <65 years (food record: 43.6%, 24 h recall: 32.0%, all $p < 0.001$). Conversely, a greater percentage of participants <65 years rated the expected burden of the RFPM/SmartIntake® (71.3%) and PortionSize® (53.1%) as low compared to those ≥65 years (RFPM/SmartIntake®: 50.3%, PortionSize®: 41.9%, all $p < 0.001$). Differences in ratings of the expected burden of the four methods between all other age categories were only significant when compared with those ≥65 years. All other comparisons of methods across age categories were not significant (all $p \geq 0.07$). We did not find any effects on preference ratings for the four methods by sex, race, education level, household income, household food security, subjective social status, BMI, or diagnosis of cardiometabolic disease (all $p \geq 0.06$; data not shown).

3.2.3. Correlation between Preference and Expected Burden of Methods

High preference was strongly correlated with low expected burden of the respective method for all methods of food intake assessment, with coefficients of $\rho = 0.85$ ($p < 0.001$) for the food record, $\rho = 0.85$ ($p < 0.001$) for the 24 h recall, $\rho = 0.82$ ($p < 0.001$) for the RFPM/SmartIntake®, and $\rho = 0.85$ ($p < 0.001$) for PortionSize®.

3.2.4. Willingness to Use Methods

The willingness to use the method over 3 days to monitor food intake was rated (8-point Likert scale) with a mean of 6.6 (SD: 2.0) for the food record, 6.1 (SD: 2.2) for the 24 h recall, 6.6 (SD: 2.0) for the RFPM/SmartIntake®, and 6.0 (SD: 2.2) for PortionSize® (Table 3), with a significant main effect ($p < 0.001$). Post hoc comparisons showed that willingness to use the food record differed from the 24 h recall ($p < 0.001$) and PortionSize® ($p < 0.001$) but not from the RFPM/SmartIntake® ($p = 0.96$). Willingness to use the 24 h recall differed from the RFPM/SmartIntake® ($p < 0.001$) but not from PortionSize® ($p = 0.23$), and willingness to use the RFPM/SmartIntake® was greater than PortionSize® ($p < 0.001$).

3.3. Preference of Methods for Alcohol Consumption Assessment

3.3.1. Preference of Methods

Figure 4A shows the percentage of participants who rated each method as their preferred method of alcohol consumption assessment. Similar to the food intake assessment, the RFPM/SmartIntake® was rated as the preferred method by the largest percentage of participants (63.3%), while the 24 h recall was rated as the preferred method by the smallest percentage of participants (33.9%). The food record and PortionSize® were rated as the preferred method by 49.3% and 53.4% respectively, with an overall difference in preference across all methods ($p < 0.001$). Pairwise comparisons showed differences in preference ratings between all methods ($p < 0.001$), except between the food record and PortionSize® ($p = 0.34$). Table 4 displays the frequencies (%) for the individual ranks (1st through 4th choice) for the different methods. There was an overall difference in preference ratings across all methods ($p < 0.001$), and pairwise comparisons showed differences between all methods ($p < 0.001$), except between the food record and PortionSize® ($p = 0.26$). We did not find any effects on preference ratings for the four methods by sex, age, race, education level, household income, household food security, subjective social status, BMI, or diagnosis of cardiometabolic disease (all $p \geq 0.13$).

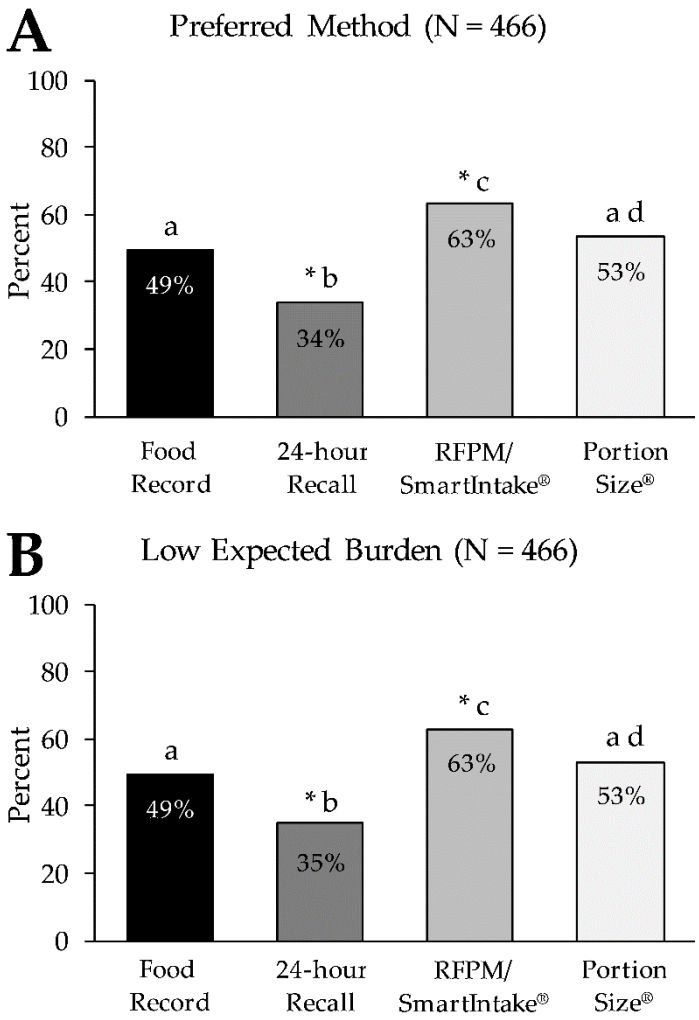


Figure 4. Percentage of participants who rated each method as their preferred method of alcohol consumption assessment (Panel **A**) and who rated the expected burden of each method as low (Panel **B**). * Denotes a significant difference in 'not preferred' for the respective method ($p < 0.001$) in the same panel. Letters (a–d) that differ from each other indicate differences between methods ($p < 0.001$) in the same panel. RFPM, Remote Food Photography Method.

Table 4. Preference, expected burden, and willingness to use the four methods of alcohol consumption assessment (N = 466).

Preference of Methods, *n* (%) [a]		
Food Record		
Most preferred (first choice)	173	(37.1)
Second-most preferred (second choice)	57	(12.2)
Second-least preferred (third choice)	128	(27.5)
Least preferred (fourth choice)	108	(23.2)
24 h Recall		
Most preferred (first choice)	54	(11.6)
Second-most preferred (second choice)	104	(22.3)
Second-least preferred (third choice)	117	(25.1)
Least preferred (fourth choice)	191	(41.0)
Remote Food Photography Method via SmartIntake® App		
Most preferred (first choice)	108	(23.2)
Second-most preferred (second choice)	187	(40.1)
Second-least preferred (third choice)	142	(30.5)
Least preferred (fourth choice)	29	(6.2)

Table 4. *Cont.*

PortionSize®		
Most preferred (first choice)	131	(28.1)
Second-most preferred (second choice)	118	(25.3)
Second-least preferred (third choice)	79	(17.0)
Least preferred (fourth choice)	138	(29.6)
Expected burden of methods, *n* (%) [b]		
Food Record		
Least burdensome (first choice)	166	(35.6)
Second-least burdensome (second choice)	64	(13.8)
Second-most burdensome (third choice)	125	(26.8)
Most burdensome (fourth choice)	111	(23.8)
24 h Recall		
Least burdensome (first choice)	56	(12.0)
Second-least burdensome (second choice)	106	(22.7)
Second-most burdensome (third choice)	118	(25.3)
Most burdensome (fourth choice)	186	(39.9)
Remote Food Photography Method via SmartIntake® App		
Least burdensome (first choice)	124	(26.6)
Second-least burdensome (second choice)	170	(36.5)
Second-most burdensome (third choice)	140	(30.0)
Most burdensome (fourth choice)	32	(6.9)
PortionSize®		
Least burdensome (first choice)	120	(25.8)
Second-least burdensome (second choice)	126	(27.0)
Second-most burdensome (third choice)	83	(17.8)
Most burdensome (fourth choice)	137	(29.4)
Willingness to use method over 3 days, mean (SD) [c]		
Food Record	6.5	(2.3)
24 h Recall	5.7	(2.7)
Remote Food Photography Method via SmartIntake® App	6.4	(2.4)
PortionSize®	6.0	(2.4)

[a] Participants were asked to rank the four methods from the most to the least preferred. [b] Likert Scale: 1 = not at all willing, 8 = very much willing. [c] Participants were asked to rank the four methods from the least to the most burdensome.

3.3.2. Expected Burden of Methods

Figure 4B shows the percentage of participants who rated the expected burden of each method of alcohol consumption assessment as low. Similar to the food intake assessment, the largest percentage of participants (63.1%) rated the expected burden of the RFPM/SmartIntake® as low, followed by PortionSize® (52.8%), the food record (49.4%), and the 24 h recall (34.7%), with an overall difference in expected burden across all methods ($p < 0.001$). Pairwise comparisons showed differences in the ratings of expected burden between all methods ($p < 0.001$), except between the food record and PortionSize® ($p = 0.43$). Table 4 displays the frequencies (%) for the individual ranks for the different methods. There was an overall difference in preference ratings across all methods ($p < 0.001$), and pairwise comparisons showed differences between all methods ($p < 0.001$), except between the food record and PortionSize® ($p = 0.23$). We did not find any effects on the ratings of the expected burden of the four methods by sex, age, race, education level, household income, household food security, subjective social status, BMI, or diagnosis of cardiometabolic disease (all $p \geq 0.06$).

3.3.3. Correlation between Preference and Expected Burden of Methods

Similar to food intake, high preference was strongly correlated with a low expected burden of the respective method for all methods of alcohol consumption assessment with coefficients of $\rho = 0.91$ ($p < 0.001$) for the food record, $\rho = 0.89$ ($p < 0.001$) for the 24 h recall, $\rho = 0.90$ ($p < 0.001$) for the RFPM/SmartIntake®, and $\rho = 0.91$ ($p < 0.001$) for PortionSize®.

3.3.4. Willingness to Use Methods

Willingness to use each method over 3 days to monitor alcohol consumption was an average of 6.5 (SD: 2.3) for the food record, 5.7 (SD: 2.7) for the 24 h recall, 6.4 (SD: 2.4) for the RFPM/SmartIntake®, and 6.0 (SD: 2.4) for PortionSize® (Table 4), with a significant main effect ($p < 0.001$). Post hoc comparisons showed that the willingness to use the food record differed from the 24 h recall and PortionSize® (all $p < 0.001$) but not from the RFPM/SmartIntake® ($p = 0.33$). Willingness to use the 24 h recall differed from the RFPM/SmartIntake® ($p < 0.001$) but not from PortionSize® ($p = 0.50$), and willingness to use the RFPM/SmartIntake® was greater than PortionSize® ($p < 0.001$).

4. Discussion

The present study used an online survey to assess the preference for and the expected burden of four different methods of food intake and alcohol consumption assessment (food record, 24-h recall, RFPM/SmartIntake®, PortionSize®) as well as the willingness to use these methods to record food intake and alcohol consumption over 3 days.

In line with our hypotheses for both food intake and alcohol consumption assessment, the RFPM/SmartIntake® was rated more preferred and less burdensome than more traditional methods (food record and 24-h recall), which is consistent with previous findings [7,9]. Additionally, the RFPM/SmartIntake® was perceived as less burdensome and more preferred compared to PortionSize®, as hypothesized. As illustrated in Figure 2; Figure 4, the graphs depicting the preference for and expected burden of the methods are virtually superimposable, and the strong correlations between high preference and low expected burden for all methods ($\rho \geq 0.82$; $p < 0.001$) support the hypothesis that expected burden influences method preference. It has previously been reported that more burdensome methods yield low adherence [1–3]. It is conceivable that methods that are more preferred by participants (due to lower perceived burden) lead to better adherence and consequently greater data quality over a longer period of time (or during repeated assessment periods) compared to less preferred methods. The difference in ratings of expected burden between the RFPM/SmartIntake® and PortionSize® could mean that the participants recognized that the self-estimation of the portion size of foods in PortionSize® would require more effort on their part, while this burden is shifted to a researcher or clinician when using RFPM/SmartIntake®. Furthermore, our hypothesis that PortionSize® would be rated more preferred than the traditional methods was only partially supported, as PortionSize® was only rated more preferred and less burdensome compared to the 24-h recall but not to the food record. This suggests that while mHealth technology certainly holds promise in reducing the burden of dietary self-monitoring [16,17], some mHealth methods are more burdensome than others, and differences in burden are detected by prospective users. Additionally, some mHealth methods are likely to be perceived as more burdensome in relation to streamlined self-report methods such as a checklist associated with a structured meal plan. The overall low preference along with the relatively high expected burden of the 24-h recall is somewhat surprising, particularly when compared to the food record. The 24-h recall, as described to participants, requires three relatively brief (20–30 min) recall interviews, whose de facto time burden is likely less than keeping a detailed food record for 3–7 days [18], which in many studies requires participants to log every eating occasion, preferably in real time and as comprehensively as possible, in order to improve weight loss outcomes [19,20]. It can be conjectured that other factors of the 24-h recall were unappealing to participants, such as the interview format, and that this influenced the ratings of burden and preference for that method. Additionally, participants did not use each method, and actual experience with the methods could influence the ratings.

The preference for and expected burden of methods was independent of sex, race, education level, household income, household food security, subjective social status, BMI, or diagnosis of cardiometabolic disease for both food intake and alcohol consumption assessment. However, we must acknowledge that our participants were predominantly women (78.5%), White (78.2%), food secure (74.0%), with overweight or obesity (71.6%),

and had a college education (60.4%). A more heterogeneous sample might have led to different results. Age affected the preference ratings of the methods for assessing food intake but not alcohol consumption. While there was a clear pattern in those <65 years with RFPM/SmartIntake® > PortionSize® > food record > 24-h recall (from the most to the least preferred), in those ≥65 years, preferences were more evenly distributed across the four methods. The food record (61.4%) was the method that was preferred by the largest percentage in those ≥65 years, followed by the RFPM/SmartIntake® (51.7%; not different from food record). The 24-h recall (45.3%) and PortionSize® (41.6%; no difference between the 2 methods) were less preferred compared to the food record. A recent cross-sectional study (N = 364) that assessed older adults' intention to use mHealth apps showed that 49.7% of the participants (mean age 75 (SD: 7) years) had no intention to use any such apps [21], which might explain why participants ≥65 in our study rated app-based methods (RFPM/SmartIntake®, PortionSize®) as relatively less preferred than traditional methods (food record, 24-h recall) compared to participants <65 years. However, while the food record was the most preferred method to assess food intake among those ≥65 years, preference for the RFPM/SmartIntake® did not differ significantly and more than half of those ≥65 years still rated the RFPM/SmartIntake® as their preferred method. The RFPM/SmartIntake® is accurate and more accurate than food records [6,9,22]; hence, the RFPM/SmartIntake® remains a viable method for participants ≥65 years of age. Nevertheless, the ability of older adults to reliably use apps needs to be evaluated before data collection. Additional support from the study staff and an easy-to-use interface of the apps may help overcome potential barriers of app-based methods and increase the acceptability and comfort of these technologies [23], which would ensure reliable use and high-quality data collection throughout the study period.

Furthermore, and in line with the overall findings on preference and expected burden of methods, the willingness to use the RFPM/SmartIntake® was greater than that for the 24-h recall and PortionSize® for assessing both food and alcohol intake. However, despite the overall greater preference and lower expected burden of the RFPM/SmartIntake® compared to the food record, the willingness to use each method for 3 days did not differ. This is interesting and suggests that while a low expected burden of a method is strongly correlated with a high preference for that method, the willingness to use the method to record food intake or alcohol consumption over 3 days is not necessarily affected by a higher expected burden in the same way as the preference for the same method. Over a longer time frame, this may be different. Nevertheless, the relatively high willingness to use the food record may have been influenced by a greater familiarity with the method compared to other methods. Food records are similar to popular diet-tracking apps such as MyFitnessPal (50 million downloads for Android in 2017 [24]) and are likely better known to the majority of the participants than the 24-h recall and especially a new, image-based app that requires the self-estimation of the portion size of foods (PortionSize®). Furthermore, while willingness differed statistically between several methods, it needs to be acknowledged that the range between the method with the highest willingness (6.6 points for food intake and 6.5 points for alcohol consumption on a 1–8 Likert scale) and the lowest willingness (6.0 points for food intake and 5.7 points for alcohol consumption) was less than 1 point and may be of little practical value. The small difference between methods along with the relatively high willingness ratings across all methods (≥6.0 points for food intake and ≥5.7 points for alcohol consumption) suggests that participants generally seem willing to use all four methods for study/research purposes, even if they rate one method to be relatively more burdensome and less preferred than another. This is encouraging as more frequent engagement with (digital) food logging methods in behavioral interventions has been reported to be associated with greater weight loss over 3 [25] and 12 [26] months. Nonetheless, the actual usage of self-report methods to assess food intake decreases over time, even when efforts are made to reduce burden by, for example, using photographic food records [27].

The findings of the study have implications in clinical research. First, when assessing food/drink intake as an outcome variable in a research study, the same assessment method must be used for all participants. Based on the results of the present study, the demographic characteristics of the study sample (e.g., older vs. younger participants) should be considered when selecting an assessment method since preference and the perceived burden of methods differed by age. This can facilitate the selection of a method that is most preferred by the majority of the study sample. Researchers might also choose to provide standardized training to participants prior to data collection in order to increase comfort with the assessment method; additionally, the need for or the intensity of the training may be influenced by the demographics of the sample. Researchers might also conduct a pilot study or needs assessment in the target population prior to starting a study to determine preferences, feasibility, and training needs. Second, the study has implications in the selection of methods for assessing food/drink intake when delivering an intervention during a clinical study. In this case, it is feasible for different participants to use different methods since the information is primarily used to facilitate intervention delivery and is not used as an outcome variable. Allowing participants to use their preferred method may increase the use of the method over time and further enhance intervention delivery.

A limitation of this analysis is that in the online survey, participants were only presented with descriptions (and illustrations) of the four methods but did not use any of the methods for the suggested period of 3 days to record food intake and alcohol consumption. While the assessment of participants' general preference and willingness to use certain methods to record food intake and alcohol consumption is important and can inform the choice of method for clinical and study settings, it is conceivable that participants' preference and willingness to use these methods might change after hands-on use for a given period. However, it has been shown that the participants who tracked their food intake with their preferred method (as indicated by the participants before the intervention) were approximately 50% more adherent to tracking food intake over 12 weeks compared with those who tracked with their non-preferred method [28]. It is therefore unlikely that the real-life use of the methods would change the pattern observed with the RFPM/SmartIntake® as the method has been found to be preferred compared to pen-and-paper records and diet recalls in two studies where participants used the methods [7,9]. Finally, we only included adults residing in the United States in this survey. While adults from all geographic regions in the United States participated (85.8% Louisiana), this limits the generalizability to only similarly developed and high-income nations.

5. Conclusions

For both food intake and alcohol consumption, the greatest percentage of participants rated the expected burden of the RFPM/SmartIntake® as low, followed by PortionSize® and food/drink records, and then the 24-h recall. Preference for these methods mirrored the ratings of expected burden, and correlations between low expected burden and high preference were strong for all methods. Because the participants' preference for a specific method as well as their expected burden of the method likely affect their compliance over time and thereby data quality, our results can be used in conjunction with existing data on the reliability and validity of these methods in order to inform the selection of assessment methods.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/nu13103340/s1, Supplementary Methods: Food Intake Assessment Preference Questionnaire, Alcohol Consumption Assessment Preference Questionnaire; Supplementary Figure S1: Percentage of participants who rated each method as their preferred method of food intake assessment in those <25 years (Panel A), 25–34 years (Panel B), 35–44 years (Panel C), 45–54 years (Panel D), 55–64 years (Panel E), and ≥65 years of age (Panel F). * Denotes a significant difference in 'not preferred' for the respective method ($p < 0.001$) in the same panel. Letters (a–d) that differ from each other indicate differences between methods ($p < 0.001$). RFPM, Remote Food Photography Method.

Author Contributions: C.H. and C.K.M. designed the Pennington Habits Survey. C.H. performed statistical analyses. C.H., N.F., J.L.D. and C.K.M. interpreted the data. C.H. drafted the manuscript and created tables and figures. N.F., J.L.D., T.L.F., C.A.M. and J.W.A. provided critical revision of the manuscript for important intellectual content. All authors have read and agreed to the published version of the manuscript.

Funding: C.H. is supported by a National Institutes of Health National Research Service Award (T32DK064584) and J.L.D. is supported by the American Heart Association (Grant No. 20POST35210907). This work was partially supported by the National Institutes of Diabetes and Digestive and Kidney Diseases grant R01 DK124558 and Louisiana State University LIFT^2 grant entitled: "PortionSize: A market-ready app to estimate and manage portion size". Further, this work was partially supported by NORC Center grant P30 DK072476 entitled "Nutrition and Metabolic Health through the Lifespan" sponsored by NIDDK. This work was also partially supported by grant # U54 GM104940 from the National Institute of General Medical Sciences of the National Institutes of Health, which funds the Louisiana Clinical and Translational Science Center. The funding sources listed above had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.

Institutional Review Board Statement: This Pennington Habits Survey was approved by the Institutional Review Board at Pennington Biomedical Research Center (PBRC, 2019-052-PBRC) and registered at ClinicalTrials.gov (NCT04150510) before the start of recruitment.

Informed Consent Statement: Interested individuals received the survey link as well as instructions that detailed the purpose of the study. Participants verified that they were adults and provided consent to participate before proceeding with the survey. Data were collected using Research Electronic Data Capture (REDCap) [10]. Participation in the survey was voluntary and participants had the option not to submit answers or to skip items if they did not wish to complete them.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author on reasonable request.

Acknowledgments: The authors would like to thank all participants for their time and commitment to the study.

Conflicts of Interest: The intellectual property surrounding the SmartIntake® and PortionSize® apps is owned by Pennington Biomedical Research Center/Louisiana State University. Author C.K.M. is the inventor of both apps, and author J.W.A. is the inventor of PortionSize®. The other authors report no conflict of interest.

References

1. Burke, L.E.; Wang, J.; Sevick, M.A. Self-Monitoring in Weight Loss: A Systematic Review of the Literature. *J. Am. Diet. Assoc.* **2011**, *111*, 92–102. [CrossRef]
2. Turner-McGrievy, G.M.; Dunn, C.G.; Wilcox, S.; Boutté, A.K.; Hutto, B.; Hoover, A.; Muth, E. Defining Adherence to Mobile Dietary Self-Monitoring and Assessing Tracking Over Time: Tracking at Least Two Eating Occasions per Day Is Best Marker of Adherence within Two Different Mobile Health Randomized Weight Loss Interventions. *J. Acad. Nutr. Diet.* **2019**, *119*, 1516–1524. [CrossRef] [PubMed]
3. Yu, Z.; Sealey-Potts, C.; Rodriguez, J. Dietary Self-Monitoring in Weight Management: Current Evidence on Efficacy and Adherence. *J. Acad. Nutr. Diet.* **2015**, *115*, 1931–1938. [CrossRef] [PubMed]
4. Holmes, B.; Dick, K.; Nelson, M. A Comparison of Four Dietary Assessment Methods in Materially Deprived Households in England. *Public Health Nutr.* **2008**, *11*, 444–456. [CrossRef]
5. Monnerie, B.; Tavoularis, L.G.; Guelinckx, I.; Hebel, P.; Boisvieux, T.; Cousin, A.; Le Bellego, L. A Cross-over Study Comparing an Online versus a Paper 7-Day Food Record: Focus on Total Water Intake Data and Participant's Perception of the Records. *Eur. J. Nutr.* **2015**, *54*, 27–34. [CrossRef] [PubMed]
6. Martin, C.K.; Correa, J.B.; Han, H.; Allen, H.R.; Rood, J.; Champagne, C.M.; Gunturk, B.K.; Bray, G.A. Validity of the Remote Food Photography Method (RFPM) for Estimating Energy and Nutrient Intake in near Real-Time. *Obesity* **2012**, *20*, 891–899. [CrossRef] [PubMed]
7. Martin, C.K.; Han, H.; Coulon, S.M.; Allen, H.R.; Champagne, C.M.; Anton, S.D. A Novel Method to Remotely Measure Food Intake of Free-Living Individuals in Real Time: The Remote Food Photography Method. *Br. J. Nutr.* **2009**, *101*, 446–456. [CrossRef]
8. Beasley, J.; Riley, W.T.; Jean-Mary, J. Accuracy of a PDA-Based Dietary Assessment Program. *Nutrition* **2005**, *21*, 672–677. [CrossRef]

9. Fazzino, T.L.; Martin, C.K.; Forbush, K. The Remote Food Photography Method and SmartIntake App for the Assessment of Alcohol Use in Young Adults: Feasibility Study and Comparison to Standard Assessment Methodology. *JMIR MHealth UHealth* **2018**, *6*, e10460. [CrossRef]
10. Harris, P.A.; Taylor, R.; Thielke, R.; Payne, J.; Gonzalez, N.; Conde, J.G. Research Electronic Data Capture (REDCap)—A Metadata-Driven Methodology and Workflow Process for Providing Translational Research Informatics Support. *J. Biomed. Inform.* **2009**, *42*, 377–381. [CrossRef] [PubMed]
11. Blumberg, S.J.; Bialostosky, K.; Hamilton, W.L.; Briefel, R.R. The Effectiveness of a Short Form of the Household Food Security Scale. *Am. J. Public Health* **1999**, *89*, 1231–1234. [CrossRef]
12. Adler, N.; Stewart, J. *The MacArthur Scale of Subjective Social Status*; MacArthur Research Network on SES & Health; University of California: San Francisco, CA, USA, 2007.
13. Coulston, A.; Boushey, C.; Ferruzzi, M.; Delahanty, L. (Eds.) *Nutrition in the Prevention and Treatment of Disease*, 4th ed.; Academic Press: Cambridge, MA, USA, 2017; ISBN 978-0-12-802947-3.
14. Yang, Y.J.; Kim, M.K.; Hwang, S.H.; Ahn, Y.; Shim, J.E.; Kim, D.H. Relative Validities of 3-Day Food Records and the Food Frequency Questionnaire. *Nutr. Res. Pract.* **2010**, *4*, 142–148. [CrossRef]
15. Martin, C.K.; Nicklas, T.; Gunturk, B.; Correa, J.B.; Allen, H.R.; Champagne, C. Measuring Food Intake with Digital Photography. *J. Hum. Nutr. Diet. Off. J. Br. Diet. Assoc.* **2014**, *27* (Suppl. 1), 72–81. [CrossRef]
16. Burke, L.E.; Styn, M.A.; Sereika, S.M.; Conroy, M.B.; Ye, L.; Glanz, K.; Sevick, M.A.; Ewing, L.J. Using MHealth Technology to Enhance Self-Monitoring for Weight Loss: A Randomized Trial. *Am. J. Prev. Med.* **2012**, *43*, 20–26. [CrossRef] [PubMed]
17. Turner-McGrievy, G.M.; Yang, C.-H.; Monroe, C.; Pellegrini, C.; West, D.S. Is Burden Always Bad? Emerging Low-Burden Approaches to Mobile Dietary Self-Monitoring and the Role Burden Plays with Engagement. *J. Technol. Behav. Sci.* **2021**, *6*, 447–455. [CrossRef]
18. Johnson, R.K. Dietary Intake—How Do We Measure What People Are Really Eating? *Obes. Res.* **2002**, *10*, 63S–68S. [CrossRef] [PubMed]
19. Harvey, J.; Krukowski, R.; Priest, J.; West, D. Log Often, Lose More: Electronic Dietary Self-Monitoring for Weight Loss. *Obesity* **2019**, *27*, 380–384. [CrossRef]
20. Peterson, N.D.; Middleton, K.R.; Nackers, L.M.; Medina, K.E.; Milsom, V.A.; Perri, M.G. Dietary Self-Monitoring and Long-Term Success with Weight Management. *Obesity* **2014**, *22*, 1962–1967. [CrossRef]
21. Askari, M.; Klaver, N.S.; van Gestel, T.J.; van de Klundert, J. Intention to Use Medical Apps Among Older Adults in the Netherlands: Cross-Sectional Study. *J. Med. Internet Res.* **2020**, *22*, e18080. [CrossRef] [PubMed]
22. Dhurandhar, N.V.; Schoeller, D.; Brown, A.W.; Heymsfield, S.B.; Thomas, D.; Sørensen, T.I.A.; Speakman, J.R.; Jeansonne, M.; Allison, D.B.; Energy Balance Measurement Working Group. Energy Balance Measurement: When Something Is Not Better than Nothing. *Int. J. Obes.* **2015**, *39*, 1109–1113. [CrossRef]
23. Bonilla, C.; Brauer, P.; Royall, D.; Keller, H.; Hanning, R.M.; DiCenso, A. Use of Electronic Dietary Assessment Tools in Primary Care: An Interdisciplinary Perspective. *BMC Med. Inform. Decis. Mak.* **2015**, *15*, 14. [CrossRef] [PubMed]
24. Ferrara, G.; Kim, J.; Lin, S.; Hua, J.; Seto, E. A Focused Review of Smartphone Diet-Tracking Apps: Usability, Functionality, Coherence With Behavior Change Theory, and Comparative Validity of Nutrient Intake and Energy Estimates. *JMIR MHealth UHealth* **2019**, *7*, e9232. [CrossRef] [PubMed]
25. Tate, D.F.; Jackvony, E.H.; Wing, R.R. A Randomized Trial Comparing Human E-Mail Counseling, Computer-Automated Tailored Counseling, and No Counseling in an Internet Weight Loss Program. *Arch. Intern. Med.* **2006**, *166*, 1620–1625. [CrossRef] [PubMed]
26. Painter, S.L.; Lu, W.; Schneider, J.; James, R.; Shah, B. Drivers of Weight Loss in a CDC-Recognized Digital Diabetes Prevention Program. *BMJ Open Diabetes Res. Care* **2020**, *8*, e001132. [CrossRef]
27. Comulada, W.S.; Swendeman, D.; Koussa, M.K.; Mindry, D.; Medich, M.; Estrin, D.; Mercer, N.; Ramanathan, N. Adherence to Self-Monitoring Healthy Lifestyle Behaviours through Mobile Phone-Based Ecological Momentary Assessments and Photographic Food Records over 6 Months in Mostly Ethnic Minority Mothers. *Public Health Nutr.* **2018**, *21*, 679–688. [CrossRef] [PubMed]
28. Shay, L.E.; Seibert, D.; Watts, D.; Sbrocco, T.; Pagliara, C. Adherence and Weight Loss Outcomes Associated with Food-Exercise Diary Preference in a Military Weight Management Program. *Eat. Behav.* **2009**, *10*, 220–227. [CrossRef] [PubMed]

Publication 10

Continuous glucose monitoring for automatic real-time assessment of eating events and nutrition: A scoping review.

Julian Brummer [1], Christina Glasbrenner [1], Sieglinde Hechenbichler Figueroa [1], Karsten Koehler [1], **Christoph Höchsmann** [1]

[1] TUM School of Medicine and Health, Technical University of Munich, Munich, Germany

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
Frontiers in Nutrition, 10: 1308348	2024	10.3389/fnut.2023.1308348	5.0 (2022)	Q1	Scopus: 0 Google Scholar: 1

Author contribution:

Last and corresponding author; developed the research question, co-conducted literature search and screening, co-drafted the manuscript, and was primarily responsible for project supervision and content.

frontiers | Frontiers in Nutrition

TYPE Systematic Review
PUBLISHED 08 January 2024
DOI 10.3389/fnut.2023.1308348

OPEN ACCESS

EDITED BY
Arpita Mukhopadhyay,
St. John's Research Institute, India

REVIEWED BY
Andrea Tumminia,
Università di Catania, Italy
Ali Cinar,
Illinois Institute of Technology, United States

*CORRESPONDENCE
Christoph Höchsmann
✉ christoph.hoechsmann@tum.de

RECEIVED 10 October 2023
ACCEPTED 13 December 2023
PUBLISHED 08 January 2024

CITATION
Brummer J, Glasbrenner C, Hechenbichler Figueroa S, Koehler K and Höchsmann C (2024) Continuous glucose monitoring for automatic real-time assessment of eating events and nutrition: a scoping review.
Front. Nutr. 10:1308348.
doi: 10.3389/fnut.2023.1308348

Continuous glucose monitoring for automatic real-time assessment of eating events and nutrition: a scoping review

Julian Brummer, Christina Glasbrenner, Sieglinde Hechenbichler Figueroa, Karsten Koehler and Christoph Höchsmann*

Department of Health and Sport Sciences, TUM School of Medicine and Health, Technical University of Munich, Munich, Germany

Background: Accurate dietary assessment remains a challenge, particularly in free-living settings. Continuous glucose monitoring (CGM) shows promise in optimizing the assessment and monitoring of ingestive activity (IA, i.e., consumption of calorie-containing foods/beverages), and it might enable administering dietary Just-In-Time Adaptive Interventions (JITAIs).

Objective: In a scoping review, we aimed to answer the following questions: (1) Which CGM approaches to automatically detect IA in (near-)real-time have been investigated? (2) How accurate are these approaches? (3) Can they be used in the context of JITAIs?

Methods: We systematically searched four databases until October 2023 and included publications in English or German that used CGM-based approaches for human (all ages) IA detection. Eligible publications included a ground-truth method as a comparator. We synthesized the evidence qualitatively and critically appraised publication quality.

Results: Of 1,561 potentially relevant publications identified, 19 publications (17 studies, total $N = 311$; for 2 studies, 2 publications each were relevant) were included. Most publications included individuals with diabetes, often using meal announcements and/or insulin boluses accompanying meals. Inpatient and free-living settings were used. CGM-only approaches and CGM combined with additional inputs were deployed. A broad range of algorithms was tested. Performance varied among the reviewed methods, ranging from unsatisfactory to excellent (e.g., 21% vs. 100% sensitivity). Detection times ranged from 9.0 to 45.0 min.

Conclusion: Several CGM-based approaches are promising for automatically detecting IA. However, response times need to be faster to enable JITAIs aimed at impacting acute IA. Methodological issues and overall heterogeneity among articles prevent recommending one single approach; specific cases will dictate the most suitable approach.

KEYWORDS

meal detection, continuous glucose monitoring, dietary assessment, healthcare technology, closed loop, sensors, meal timing

1 Introduction

Nutrition has a major impact on people's health and well-being (1–8). However, accurately assessing nutrition and dietary intake remains challenging, with the most precise tools often involving high costs, participant and staff burden, or privacy issues (9–14). Yet, valid and reliable measurement of dietary behavior is essential to accurately detect changes in research settings and guide patient counseling in clinical practice (e.g., weight loss programs). Technological advances in recent years have led to new approaches for accurately assessing dietary intake that try to overcome some of the shortcomings of traditional dietary assessment methods (9, 15–18).

An attractive technology-based option for assessing the consumption of calorie-containing foods and beverages (ingestive activity, IA) is continuous glucose monitoring (CGM). CGM involves using a sensor that measures glucose concentrations in the interstitial fluid (19, 20) as a proxy for blood glucose levels (20, 21). CGM has become an important tool in diabetes care (19, 22–25). For instance, it is an integral component of artificial pancreas (AP) systems designed to automate and improve blood glucose regulation in individuals with type 1 diabetes mellitus (T1DM) via the utilization of CGM, an insulin infusion pump, and a control algorithm (26). Beyond diabetes management, CGM is gaining popularity for use in healthy individuals and athletes (20, 27). Several CGM devices show satisfactory accuracy data (20, 28, 29).

The automatic and (near-)real-time detection of IA via CGM could offer benefits in (clinical) practice, including a reduced participant and staff burden. In addition, interventionists could monitor meal plan adherence more closely and detect deviations from intervention goals as they occur. Consequently, targeted and personalized countermeasures could be deployed proactively. One particularly useful approach would be CGM-based detection of IA in the context of Just-In-Time Adaptive Interventions (JITAIs). JITAIs aim to exploit the full potential of remote monitoring combined with delivering intervention content in the moment/context when it is most needed and the patient is likely to be (most) receptive (30). Preliminary results show promising effects of JITAIs on predicting and preventing dietary lapses (31). If detection times of the CGM-based approaches in question were extremely short (e.g., less than a few minutes), JITAIs could aim at acutely impacting IA (e.g., sending a prompt asking a person to terminate a meal). If detection times were relatively short (e.g., less than an hour), JITAIs could aim at altering subsequent IA (e.g., a dinner meal). In both cases, information on IA would be much more readily available than with traditional dietary assessment methods (e.g., 24-h recalls).

However, there may also be challenges associated with the use of CGM for the automatic monitoring of IA. On the one hand, there are system-inherent challenges. For example, postprandial rises in blood glucose vary in timing and extent depending on meal composition, meal quantity, inter-individual variability, and many other factors (32–40). Further, there is a delay between interstitial fluid and blood glucose concentrations (20, 41, 42). On the other hand, blood glucose levels are not only influenced by IA but also by other factors such as physical activity, stress, and diurnal fluctuations (20, 41, 43–55). Thus, false positive detections (e.g., erroneously flagging a meal due to glucose increases caused by stress) and false negative detections (e.g., erroneously *not* flagging a meal because other factors render the glucose response too flat) might occur.

In the past years, research examining the use of CGM for the automatic detection of IA has accumulated. Recent publications reviewed options for the automatic detection of IA using wearable–/sensor-based methods (15–18), but they did not specifically address CGM. The present article aims to close this research gap and answer the following guiding questions:

1 Which approaches using CGM for the automatic detection of IA in (near-)real-time have been investigated, and have these approaches relied solely on CGM or also used other data (e.g., sensors/wearables)?
2 How accurate are these approaches in detecting IA?
3 Can these approaches be used in the context of JITAIs?

2 Methods

The reporting of this review is based on the updated Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) guideline (56).

2.1 Search strategy

The primary systematic literature search was conducted on 09 September 2022, using the IEEE Xplore, PubMed, Scopus, and Web of Science databases. An identical supplementary search was conducted on 02 October 2023. The search term was developed and refined by two authors (JB, CH) to capture all relevant publications, and the search term contained: (intake OR uptake OR eating OR ingest* OR meal OR drink* OR beverage OR consum* OR oral) AND (monitor* OR assess* OR detect* OR estimat* OR measur* OR sens*) AND ("continuous glucose monitoring" OR "real time continuous glucose monitoring" OR "real-time continuous glucose monitoring" OR "flash glucose monitoring" OR "intermittently scanned continuous glucose monitoring" OR CGM OR rtCGM OR isCGM OR "artificial pancreas" OR "artificial beta cell*" OR "artificial beta-cell*" OR "artificial β-cell*" OR "artificial β cell*") AND (algorithm OR "deep learning" OR "machine learning" OR "neural network*" OR AI OR "artificial intelligence"). Because a fully-closed-loop AP system must first detect meals to adequately manage the following increases in glucose by delivering insulin to the patient (57), the search also included AP systems.

JB conducted the database searches and removed duplicates for the primary and supplementary searches. Two authors (JB, CH) screened the titles and abstracts against the predefined eligibility. In discrepancies, a consensus was reached via discussions and ineligible publications were discarded. For the primary search, JB screened the full texts of the remaining publications for eligibility and consulted with CH, who then independently screened these full texts in cases of uncertainty. In addition, one other author (SHF) conducted independent cross-checks for a randomly selected 20% of the full texts. For the supplementary search, two authors (JB, CH) independently screened the full texts of the remaining publications. Again, in case of discrepancies, a consensus was reached via discussions, and subsequently, ineligible publications were discarded. Another author (CG) was consulted for her technical/mathematical

expertise during the screening process. JB hand-searched the reference lists of eligible publications for any additional relevant literature. In several cases, the corresponding authors of articles were contacted, e.g., to receive full texts or raw data or to clarify results.

2.2 Eligibility criteria

We included publications if the following inclusion criteria were met: (1) publications were written in English or German and published until 2 October 2023; (2) publications are original articles published in peer-reviewed scientific journals or conference papers; (3) at least one performance measure of the automatic detection of IA is reported explicitly. For example, the accuracy was calculated by comparing the CGM-*based* (did not have to exclusively rely on CGM as input) approach against a ground-truth method (e.g., self-reported or observed IA); (4) a CGM-based approach was used to detect IA *in vivo* in free-living, semi free-living, or laboratory settings. This also included trials of AP systems if criterion 3 was met; (5) only the most recent publication on a specific approach by a particular research group was included if it supersedes preceding publications.

We excluded publications for the following exclusion criteria: (1) the approach was not tested in human participants (e.g., *in silico* studies); (2) no outcome results were reported (e.g., study protocol publications); (3) outcomes did not include an explicit performance measure describing the results of the automatic detection of IA (e.g., only figures showcasing the CGM trends over time); (4) the methodology was described without sufficient detail. We did not apply restrictions regarding publication date or participant age.

2.3 Data extraction

The following information was extracted: (1) first author and publication year; (2) a summary of the study; (3) sample size and, if available, sex and age of participants; (4) participants' diabetes status [no diabetes, prediabetes, T1DM or type 2 diabetes mellitus (T2DM)]; (5) scope of the study (duration/number of IA events) and, if available, information on the IA events (e.g., meal composition); (6) ground-truth/criterion method(s); (7) performance measure(s); (8) details on the CGM device and if applicable other relevant devices used in the study.

JB extracted relevant information from the original publications, and in cases of uncertainty, the respective publications were double-checked by CH. Two other authors (CG, SHF) also double-checked the extracted information. One other author (CG) further extracted technical details of the tested approaches.

2.4 Data synthesis

We synthesized the evidence qualitatively, focusing on answering the three research questions outlined above. Although using an explicit cutoff (e.g., ≥80% F1-score or accuracy) is desirable for performance evaluation and has been used in a related review (16), this approach was not feasible, as only a few publications reported accuracy and/or F1-score values.

Furthermore, we appraised the included publications critically. We considered the following aspects of being of concern: (1) error-prone methods for identifying the ground truth of IA [e.g., self-reported IA (58) or retrospective identification from CGM data, whereas inpatient settings with observed IA were generally assumed to be less error-prone]; (2) a sample consisting exclusively of individuals with diabetes as this might limit generalizability to non-diabetic populations; (3) meal announcement/meal-accompanying insulin boluses, as there might be an interference with the (early) postprandial blood glucose levels that are relevant for the automatic detection of IA; (4) algorithm inputs other than CGM since ultimately a CGM-only approach would be desirable to minimize costs and effort.

3 Results

The literature search identified a total of 1,561 potentially relevant publications. Nineteen publications reporting data from 17 studies (for 2 studies, 2 publications each were relevant, see Table 1), including 311 participants, met the inclusion criteria (59–66, 68, 70, 71, 73–75, 82–85, 87). Figure 1 shows the process of the literature search, screening, and selection in a PRISMA-style flow diagram (88).

Many of the screened publications were excluded from the present review because they did not include details on the detection of IA but instead focused on measures of glycemic control (e.g., time in specific glucose ranges). Further, some publications were excluded because they included graphic CGM data with IA marked as such but did not provide quantitative data on the detection of IA. Another common reason for exclusion was investigation *in silico*, often using virtual patients with T1DM.

We were unable to retrieve the full text of one publication despite several efforts to contact the authors directly. This publication was excluded; however, it was considered likely ineligible based on its abstract.

3.1 Study characteristics

The included publications were published between 2008 and 2023 and reported an average sample size of 18.3 ($SD = 15.1$) participants. Table 1 provides an overview of the included publications and the extracted information. The publications covered a wide age range, including pediatric (62, 73), adolescent (62, 64, 65), and adult (59–61, 63, 66, 68, 70, 71, 74, 75, 82–85, 87) populations. Fourteen publications included participants with T1DM (59, 62–66, 68, 70, 71, 73, 74, 84, 85, 87), one publication included a sample of participants with T1DM or T2DM (75), one publication included participants with prediabetes or moderately controlled T2DM (83) and three publications included participants without diabetes (60, 61, 82). Publications included both controlled/inpatient (59, 62–65, 68, 74, 82, 85) and free-living settings (60, 61, 66, 70, 71, 73, 75, 83, 87).

In several aspects, there was substantial heterogeneity among the included publications. First, the number and type of performance metrics reported for the tested approaches differed substantially. Commonly reported performance metrics included the number of true and/or false positives and/or negatives (including frequencies per day and rates) (64, 66, 68, 70, 73, 75, 84, 85, 87), sensitivity (60, 61, 70,

TABLE 1 Characteristics of included publications.

Author, year	Study summary	*N* (male/female), age (if available)	Diabetes status	Scope of the study/information on the assessed eating events	Ground-truth method(s)	Performance measure(s)[a]	Details on the CGM device and other input devices
Atlas et al. (59)	Pilot feasibility clinical trial investigating the performance of the MD-Logic Artificial Pancreas (MDLAP) System, a fully CL system utilizing a patient's diabetes treatment management in conjunction with a fuzzy logic-based control-to-range module and a control-to-target module to control blood glucose levels.	7 (2/5); 23.9 ± 3.4 years, range 19–30 years	T1DM	8 h CL sessions in a resting state, with 9 fasting sessions ($n = 6$) and 3 meal challenge sessions ($n = 2$; mixed meal with 40–60 g of CHO after 8 h fast) Two participants completed 1 additional 24 h CL session each, in which mixed meals (CHO content for each meal was 17.5–70 g) were consumed at 1,930 h, 0800 h, and 1,300 h, with participants entering the sessions after ≥3 h of fasting.	Inpatient study	Overall mean detection time: 23 min after consumption	CGM: Freestyle Navigator (Abbott Diabetes Care, Alameda, CA) or STS-Seven System (DexCom, San Diego, CA); sampling rate: 5 min
Bertrand et al. (60)[b]	Comparison of eating activity detection systems using (1) wearable wristbands vs. (2) wearable wristbands + CGM. Three machine learning algorithms were applied for the classification of eating and non-eating events: support vector machine (SVM), random forest (RF), and extreme gradient boosting tree (XGB). For each algorithm, one model based on the CGM data was compared to a model without CGM data.	10 (5/5); range 19–51 years	Healthy, non-diabetic	Up to 2 weeks	An app ("aTimeLogger") was used to log the ground truth	Mean (standard deviation)[c]: MCC_{XGB}: 0.35 (0.10) – MCC_{SVM}: 0.37 (0.06) $F1\text{-}score_{XGB}$: 0.49 (0.10) – $F1\text{-}score_{SVM}$: 0.50 (0.08) $Sensitivity_{XGB}$: 0.67 (0.19) – $Sensitivity_{SVM}$: 0.63 (0.20) $Specificity_{XGB}$: 0.74 (0.07) – $Specificity_{SVM}$: 0.77 (0.14) $Precision_{XGB}$: 0.41 (0.10) – $Precision_{SVM}$: 0.46 (0.11)	CGM: FreeStyle Libre 2 Wearable wristbands measuring steps and heart rate: Fitbit Charge 3 (dominant hand) and Mi Band 4 (non-dominant hand)
Bertrand et al. (61)[b]	This publication used the same dataset as Bertrand et al. (60). Two tree-based ensemble learning algorithms were used: random forest (RF) and extreme gradient boosting tree (XGB). Compared to Bertrand et al. (60) different resampling techniques were investigated for their performance in detecting eating activities vs. non-eating activities: no resampling (-N), random up-sampling (-U), random down-sampling (-D), and SMOTE resampling (-S).The combination of the two machine learning algorithms and the different resampling techniques resulted in eight classification models that were compared.	10 (5/5), average age: 32 years	Healthy, non-diabetic	Up to 14 days; in total, 1,361 activity events were collected	Free-living environment ➔ participants used an app ("aTimeLogger") to log the ground truth of their activities	Mean (standard deviation)[d]: $MCC_{XGB\text{-}N}$: 0.34 (0.13) – $MCC_{XGB\text{-}U}$: 0.38 (0.12); $F1\text{-}score_{XGB\text{-}N}$: 0.33 (0.16) – $F1\text{-}score_{RF\text{-}S}$: 0.51 (0.10); $Sensitivity_{XGB\text{-}N}$: 0.23 (0.14) – $Sensitivity_{XGB\text{-}D}$: 0.67 (0.19); $Specificity_{XGB\text{-}D}$: 0.74 (0.07) – $Specificity_{XGB\text{-}N}$: 0.98 (0.03); $Precision_{XGB\text{-}D}$: 0.41 (0.10) – $Precision_{XGB\text{-}N}$: 0.80 (0.21); $Accuracy_{XGB\text{-}D}$: 0.73 (0.03) – $Accuracy_{XGB\text{-}N}$: 0.82 (0.05). RF-S was deemed to have the best overall performance of the eight models.	CGM: a FreeStyle Libre 2 Wearables: Mi Band 4 (non-dominant wrist), FitbitCharge 3 (dominant wrist)

(Continued)

10.3389/fnut.2023.1308348

TABLE 1 (Continued)

Author, year	Study summary	*N* (male/female), age (if available)	Diabetes status	Scope of the study/information on the assessed eating events	Ground-truth method(s)	Performance measure(s)[a]	Details on the CGM device and other input devices
Dassau et al. (62)	Using data from a pilot study, four different detection methods were compared: backward difference (BD), Kalman filter estimation (Kalman), combination of BD and Kalman (BD + Kalman), and second derivative of glucose (G"). Central aim was to reduce FP detections to reduce the risk for erroneous insulin injections in the context of an AP. To do so, a voting algorithm was implemented, using either a two-out-of-three (BD, BD + Kalman, and G") or three-out-of-four (BD, Kalman, B + Kalman, and G") scheme to check for concordance in meal detections by the above-mentioned methods. Importantly, insulin meal boluses were purposefully delayed by 1 h and thus did not confound the postprandial BG changes.	17 (40% girls); 11 ± 4 years, range 4–17 years	T1DM	17 breakfast meals (one per participant) with an average of 56 g of CHO (range: 22–105 g of CHO), the content of which was decided upon by the participants	Inpatient study	ΔT (min) = Detection time from the onset of the meal; ΔG (mg/dL) = Difference in the glucose level when detection took place minus the preprandial value.	CGM: FreeStyle Navigator (Abbott Diabetes Care, Alameda, CA) with a sampling rate of 1 min
						Average values for the four different detection methods: $\Delta T_{BD} = 29$ min, $\Delta G_{BD} = 13$ mg/dL; $\Delta T_{Kalman} = 35$ min, $\Delta G_{Kalman} = 30$ mg/dL; $\Delta T_{BD+Kalman} = 31$ min, $\Delta G_{BD+Kalman} = 18$ mg/dL; $\Delta T_{G"} = 30$ min, $\Delta G_{G"} = 16$ mg/dL.	
						ΔG using the Kalman algorithm was statistically significantly higher compared with ΔG using the other methods ($p < 0.001$).	
						Average values for the different voting schemes: $\Delta T_{two\text{-}out\text{-}of\text{-}three} = 30$ min, $\Delta G_{two\text{-}out\text{-}of\text{-}three} = 15 \pm 10$ mg/dL; $\Delta T_{three\text{-}out\text{-}of\text{-}four} = 32$ min, $\Delta G_{three\text{-}out\text{-}of\text{-}four} = 21 \pm 9$ mg/dL.	

(Continued)

10.3389/fnut.2023.1308348

TABLE 1 (Continued)

Author, year	Study summary	*N* (male/female), age (if available)	Diabetes status	Scope of the study/information on the assessed eating events	Ground-truth method(s)	Performance measure(s)[a]	Details on the CGM device and other input devices
Dovc et al. (63)	Double-blind, randomized, two-period crossover study on the safety and efficacy of fully CL insulin therapy/glucose control using two different insulin solutions (faster vs. standard insulin aspart). Atlas et al.'s (59) fuzzy logic-based control algorithm was used (see above).	20 (9/11); 21.3 ± 2.3 years	T1DM	Two 27-h (1,500 h on the first day to 1,800 h the next day) CL inpatient stays with meals that were unannounced to and thus uncovered by the fully CL device. Standardized and identical meals were given on both study visits. Main meals contained ~1 g of CHO/kg of body mass and snacks about half of this amount. Macronutrient distribution was about 50% CHO, 20% proteins, and 30% fats (<10% saturated fats). Meals were scheduled at: ~1,500 h (snack); 1 h after the end of an exercise protocol between 1,900 and 2,000 h (dinner); 0800 h (breakfast); 1,200 h on the next day (lunch).	Inpatient study	Median time of delivered prandial bolus was 38.4 min (32.7, 55.8) for meals in the faster insulin aspart arm and 30.1 min (26.9, 54.6) in the standard insulin aspart arm ($p = 0.388$).	CGM: Enlite II sensor (Medtronic Diabetes); CL algorithm: DreaMed GlucoSitter (DreaMed Diabetes, Petah Tikva, Israel)
El Fathi et al. (64)[e]	Preliminary results from a randomized three-way experiment on the safety and efficacy of CL insulin delivery with or without a meal detection module (an adaptive model-based meal detection algorithm) versus conventional pump therapy after a missed insulin bolus. The data stem from the same study as Palisaitis et al. (65).	4 adolescents	T1DM	Per participant, three 9-h inpatient visits with one uncovered lunch meal with 60 g of CHO per visit were conducted ➔ 4 participants x 3 visits x 9 h = 108 h of data	Inpatient study	Comparison of the incremental AUCs after the missed insulin bolus across the three conditions: conventional pump therapy = reference standard (29.6 ± 6.5 h mmol/l), CL without meal detection: −16% incremental AUC (24.8 ± 11.5 h mmol/l), CL with meal detection: −39% incremental AUC (18.0 ± 2.7 h mmol/l); The collected data were also used to run the meal detection algorithm offline over the 108 h (4 patients × 3 visits × 9 h) of clinical data: 12/12 unannounced meals detected successfully; no FPs; time until meal detection = 35 [30–40] min; glucose increase at meal detection time = 2.89 ± 1.72 mmoL/L	NIA

(Continued)

10.3389/fnut.2023.1308348

TABLE 1 (Continued)

Author, year	Study summary	*N* (male/female), age (if available)	Diabetes status	Scope of the study/information on the assessed eating events	Ground-truth method(s)	Performance measure(s)[a]	Details on the CGM device and other input devices
Faccioli et al. (66) [supplemented with information from Anderson et al. (67)]	Data from a multicenter clinical trial on the feasibility of a long-term automated insulin delivery system were used for a retrospective evaluation of a super-twisting-based meal detector. 14 days of SAP therapy under free-living conditions preceded the main phase of the study; these data were used for the evaluation of the meal detector. Due to the use of SAP therapy, manual meal insulin boluses were given, and the results need to be interpreted taking this into consideration.	30 (17/13); median age 44 years, range 18–66 years	T1DM	14 days	Patient-reported meal times. Of note, 11/30 participants had <20 registered meals for 14 days. Since some missed meal announcements occurred, the authors only selected portions of data with certain meal information.	All values refer to median (interquartile range): TP = 16 (10); FN = 6 (4); FP = 7 (3); Recall = 70% (13%); Precision = 73% (26%); F1-score = 68% (16%); FP per day = 1.4 (1.4); CHO content related to FNs = 32 g (32 g); detection time = 45 min (45 min)	CGM: DexCom G4 (DexCom, Inc., San Diego, CA, United States); sampling time of 5 min
Fushimi et al. (68) [supplemented with information from Sánchez-Peña et al. (69)]	CGM data obtained during a clinical trial were used to evaluate the Automatic Regulation of Glucose (ARG) algorithm with an additional automatic switching signal generator (SSG), i.e., a meal detection module. Importantly, in the clinical trial meal announcement was used, so the results need to be interpreted in light of this potential bias.	5 (2/3), 43 ± 6 years, range 32–48 years	T1DM	Five standardized meals per participant: one breakfast, one lunch, one afternoon snack, two dinners. Breakfast and afternoon snack: a cup of tea or coffee, two slices of whole-meal bread or five crackers, diet jam, spreadable cheese (≈28 g of CHO). Dinners: whole pasta, lean meat, fresh fruit (≈55 g of CHO). Lunch: same as dinners, but mashed potatoes instead of whole pasta (≈55 g CHO). One meal was excluded due to pump occlusion ➔ 24 eligible meals	Inpatient study	2 FPs (8.3%); 2 FNs (8.3%); efficiency = 83.3%	CGM: Dexcom G4 sensor (Dexcom, San Diego, CA), sampling rate: 5 min
Godoy et al. (70)	A feedback scheme-based meal detection and CHO estimation algorithm was developed and evaluated retrospectively on a clinical dataset. Of note, insulin boluses were used as another algorithm input. The results need to be interpreted considering this.	11 adults	T1DM	5 days, whereby the first 3 days were used for identification/calibration and the following 2 days were used for the validation of the proposed model	Free-living data with CGM measurements, insulin pump recordings, participant-recorded CHO estimates, etc.	184 TPs; 263 TNs; 9 FPs; 2 FNs; 98.92% sensitivity; 96.69% specificity; 97.60% accuracy[f]; Mean time gap (estimated meal onset time – actual meal onset time) = 9.0 min and 25 min delay time	Insulin pump MiniMed 640G; CGM sampling time = 5 min, but up-sampled to 1 min to increase the detection sensitivity

(Continued)

10.3389/fnut.2023.1308348

TABLE 1 (Continued)

Author, year	Study summary	*N* (male/female), age (if available)	Diabetes status	Scope of the study/information on the assessed eating events	Ground-truth method(s)	Performance measure(s)[a]	Details on the CGM device and other input devices
Hoyos et al. (71) [supplemented with information from Aleppo et al. (72)]	Data from a study assessing the reliability of CGM measurements were used to compare two scenarios: one scenario with the original meal events announced by the participants and one with the meal events generated automatically by the super-twisting-based meal detector introduced in Faccioli et al. (66). "An unsupervised clustering algorithm based on Fuzzy C-Means was applied to classify event-to-event segments of CGM data. Events defining data partitioning were automatically generated based on: (1) an automatic meal detection algorithm (for day periods) and (2) time of day (for night periods)." (p. 576)	44 adults	T1DM	26-week study; only participants with an average of 3 to 5 reported meals per day were considered	Free-living data with CGM measurements, insulin pump recordings, etc.	Results (M ± SD) for automatically detected meals: Number of clusters (c*) = 8.09 ± 1.67; Fukuyama-Sugeno index (V_{FS}) = −16,893 ± 5,838; Compactness (V_{com}) = 0.236 ± 0.063; Variance (V_{var}) = 966 ± 653.4; Time in range = 45.2 ± 15% There was lower variance in the clusters of the automatically detected meals as compared to the announced meals, thereby underscoring the algorithm's value.	CGM: Dexcom G4 Platinum; Sampling time = 5 min
Kölle et al. (73)	Four different detection methods were retrospectively compared using a clinical dataset: two methods based on the classification of horizons (classification of estimated R_α horizons [LDA R_α] and classification of CGM horizons [LDA CGM], respectively) and two methods based on threshold violations (threshold on current R_α estimate [Threshold] and the previously published Glucose Rate Increase Detector algorithm (79), respectively). Note that often meals were accompanied by insulin boluses, so, again, results need to be interpreted in light of this.	12 (8/4); 7.3 ± 4.7 years	T1DM	492 of 849 identified meals were included, whereby the authors focused on meals that would necessitate automatic meal detection (e.g., larger meals); meals were divided into categories of pre-meal, at-meal, post-meal and no insulin bolus	Two experienced diabetologists independently retrospectively identified meals from free-living CGM data and logged information from insulin infusion pumps.	Averages across 10 cross-validated Monte Carlo runs: LDA R_α: sensitivity = 0.92; 1.50 FPs/day; time of detection after meal start = 18.59 min LDA CGM: sensitivity = 0.90; 1.37 FPs/day; time of detection after meal start = 11.78 min Threshold: sensitivity = 0.64; 1.28 FPs/day; time of detection after meal start = 32.67 min GRID: sensitivity = 0.21; 2.81 FPs/day; time of detection after meal start = 42.53 min	CGM: Medtronic Enlite 2

(Continued)

10.3389/fnut.2023.1308348

TABLE 1 (Continued)

Author, year	Study summary	*N* (male/female), age (if available)	Diabetes status	Scope of the study/information on the assessed eating events	Ground-truth method(s)	Performance measure(s)[a]	Details on the CGM device and other input devices
Mosquera-Lopez et al. (74)	A single-center, randomized crossover trial was conducted to compare postprandial (4 h) glucose control following unannounced meals using a hybrid model predictive control algorithm and a newly developed robust artificial pancreas (RAP) system (i.e., two intervention visits per participant). The RAP system used machine learning for automated meal detection and meal size estimation. CGM and insulin data were used.	15 (6/9) participants enrolled (age: 37.6 ± 10.4 years), 2 participants withdrew from the study ➔ 13 included in analysis	T1DM	Only the intervention visit involving the RAP system was used to determine meal detection performance; this visit involved a total of 24 participant-chosen study meals with a CHO content of 45-66 g.	Participant confirmation of alerts sent out by the meal detection system; data on mealtime records entered into a cloud-based database by a study investigator.	Sensitivity = 83.3% (95% CI 62.6–95.2%); false discovery rate = 16.6% (95% CI 4.7–37.4%); detection time (M ± SD) = 25.9 ± 0.9 min	CGM: Dexcom G6 (DexCom, Inc., San Diego, CA, United States); Insulin pump: Omnipod (Omnipod Insulet Corporation, Acton, MA, United States)
Ornetzeder et al., 2019 (75) [supplemented with information from Zschornack et al. (76)]	Three previously published meal detection algorithms (79–81) were compared by using data from two clinical trials, one with participants with T1DM and one with insulin-treated participants with T2DM. Importantly, in both datasets insulin meal boluses were given, so the results need to be interpreted in light of this. Furthermore, small meals (<20 g of CHO) were treated differently from larger meals in that they did not contribute to FN and FP counts.	10 ➔ 5 (3/2) participants with T1DM (mean age: 48 years) and 5 (4/1) participants with T2DM (mean age: 65 years); both samples were random subsamples of the respective study samples	T1DM, T2DM	T1DM participants: datasets with 7 days per participant but the first 48 h after the insertion of the CGM sensor were not considered for the performance assessment; T2DM participants: datasets with 2–3 days per participant	T1DM: no further information	T1DM (fixed parameters; all means): $\Delta T_{Harvey} = 19.1$ min; $\Delta T_{Samadi} = 12.7$ min; $\Delta T_{Ramkissoon} = 19.6$ min; FP/day$_{Harvey} = 0.6$; FP/day$_{Samadi} = 0.9$; FP/day$_{Ramkissoon} = 0.4$; Sensitivity$_{Harvey} = 77.0\%$; Sensitivity$_{Samadi} = 73.7\%$; Sensitivity$_{Ramkissoon} = 75.0\%$	T1DM: prototype CGM system using a sensor in an early development phase (Roche Diagnostics GmbH, Mannheim, Germany)
					T2DM: outpatient study	T2DM (fixed parameters; all means): $\Delta T_{Harvey} = 27.8$ min; $\Delta T_{Samadi} = 24.6$ min; $\Delta T_{Ramkissoon} = 26.9$ min; FP/day$_{Harvey} = 1.3$; FP/day$_{Samadi} = 1.3$; FP/day$_{Ramkissoon} = 0.9$; Sensitivity$_{Harvey} = 70.5\%$; Sensitivity$_{Samadi} = 67.9\%$; Sensitivity$_{Ramkissoon} = 70.7\%$	
						T2DM (patient-specific parameters; all means): $\Delta T_{Harvey} = 30.7$ min; $\Delta T_{Samadi} = 30.5$ min; $\Delta T_{Ramkissoon} = 26.1$ min; FP/day$_{Harvey} = 1.0$; FP/day$_{Samadi} = 1.1$; FP/day$_{Ramkissoon} = 0.4$; Sensitivity$_{Harvey} = 79.9\%$; Sensitivity$_{Samadi} = 80.2\%$; Sensitivity$_{Ramkissoon} = 67.3\%$	T2DM: CGM sampling rates: 1 min and 5 min for the T1DM and T2DM datasets, respectively
					For both trials CGM data and information on the ingested amount of CHO, meal timing, and insulin were recorded	ΔT was defined as the time between meal ingestion and the detection event	

(Continued)

10.3389/fnut.2023.1308348

TABLE 1 (Continued)

Author, year	Study summary	*N* (male/female), age (if available)	Diabetes status	Scope of the study/information on the assessed eating events	Ground-truth method(s)	Performance measure(s)[a]	Details on the CGM device and other input devices
Palacios et al. (82)	Data was collected in a cross-over study on post-resistance exercise nutrient timing comparing immediate vs. 3-h delayed post-exercise nutrition. Tree-based machine learning models (random forest model and gradient boosting machines) using a cold-start and a non-cold-start approach were applied, respectively. Six primary variables were used for the machine learning model: glucose, heart rate, physical activity, core temperature, skin temperature, and respiration rate. As physical activity was found to not aid in prediction it was removed and new lag variables were included.	9 (9/0); 24.3 ± 4.6 years	Healthy, non-diabetic	48 h; Total daily energy intake had an approximate macronutrient distribution of 52% CHO, 32% fat, 16% protein; meals were consumed at approximately 08:00, 11:20, 16:00, and 18:00 on both days.	48-h inpatient study in a whole-room calorimeter ➔ exact mealtimes were recorded	(1) area under the receiver operating characteristic curve (AUC-ROC); (2) area under the precision-recall curve (AUC-PR) ➔ cold-start: k = 110 min: AUC-ROC = 0.891; AUC-PR = 0.803; non-cold-start: k = 20 min, AUC-ROC = 0.996; AUC-PR = 0.964 k = minimal window size	CGM: iPRO Professional CGM System (Medtronic MiniMed, Inc., Northridge, CA) placed on the abdomen; Equivital Sensor Electronics Module (Equivital I: Hidalgo Ltd., Cambridge, United Kingdom) combined with a heat-sensitive transmitter (pill) to assess heart rate, heat flux, core body temperature; To match the calorimeter data, the Equivital and CGM data were resampled using cubic spline interpolation
Palisaitis et al. (65)[e]	Randomized, three-way, crossover trial comparing an AP system equipped with a meal detection algorithm (AP + MDA) with the AP alone and conventional pump therapy (CSII) in controlling blood glucose levels after a meal without accompanying insulin bolus. The data is from the same study as El Fathi et al. (64).	11; 14.9 ± 1.3 years	T1DM	One self-selected lunch meal which was standardized between interventions for each participant ➔ Mixed meals with 55-65 g of CHO served 4 h after the start of the intervention	Inpatient setting or at home with a member of the research staff with three 9-h interventions from 0800 to 1,700 or 0900 to 1,800	Median meal detection time in the AP + MDA condition: 40.0 min (interquartile range 40.0–57.5 min) after consumption of the meal; Incremental glucose from the start of the meal until time of meal detection: 2.6 mmol/L [2.4–4.8], and a rate of change of 10.1 [7.3–12.5] mmol/L/h	CGM: Dexcom G5

(Continued)

10.3389/fnut.2023.1308348

TABLE 1 (Continued)

Author, year	Study summary	*N* (male/female), age (if available)	Diabetes status	Scope of the study/information on the assessed eating events	Ground-truth method(s)	Performance measure(s)[a]	Details on the CGM device and other input devices
Popp et al. (83)	Self-reported (SR; using a smartphone app) timing of eating occasions (consumption of foods and beverages >0 kcal) was compared to objective assessment methods, i.e., a wrist-motion-based (WM) classifier using an ActiGraph worn on participants' dominant wrists and a simulation-based explanation system using CGM. The data come from an ancillary study of a weight loss intervention study.	31 completers; 62% females; age: 59 ± 11 years	Prediabetes/ moderately controlled T2DM	10 days	Free-living; Date- and time-stamped eating occasions were entered into a smartphone app; herein, we assume that these serve as the ground-truth method.	CGM method found the longest eating window and the largest number of eating occasions per day.	ActiGraph: ActiGraph GT9X-BT (Pensacola, FL, United States).
						Pearson's correlations: first eating occasion identified by SR and CGM: r = 0.534, $p = 0.01$; first eating occasion identified by CGM and WM: r = 0.325, $p = 0.004$; eating midpoint identified by CGM and WM: r = 0.253, $p = 0.03$.	
						Overlap between methods: Tolerance windows of ±0, 5, and 10 min: <40% of eating occasions identified by both WM and CGM; tolerance windows of ±30, 60, 120 min: overlap between SR and CGM: 55 to 80% of eating occasions; overlap between WM and CGM: ~23% regardless of the tolerance window used.	CGM: Abbott Freestyle Libre Pro (Abbott Park, IL, United States) providing 15-min average glucose values
						% of meals identified by all methods: Tolerance windows of ±0, 10 and 15 min: no matching meals identified by all methods; tolerance window of 30 min: 4% of SR meals were also detected by CGM and WM; tolerance window of both 60 and 120 min: 7% overlap; overlap of the three methods over 3 days was found in only one participant	

(Continued)

10.3389/fnut.2023.1308348

TABLE 1 (Continued)

Author, year	Study summary	*N* (male/female), age (if available)	Diabetes status	Scope of the study/information on the assessed eating events	Ground-truth method(s)	Performance measure(s)[a]	Details on the CGM device and other input devices
Samadi et al. (84)	Retrospective evaluation of meal detection and CHO estimation algorithms using clinical data collected in CL experiments using the integrated multivariable adaptive AP system (IMA-AP). Their approach relies on the qualitative analysis of the glucose trajectory and preceding insulin infusion data to detect disturbances and estimate their magnitude by estimating the amount of ingested CHO. Importantly, in these experiments, no meal announcement-based feedforward meal bolusing was used, so the data do not include manual meal-time insulin boluses.	11; 18–35 years	T1DM	117 meals/snacks (7–9 meals and a maximum of 6 snacks per participant) which are distinguished by a CHO threshold of 35 g	NIA	Detection rates (sensitivity): 93.5% (86/92) for meals and 68.0% (17/25) for snacks; FP rate 20.8% (27 FPs and 103 TPs); this equates to 1.05 FPs per day.	CGM: sampling time: 5 min
						Higher probability of detection with higher CHO contents.	
						For detected meals and snacks the increase in glucose from consumption until detection is on average 8.8 ± 21.3 mg/dL (in median ± mean absolute deviation [MAD] as 10.0 ± 14.4 mg/dL) for detected meals and snacks	Real-time data on biometric variables: BodyMedia SenseWear armband and Zephyr chest-band (Bioharness-3; Zephyr Technology, Annapolis, MD)
						Detection time (time from start of the meal to when the algorithm first reports a CHO estimate): 34.8 ± 22.8 min (in median ± MAD as 30.0 ± 16.0 min).	
Turksoy et al. (85) [supplemented with information from Turksoy et al. (86)]	Data obtained during AP trials without meal announcement were used to test a new meal detection approach that requires only CGM data. The meal detection algorithm was meant to be integrated into the integrated multivariable adaptive AP (IMA-AP).	9 (9/0); mean age 18.3 years	T1DM	32-h CL sessions were conducted with each participant including breakfast, lunch, dinner and a snack as well as additional snacks if requested by participants; 63 dietary events (50 main meals and 13 snacks) ➔ Foods were selected based on subjects' personal requirements and there was no limitation on food intake; M = 44 ± 9.38 g of CHO, whereby main meals were higher in CHO than the snacks.	Inpatient study	61/63 (96.8%) meals/snacks detected successfully; 2 FNs; 1 FP; For the events that were detected successfully the average change in glucose from the start of the meal until the time of the meal detection is 16 ± 9.42 mg/dL.	CGM: Guardian REAL-time CGM (Medtronics, Northridge, CA, United States); sampling time of 5 min, but interpolations used to generate 1-min sampled data

(Continued)

10.3389/fnut.2023.1308348

TABLE 1 (Continued)

Author, year	Study summary	*N* (male/female), age (if available)	Diabetes status	Scope of the study/information on the assessed eating events	Ground-truth method(s)	Performance measure(s)[a]	Details on the CGM device and other input devices
Weimer et al. (87)	Evaluation of a physiological parameter-invariant (PAIN)-based meal detector against three established meal-detection algorithms (62, 79) (89) on a clinical dataset. Importantly, participants used insulin therapy during the monitoring period.	61; 45.7 ± 15.3 years	T1DM	Average duration of monitoring: 17 days	Patient-reported mealtimes (time of inputting meal information into the insulin pump)	Detection rate (i.e., correctly detecting meals within 2 h of the patient-reported mealtime) of the detectors based on operating points that are closest to 2 FPs/day (i.e., relative sensitivity for a standardized specificity): PAIN = 86.9% at 2.01 FPs/day; Dassau et al. = 74.1% at 1.99 FPs/day; Lee & Bequette = 73.4% at 1.99 FPs/day; Harvey et al. = 79.4% at 1.97 FPs/day With detection rates ≥ 55% and ≤3.7 FPs/day the PAIN-based detector performs more reliably across individuals than the other approaches (based on $n = 53$ since participants with <10 reported meals were excluded from this analysis). The other approaches show lower detection rates (on average and worst-case). The other approaches have lower average FP rates but display greater variance and higher FP rates in the worst-case scenarios.	CGM: 5-min CGM readings

Publications included in the review, ordered by the first author's name. Some publications also include further information (e.g., in silico validation data, CHO content estimation), which are not discussed here. [a]Only performance measures relevant for our research questions are reported herein. [b]Bertrand et al. (60) and Bertrand et al. (61) are reports of the same study. [c]The authors of this paper report many different performance metrics for each algorithm; for space-saving reasons we only report the lowest and highest values for each performance metric here. These values refer to the models that included CGM data. [d]The authors of this paper report many different performance metrics for each algorithm; for space-saving reasons we only report the lowest and highest means (and the according standard deviations) for each performance metric here. [e]El Fathi et al. (64) and Palisaitis et al. (65) are reports of the same study. [f]We calculated sensitivity, specificity, and accuracy from the total number of TPs, TNs, FPs, and FNs, respectively, and thus, the values refer to the total sensitivity, specificity, and accuracy across the sample. AP, artificial pancreas; AUC, area under the curve; BG, blood glucose; CGM, continuous glucose monitoring; CHO, carbohydrates; CI, confidence interval; CL, closed-loop; CSII, continuous subcutaneous insulin infusion; FN(s), false negative(s); FP(s), false positive(s); LDA, linear discriminant analysis; M, mean; MCC, Matthew's correlation coefficient; NIA, no information were available; SAP, sensor-augmented pump; SD, standard deviation; T1DM, type 1 diabetes mellitus; T2DM, type 2 diabetes mellitus; TN(s), true negative(s); TP(s), true positive(s).

10.3389/fnut.2023.1308348

73–75, 84, 87), specificity (60, 61, 70), accuracy (61, 70), precision (60, 61, 66), F1-score (60, 61, 66), Matthew's correlation coefficient (60, 61), Pearson's correlations (83), detection time or time until an insulin bolus was delivered (59, 62–66, 70, 73–75, 84), change in glucose concentrations (62, 65, 84, 85), and area under the curve (82). However, even when the same metrics were reported, their definition was sometimes inconsistent across publications. For instance, the detection window for true positive detections ranged from 60 to 180 min, depending on the publication (66, 73, 75). In addition, the general study setup varied between publications, including differences in the sample composition, the use of meal announcement/meal-accompanying insulin boluses, the ground-truth method used for identifying IA, the devices used, and the scope of the data collection (Table 1).

Table 2 shows the result of the critical appraisal of all included publications. There were some concerns regarding the applied methodology for all publications; these concerns were substantial for most publications. In 16/19 (84.2%) publications, the sample consisted exclusively of individuals with (pre)diabetes (59, 62–66, 68, 70, 71, 73–75, 83–85, 87). Further, 9/19 (47.4%) publications used error-prone methods for measuring the ground truth of IA, mostly self-reported IA (60, 61, 66, 70, 71, 73, 75, 83, 87). Moreover, 7/19 (36.8%) publications used meal announcements and/or meal-accompanying insulin boluses (66, 68, 70, 71, 73, 75, 87). Finally, 8/19 (42.1%) publications utilized other inputs besides CGM, e.g., heart rate or the insulin sensitivity factor (59–61, 63, 64, 70, 74, 82). Overall, 15/19 (78.9%) publications elicited methodological concerns in two or more appraised domains (59–61, 63, 64, 66, 68, 70, 71, 73–75, 83, 84, 87), and all publications had methodological concerns in at least one of the appraised domains (59–66, 68, 70, 71, 73–75, 82–85, 87).

3.2 Overview of detection approaches

Our review identified a wide range of methods to automatically detect IA, including fuzzy logic (59, 63), model predictive control (74), support vector machine (60), random forest (60, 61, 82), (extreme) gradient boosting trees (60, 61, 82), backward difference (62), Kalman filter estimation (62), second derivative of glucose (62), Kalman filters (64, 68), switching signal generator (68), simulation-based explanation (83), classification of horizons (73), analysis of the glucose trajectory (84), pattern recognition using linear discriminant analysis (73), and threshold violation-based approaches (73). Further, adaptive model-based (64), super-twisting-based (66, 71), feedback scheme-based (70), and physiological parameter-invariant-based (87) meal detection approaches were applied.

The reviewed approaches used different inputs to automatically detect IA. As summarized in Table 2, some methods relied solely on CGM as an input (62, 65, 66, 68, 71, 73, 75, 83–85, 87). Others also included data from insulin treatment or other sensor systems (e.g., accelerometry, photoplethysmography, temperature sensors; see Table 1) (59–61, 63, 64, 70, 74, 82).

3.3 Performance of the approaches

We identified several CGM-based approaches for the automatic detection of IA that achieved high values in the respective performance metrics (Table 1). However, the substantial heterogeneity in the applied methodology and reporting of results needs to be considered.

For instance, Godoy and colleagues achieved 98.9% sensitivity, 96.7% specificity, and 97.6% accuracy with their feedback scheme-based algorithm (70). Notably, the algorithm uses certain patient-specific parameters, such as the insulin sensitivity factor derived from participants' usual diabetes treatment (70). Similarly, the algorithm by El Fathi et al. successfully detected 12/12 meals without any false positives and a detection time of 35 min (64). In two publications using the same dataset, Bertrand et al. investigated several IA detection approaches in individuals without diabetes (60, 61). A range of performance metrics is reported in both publications. In the first publication, the highest achieved mean sensitivity was 66.8%, and the highest achieved mean specificity was 77.3%, for example (60). In the second publication, the highest achieved mean sensitivity was 66.8%, and the highest achieved mean specificity was 97.7% (61). Importantly, in both publications, the models did not exclusively rely on CGM as input (60, 61). Similarly, Palacios et al. had a sample of individuals without diabetes (82). However, their models, too, did not exclusively rely on CGM as input, but also utilized other physiological variables such as heart rate and skin temperature (82). Palacios et al. reported the area under the receiver operating characteristic curve (AUC-ROC) and the area under the precision-recall curve (AUC-PR) (82). For cold-start cases with a window size of $k = 110$ min, they reported an AUC-ROC of 89.1% and an AUC-PR value of 80.3% (82). For non-cold-start cases and $k = 20$ min, the AUC-ROC was 99.6%, and the AUC-PR was 96.4% (82).

The performance of CGM-only approaches, which hold particularly great value for practical applications, varied substantially. Sensitivities varied between 20.8% (73) and 96.8% (85). Where reported, average false positive detections ranged from 0.4 (75) to 2.8 (73) IA events per day. Selected publications further reported a false positive rate of 20.8% (84), a false discovery rate of 16.6% (74), and a median precision value of 73.0% (66). Moreover, publications reported detection times between 11.8 min (mean) (73) and 45.0 min (median) (66). Importantly, all CGM-only approaches were tested on samples consisting exclusively of individuals with (pre)diabetes, some of which also used meal announcements/insulin boluses. A detailed description of the performance metrics for each of the included publications is provided in Table 1.

3.4 Detection times

The detection time is the relevant metric to evaluate whether the identified CGM-based IA detection approaches could be used in the context of JITAIs.

A detection time measure was reported in 11/19 (57.9%) publications (59, 62–66, 70, 73–75, 84). The detection time was commonly defined as the time between the start of the IA (i.e., typically a meal) and its (automatic) detection by the CGM-based approach. Mean (59, 62, 64, 70, 73–75, 84) and median detection times (65, 66, 84) were reported, thus impeding direct comparisons. One publication reported the median time of the delivered prandial insulin boluses (63).

Overall, the reported detection times varied between 9.0 min (mean) (70) and 45.0 min (median) (66), with most values falling into the 20-to-40-min range (Table 1).

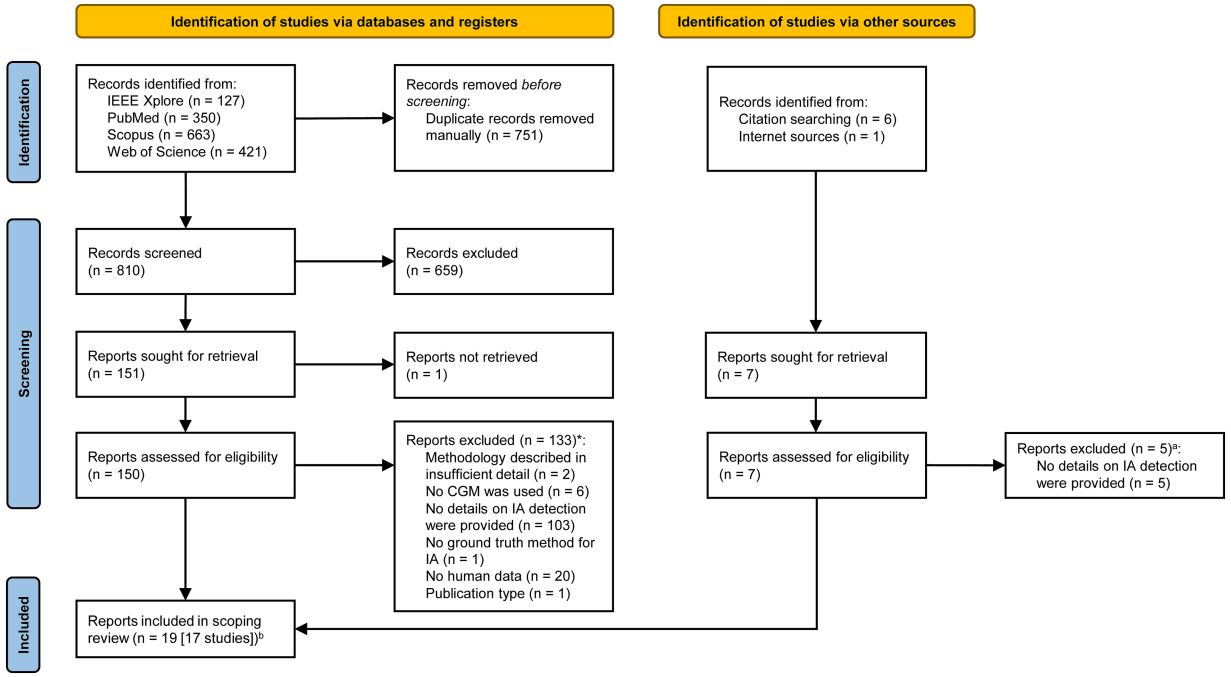


FIGURE 1
PRISMA Flow diagram (88). [a]In some cases, more than one reason led to the exclusion of a publication; here, only the primary reason is listed for each publication. [b]In two cases, two publications of the same study were included. CGM, continuous glucose monitoring; IA, ingestive activity.

4 Discussion

The primary objective of this review was to examine whether CGM can be used to automatically detect IA in (near-)real time. In sum, there are various promising approaches that show satisfactory to excellent performance on measures such as sensitivity and specificity. However, the performance of CGM-based methods for automatically detecting IA varies. Similarly, detection times vary, but currently, they appear too long to administer JITAIs for acutely altering IA. Methodological issues and overall heterogeneity among publications make it difficult to recommend the best-performing approach.

4.1 Which approaches using CGM for the automatic detection of IA in (near-) real-time have been investigated, and have these approaches relied solely on CGM or also used other data (e.g., sensors/ wearables)?

Our results indicate that both CGM-only approaches and those supplemented with other input data (e.g., accelerometry, photoplethysmography, temperature sensors) have been tested. Moreover, various algorithms have been used to detect IA. Since approaches using different sensor modalities and/or programming methods were successful at automatically detecting IA, it is evident that various solutions can be used for automated, CGM-based IA detection.

4.2 How accurate are these approaches in detecting IA?

Our review showed that the performance evaluation of any single approach depends on the respective case and priorities. For example, if the goal is to combine a CGM-based approach with smartphone prompts to enable comprehensive diet logs, the method should have high sensitivity to avoid missing a potential IA (false negative). In this case, specificity would only play a minor role as nothing is lost by sending a prompt in response to a false positive detection – the prompt can remain unanswered by the patient/participant. In contrast, when the goal is to use the CGM-based approach as a stand-alone IA assessment tool, high specificity would be critical to avoid artificial inflation of the number of daily meals, for example. Thus, a single best approach for all scenarios could not be identified. The substantial heterogeneity of the applied methods and reporting of results, including the broad range of the number and type of reported performance metrics and their varying definitions, made it difficult to compare the performance of the different approaches.

However, collectively, our results demonstrate that there are indeed several relatively well-performing CGM-based approaches for the automatic detection of IA. One example is the feedback scheme-based algorithm by Godoy et al., which achieved near-perfect sensitivity, specificity, and accuracy (70). However, this algorithm relies on several patient-specific parameters as input that are derived from participants' usual diabetes treatment (70). Thus, it remains to be determined whether this approach could be adapted to work equally well in individuals without diabetes, for whom these data are not routinely assessed. Similarly, methodological issues further limiting studies' internal and/or external validity pertain to using meal

TABLE 2 Critical appraisal of included publications.

Publication	Error-prone IA ground-truth method	Sample with (pre)diabetes	Meal announcement/ insulin boluses	Other inputs in addition to CGM	Overall rating
Atlas et al. (59)	x	✓	x	✓	★★☆☆
Bertrand et al. (60)	✓	x	x	✓	★★☆☆
Bertrand et al. (61)	✓	x	x	✓	★★☆☆
Dassau et al. (62)	x	✓	x	x	★★★☆
Dovc et al. (63)	x	✓	x	✓	★★☆☆
El Fathi et al. (64)	x	✓	x	✓	★★☆☆
Faccioli et al. (66)	✓	✓	✓	x	★☆☆☆
Fushimi et al. (68)	x	✓	✓	x	★★☆☆
Godoy et al. (70)	✓	✓	✓	✓	☆☆☆☆
Hoyos et al. (71)	✓	✓	✓	x	★☆☆☆
Kölle et al. (73)	✓	✓	✓	x	★☆☆☆
Mosquera-Lopez et al. (74)	x	✓	x	✓	★★☆☆
Ornetzeder et al. (75)	✓	✓	✓	x	★☆☆☆
Palacios et al. (82)	x	x	x	✓	★★★☆
Palisaitis et al. (65)	x	✓	x	x[a]	★★★☆
Popp et al.(83)	✓[b]	✓	x	x	★★☆☆
Samadi et al.(84)	?	✓	x	x	★★☆☆
Turksoy et al. (85)	x	✓	x	x	★★★☆
Weimer et al. (87)	✓	✓	✓	x	★☆☆☆

x, no; ✓, yes;?, no information available; CGM, continuous glucose monitoring; IA, ingestive activity. "x" in all columns would signal minimal methodological concerns and thus an overall rating of four stars in the rightmost column; the more "✓," the greater the methodological concerns, and consequently, the fewer stars are awarded in the rightmost column. Note that for the overall rating, no available information ("?") was treated as eliciting methodological concerns ("✓"). [a]While insulin pump data were also used in this paper, the meal detection algorithm relies only on glucose data as input. [b]The authors state that self-reported IA is not seen as the ground-truth method in their work; however, in the absence of direct observation of IA in this study, herein, we assume self-reported IA as the ground-truth.

announcements or insulin boluses and focusing on samples with diabetes in the reviewed studies. All included publications suffered at least one such methodological limitation (Table 2).

Several reviewed articles reported solutions that relied solely on CGM as input for their IA detection algorithms. Performance among these approaches varied, but sensitivities ≥90% were achieved by several groups (73, 84, 85), and false positive occurrences < 1 per day were reported (75). This suggests that inputs other than CGM are not necessary to achieve excellent performance in automatically detecting IA.

Of note, some algorithms that incorporated inputs other than CGM might also work with CGM as their only input for the specific goal of IA detection. For example, in Bertrand et al.'s machine-learning algorithms, data from two wrist-worn activity trackers were incorporated in addition to the CGM data (61). However, the 20 most important features were derived from the CGM data (61). Hence, it is likely that an adaptation of their algorithm that relies exclusively on the CGM data as input might also achieve good – albeit likely *somewhat* worse – IA detection performance. Similar cases can be made for other publications in which insulin data were used as input in addition to the CGM data (59, 63, 64). These results suggest that it is possible to automatically detect IA using CGM-based and even CGM-only algorithms.

4.3 Can these approaches be used in the context of JITAIs?

Generally, to successfully administer a dietary JITAI, IA must be detected in (near-)real-time. However, precisely how short the detection would have to be depends on the specific goal, as outlined before. Detection times as fast as 9.0 min were reported (70), but most approaches needed 20 to 40 min to detect IA (Table 1). This can generate feedback on IA much faster than traditional dietary assessment methods, such as 24-h recalls, thus creating opportunities for earlier intervention. For instance, detecting deviations from a standardized study procedure (e.g., when IA is detected in a fasting window) is likely possible. Further, when deviations from a specific meal plan (e.g., low carbohydrate) are detected, the plan could be adjusted for the subsequent meals on the same day. Automated meal detection could further trigger behavioral intervention prompts regarding portion size and eating rate (i.e., reminders to eat more slowly) for future meals. However, in most cases, detection times are too long to modify/influence IA truly in the moment it occurs (e.g., a participant on a ketogenic diet has likely already finished a carbohydrate-rich meal by the time it is detected). Regardless, it is debatable if that is really the goal and what intervening *during* an eating event would look like.

4.4 Implications for clinical and research practice

Our review shows that several CGM-based options for the automatic detection of IA exist. Ultimately, the specific use case will dictate the most suitable approach. Different approaches might be appropriate depending on factors such as the budget, population, targeted level of wearing comfort, and goal of the automatic IA detection.

Notably, other innovative methods for the automatic detection of IA, such as those using wearable-, sensor-, and image-based methods (9, 15–18), are also promising. These methods may even be superior to CGM-based approaches regarding detection times. Wang and colleagues identified several devices that can quickly detect IA (16), such as a headband device that can detect eating events via chewing sounds within only 3 min (16, 90). Similarly, a pilot study by Kumar and colleagues investigating the use of abdominal sounds to detect IA found an average detection time of only 4.3 min (91). It has even been demonstrated that eating events can be predicted ahead of time (16, 92). Yang et al. used a camera, a GPS device, and an accelerometer to predict eating and food-purchasing events up to 4 min in advance (92). The authors found that a trained gradient-boosting model achieved a mean accuracy of 72.9% in predicting eating events 0–4 min in advance (92). This highlights that different methodologies might have inherent strengths and limitations. The suboptimal detection times might be considered an inherent limitation of CGM-based approaches. Recent advances have tried to solve the CGM-inherent lag time issue (93), but more research is needed. It remains to be seen whether these limitations inherent to using CGM for automatically detecting IA can be overcome. On the other hand, one key benefit of using CGM might be its unobtrusiveness, which could facilitate its acceptance in practice. This unobtrusiveness contrasts many other, more obtrusive approaches such as glasses and camera-based methods (9, 16, 18).

A promising prospect might be to use a sensitive CGM-based approach that sends a prompt to the patient/participant asking them to log IA in case of a true positive detection. Thus, the CGM-based approach would serve as an automated reminder. That way, a false positive detection does not automatically lead to erroneous IA information but needs to be verified by the person. In this context, the suboptimal detection times also likely would be acceptable.

4.5 Limitations and directions for future research

4.5.1 Sample characteristics

Unsurprisingly, most publications included samples with diabetes, as the primary use case for automated IA detection is AP systems. However, to examine the potential of CGM-based approaches for detecting IA in various populations, more research in more diverse populations, including healthy individuals, should be conducted. This is particularly important as the generalizability of previous findings to non-diabetic individuals is likely limited, for instance, due to the usually far lower variations in blood glucose levels in persons without diabetes (77) as compared to persons with diabetes (78). Thus, there may likely be systematic differences in the performance of such approaches in individuals with diabetes compared to those without diabetes. Moreover, in many studies, meals were announced to the system, and/or manual insulin boluses accompanied the registered meals. For example, Ornetzeder and colleagues evaluated the detection performance of three previously published algorithms (79–81) using meals accompanied by insulin boluses (75). While the resulting performance metrics of this publication and similar others are promising, they need to be interpreted considering the applied insulin boluses. Ornetzeder et al. argue that this potential distortion was deemed acceptable due to a lack of alternative, insulin bolus-free datasets and the time it takes for the administered insulin to achieve its peak action (75). However, it is still possible that the results of CGM-based IA detection approaches might differ in scenarios without exogenous insulin infusions. Specifically, the administered insulin might flatten the blood glucose excursions from the meal's start, making its automatic detection less likely. In line with this, Faccioli et al. state that some of their false negatives might have been related to the attenuated postprandial CGM curves following the administration of meal-accompanying insulin therapy (66). At the same time, it should be considered that the postprandial glucose excursions of individuals with insulin-dependent diabetes would be much more pronounced without insulin treatment than in non-diabetic individuals (82). As such, it could be argued that by administering meal boluses, the postprandial glucose excursions of individuals with diabetes more closely approximate those of individuals without diabetes. Direct evidence is, of course, still necessary to increase confidence in any conclusions. Thus, future studies should ultimately enroll more individuals without diabetes.

4.5.2 Research focus

Moreover, it also needs to be considered that for the initialization of closed-loop systems, background information (e.g., treatment management, physical characteristics of the patient) is typically provided to the system (59). This information may only sometimes be readily available in other contexts. In addition, the goals of algorithms geared toward use in closed-loop/AP systems might differ from approaches aimed at the use for automatic detection of IA in general. For instance, in their AP-oriented work, Kölle et al. focused on glucose excursions caused by larger meals because smaller meals or snacks, which do not cause a substantial increase in blood glucose levels, do not necessarily need to be detected and trigger an insulin bolus to ensure adequate glucose control (73). Yet, in a scenario where the automatic detection of IA via CGM is meant to provide information on *any* IA – irrespective of its size – this argument does not hold up. This example highlights the potential differences in the setup of algorithms depending on the goal.

Taken together, fundamentally different circumstances and goals may be pursued, and thus, algorithms may be constructed differently, depending on the research question. Consequently, it might be possible to further optimize algorithms to automatically detect IA in research or clinical settings other than closed-loop/AP systems.

4.5.3 Comparability of approaches

There was substantial heterogeneity in how the performance of the investigated approaches was evaluated across the reviewed publications. Thus, as noted by others (66), a direct comparison between the approaches is difficult due to differences in the utilized datasets, preprocessing, and evaluation methods. Differences like these ultimately hamper the search for the best-performing

approaches. Performance metrics reported in publications should include at least the following measures: the number of true positives, false positives, and false negatives, which can be used to calculate important metrics such as sensitivity and precision; the detection time, defined as the time from the start of the IA to the time the algorithm detects the IA, whose reporting allows researchers and practitioners to judge whether a specific approach could be used to administer JITAIs, for example. A short detection time followed by a prompt could also allow for more immediate self-reported IA. More accurate self-reports could be the consequence due to diminished recall bias.

4.5.4 Future avenues

In general, more research should be dedicated to using CGM for the specific goal of automatically detecting IA in a broad range of populations, particularly in individuals without diabetes. Such approaches have several potential benefits, but prior research has mainly focused on using CGM for diabetes care and AP systems. However, as explained, algorithms will likely be constructed differently for the specific goal of automatically detecting IA. Moreover, previous findings will have to be replicated and extended in non-diabetic samples to overcome the currently limited generalizability.

Depending on the use case, several advancements would be necessary to rely exclusively on a CGM-based/CGM-only approach for the remote monitoring of IA. To fully automate the logging of IA times in a reliable manner, most systems would have to be even more accurate than they currently are.

If the goal is to further automate IA timing and effectively log macronutrient intake, approaches would have to incorporate specific algorithms for this task. Several publications explored whether estimating macronutrients from CGM data is possible. For instance, Samadi et al. estimated the carbohydrate content of meals (84). Results were promising, with 64.1% of the detected IA events having an absolute carbohydrate estimation error of less than 25 g (84).

Similarly, if the goal is to administer JITAIs to impact acute IA, detection times would have to decrease further. However, as mentioned above, the lag time-caused suboptimal detection times might have to be considered an inherent limitation of CGM-based approaches. Only if future studies succeed at further reducing detection times will the application of CGM-based approaches for dietary JITAIs aiming to alter IA in the moment in the truest sense of the word become possible. This is especially true for cases in which meals are followed by only small and/or delayed postprandial glucose excursions (e.g., after high-fat meals) or when meals contain only a small amount of carbohydrates, as (timely) detection appears difficult here (66, 74). It would also be necessary to explicitly test the detection performance in cases of such challenging IA (e.g., ketogenic diets). Empirical data on such cases might enable the prediction of in which settings CGM-based approaches can be used for successfully detecting IA (e.g., only in contexts where at least moderate amounts of carbohydrates are consumed).

We advise that future studies use different approaches on the same dataset, providing comprehensive CGM and objective IA data, and then compare their performance using the abovementioned metrics. A starting point could be to compare the CGM-only approaches highlighted in Table 2. Such a fair and standardized comparison could further illuminate the currently most promising approach(es).

While CGM-only approaches are highly attractive because they only necessitate one single sensor (i.e., the CGM), multi-sensor solutions also hold great potential and should thus be further investigated. Specifically, combining the strengths of different sensors (e.g., CGM and wristbands) may yield superior results as compared to relying on only one sensor, although this remains to be determined empirically.

Lastly, similar to others (10), we strongly advise that researchers use interdisciplinary collaborations to develop new CGM-based dietary monitoring tools to combine technological and biological/nutritional expertise. Interdisciplinary collaborations should ensure that the resulting tools are useful and optimized from both perspectives.

4.6 Conclusion

Based on an exhaustive and systematic literature search, this scoping review shows that it is possible to automatically detect IA using CGM-based approaches. Despite methodological issues and substantial overall heterogeneity among publications, CGM-based dietary monitoring might complement clinical and research practice.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

JB: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Visualization. CG: Methodology, Writing – review & editing. SHF: Methodology, Writing – review & editing. KK: Conceptualization, Writing – review & editing. CH: Conceptualization, Methodology, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was funded internally by the Technical University of Munich. CH was supported by a Research Fellowship from the Alexander von Humboldt Foundation. These funding sources had no role in the design, execution, analyses, interpretation of the data, or decision to submit results. Where reported, most reviewed publications received funding primarily from national agencies. Some publications have also reported funding from non-profit organizations and industry partners.

Acknowledgments

The authors want to thank the authors of Bertrand et al. (60, 61), El Fathi et al. (64), and Palisaitis et al. (65) for kindly providing additional information that helped with the interpretation of their results.

Conflict of interest

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

CH and KK were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

1. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the global burden of disease study 2017. *Lancet*. (2019) 393:1958–72. doi: 10.1016/S0140-6736(19)30041-8

2. Cena H, Calder PC. Defining a healthy diet: evidence for the role of contemporary dietary patterns in health and disease. *Nutrients*. (2020) 12:334. doi: 10.3390/nu12020334

3. Barbaresko J, Lellmann AW, Schmidt A, Lehmann A, Amini AM, Egert S, et al. Dietary factors and neurodegenerative disorders: an umbrella review of Meta-analyses of prospective studies. *Adv Nutr*. (2020) 11:1161–73. doi: 10.1093/advances/nmaa053

4. Neuenschwander M, Ballon A, Weber KS, Norat T, Aune D, Schwingshackl L, et al. Role of diet in type 2 diabetes incidence: umbrella review of meta-analyses of prospective observational studies. *BMJ*. (2019) 366:l2368. doi: 10.1136/bmj.l2368

5. Papadimitriou N, Markozannes G, Kanellopoulou A, Critselis E, Alhardan S, Karafousia V, et al. An umbrella review of the evidence associating diet and cancer risk at 11 anatomical sites. *Nat Commun*. (2021) 12:4579. doi: 10.1038/s41467-021-24861-8

6. Willett WC, Stampfer MJ. Current evidence on healthy eating. *Annu Rev Public Health*. (2013) 34:77–95. doi: 10.1146/annurev-publhealth-031811-124646

7. English LK, Ard JD, Bailey RL, Bates M, Bazzano LA, Boushey CJ, et al. Evaluation of dietary patterns and all-cause mortality: a systematic review. *JAMA Netw Open*. (2021) 4:e2122277. doi: 10.1001/jamanetworkopen.2021.22277

8. Schwingshackl L, Schwedhelm C, Hoffmann G, Lampousi AM, Knüppel S, Iqbal K, et al. Food groups and risk of all-cause mortality: a systematic review and meta-analysis of prospective studies. *Am J Clin Nutr*. (2017) 105:1462–73. doi: 10.3945/ajcn.117.153148

9. Höchsmann C, Martin CK. Review of the validity and feasibility of image-assisted methods for dietary assessment. *Int J Obes*. (2020) 44:2358–71. doi: 10.1038/s41366-020-00693-2

10. Klurfeld DM, Hekler EB, Nebeker C, Patrick K, Khoo CSH. Technology innovations in dietary intake and physical activity assessment: challenges and recommendations for future directions. *Am J Prev Med*. (2018) 55:e117–22. doi: 10.1016/j.amepre.2018.06.013

11. McClung HL, Ptomey LT, Shook RP, Aggarwal A, Gorczyca AM, Sazonov ES, et al. Dietary intake and physical activity assessment: current tools, techniques, and Technologies for use in adult populations. *Am J Prev Med*. (2018) 55:e93–e104. doi: 10.1016/j.amepre.2018.06.011

12. Ravelli MN, Schoeller DA. An objective measure of energy intake using the principle of energy balance. *Int J Obes*. (2021) 45:725–32. doi: 10.1038/s41366-021-00738-0

13. Webb GP. Methods of nutritional assessment and surveillance In: *Nutrition: Maintaining and improving health. 5th*. Boca Raton, FL: CRC Press (2019). 47–87.

14. Wong WW, Roberts SB, Racette SB, das SK, Redman LM, Rochon J, et al. The doubly labeled water method produces highly reproducible longitudinal results in nutrition studies. *J Nutr*. (2014) 144:777–83. doi: 10.3945/jn.113.187823

15. Bell BM, Alam R, Alshurafa N, Thomaz E, Mondol AS, de la Haye K, et al. Automatic, wearable-based, in-field eating detection approaches for public health research: a scoping review. *NPJ Digit Med*. (2020) 3:38. doi: 10.1038/s41746-020-0246-2

16. Wang L, Allman-Farinelli M, Yang JA, Taylor JC, Gemming L, Hekler E, et al. Enhancing nutrition care through real-time, sensor-based capture of eating occasions: a scoping review. *Front Nutr*. (2022) 9:852984. doi: 10.3389/fnut.2022.852984

17. Hassannejad H, Matrella G, Ciampolini P, De Munari I, Mordonini M, Cagnoni S. Automatic diet monitoring: a review of computer vision and wearable sensor-based methods. *Int J Food Sci Nutr*. (2017) 68:656–70. doi: 10.1080/09637486.2017.1283683

18. Vu T, Lin F, Alshurafa N, Xu W. Wearable food intake monitoring technologies: a comprehensive review. *Computers*. (2017) 6:4. doi: 10.3390/computers6010004

19. Bellido V, Aguilera E, Cardona-Hernandez R, Diaz-Soto G, González Pérez de Villar N, Picón-César MJ, et al. Expert recommendations for using time-in-range and other continuous glucose monitoring metrics to achieve patient-centered glycemic control in people with diabetes. *J Diabetes Sci Technol*. (2023) 17:1326–36. doi: 10.1177/19322968221088601

20. Holzer R, Bloch W, Brinkmann C. Continuous glucose monitoring in healthy adults—possible applications in health care, wellness, and sports. *Sensors*. (2022) 22:2030. doi: 10.3390/s22052030

21. Zheng M, Ni B, Kleinberg S. Automated meal detection from continuous glucose monitor data through simulation and explanation. *J Am Med Inform Assoc*. (2019) 26:1592–9. doi: 10.1093/jamia/ocz159

22. Lal RA, Ekhlaspour L, Hood K, Buckingham B. Realizing a closed-loop (artificial pancreas) system for the treatment of type 1 diabetes. *Endocr Rev*. (2019) 40:1521–46. doi: 10.1210/er.2018-00174

23. American Diabetes Association Professional Practice Committee. 6. Glycemic targets: standards of medical Care in Diabetes—2022. *Diabetes Care*. (2022) 45:S83–96. doi: 10.2337/dc22-S006

24. Aleppo G, Webb K. Continuous glucose monitoring integration in clinical practice: a stepped guide to data review and interpretation. *J Diabetes Sci Technol*. (2019) 13:664–73. doi: 10.1177/1932296818813581

25. Pease A, Lo C, Earnest A, Kiriakova V, Liew D, Zoungas S. The efficacy of Technology in Type 1 diabetes: a systematic review, network Meta-analysis, and narrative synthesis. *Diabetes Technol Ther*. (2020) 22:411–21. doi: 10.1089/dia.2019.0417

26. Renard E. Automated insulin delivery systems: from early research to routine care of type 1 diabetes. *Acta Diabetol*. (2022) 60:151–61. doi: 10.1007/s00592-022-01929-5

27. Ishihara K, Uchiyama N, Kizaki S, Mori E, Nonaka T, Oneda H. Application of continuous glucose monitoring for assessment of individual carbohydrate requirement during ultramarathon race. *Nutrients*. (2020) 12:1121. doi: 10.3390/nu12041121

28. Alva S, Bailey T, Brazg R, Budiman ES, Castorino K, Christiansen MP, et al. Accuracy of a 14-day factory-calibrated continuous glucose monitoring system with advanced algorithm in pediatric and adult population with diabetes. *J Diabetes Sci Technol*. (2022) 16:70–7. doi: 10.1177/1932296820958754

29. Freckmann G, Link M, Kamecke U, Haug C, Baumgartner B, Weitgasser R. Performance and usability of three Systems for Continuous Glucose Monitoring in direct comparison. *J Diabetes Sci Technol*. (2019) 13:890–8. doi: 10.1177/1932296819826965

30. Schembre SM, Liao Y, Robertson MC, Dunton GF, Kerr J, Haffey ME, et al. Just-in-time feedback in diet and physical activity interventions: systematic review and practical design framework. *J Med Internet Res*. (2018) 20:e106. doi: 10.2196/jmir.8701

31. Forman EM, Goldstein SP, Zhang F, Evans BC, Manasse SM, Butryn ML, et al. OnTrack: development and feasibility of a smartphone app designed to predict and prevent dietary lapses. *Transl Behav Med*. (2019) 9:236–45. doi: 10.1093/tbm/iby016

32. American Diabetes Association. Postprandial Blood Glucose. *Diabetes Care*. (2001) 24:775–8. doi: 10.2337/diacare.24.4.775

33. Bell KJ, Smart CE, Steil GM, Brand-Miller JC, King B, Wolpert HA. Impact of fat, protein, and glycemic index on postprandial glucose control in type 1 diabetes: implications for intensive diabetes Management in the Continuous Glucose Monitoring era. *Diabetes Care*. (2015) 38:1008–15. doi: 10.2337/dc15-0100

34. Gingras V, Bonato L, Messier V, Roy-Fleming A, Smaoui MR, Ladouceur M, et al. Impact of macronutrient content of meals on postprandial glucose control in the context of closed-loop insulin delivery: A randomized cross-over study. *Diabetes Obes Metab*. (2018) 20:2695–9. doi: 10.1111/dom.13445

35. Furthner D, Lukas A, Schneider AM, Mörwald K, Maruszczak K, Gombos P, et al. The role of protein and fat intake on insulin therapy in Glycaemic control of Paediatric type 1 diabetes: a systematic review and research gaps. *Nutrients*. (2021) 13:3558. doi: 10.3390/nu13103558

36. Pinsker JE, Lee JB, Dassau E, Seborg DE, Bradley PK, Gondhalekar R, et al. Randomized crossover comparison of personalized MPC and PID control algorithms for the artificial pancreas. *Diabetes Care*. (2016) 39:1135–42. doi: 10.2337/dc15-2344

37. Shilo S, Godneva A, Rachmiel M, Korem T, Kolobkov D, Karady T, et al. Prediction of personal glycemic responses to food for individuals with type 1 diabetes through integration of clinical and microbial data. *Diabetes Care*. (2022) 45:502–11. doi: 10.2337/dc21-1048

38. Rein M, Ben-Yacov O, Godneva A, Shilo S, Zmora N, Kolobkov D, et al. Effects of personalized diets by prediction of glycemic responses on glycemic control and metabolic health in newly diagnosed T2DM: a randomized dietary intervention pilot trial. *BMC Med*. (2022) 20:56. doi: 10.1186/s12916-022-02254-y

39. Rossetti P, Ampudia-Blasco FJ, Laguna A, Revert A, Vehì J, Ascaso JF, et al. Evaluation of a novel continuous glucose monitoring-based method for mealtime

insulin dosing--the iBolus--in subjects with type 1 diabetes using continuous subcutaneous insulin infusion therapy: a randomized controlled trial. *Diabetes Technol Ther*. (2012) 14:1043–52. doi: 10.1089/dia.2012.0145

40. Hall KD, Guo J, Courville AB, Boring J, Brychta R, Chen KY, et al. Effect of a plant-based, low-fat diet versus an animal-based, ketogenic diet on ad libitum energy intake. *Nat Med*. (2021) 27:344–53. doi: 10.1038/s41591-020-01209-1

41. Zaharieva DP, Turksoy K, McGaugh SM, Pooni R, Vienneau T, Ly T, et al. Lag time remains with newer real-time continuous glucose monitoring technology during aerobic exercise in adults living with type 1 diabetes. *Diabetes Technol Ther*. (2019) 21:313–21. doi: 10.1089/dia.2018.0364

42. Cobelli C, Renard E, Kovatchev B. Artificial pancreas: past, present, future. *Diabetes*. (2011) 60:2672–82. doi: 10.2337/db11-0654

43. Armario A, Marti O, Molina T, de Pablo J, Valdes M. Acute stress markers in humans: response of plasma glucose, cortisol and prolactin to two examinations differing in the anxiety they provoke. *Psychoneuroendocrinology*. (1996) 21:17–24. doi: 10.1016/0306-4530(95)00048-8

44. García-García F, Kumareswaran K, Hovorka R, Hernando ME. Quantifying the acute changes in glucose with exercise in type 1 diabetes: a systematic review and Meta-analysis. *Sports Med*. (2015) 45:587–99. doi: 10.1007/s40279-015-0302-2

45. Park SH, Yao J, Chua XH, Chandran SR, Gardner DSL, Khoo CM, et al. Diet and physical activity as determinants of continuously measured glucose levels in persons at high risk of type 2 diabetes. *Nutrients*. (2022) 14:366. doi: 10.3390/nu14020366

46. Bennetsen SL, Feineis CS, Legaard GE, Lyngbæk MPP, Karstoft K, Ried-Larsen M. The impact of physical activity on glycemic variability assessed by continuous glucose monitoring in patients with type 2 diabetes mellitus: a systematic review. *Front Endocrinol*. (2020) 11:486. doi: 10.3389/fendo.2020.00486

47. Colberg SR, Sigal RJ, Yardley JE, Riddell MC, Dunstan DW, Dempsey PC, et al. Physical activity/exercise and diabetes: a position statement of the American Diabetes Association. *Diabetes Care*. (2016) 39:2065–79. doi: 10.2337/dc16-1728

48. Wiesli P, Schmid C, Kerwer O, Nigg-Koch C, Klaghofer R, Seifert B, et al. Acute psychological stress affects glucose concentrations in patients with type 1 diabetes following food intake but not in the fasting state. *Diabetes Care*. (2005) 28:1910–5. doi: 10.2337/diacare.28.8.1910

49. Gonder-Frederick LA, Grabman JH, Kovatchev B, Brown SA, Patek S, Basu A, et al. Is psychological stress a factor for incorporation into future closed-loop systems? *J Diabetes Sci Technol*. (2016) 10:640–6. doi: 10.1177/1932296816635199

50. Sapolsky RM, Romero LM, Munck AU. How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocr Rev*. (2000) 21:55–89. doi: 10.1210/edrv.21.1.0389

51. Bequette BW. Challenges and recent Progress in the development of a closed-loop artificial pancreas. *Annu Rev Control*. (2012) 36:255–66. doi: 10.1016/j.arcontrol.2012.09.007

52. Colmegna P, Garelli F, De Battista H, Sánchez-Peña R. Automatic regulatory control in type 1 diabetes without carbohydrate counting. *Control Eng Pract*. (2018) 74:22–32. doi: 10.1016/j.conengprac.2018.02.003

53. Reddy R, Wittenberg A, Castle JR, el Youssef J, Winters-Stone K, Gillingham M, et al. Effect of aerobic and resistance exercise on glycemic control in adults with type 1 diabetes. *Can J Diabetes*. (2019) 43:406–414.e1. doi: 10.1016/j.jcjd.2018.08.193

54. Cinar A, Turksoy K. Multivariable control of glucose concentration In: *Advances in artificial pancreas systems*. Cham: Springer International Publishing (2018). 65–82.

55. Maahs DM, DeSalvo D, Pyle L, Ly T, Messer L, Clinton P, et al. Effect of acetaminophen on CGM glucose in an outpatient setting. *Diabetes Care*. (2015) 38:e158–9. doi: 10.2337/dc15-1096

56. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. (2018) 169:467–73. doi: 10.7326/M18-0850

57. Diamond T, Cameron F, Bequette BW. A new meal absorption model for artificial pancreas systems. *J Diabetes Sci Technol*. (2022) 16:40–51. doi: 10.1177/1932296821990111

58. Ravelli MN, Schoeller DA. Traditional self-reported dietary instruments are prone to inaccuracies and new approaches are needed. *Front Nutr*. (2020) 7:90. doi: 10.3389/fnut.2020.00090

59. Atlas E, Nimri R, Miller S, Grunberg EA, Phillip M. MD-logic artificial pancreas system. *Diabetes Care*. (2010) 33:1072–6. doi: 10.2337/dc09-1830

60. Bertrand L., Cleyet-Marrel N., Liang Z. The role of continuous glucose monitoring in automatic detection of eating activities. In *2021 IEEE 3rd global conference on life sciences and technologies (LifeTech)*. Nara, Japan: IEEE; (2021). p. 313–314.

61. Bertrand L, Cleyet-Marrel N, Liang Z. Recognizing eating activities in free-living environment using consumer wearable devices. In: *Engineering Proceedings MDPI*; (2021). p. 58.

62. Dassau E, Bequette BW, Buckingham BA, Doyle FJ. Detection of a meal using continuous glucose monitoring. *Diabetes Care*. (2008) 31:295–300. doi: 10.2337/dc07-1293

63. Dovc K, Piona C, Yeşiltepe Mutlu G, Bratina N, Jenko Bizjan B, Lepej D, et al. Faster compared with standard insulin Aspart during day-and-night fully closed-loop insulin therapy in type 1 diabetes: a double-blind randomized crossover trial. *Diabetes Care*. (2020) 43:29–36. doi: 10.2337/dc19-0895

64. El Fathi A, Palisaitis E, Boulet B, Legault L, Haidar A. An Unannounced Meal Detection Module for artificial pancreas control systems. In: *2019 American control conference (ACC)*. Philadelphia, PA, USA: IEEE; (2019). p. 4130–4135.

65. Palisaitis E, El Fathi A, von Oettingen JE, Haidar A, Legault L. A meal detection algorithm for the artificial pancreas: a randomized controlled clinical trial in adolescents with type 1 diabetes. *Diabetes Care*. (2021) 44:604–6. doi: 10.2337/dc20-1232

66. Faccioli S, Sala-Mira I, Díez JL, Facchinetti A, Sparacino G, del Favero S, et al. Super–twisting-based meal detector for type 1 diabetes management: improvement and assessment in a real-life scenario. *Comput Methods Prog Biomed*. (2022) 219:106736. doi: 10.1016/j.cmpb.2022.106736

67. Anderson SM, Raghinaru D, Pinsker JE, Boscari F, Renard E, Buckingham BA, et al. Multinational home use of closed-loop control is safe and effective. *Diabetes Care*. (2016) 39:1143–50. doi: 10.2337/dc15-2468

68. Fushimi E, Colmegna P, De Battista H, Garelli F, Sanchez-Pena R. Unannounced meal analysis of the ARG algorithm. In: *2019 American control conference (ACC)*. Philadelphia, PA, USA: IEEE; (2019) p. 4740–4745.

69. Sánchez-Peña R, Colmegna P, Garelli F, De Battista H, García-Violini D, Moscoso-Vásquez M, et al. Artificial pancreas: clinical study in Latin America without Premeal insulin boluses. *J Diabetes Sci Technol*. (2018) 12:914–25. doi: 10.1177/1932296818786488

70. Godoy JL, Sereno JE, Rivadeneira PS. Meal detection and carbohydrate estimation based on a feedback scheme with application to the artificial pancreas. *Biomed Signal Process Control*. (2021) 68:102715. doi: 10.1016/j.bspc.2021.102715

71. Hoyos J. D., Laguna A. J., Carpintero A. R., Sala-Mira I., Diez J. L., Bondia J. Characterization of glycemic patterns in type 1 diabetes without insulin or meal input data. In: 2022 10th international conference on systems and control (ICSC). Marseille, France: IEEE; (2022). p. 576–581.

72. Aleppo G, Ruedy KJ, Riddlesworth TD, Kruger DF, Peters AL, Hirsch I, et al. REPLACE-BG: a randomized trial comparing continuous glucose monitoring with and without routine blood glucose monitoring in adults with well-controlled type 1 diabetes. *Diabetes Care*. (2017) 40:538–45. doi: 10.2337/dc16-2482

73. Kölle K, Biester T, Christiansen S, Fougner AL, Stavdahl O. Pattern recognition reveals characteristic postprandial glucose changes: non-individualized meal detection in diabetes mellitus type 1. *IEEE J Biomed Health Inform*. (2020) 24:594–602. doi: 10.1109/JBHI.2019.2908897

74. Mosquera-Lopez C, Wilson LM, El Youssef J, Hilts W, Leitschuh J, Branigan D, et al. Enabling fully automated insulin delivery through meal detection and size estimation using artificial intelligence. *NPJ Digit Med*. (2023) 6:39. doi: 10.1038/s41746-023-00783-1

75. Ornetzeder C, Reiterer F, Christensen MB, Norgaard K, Freckmann G, del Re L. Feasibility of fully closed loop insulin delivery in type 2 diabetes. In: *2019 IEEE conference on control technology and applications (CCTA)*. Hong Kong, China: IEEE; (2019). p. 906–913.

76. Zschornack E, Schmid C, Pleus S, Link M, Klötzer HM, Obermaier K, et al. Evaluation of the performance of a novel system for continuous glucose monitoring. *J Diabetes Sci Technol*. (2013) 7:815–23. doi: 10.1177/193229681300700403

77. Shah VN, DuBose SN, Li Z, Beck RW, Peters AL, Weinstock RS, et al. Continuous glucose monitoring profiles in healthy nondiabetic participants: a multicenter prospective study. *J Clin Endocrinol Metab*. (2019) 104:4356–64. doi: 10.1210/jc.2018-02763

78. Gubitosi-Klug RA, Braffett BH, Bebu I, Johnson ML, Farrell K, Kenny D, et al. Continuous glucose monitoring in adults with type 1 diabetes with 35 years duration from the DCCT/EDIC study. *Diabetes Care*. (2022) 45:659–65. doi: 10.2337/dc21-0629

79. Harvey RA, Dassau E, Zisser H, Seborg DE, Doyle FJ. Design of the Glucose Rate Increase Detector: a meal detection module for the health monitoring system. *J Diabetes Sci Technol*. (2014) 8:307–20. doi: 10.1177/1932296814523881

80. Samadi S, Turksoy K, Hajizadeh I, Feng J, Sevil M, Cinar A. Meal detection and carbohydrate estimation using continuous glucose sensor data. *IEEE J Biomed Health Inform*. (2017) 21:619–27. doi: 10.1109/JBHI.2017.2677953

81. Ramkissoon C, Herrero P, Bondia J, Vehi J. Unannounced meals in the artificial pancreas: detection using continuous glucose monitoring. *Sensors*. (2018) 18:884. doi: 10.3390/s18030884

82. Palacios V, Woodbridge DMK, Fry JL. Machine learning-based meal detection using continuous glucose monitoring on healthy participants: an objective measure of participant compliance to protocol. In: *2021 43rd annual international conference of the IEEE engineering in Medicine & Biology Society (EMBC)*. Mexico: IEEE; (2021). p. 7032–7035.

83. Popp CJ, Wang C, Hoover A, Gomez LA, Curran M, St-Jules DE, et al. Objective determination of eating occasion timing (OREO): combining self-report, wrist motion, and continuous glucose monitoring to detect eating occasions in adults with pre-diabetes and obesity. *J Diabetes Sci Technol*. (2023):19322968231197205. doi: 10.1177/19322968231197205

84. Samadi S, Rashid M, Turksoy K, Feng J, Hajizadeh I, Hobbs N, et al. Automatic detection and estimation of unannounced meals for multivariable artificial pancreas system. *Diabetes Technol Ther*. (2018) 20:235–46. doi: 10.1089/dia.2017.0364

85. Turksoy K, Samadi S, Feng J, Littlejohn E, Quinn L, Cinar A. Meal detection in patients with type 1 diabetes: a new module for the multivariable adaptive artificial pancreas control system. *IEEE J Biomed Health Inform*. (2016) 20:47–54. doi: 10.1109/JBHI.2015.2446413

86. Turksoy K, Quinn LT, Littlejohn E, Cinar A. An integrated multivariable artificial pancreas control system. *J Diabetes Sci Technol*. (2014) 8:498–507. doi: 10.1177/1932296814524862

87. Weimer J, Chen S, Peleckis A, Rickels MR, Lee I. Physiology-invariant meal detection for type 1 diabetes. *Diabetes Technol Ther*. (2016) 18:616–24. doi: 10.1089/dia.2015.0266

88. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. Statement: an updated guideline for reporting systematic reviews. *BMJ*. (2020) 2021:n71. doi: 10.1136/bmj.n71

89. Lee H, Wayne Bequette, A closed-loop artificial pancreas based on model predictive control: Human-friendly identification and automatic meal disturbance rejection. *Biomed. Signal Process. Control*. (2009) 4:347–54. doi: 10.1016/j.bspc.2009.03.002

90. Bi S, Wang T, Tobias N, Nordrum J, Wang S, Halvorsen G, et al. Auracle: detecting eating episodes with an ear-mounted sensor. *Proc ACM Interact Mob Wearable Ubiquitous Technol*. (2018) 2:1–27. doi: 10.1145/3264902

91. Kumar TS, Soiland E, Stavdahl O, Fougner AL. Pilot study of early meal onset detection from abdominal sounds. In: *2019 E-health and bioengineering conference (EHB)*. Iasi, Romania: IEEE; (2019). p. 1–4.

92. Yang JA, Wang J, Nakandala S, Kumar A, Jankowska MM. Predicting eating events in free living individuals. In *2019 15th international conference on eScience (eScience)*. San Diego, CA, USA: IEEE; (2019). p. 627–629.

93. Halvorsen M, Benam KD, Khoshamadi H, Fougner AL. Blood glucose level prediction using subcutaneous sensors for in vivo study: compensation for measurement method slow dynamics using Kalman filter approach. In *2022 IEEE 61st conference on decision and control (CDC)*. Cancun, Mexico: IEEE; (2022). p. 6034–6039.

PART III

Other Work

Publication 11

Effects of a 2-year primary care lifestyle intervention on cardiometabolic risk factors: A cluster-randomized trial.

Christoph Höchsmann [1], James L Dorling [1], Corby K Martin [1], Robert L Newton Jr [1], John W Apolzan [1], Candice A Myers [1], Kara D Denstel [1], Emily F Mire [1], William D Johnson [1], Dachuan Zhang [1], Connie L Arnold [2], Terry C Davis [2], Vivian Fonseca [3], Carl J Lavie [4], Eboni G Price-Haywood [5], Peter T Katzmarzyk [1] for the PROPEL Research Group

[1] Pennington Biomedical Research Center, Baton Rouge, LA, USA

[2] Louisiana State University Health Sciences Center, Shreveport, LA, USA

[3] Department of Medicine, Tulane University Health Sciences Center, New Orleans, LA, USA

[4] John Ochsner Heart and Vascular Institute, New Orleans, LA, USA

[5] Center for Outcomes and Health Services Research, New Orleans, LA, USA

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
Circulation, 143(12): 1202–1214	2021	10.1161/CIRCULATIONAHA.120.051328	39.9 (2021)	Q1	Scopus: 27 Google Scholar: 31

Author contribution:

First author; oversaw intervention sessions at six participating clinics and led the weekly meetings with the interventionists during the last six months of the study, co-developed the research question and statistical model, drafted the manuscript, and created tables and figures.

Circulation

ORIGINAL RESEARCH ARTICLE

Effects of a 2-Year Primary Care Lifestyle Intervention on Cardiometabolic Risk Factors

A Cluster-Randomized Trial

BACKGROUND: Intensive lifestyle interventions (ILIs) are the first-line approach to effectively treat obesity and manage associated cardiometabolic risk factors. Because few people have access to ILIs in academic health centers, primary care must implement similar approaches for a meaningful effect on obesity and cardiometabolic disease prevalence. To date, however, effective lifestyle-based obesity treatment in primary care is limited. We examined the effectiveness of a pragmatic ILI for weight loss delivered in primary care among a racially diverse, low-income population with obesity for improving cardiometabolic risk factors over 24 months.

METHODS: The PROPEL trial (Promoting Successful Weight Loss in Primary Care in Louisiana) randomly allocated 18 clinics equally to usual care or an ILI and subsequently enrolled 803 (351 usual care, 452 ILI) adults (67% Black, 84% female) with obesity from participating clinics. The usual care group continued to receive their normal primary care. The ILI group received a 24-month high-intensity lifestyle-based obesity treatment program, embedded in the clinic setting and delivered by health coaches in weekly sessions initially and monthly sessions in months 7 through 24.

RESULTS: As recently demonstrated, participants receiving the PROPEL ILI lost significantly more weight over 24 months than those receiving usual care (mean difference, −4.51% [95% CI, −5.93 to −3.10]; $P<0.01$). Fasting glucose decreased more in the ILI group compared with the usual care group at 12 months (mean difference, −7.1 mg/dL [95% CI, −12.0 to −2.1]; $P<0.01$) but not 24 months (mean difference, −0.8 mg/dL [95% CI, −6.2 to 4.6]; $P=0.76$). Increases in high-density lipoprotein cholesterol were greater in the ILI than in the usual care group at both time points (mean difference at 24 months, 4.6 mg/dL [95% CI, 2.9–6.3]; $P<0.01$). Total:high-density lipoprotein cholesterol ratio and metabolic syndrome severity (z score) decreased more in the ILI group than in the usual care group at both time points, with significant mean differences of the change of −0.31 (95% CI, −0.47 to −0.14; $P<0.01$) and −0.21 (95% CI, −0.36 to −0.06; $P=0.01$) at 24 months, respectively. Changes in total cholesterol, low-density lipoprotein cholesterol, triglycerides, and blood pressure did not differ significantly between groups at any time point.

CONCLUSIONS: A pragmatic ILI consistent with national guidelines and delivered by trained health coaches in primary care produced clinically relevant improvements in cardiometabolic health in an underserved population over 24 months.

REGISTRATION: URL: https://www.clinicaltrials.gov; Unique identifier: NCT02561221.

Christoph Höchsmann, PhD
James L. Dorling, PhD
Corby K. Martin, PhD
Robert L. Newton Jr, PhD
John W. Apolzan, PhD
Candice A. Myers, PhD
Kara D. Denstel, MPH
Emily F. Mire, MS
William D. Johnson, PhD
Dachuan Zhang, MAppStats
Connie L. Arnold, PhD
Terry C. Davis, PhD
Vivian Fonseca, MD
Carl J. Lavie, MD
Eboni G. Price-Haywood, MD
Peter T. Katzmarzyk, PhD
For the PROPEL Research Group

Key Words: cardiovascular diseases ■ life style ■ obesity ■ primary health care

Sources of Funding, see page 1213

https://www.ahajournals.org/journal/circ

Clinical Perspective

What Is New?

- The PROPEL trial (Promoting Successful Weight Loss in Primary Care in Louisiana) trial examined the effectiveness of an intensive lifestyle intervention for weight loss delivered by trained health coaches in primary care among a racially diverse, low-income population with obesity for improving cardiometabolic risk factors over 24 months.
- Although pragmatic, the PROPEL intensive lifestyle intervention is consistent with national guidelines and demonstrated clinically relevant improvements in high-density lipoprotein cholesterol, total:high-density lipoprotein cholesterol ratio, metabolic syndrome severity (12 and 24 months), and fasting glucose (12 months only).
- The PROPEL model is a viable option to deliver effective obesity and cardiometabolic risk factor treatment in primary care.

What Are the Clinical Implications?

- The present results underline the cardiometabolic effectiveness of a comprehensive weight loss intervention model delivered by trained health coaches in a primary care setting.
- The collaborative care approach of the PROPEL model likely offers more successful primary care–based obesity treatment than the existing Centers for Medicare and Medicaid Services–supported model, which relies solely on primary care practitioners for obesity treatment with, to date, limited success.
- Broader implementation of the PROPEL model would particularly allow underserved populations to receive effective obesity and concomitant cardiometabolic disease risk treatment and thereby potentially contribute to reducing health inequities.

In 2017 to 2018, the prevalence of obesity among US adults was an estimated 42.4%, with severe, class III obesity being 9.2%.[1] Obesity is associated with serious chronic health risks such as cardiovascular disease (CVD), type 2 diabetes, several cancers, depression, and premature mortality,[2] presenting substantial public health and economic burden in many countries.[3] Certain sociodemographic groups are particularly affected by the obesity epidemic and consequently are at the highest risk for deleterious health effects. Black adults had an estimated prevalence of obesity as high as 49.6% in 2017 to 2018; among all adults, those 40 to 59 years of age had the highest prevalence of severe obesity (body mass index ≥40 kg/m^2; 11.5%).[1] Along with race/ethnicity, other social determinants of health such as lower levels of education, income, and food security intersect to increase risk of obesity and related comorbidities.[4–6] Such inequities are further sustained by policies (eg, health care access and affordability), systems (eg, racism, discrimination, segregation), and environments (eg, community, neighborhood), which also lead to higher obesity levels in racial/ethnic minorities.[7,8] Such health disparities are particularly apparent in states such as Louisiana, where the median household income and food security rates are among the lowest in the United States.[9,10]

Intensive lifestyle interventions (ILIs) are the first-line approach to promote weight loss and to effectively treat obesity and manage associated health risks, as outlined in the 2013 AHA [American Heart Association]/ACC [American College of Cardiology]/TOS [The Obesity Society] Guideline for the Management of Overweight and Obesity in Adults.[11,12] Large trials such as the DPP (Diabetes Prevention Program)[13] and Look AHEAD (Action for Health in Diabetes)[14] have shown that high-intensity lifestyle interventions conducted in academic health centers can induce weight loss of 5.8% and 6.4% over 2 years, respectively, and that these weight reductions are accompanied by health-beneficial changes in blood pressure (BP), glucose control, and dyslipidemia.[15–17]

However, only a small proportion of the population has access to ILIs in academic health centers. Therefore, uptake of similar approaches by primary care, the cornerstone of medical care in the United States, is imperative for a meaningful effect on the global obesity prevalence and for achieving public health goals to reduce health inequities.[18] Despite recommendations by the US Preventive Services Task Force that physicians offer intensive multicomponent behavioral interventions to individuals with obesity,[19] to date, effective lifestyle-based obesity treatment in primary care is often lacking, and long-term success for weight loss and improvement in cardiometabolic risk factors is consequently limited.[20] This is attributable in part to infrequent treatment contacts, time constraints during visits, and primary care practitioners' lack of training in behavior therapy and nutrition education.[20,21] Weight loss interventions based in primary care have produced a range of weight loss (1–2 kg in low-intensity interventions to 4–7 kg in higher-intensity interventions).[20,22] For example, in the POWER trials (Practice-Based Opportunities for Weight Reduction), greater weight loss was experienced in the Baltimore (5.2%) and Philadelphia (4.7%) trials compared with the Boston trial (1.7%).[23–25] The Boston sample consisted predominantly of Black participants with low annual income, and the intervention consisted of monthly counseling telephone calls in the first 12 months, followed by bimonthly calls in the remaining 12 months (18 total calls), in addition to self-monitoring through a website or through an interactive voice response system.[25]

We conducted a 24-month cluster-randomized trial in which a high-intensity lifestyle intervention was delivered face to face by trained health coaches embedded within primary care clinics among an underserved population with obesity.[26] The primary outcome measure of the trial was percentage change in body weight from baseline to month 24. The ILI group lost significantly more weight than the usual care group, with a mean difference of −4.51% (95% CI, −5.93 to −3.10) between the groups (P<0.01).[27] Similar to the demonstrated effectiveness for weight loss of the trial, we hypothesized that participants receiving the ILI would show improvements in cardiometabolic risk factors relative to those receiving usual care.

METHODS

Design and Participants

The PROPEL trial (Promoting Successful Weight Loss in Primary Care in Louisiana) was a 24-month high-intensity lifestyle intervention delivered within primary care clinics among a racially diverse, low-income population with obesity. The PROPEL trial was conducted between April 2016 and September 2019, and the study protocol and all procedures were approved by Pennington Biomedical Research Center's Institutional Review Board. The data that support the findings of this study are available from the corresponding author on reasonable request. Study materials and the statistical analysis plan are available as supplementary material associated with the PROPEL primary outcome article.[27] The trial design and rationale have been published in detail.[26] In brief, 18 primary care clinics across Louisiana were randomly allocated in equal numbers to either an ILI group or a usual care group. Participants were recruited from the participating clinics and deemed eligible if they were 20 to 75 years old, had a body mass index of 30 to 50 kg/m^2, and were patients at a participating clinic. Participants were excluded if they were currently participating in a weight loss program, had used weight loss medication, had undergone bariatric surgery within the last 2 years, or had lost >10 lb of weight within the last 6 months. The complete list of eligibility criteria is given elsewhere.[26] All participants provided written informed consent before enrollment in the study.

Intervention

Participants at clinics allocated to the ILI group received a comprehensive, high-intensity lifestyle intervention program[26] based on DPP,[28] Look AHEAD,[29] and CALERIE[30] (Comprehensive Assessment of Long-Term Effects of Reducing Intake of Energy Phase) and consistent with the 2013 AHA/ACC/TOS guidelines.[12] The regimen was adapted to be literacy and culturally appropriate for a low-income target population; as reported previously, 31% of PROPEL participants scored ≤6 on the Rapid Estimate of Adult Literacy in Medicine Short Form health literacy assessment at baseline, corresponding to less than a ninth-grade education and indicating limited health literacy.[26] The ILI program was embedded in the primary care clinics and consisted of weekly sessions with trained health coaches (16 face-to-face sessions and 6 via phone) during the first 6 months and at least monthly sessions for the remaining 18 months, alternating between face-to-face and phone sessions. All health coaches had higher education degrees in nutrition, physical activity, or behavioral medicine and received further training in the management of obesity and related comorbidities, fundamentals of health literacy, and patient communication and education before the start of the intervention. During the intervention, the health coaches worked with participants to meet the predefined individual goal of 10% weight loss by coaching them to develop and adhere to personalized action plans focusing on changes in eating, diet, and physical activity behavior. To monitor weight loss progress and to promote intervention fidelity, participants were encouraged to measure their weight daily with the provided BodyTrace scale (BodyTrace Inc, Palo Alto, CA), which transmitted weight data wirelessly and in real time to a computer tracking system. This system plotted the weight data onto a personalized weight graph, which was available via a website at any time, and allowed participants and health coaches to detect deviations from the intended weight loss progress quickly.[31,32] If deviations occurred, the personalized action plans were adjusted, using additional components of the toolbox approach (ie, tailored behavioral, nutritional, and physical activity strategies) that has been shown to improve intervention efficacy in previous clinical trials.[28–30] Primary care providers in ILI clinics received a series of webinars on lifestyle weight management practices, lipid management, and ways to improve communication with patients with low health literacy or obesity throughout the intervention period.

Participants at clinics allocated to the usual care group continued to receive their normal care from their primary care provider during the 24-month intervention period. In addition, they received several newsletters on topics such as the importance of sleep for health, tips for limiting sitting time, brain and memory health, and smoking cessation. Primary care providers in usual care clinics received information on the Centers for Medicare & Medicaid Services approach for intensive behavioral therapy for obesity[33] via a presentation by PROPEL staff at baseline and an informational brochure on the same topic annually.

Outcome Measures

Fasting Blood Glucose and Lipids

Fasting blood glucose (FBG) and lipids (total cholesterol, high-density lipoprotein cholesterol [HDL-C], low-density lipoprotein [LDL] cholesterol [LDL-C], and triglycerides) were measured at baseline, month 12, and month 24 using fingerstick blood samples and the Cholestech LDX Analyzer (Alere Inc, Waltham, MA). Participants were instructed to arrive at each study visit after an overnight fast (≥10 h), and the Cholestech Analyzer was calibrated daily before analyzing participant blood samples using standard controls. Furthermore, we calculated non–HDL-C (total cholesterol–HDL-C) levels and the total:HDL-C ratio, with the former representing a good surrogate measure of apolipoprotein B because it includes all atherogenic lipoproteins such as LDL, very-low-density lipoprotein, and intermediate-density lipoproteins.[34] Both non–HDL-C and the total:HDL-C ratio have previously been shown to be strongly associated with long-term risk of atherosclerotic CVD.[35,36]

Blood Pressure

Resting systolic BP (SBP) and diastolic BP (DBP) were measured with a validated automated BP monitor (Model HEM-907XL,

OMRON Corp, Kyoto, Japan) at baseline and all follow-up visits after 5 minutes of seated rest. At each time point, 2 measurements were taken with 1 minute between measurements. If the 2 measurements differed by >20 mm Hg (SBP) or 10 mm Hg (DBP), a third measurement was obtained, and the mean of the 2 closest measurements was used for analysis. In addition, mean arterial pressure (MAP) was calculated as MAP=DBP+⅓(SBP−DBP).

Metabolic Syndrome Severity z Score

In addition to the prespecified outcome measures, we calculated metabolic syndrome severity *z* score (MetS-Z) values for all assessment visits, as described previously.[37] In contrast to the traditional binary metabolic syndrome (MetS) classification, a continuous MetS severity score allows better detection of a worsening or improving condition over time.[37] In brief, the MetS-Z was derived from the 5 traditional components of the MetS (waist circumference, SBP, FBG, fasting HDL-C, and fasting triglycerides) using the 1999 to 2010 NHANES (National Health and Nutrition Examination Survey) data for adults 20 to 64 years of age via a factor analysis approach.[37] To take differences in MetS criteria by race/ethnicity into account, different equations for computing the MetS-Z were generated for each of 6 subgroups (non-Hispanic White, non-Hispanic Black, and Hispanic men and women).[37] The resulting MetS-Z acts as a continuous biomarker of MetS severity[38] that correlates with other established MetS risk markers such as insulin and adiponectin.[39] In addition, MetS-Z is associated with long-term type 2 diabetes[39–41] and CVD risk,[39,42,43] even when the individual components of MetS are included in the model.[40,43]

Statistical Analyses

The PROPEL trial was powered for the primary outcome (mean percent weight loss from baseline to month 24), and the total sample size provided at least 97% power to detect a mean difference of 3.5% in the primary outcome between the ILI group and the usual care group at 24 months.[26] The cardiometabolic risk factors reported herein are secondary outcomes. All outcomes were analyzed at all available time points in the context of repeated-measures linear mixed-effects multilevel models, which included random cluster (clinic) effects. Covariates in the models included study arm, assessment time point, and their interaction, as well as age, sex, and race. Furthermore, binary medication use variables for hypertension, diabetes, and high cholesterol (use versus no use) at each time point were entered into the respective models as additional covariates. Clinic-level covariates (total number of patients, percentage of participants who were Black, percentage of patients on Medicaid) were nonsignificant, and results of analyses that included these additional covariates did not differ meaningfully; therefore, the models without these clinic-level covariates are reported. We conducted intention-to-treat analyses that included all participants as randomized, regardless of the number of assessments obtained, and used restricted maximum likelihood incorporating all available data. Missing values were assumed to be missing at random. To examine the heterogeneity of treatment effects, we conducted 3 prespecified subgroup analyses (Black versus other races, women versus men, and younger [21–42 years] versus middle-aged [43–56 years] and older [57–74 years] adults), as well as 2 explorative subgroup analyses (participants with diabetes versus without diabetes and with hypertension versus without hypertension).

In additional analyses, including only participants allocated to the ILI group, we used mixed linear regression models to estimate the effect of percent weight change (the primary outcome of the PROPEL trial[26,27]) on change in cardiometabolic risk factors. The random clustering effects of clinics were taken into account. Changes in weight and cardiometabolic risk factors at all available time points were entered into the models, allowing for different slopes at each time point. The regression model of predicted cardiometabolic risk factors for the *ith* ILI participant in clinic j at time k can be expressed as $\Delta cardiovascular\ risk\ factor_{ijk} = \hat{\alpha}_k + \hat{\beta}_k \times percent\ weight\ change_{ijk} + \hat{\gamma}_j$. Furthermore, we analyzed change in cardiometabolic outcomes in the ILI group by categories of weight loss (<5%, 5%–<10%, ≥10%).

All analyses were conducted with SAS version 9.4 (SAS Institute Inc, Cary, NC) for Windows with the significance level set to 0.05 (2 sided).

RESULTS

Participants

The PROPEL trial enrolled a total of 803 (67% Black; 84% female) adults with a mean age of 49.4 years (SD, 13.1 years) and a mean body mass index of 37.2 kg/m^2 (SD, 4.7 kg/m^2) into the ILI (n=452) and the usual care (n=351) groups. Baseline characteristics are displayed in Table 1. There was a greater proportion of Black individuals, women, and participants with diabetes in the ILI group, and participants in the ILI group had significantly lower FBG, total:HDL-C, and MetS-Z values and higher HDL-C values at baseline compared with those in usual care (all $P \leq 0.03$). One hundred thirty-three participants (16.6%) were lost to follow-up at 24 months for various reasons. Eighteen clinics (9 in each group) and 803 participants (452 in ILI and 351 in usual care) were included in the primary analysis (Figure 1).

Change in Outcome Measures

As recently demonstrated,[27] the ILI group (–4.99% [95% CI, –6.02% to –3.96%]) lost more weight than the usual care group (–0.48% [95% CI, –1.57% to 0.61%]), with a mean difference of –4.51% (95% CI, –5.93% to –3.10%) between the groups ($P<0.01$). Table 2 shows the change in cardiometabolic risk factors over 2 years. FBG decreased from baseline to 12 months (–4.5 mg/dL [standard error (SE), 2.1 mg/dL]; $P=0.04$) but not 24 months (–0.8 mg/dL [SE, 2.1 mg/dL]; $P=0.70$) in the ILI group. In the usual care group, FBG did not change significantly at either time point (all $P>0.20$), leading to a significant mean difference of –7.1 mg/dL (SE, 2.4 mg/dL; $P<0.01$) at 12 months and –0.8 mg/dL (SE, 2.5 mg/

Table 1. Participant Characteristics at Baseline

	UC	ILI	All
Participants, n (%)	351 (43.7)	452 (56.3)	803
Race, n (%)			
Black	208 (59.3)	332 (73.5)*	540 (67.3)
White	113 (32.2)	95 (21.0)*	208 (25.9)
Other	30 (8.5)	25 (5.5)	55 (6.8)
Sex, n (%)			
Male	71 (20.2)	54 (11.9)*	125 (15.6)
Female	280 (79.8)	398 (88.1)*	678 (84.4)
Diabetes, n (%)	104 (29.6)	103 (22.8)*	207 (25.8)
Hypertension, n (%)	196 (55.8)	237 (52.4)	433 (53.9)
Hypercholesterolemia, n (%)	257 (73.2)	306 (67.7)	563 (70.1)
Age, y	50.2±13.6	48.8±12.7	49.4±13.1
Body weight, kg	102.7±17.0	101.6±16.4	102.1±16.7
Body mass index, kg/m^2	37.2±4.8	37.3±4.6	37.2±4.7
Waist circumference, cm†	113.9±12.6	113.1±12.4	113.4±12.5
FBG, mg/dL‡	112.3±40.2	106.4±31.9*	109.0±35.8
Total cholesterol, mg/dL§	180.0±36.7	179.6±37.5	179.8±37.1
LDL-C, mg/dL‖	106.7±31.5	105.7±32.8	106.2±32.2
HDL-C, mg/dL#	47.8±14.4	50.5±14.4*	49.3±14.4
Non–HDL-C, mg/dL**	132.0±35.6	128.5±37.1	130.1±36.4
Total:HDL-C ratio††	4.04±1.40	3.80±1.38*	3.91±1.39
Triglycerides, mg/dL‡‡	131.6±69.4	125.2±72.8	128.0±71.3
SBP, mm Hg	122.6±16.5	123.1±16.3	122.9±16.4
DBP, mm Hg	78.4±10.6	79.7±10.6	79.1±10.6
MAP, mm Hg	93.1±11.4	94.2±11.4	93.7±11.4
MetS-Z§§	1.05±1.18	0.87±0.96*	0.95±1.06

Values are mean±SD when appropriate.

DBP indicates diastolic blood pressure; FBG, fasting blood glucose; HDL-C, high-density lipoprotein cholesterol; ILI, intensive lifestyle intervention; LDL-C, low-density lipoprotein cholesterol; MAP, mean arterial pressure; MetS-Z, metabolic syndrome *z* score; SBP, systolic blood pressure; and UC, usual care.

*Significantly different from UC (P<0.05). We used χ^2 to test for differences among the groups on categorical variables and *t* test to test for differences between ILI and UC on baseline values of continuous variables.

†Data available for 349 of 351 participants (UC), 451 of 452 participants (ILI), and 800 of 803 total.

‡Data available for 344 of 351 participants (UC), 439 of 452 participants (ILI), and 783 of 803 total.

§Data available for 340 of 351 participants (UC), 434 of 452 participants (ILI), and 774 of 803 total.

‖Data available for 328 of 351 participants (UC), 402 of 452 participants (ILI), and 730 of 803 total.

#Data available for 343 of 351 participants (UC), 438 of 452 participants (ILI), and 781 of 803 total.

**Data available for 339 of 351 participants (UC), 432 of 452 participants (ILI), and 771 of 803 total.

††Data available for 339 of 351 participants (UC), 432 of 452 participants (ILI), and 771 of 803 total.

‡‡Data available for 333 of 351 participants (UC), 411 of 452 participants (ILI), and 744 of 803 total.

§§Data available for 306 of 351 participants (UC), 393 of 452 participants (ILI), and 699 of 803 total.

dL; P=0.76) at 24 months in favor of the ILI group. HDL-C increased in the ILI group by 4.7 mg/dL (SE, 0.6 mg/dL; P<0.01) at 12 months and by 4.3 mg/dL (SE, 0.7 mg/dL; P<0.01) at 24 months but did not change in the usual care group at either time point. The mean difference in HDL-C between the 2 groups was 4.1 mg/dL (SE, 0.8 mg/dL; P<0.01) at 12 months and 4.6 mg/dL (SE, 0.8 mg/dL; P<0.01) at 24 months, both in favor of the ILI group. Similarly, we found a significant mean difference in total:HDL-C ratio between the 2 groups at both time points with a reduction in the ILI group relative to the usual care group of –0.29 (SE, 0.07 mg/dL; P<0.01) at 12 months and of –0.31 (SE, 0.08 mg/dL; P<0.01) at 24 months. Furthermore, MetS-Z values were decreased in the ILI group at 12 months (–0.35 [SE, 0.06]; P<0.01) and 24 months (–0.20 [SE, 0.06]; P<0.01), whereas they did not change in the usual care group at either time point. The mean difference of the change in MetS-Z between the groups was –0.40 (SE, 0.07; P<0.01) at 12 months and –0.21 (SE, 0.07; P=0.01) at 24 months, both in favor of the ILI group. There were no significant differences in total cholesterol, LDL-C, non–HDL-C, triglycerides, SBP, DBP, or MAP between the 2 groups at any time point (all P≥0.11). We conducted 2 sensitivity analyses: (1) excluding all data for time points at which participants reported not having taken their BP or diabetes medication before the study visit (Table I in the Data Supplement) and (2) excluding all outliers (±3 SD of the change in the respective outcome) from the analysis (Table II in the Data Supplement). For BP and FBG, only those outliers who also reported not having taken their respective medication before the study visit were excluded. In general, the results for the 2 sensitivity analyses did not differ meaningfully from the main analysis (Table 2); however, in contrast to the main analysis, sensitivity analysis 2 yielded a significant increase in total cholesterol in the ILI group relative to the usual care group with a mean difference of 6.6 mg/dL (SE, 2.0 mg/dL; P<0.01) at 24 months.

Subgroup Analyses

A potential race effect was found for FBG at 12 months in that a significant difference between usual care and ILI was shown for Black participants but not for other races (Table III in the Data Supplement). Women and men responded overall similarly to the intervention; however, for HDL-C, the difference between usual care and ILI was nearly twice as large in men compared with women (both time points), and for FBG at 12 months, the difference between groups was 5 times larger in men than in women (Table IV in the Data Supplement). For total:HDL-C ratio and Mets-Z (both at 24 months), a significant difference between groups was found for women but not for men. When these sex-based

Circulation. 2021;143:1202–1214. DOI: 10.1161/CIRCULATIONAHA.120.051328

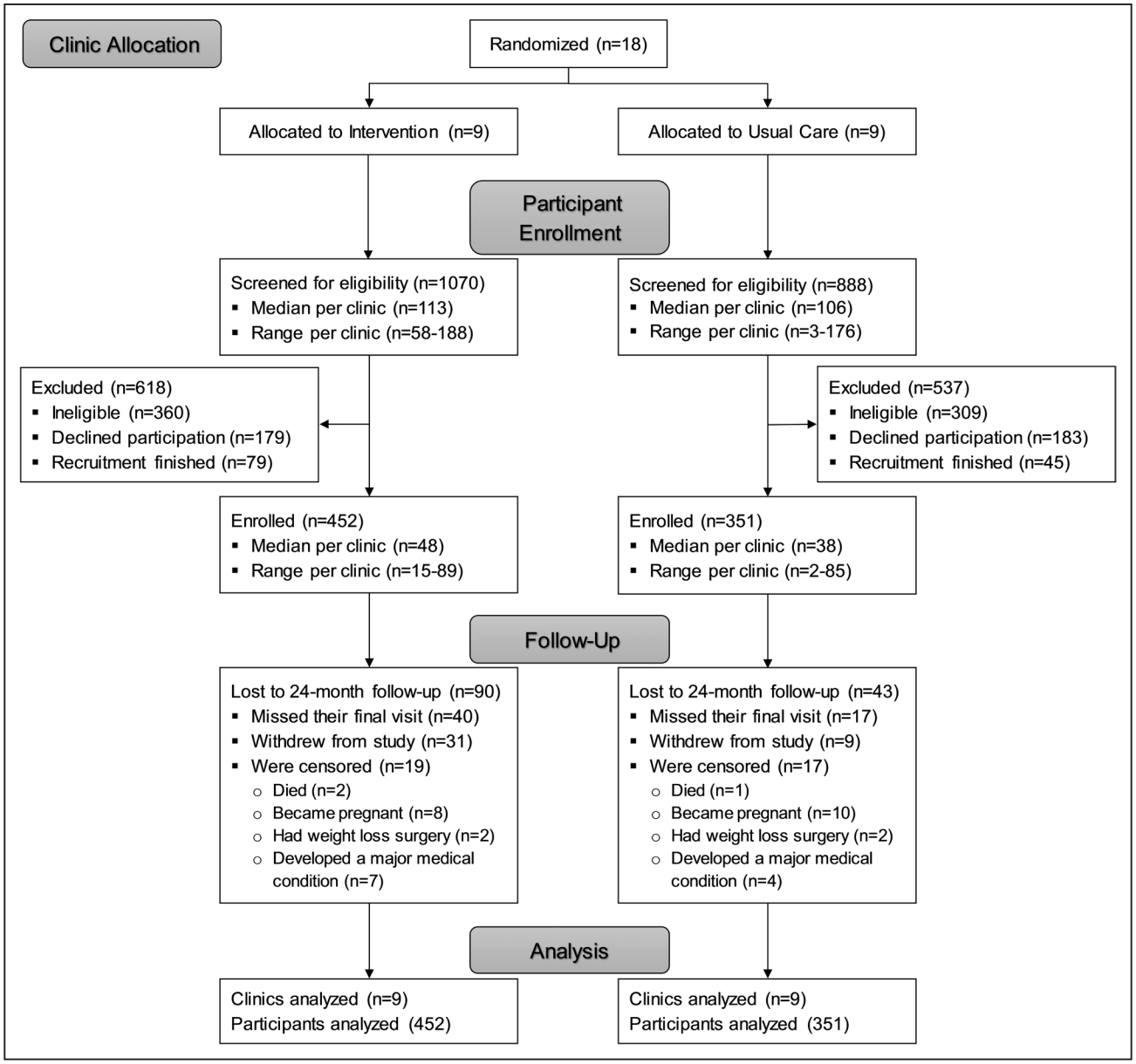


Figure 1. Participant flow through the PROPEL trial (Promoting Successful Weight Loss in Primary Care in Louisiana).

subgroup analyses are interpreted, the markedly smaller sample of men (20.2% men in usual care, 11.9% men in ILI, 15.6% men overall) should be noted, which likely contributed to the substantial variability in the data for men for all outcomes. Furthermore, there was a potential age effect for FBG with a significant reduction in ILI compared with usual care only in older adults (Table V in the Data Supplement), noting that FBG in younger participants in both groups increased substantially by ≈20 mg/dL (P<0.01 for all) at both time points (data not shown). Similarly, only older adults in ILI had decreased total:HDL-C ratio and MetS-Z significantly at both time points compared with those in usual care, showing a substantial reduction in MetS-Z compared with the other age groups. A potential diabetes-dependent treatment effect was found; only participants with diabetes showed a significant difference in FBG between usual care and ILI at 12 months (Table VI in the Data Supplement), which was likely driven by a substantial decrease of 15.6 mg/dL (SE, 7.5 mg/dL; P=0.05; data not shown) in the ILI group. At 24 months, the mean difference in FBG between usual care and ILI among those with diabetes was no longer significant. Furthermore, significant reductions in total:HDL-C ratio and MetS-Z in the ILI group compared with the usual care group at both time points were found only for individuals without diabetes. A significant difference in SBP, DBP, and MAP between usual care and ILI was found only for individuals without hypertension and only for some but not all of the time points (Table VII in the Data Supplement).

The regression models revealed significant direct associations between percent weight change and

Table 2. Change in Cardiometabolic Risk Factors Over 2 Years

	UC	ILI	Difference	*P* value
FBG, mg/dL				
At 12 mo	2.6 (−1.5 to 6.7)	−4.5 (−8.9 to −0.1)	−7.1 (−12.0 to −2.1)	<0.01*
At 24 mo	0.0 (−4.4 to 4.3)	−0.8 (−5.4 to 3.7)	−0.8 (−6.2 to 4.6)	0.76
Total cholesterol, mg/dL				
At 12 mo	0.8 (−3.2 to 4.9)	2.9 (−1.2 to 7.0)	2.0 (−3.0 to 7.0)	0.40
At 24 mo	0.8 (−3.6 to 5.1)	5.2 (0.8 to 9.5)	4.4 (−1.1 to 9.8)	0.11
LDL-C, mg/dL				
At 12 mo	1.3 (−2.2 to 4.7)	1.2 (−2.5 to 4.9)	−0.1 (−4.4 to 4.2)	0.97
At 24 mo	1.9 (−1.9 to 5.7)	3.5 (−0.4 to 7.4)	1.6 (−3.2 to 6.4)	0.49
HDL-C, mg/dL				
At 12 mo	0.6 (−0.7 to 1.9)	4.7 (3.3 to 6.0)	4.1 (2.4 to 5.7)	<0.01*
At 24 mo	−0.3 (−1.7 to 1.1)	4.3 (2.9 to 5.7)	4.6 (2.9 to 6.3)	<0.01*
Non–HDL-C, mg/dL				
At 12 mo	1.0 (−3.0 to 5.0)	−0.5 (−4.5 to 3.6)	−1.4 (−6.4 to 3.6)	0.55
At 24 mo	1.7 (−2.7 to 6.1)	1.9 (−2.5 to 6.2)	0.2 (−5.4 to 5.8)	0.95
Total:HDL-C ratio				
At 12 mo	0.01 (−0.11 to 0.13)	−0.28 (−0.41 to −0.16)	−0.29 (−0.44 to −0.14)	<0.01*
At 24 mo	0.11 (−0.03 to 0.24)	−0.20 (−0.34 to −0.06)	−0.31 (−0.47 to −0.14)	<0.01*
Triglycerides, mg/dL				
At 12 mo	−0.2 (−11.2 to 10.8)	−7.8 (−18.9 to 3.3)	−7.6 (−21.4 to 6.3)	0.26
At 24 mo	−3.6 (−14.5 to 7.4)	−9.3 (−20.2 to 1.7)	−5.7 (−19.4 to 8.0)	0.39
SBP, mm Hg				
At 6 mo	1.2 (−1.5 to 4.0)	−0.2 (−2.8 to 2.4)	−1.4 (−4.1 to 2.2)	0.42
At 12 mo	2.1 (−0.7 to 4.9)	0.4 (−2.3 to 3.0)	−1.8 (−4.4 to 2.0)	0.33
At 18 mo	1.1 (−1.8 to 4.0)	−0.2 (−2.9 to 2.5)	−1.3 (−4.1 to 2.5)	0.48
At 24 mo	0.4 (−2.5 to 3.3)	1.9 (−0.8 to 4.7)	1.6 (−1.3 to 5.3)	0.41
DBP, mm Hg				
At 6 mo	0.2 (−1.6 to 2.1)	−0.9 (−2.7 to 0.8)	−1.2 (−3.5 to 1.2)	0.32
At 12 mo	0.2 (−1.7 to 2.1)	−1.3 (−3.1 to 0.4)	−1.5 (−3.9 to 0.9)	0.21
At 18 mo	−0.7 (−2.6 to 1.1)	−1.8 (−3.6 to 0.0)	−1.1 (−3.5 to 1.4)	0.37
At 24 mo	−0.6 (−2.5 to 1.3)	−0.6 (−2.4 to 1.2)	0.0 (−2.4 to 2.5)	0.97
MAP, mm Hg				
At 6 mo	0.6 (−1.5 to 2.6)	−0.7 (−2.6 to 1.3)	−1.2 (−3.9 to 1.4)	0.35
At 12 mo	0.8 (−1.3 to 2.9)	−0.8 (−2.8 to 1.2)	−1.6 (−4.3 to 1.1)	0.24
At 18 mo	−0.1 (−2.3 to 2.0)	−1.3 (−3.3 to 0.7)	−1.2 (−3.9 to 1.6)	0.40
At 24 mo	−0.3 (−2.5 to 1.9)	0.3 (−1.8 to 2.3)	0.6 (−2.2 to 3.4)	0.69
MetS-Z				
At 12 mo	0.05 (−0.07 to 0.17)	−0.35 (−0.48 to −0.23)	−0.40 (−0.54 to −0.26)	<0.01*
At 24 mo	0.01 (−0.12 to 0.13)	−0.20 (−0.33 to −0.07)	−0.21 (−0.36 to −0.06)	0.01*

Values are mean (95% CI).

DBP indicates diastolic blood pressure; FBG, fasting blood glucose; HDL-C, high-density lipoprotein cholesterol; ILI, intensive lifestyle intervention; LDL-C, low-density lipoprotein cholesterol; MAP, mean arterial pressure; MetS-Z, metabolic syndrome *z* score; SBP, systolic blood pressure; and UC, usual care.

*Statistically significant ($P<0.05$).

change in FBG (month 12: B=0.59, SE=0.20, $P<0.01$; month 24: B=0.80, SE=0.19, $P<0.01$; Figure 2A), MetS-Z (month 12: B=0.04, SE=0.01, $P<0.01$; month 24: B=0.05, SE=0.01, $P<0.01$; Figure 2C), and total:HDL-C ratio (month 12: B=0.02, SE=0.01, $P<0.01$; month 24: B=0.03, SE=0.01, $P<0.01$; Figure 2D). Furthermore, we

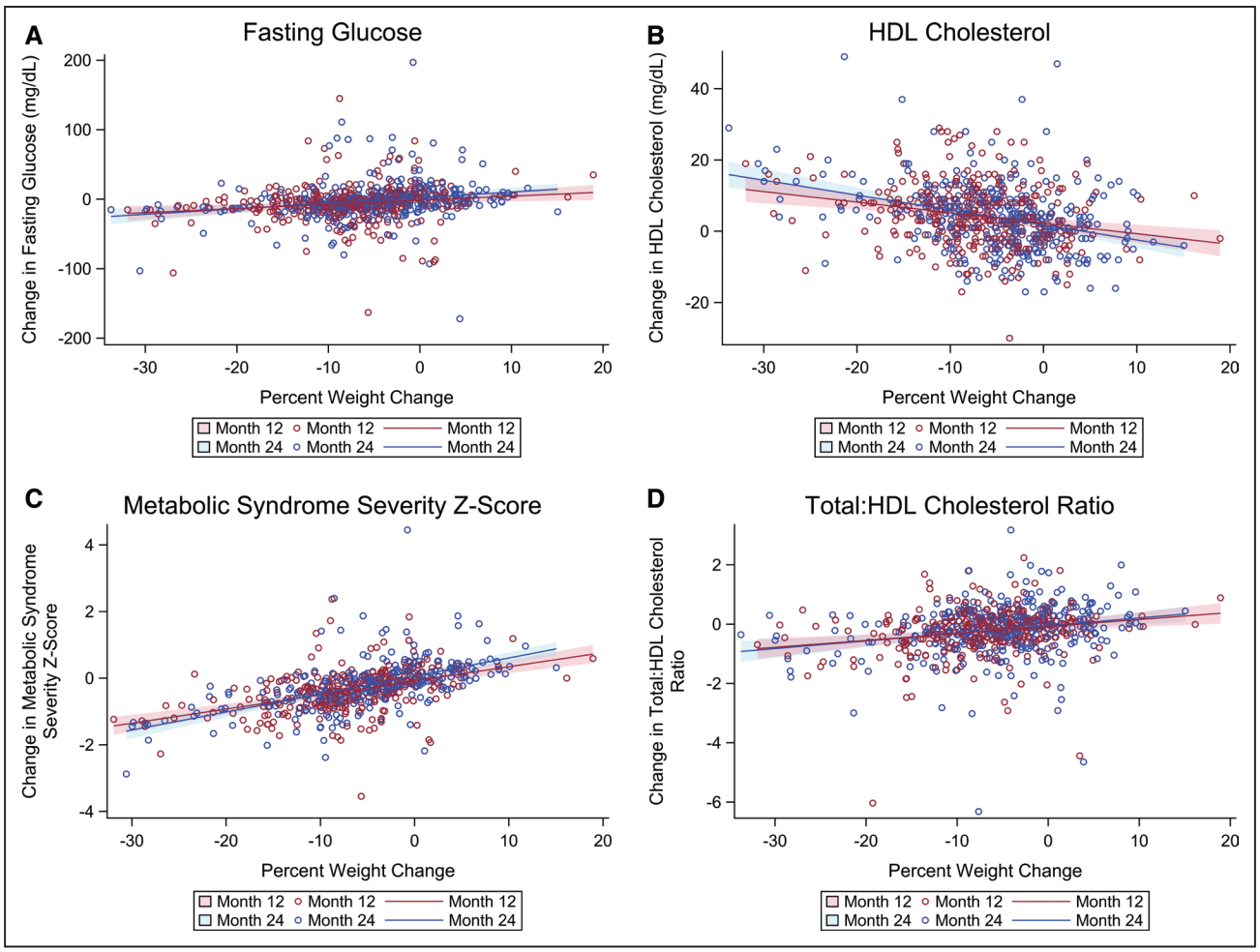


Figure 2. Association between percent weight change and change in fasting glucose levels (A), high-density lipoprotein (HDL) cholesterol (B), metabolic syndrome severity z score (C), and total:HDL cholesterol ratio (D).

Mixed linear regression models included weight change and change in the respective cardiometabolic risk factor for all available time points, producing different slopes and 95% CIs (shaded area) of the regression lines at each time point. To improve readability, the random clustering effects of the clinics are not incorporated into the graphs because this would yield 9 regression lines (1 for each clinic) for each time point.

found a significant inverse association between percent weight change and the change in HDL-C (month 12: B=–0.30, SE=0.07, *P*<0.01; month 24: B=–0.42, SE=0.06, *P*<0.01; Figure 2B). The significant associations were consistent across time. To improve readability, the random clustering effects of the clinics, yielding 9 regression lines (1 for each clinic) for each time point, are not incorporated into Figure 2. For completeness, Figure I in the Data Supplement illustrates the associations presented in Figure 2 with the inclusion of the random clustering effects of the clinics. Associations between percent weight change and changes in DBP, MAP, and triglycerides were not consistent across time, and changes in SBP, total cholesterol, LDL-C, and non–HDL-C were not predicted by percent weight change at any time point. Table 3 additionally shows the change in cardiometabolic risk factors in the ILI group over 2 years by categories of weight loss. Although participants with <5% weight loss showed no significant changes in any of the outcomes at any time point (except HDL-C at 12 months), weight loss ≥5% and particularly ≥10% was associated with significant improvements in FBG, HDL-C, total:HDL-C ratio, triglycerides, and Mets-Z, with changes in outcomes seen at ≥10% weight loss generally exceeding those at ≥5%–<10% weight loss. Effects for other outcomes were not consistent across time.

DISCUSSION

The present results show that a high-intensity lifestyle-based obesity treatment program, consistent with the *2013* AHA/ACC/TOS guidelines and delivered in primary care among an underserved population, elicits significant improvements in several cardiometabolic outcomes, highlighting the clinical relevance of the PROPEL intervention. Specifically, the PROPEL intervention led to reductions in FBG over 12 months, which were, however, not sustained over 24 months. Furthermore, the ILI group showed beneficial increases in HDL-C at 12 and

Table 3. Change in Cardiometabolic Risk Factors in the ILI Group Over 2 Years, by Categories of Weight Loss

	<5% Weight loss	5%–<10% Weight loss	≥10% Weight loss
FBG, mg/dL			
At 12 mo	−0.5 (−6.5 to 5.5)	−10.5 (−16.0 to −5.1)*	−7.2 (−11.3 to −3.0)*
At 24 mo	3.4 (−2.1 to 8.9)	−5.1 (−11.2 to 1.0)	−10.9 (−15.7 to −6.0)*
Total cholesterol, mg/dL			
At 12 mo	1.0 (−4.4 to 6.4)	5.6 (−0.4 to 11.6)	3.7 (−2.3 to 9.6)
At 24 mo	3.7 (−1.1 to 8.5)	4.7 (−1.9 to 11.4)	8.1 (1.1 to 15.0)*
LDL-C, mg/dL			
At 12 mo	−0.3 (−5.3 to 4.7)	4.6 (−0.5 to 9.7)	−0.5 (−7.4 to 6.4)
At 24 mo	3.7 (−0.7 to 8.1)	1.9 (−3.7 to 7.6)	1.0 (−6.8 to 8.9)
HDL-C, mg/dL			
At 12 mo	2.1 (0.2 to 3.9)*	4.4 (2.7 to 6.2)*	7.5 (5.7 to 9.4)*
At 24 mo	1.6 (−0.1 to 3.3)	4.3 (2.4 to 6.3)*	9.4 (7.2 to 11.6)*
Non–HDL-C, mg/dL			
At 12 mo	−0.7 (−5.6 to 4.3)	1.9 (−3.3 to 7.0)	−4.1 (−9.9 to 1.8)
At 24 mo	3.0 (−1.5 to 7.5)	0.0 (−5.8 to 5.7)	−4.0 (−11.1 to 3.0)
Total:HDL-C ratio			
At 12 mo	−0.13 (−0.30 to 0.05)	−0.18 (−0.33 to −0.03)*	−0.48 (−0.65 to −0.31)*
At 24 mo	−0.02 (−0.18 to 0.14)	−0.25 (−0.42 to −0.08)*	−0.52 (−0.72 to −0.31)*
Triglycerides, mg/dL			
At 12 mo	−1.6 (−19.6 to 16.5)	−7.8 (−18.3 to 2.7)	−19.5 (−33.5 to −5.6)*
At 24 mo	−0.7 (−17.4 to 16.0)	−12.0 (−23.4 to −0.5)*	−25.3 (−41.0 to −9.7)*
SBP, mm Hg			
At 6 mo	2.0 (−1.2 to 5.1)	−0.5 (−3.2 to 2.2)	−2.1 (−6.2 to 1.9)
At 12 mo	1.0 (−2.1 to 4.2)	0.4 (−2.4 to 3.2)	−0.6 (−4.8 to 3.5)
At 18 mo	0.2 (−2.8 to 3.2)	0.2 (−2.7 to 3.1)	−2.3 (−6.6 to 2.0)
At 24 mo	2.9 (0.0 to 5.7)	−1.1 (−4.2 to 1.9)	1.6 (−2.9 to 6.1)
DBP, mm Hg			
At 6 mo	1.2 (−0.8 to 3.2)	−1.5 (−3.5 to 0.6)	−3.0 (−5.4 to −0.6)*
At 12 mo	−0.5 (−2.5 to 1.6)	−1.5 (−3.6 to 0.6)	−2.3 (−4.8 to 0.2)
At 18 mo	−1.2 (−3.1 to 0.8)	−1.2 (−3.4 to 1.0)	−3.9 (−6.5 to −1.3)*
At 24 mo	0.6 (−1.2 to 2.5)	−2.9 (−5.2 to −0.6)*	−1.4 (−4.1 to 1.3)
MAP, mm Hg			
At 6 mo	1.4 (−0.8 to 3.7)	−1.1 (−3.3 to 1.0)	−2.7 (−5.6 to 0.1)
At 12 mo	0.0 (−2.3 to 2.3)	−0.9 (−3.1 to 1.3)	−1.8 (−4.7 to 1.2)
At 18 mo	−0.7 (−2.9 to 1.5)	−0.8 (−3.0 to 1.5)	−3.4 (−6.5 to −0.4)*
At 24 mo	1.4 (−0.7 to 3.5)	−2.3 (−4.7 to 0.0)	−0.4 (−3.6 to 2.7)
MetS-Z			
At 12 mo	−0.07 (−0.20 to 0.06)	−0.49 (−0.63 to −0.34)*	−0.67 (−0.80 to −0.54)*
At 24 mo	0.11 (−0.01 to 0.23)	−0.41 (−0.57 to −0.24)*	−0.79 (−0.94 to −0.64)*

Values are mean (95% CI).

DBP indicates diastolic blood pressure; FBG, fasting blood glucose; HDL-C, high-density lipoprotein cholesterol; ILI, intensive lifestyle intervention; LDL-C, low-density lipoprotein cholesterol; MAP, mean arterial pressure; MetS-Z, metabolic syndrome *z* score; and SBP, systolic blood pressure.

*Significant ($P<0.05$) change from baseline.

24 months compared with the usual care group that remained relatively unchanged at both time points. This increase in HDL-C likely drove the significant beneficial decrease in total:HDL-C ratio in the ILI group compared with the usual care group at both time points, particularly because total cholesterol, LDL-C, and non–HDL-C

did not change during the intervention in either group. The PROPEL intervention yielded improvements in specific disease risk over time, as demonstrated by the significant decrease in MetS-Z in the ILI group compared with the usual care group at 12 and 24 months. Participants in the ILI group with more weight loss during the 24-month intervention period showed greater improvements in FBG, HDL-C, total:HDL-C ratio, triglycerides, and MetS-Z than those with less weight loss, as shown by the regression model and the analysis of change in cardiometabolic outcomes by weight loss category.

Similar to our results, the primary care–based behavioral lifestyle intervention of the POWER-UP trial (Improving the Management of Obesity in Primary Care Practice), which showed a mean weight loss of 4.7% at 24 months (PROPEL, 5.0% at 24 months[27]), found a significant difference in FBG between the enhanced lifestyle intervention group and the control group at 12 but not 24 months.[23] However, POWER-UP found a decrease in FBG in the intervention group at both time points, whereas we saw significant reductions in the ILI group at 12 months only. FBG in the control groups of both trials remained relatively stable throughout the entire study.

The lack of change in FBG from baseline to 24 months as observed for the ILI group in PROPEL is comparable to the DPP (mean weight loss of 5.8% at 24 months[13]), which was conducted in academic health centers. DPP likewise did not find changes in FBG from baseline to 24 months in the lifestyle intervention group (baseline values, ≈106 mg/dL, comparable to PROPEL); however, the DPP found a significant mean difference of ≈5 mg/dL between the groups at 24 months.[15] Look AHEAD, also conducted in academic health centers (mean weight loss, 6.4% at 24 months[14]), found decreases in FBG at 12 months in both the lifestyle intervention group and control group, with a significantly greater decrease in the intervention group (mean difference, ≈14 mg/dL), however.[44] It is noteworthy that Look AHEAD specifically enrolled individuals with type 2 diabetes who were ≈10 years older than PROPEL participants on average. Further, improvements in HDL-C in the ILI group relative to the usual care group in PROPEL were approximately twice as large compared with the observed increases in Look AHEAD at 12 and 24 months (≈2 mg/dL for both time points)[17,44]; the DPP and POWER-UP trials did not find a significant difference between the intervention and respective control groups at any time point.[16,23] The improvements in HDL-C are of particular interest because HDL-C is less affected by statins and nonstatin cholesterol-lowering drugs than LDL-C and all apolipoprotein B-containing lipoproteins,[45,46] and changes in HDL-C are consequently more likely to be attributable to the ILI program.[47] Although these results are undeniably positive, it has to be emphasized that HDL-C is more a marker than a mediator of cardiovascular risk. Although higher HDL-C is generally strongly associated with a lower CVD risk in epidemiological studies, in intervention trials, increases in HDL-C through pharmacological means or because of weight loss or exercise do not consistently lead to improvements in hard CVD end points.[48,49] It has been suggested that the total:HDL-C ratio may be a superior predictor of CVD event risk compared with classic lipid parameters,[50,51] and the beneficial decreases in this important cumulative index of the atherogenic risk in our study therefore underline the clinical relevance of the present findings. The significant reductions in MetS-Z at both time points in the ILI group compared with the usual care group additionally demonstrate the cardiometabolic risk reduction that was achieved by the high-intensity lifestyle intervention in PROPEL. The effect was comparable, albeit slightly attenuated, compared with the lifestyle intervention group in DPP, which showed decreases of 0.40 (PROPEL, 0.35) and 0.31 (PROPEL, 0.20) at 12 and 24 months, respectively.[38] Decreases of the magnitude reported for the DPP lifestyle intervention group are associated with a significantly reduced 5-year risk of diabetes (hazard ratio, 0.57),[38] and a similar or slightly blunted protective effect can consequently be assumed for the PROPEL ILI group. Although the demonstrated improvements in cardiometabolic risk factors speak to the success of the PROPEL ILI program and sustained improvements in these risk factors are likely beneficial for long-term CVD risk, it has to be acknowledged that, to date, defacto reductions in CVD events through ILIs have not yet been shown.[16,17]

It is further noteworthy that despite the clinically relevant weight loss and accompanying improvements in metabolic parameters in PROPEL, there were no significant improvements in BP (SBP, DPB, or MAP) at any time point. This is different from large trials conducted in academic health centers such as Look AHEAD (average baseline BP, 129/70 mm Hg)[17,44] and DPP (average baseline BP, 124/78 mm Hg),[16] which both showed significant reductions in SBP and DBP between the lifestyle intervention group and control group at 12 and 24 months. Similar to our results, POWER-UP (average baseline BP, 121/76 mm Hg), which likewise was conducted in primary care clinics, did not find significant changes in BP between the intervention group and the respective control group. The specific reason for this lack of effect in PROPEL is of course conjecture; however, >50% of participants were prescribed antihypertensive medication at baseline and throughout the trial. The concurrent drug treatment, along with the relatively normal mean BP values at baseline (123/80 mm Hg in ILI, 123/78 mm Hg in usual care), indicating a predominantly well-controlled BP, possibly masked any BP-lowering effect achieved by the PROPEL intervention. It is further surprising that we did not find any improvements in LDL-C and non–HDL-C after the

24-month intervention, especially because ≈70% of participants had hypercholesterolemia at baseline. Akin to BP, however, a substantial proportion of participants (≈30%) was prescribed cholesterol-lowering medication at baseline and throughout the trial. This likely contributed to relatively stable values in all apolipoprotein B–containing lipoproteins that may have overshadowed any potentially cholesterol-lowering effect elicited by the PROPEL intervention.

There were no race differences in the change in cardiometabolic risk factors over 24 months. This is interesting because the main outcomes article showed less weight loss in Black individuals compared with individuals of other races,[27] and an attenuated response in cardiometabolic risk factors would consequently be conceivable. Similarly, in this pragmatic trial, there were no notable differences in outcomes between women and men, participants with diabetes and without diabetes, or participants with hypertension and without hypertension over 24 months.

We did, however, find an age effect for FBG and Mets-Z with significant improvements in the ILI group compared with the usual care group at both time points only in older adults. This is congruent with findings from DPP[52] and suggests that the effectiveness of the ILI of the trial in preventing the deterioration of glucose tolerance is enhanced in older adults (≥60 years).

There is a gap between obesity management guidelines and what is currently implemented in primary care. Treatment models based on the AHA/ACC/TOS guidelines adapted to real-life settings that add effective delivery methods for obesity treatment in primary care are needed. Results from our trial demonstrate that significant weight loss in primary care is possible by the addition of a health coach to the collaborative care team. The PROPEL approach is scalable and likely achievable in most primary care settings. However, for broad implementation of such approaches and to allow underserved populations in particular to receive effective obesity and concomitant disease risk treatment more easily, the Centers for Medicare & Medicaid Services reimbursement regulations need to be extended to include coverage of health coach–delivered obesity treatment in primary care, as used in PROPEL. To date, the Centers for Medicare & Medicaid Services reimburses intensive behavioral therapy for obesity only if delivered by a primary care practitioner in the context of a hospital, clinic, or physician's office,[33] making the implementation of approaches to reduce health inequities unnecessarily difficult.

Strengths and Limitations

A major strength of PROPEL is its sample, consisting of a racially diverse, low-income population that typically lacks access to effective weight loss treatment in clinical research or primary care. The minimal inclusion/exclusion criteria of the trial allow broad generalizability to other underserved populations across the United States. Furthermore, the present results underline the effectiveness of a comprehensive and scalable weight loss and cardiometabolic risk factor treatment model that applies to many primary care settings. Last, the cluster-randomized design of the trial minimized contamination effects between the 2 study arms. A limitation of the trial, as is often the case in lifestyle interventions,[53] is that the sample was mostly women, which limits the generalizability of the present results for both sexes. A further limitation is that although the Cholestech LDX Analyzer measures glucose, total cholesterol, HDL-C, and triglycerides, the LDL-C levels are calculated from the total cholesterol, HDL-C, and triglyceride test results. To address this limitation, we calculated non–HDL-C, which represents the cholesterol concentration of all atherogenic lipoproteins.[36] Although we accounted for BP, glucose, and cholesterol medication use (use versus no use) at each time point, we were unable to measure changes in dose over time and medication adherence. These shortcomings may have influenced our results.

Conclusions

A pragmatic high-intensity lifestyle-based obesity treatment program, consistent with the 2013 AHA/ACC/TOS guidelines and delivered by trained health coaches in primary care, yielded significant improvements in several cardiometabolic risk markers in an underserved population over 24 months, which could translate into long-term reduction in CVD risk.

ARTICLE INFORMATION

Received September 3, 2020; accepted December 7, 2020.

The Data Supplement is available with this article at https://www.ahajournals.org/doi/suppl/10.1161/circulationaha.120.051328.

Correspondence

Peter T. Katzmarzyk, PhD, Pennington Biomedical Research Center, 6400 Perkins Road, Baton Rouge, LA 70808. Email peter.katzmarzyk@pbrc.edu

Affiliations

Pennington Biomedical Research Center, Baton Rouge, LA (C.H., J.L.D., C.K.M., R.L.N., J.W.A., C.A.M., K.D.D., E.F.M., W.D.J., D.Z., P.T.K.). Department of Medicine and Feist-Weiller Cancer Center, Louisiana State University Health Sciences Center, Shreveport (C.L.A., T.C.D.). Department of Medicine, Section of Endocrinology, Tulane University Health Sciences Center, New Orleans, LA (V.F.). Department of Cardiovascular Diseases, John Ochsner Heart and Vascular Institute, New Orleans, LA (C.J.L.). Ochsner Clinic Foundation, Center for Outcomes and Health Services Research, New Orleans, LA (E.G.P.-H.).

Acknowledgments

The authors acknowledge the patients and stakeholders who are members of the PROPEL Patient Advisory Boards, Community Monitoring Board, and the Project Management Committee (Chris Lodge and Ava Zebrick, MSHCM), who significantly affected the trial design and conduct. The authors are also

indebted to the PROPEL patients, assessment technicians, and health coaches, without whom this study would not have been possible. The authors gratefully acknowledge the contributions of Willie C. White III, MPH, and the David Raines Community Health Centers; Dr Gary Wiltz and the Teche Action Clinic sites; Michael G. Griffin, Dr Robert Post, and the Daughters of Charity Services of New Orleans; and the Ochsner Health System and Access Health Louisiana clinic sites. The authors also thank the Data and Safety Monitoring Board for monitoring patient safety and the overall conduct of the trial: Robert Ross, PhD (chair), John Lefante, PhD, Michael Rolfsen, MD, and Chris Lodge. The PROPEL Research Group includes: Pennington Biomedical Research Center: Peter T. Katzmarzyk, PhD (principal investigator), Robert L. Newton Jr, PhD (outcomes assessment director), Corby K. Martin, PhD (intervention director), John W. Apolzan, PhD (intervention codirector), William Johnson, PhD (biostatistician), Kara D. Denstel, MPH (project manager), Emily F. Mire, MS (data manager), Robert K. Singletary Jr, MHS, Cheryl Lewis, MPH, Phillip Brantley, PhD, Ronald Horswell, PhD, Betty Kennedy, PhD, Dachuan Zhang, MAppStats, Stephanie Authement, RD, LDN, MS, Shiquita Brooks, RDN, LDN, Danielle S. Burrell, MEd, MCHES, Leslie Forest-Everage, MA, Angelle Graham Ullmer, RDN, LDN, MS, Laurie Murphy, RDN, LDN, Cristalyn Reynolds, MA, Kevin Sanders, MS, RDN, LDN, Stephen Bower, MS, Daishaun Gabriel, MHA, Hillary Gahagan, MPH, Tabitha K. Gray, MA, Jill Hancock, MPH, Marsha Herrera, Brittany Molinere, Georgia Morgan, MA, Brittany Neyland, Stephanie Rincones, Deanna Robertson, MA, Ekambi Shelton, MPH, Russell J. Tassin, MS, Kaili Williams; Louisiana State University Health Sciences Center at New Orleans: Benjamin F. Springgate, MD; Louisiana State University Health Sciences Center at Shreveport: Terry C. Davis, PhD, Connie L. Arnold, PhD; Ochsner Health System: Eboni Price-Haywood, MD, Carl J. Lavie, MD, Jewel Harden-Barrios, MEd; Tulane University Medical School: Vivian A. Fonseca, MD, Tina K. Thethi, MD (MEDICAL Monitor), Jonathan Gugel, MD; Xavier University: Kathleen B. Kennedy, PhD, Daniel F. Sarpong, PhD, Amina D. Massey. Study data were collected and managed with REDCap (Research Electronic Data Capture) electronic data capture tools hosted at Pennington Biomedical. REDCap is a secure, web-based application designed to support data capture for research studies, providing an intuitive interface for validated data entry, audit trails for tracking data manipulation and export procedures, automated export procedures for seamless data downloads to common statistical packages, and procedures for importing data from external sources.

Sources of Funding

Research reported in this article was funded through a Patient-Centered Outcomes Research Institute award (OB-1402-10977). The statements in this article are solely the responsibility of the authors and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute, its Board of Governors, or its Methodology Committee. Additional support was provided by 1 U54 GM104940 from the National Institute of General Medical Sciences of the National Institutes of Health, which funds the Louisiana Clinical and Translational Science Center, and Nutrition Obesity Research Centers grant P30DK072476 titled "Nutrition and Metabolic Health Through the Lifespan" sponsored by National Institute of Diabetes and Digestive and Kidney Diseases. Dr Höchsmann receives funding through an National Institutes of Health National Institute of Diabetes and Digestive and Kidney Diseases National Research Service Award (T32DK064584). Dr Dorling is funded by the American Heart Association (grant 20POST35210907). The authors also thank Health and Nutrition Technology (Carmel, CA) for providing the HealthOne formula and Nutrisystem (Fort Washington, PA) for providing the meal replacements used in the study.

Disclosures

Louisiana State University, Pennington Biomedical Research Center, and Montclair State University have interest in the intellectual property surrounding the weight graph that was used in the intervention, and Dr Martin, among others, is an inventor of the technology. Licensing of that technology results in financial benefits to Louisiana State University, Pennington Biomedical Research Center, Montclair State University, and the inventors. Dr Martin further reports grants from WW, Egg Board, Academy of Nutrition and Dietetics, and ABGIL, as well as personal fees from NaturallySlim and ACAP, Health WW, Zafgen, Florida Hospital, Metagenic, Gila Therapeutics, Academy of Nutrition and Dietetics, ABGIL, Adaptic, EHE health, and OpenFit. Dr Fonseca reports grants from Bayer, Boehringer Ingelheim, and Gilead; personal fees from Takeda, Novo Nordisk, Sanofi-Aventis, Eli Lilly, Abbott, AstraZeneca, and Intarcia; and stock options from Microbiome Technologies, Insulin Algorithms, BRAVO4Health, and Amgen. Dr Lavie reports personal fees from Amgen, AstraZeneca, Esperion, Sanofi, and Regeneron. The work related to this article occurred before Dr Price-Haywood's appointment to the Patient-Centered Outcomes Research Institute Board of Governors. The other authors report no conflicts.

Supplemental Materials

Supplemental Tables I–VII
Supplemental Figure I

REFERENCES

1. Hales CM, Carroll MD, Fryar CD, Ogden CL. *Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017–2018.* National Center for Health Statistics; 2020. NCHS Data Brief, no. 360.
2. Kopelman P. Health risks associated with overweight and obesity. *Obes Rev.* 2007;8(suppl 1):13–17. doi: 10.1111/j.1467-789X.2007.00311.x
3. Tremmel M, Gerdtham U-G, Nilsson PM, Saha S. Economic burden of obesity: a systematic literature review. *Int J Environ Res Public Health.* 2017;14:435. doi: 10.3390/ijerph14040435
4. Ogden CL. Prevalence of obesity among adults, by household income and education–United States, 2011–2014. *MMWR Morb Mortal Wkly Rep.* 2017;66:1369–1373. doi: 10.15585/mmwr.mm6650a1
5. Seligman HK, Laraia BA, Kushel MB. Food insecurity is associated with chronic disease among low-income NHANES participants. *J Nutr.* 2010;140:304–310. doi: 10.3945/jn.109.112573
6. Havranek EP, Mujahid MS, Barr DA, Blair IV, Cohen MS, Cruz-Flores S, Davey-Smith G, Dennison-Himmelfarb CR, Lauer MS, Lockwood DW, et al; American Heart Association Council on Quality of Care and Outcomes Research, Council on Epidemiology and Prevention, Council on Cardiovascular and Stroke Nursing, Council on Lifestyle and Cardiometabolic Health, and Stroke Council. Social determinants of risk and outcomes for cardiovascular disease: a scientific statement from the American Heart Association. *Circulation.* 2015;132:873–898. doi: 10.1161/CIR.0000000000000228
7. Kumanyika SK. A framework for increasing equity impact in obesity prevention. *Am J Public Health.* 2019;109:1350–1357. doi: 10.2105/AJPH.2019.305221.
8. Muncan B. Cardiovascular disease in racial/ethnic minority populations: illness burden and overview of community-based interventions. *Public Health Rev.* 2018;39:32. doi: 10.1186/s40985-018-0109-4
9. Guzman G. *Household income: 2018: American community survey briefs*. US Census Bureau. 2019. Accessed March 31, 2020. https://www.census.gov/content/dam/Census/library/publications/2019/acs/acsbr18-01.pdf
10. Coleman-Jensen A, Rabbitt MP, Gregory CA, Singh A. Household food security in the United States in 2017: ERR-256. US Department of Agriculture. 2018. Accessed March 31, 2020. https://www.ers.usda.gov/webdocs/publications/90023/err-256.pdf
11. Jensen MD, Ryan DH. New obesity guidelines: promise and potential. *JAMA.* 2014;311:23–24. doi: 10.1001/jama.2013.282546
12. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, Hu FB, Hubbard VS, Jakicic JM, Kushner RF, et al; American College of Cardiology/American Heart Association Task Force on Practice Guidelines; Obesity Society. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. *Circulation.* 2014;129(suppl 2):S102–S138. doi: 10.1161/01.cir.0000437739.71477.ee
13. Hamman RF, Wing RR, Edelstein SL, Lachin JM, Bray GA, Delahanty L, Hoskin M, Kriska AM, Mayer-Davis EJ, Pi-Sunyer X, et al. Effect of weight loss with lifestyle intervention on risk of diabetes. *Diabetes Care.* 2006;29:2102–2107. doi: 10.2337/dc06-0560
14. Wadden TA, Neiberg RH, Wing RR, Clark JM, Delahanty LM, Hill JO, Krakoff J, Otto A, Ryan DH, Vitolins MZ; Look AHEAD Research Group. Four-year weight losses in the Look AHEAD study: factors associated with long-term success. *Obesity (Silver Spring).* 2011;19:1987–1998. doi: 10.1038/oby.2011.230
15. The DPP Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med.* 2002;346:393–403. doi: 10.1056/NEJMoa012512.
16. Orchard TJ, Temprosa M, Barrett-Connor E, Fowler S, Goldberg R, Mather K, Marcovina S, Montez M, Ratner R, Saudek C, et al. Long-term effects of the Diabetes Prevention Program interventions on cardiovascular risk factors: a report from the DPP Outcomes Study. *Diabet Med.* 2013;30:46–55.doi: 10.1111/j.1464-5491.2012.03750.x.

17. Look AHEAD Research Group, Wing RR, Bolin P, Brancati FL, Bray GA, Clark JM, Coday M, Crow RS, Curtis JM, Egan CM, Espeland MA, et al. Cardiovascular effects of intensive lifestyle intervention in type 2 diabetes. *N Engl J Med.* 2013;369:145–154. doi: 10.1056/NEJMoa1212914.
18. US Department of Health and Human Services. Healthy People 2020 Framework: The Vision, Mission, and Goals of Healthy People 2020. 2011. Accessed May 3, 2019. https:health.gov/healthypeople
19. Moyer VA, U.S. Preventive Services Task Force. Screening for and management of obesity in adults: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med.* 2012;157:373–378. doi: 10.7326/0003-4819-157-5-201209040-00475
20. Carvajal R, Wadden TA, Tsai AG, Peck K, Moran CH. Managing obesity in primary care practice: a narrative review. *Ann NY Acad Sci.* 2013;1281:191–206. doi: 10.1111/nyas.12004
21. Pagoto SL, Pbert L, Emmons K; Society of Behavioral Medicine Public Policy Leadership Group. The Society of Behavioral Medicine position statement on the CMS decision memo on intensive behavior therapy for obesity. *Transl Behav Med.* 2012;2:381–383. doi: 10.1007/s13142-012-0168-x
22. Tronieri JS, Wadden TA, Chao AM, Tsai AG. Primary care interventions for obesity: review of the evidence. *Curr Obes Rep.* 2019;8:128–136. doi: 10.1007/s13679-019-00341-5
23. Wadden TA, Volger S, Sarwer DB, Vetter ML, Tsai AG, Berkowitz RI, Kumanyika S, Schmitz KH, Diewald LK, Barg R, et al. A two-year randomized trial of obesity treatment in primary care practice. *N Engl J Med.* 2011;365:1969–1979. doi: 10.1056/NEJMoa1109220
24. Appel LJ, Clark JM, Yeh HC, Wang NY, Coughlin JW, Daumit G, Miller ER 3rd, Dalcin A, Jerome GJ, Geller S, et al. Comparative effectiveness of weight-loss interventions in clinical practice. *N Engl J Med.* 2011;365:1959–1968. doi: 10.1056/NEJMoa1108660
25. Bennett GG, Warner ET, Glasgow RE, Askew S, Goldman J, Ritzwoller DP, Emmons KM, Rosner BA, Colditz GA; Be Fit, Be Well Study Investigators. Obesity treatment for socioeconomically disadvantaged patients in primary care practice. *Arch Intern Med.* 2012;172:565–574. doi: 10.1001/archinternmed.2012.1
26. Katzmarzyk PT, Martin CK, Newton RL Jr, Apolzan JW, Arnold CL, Davis TC, Denstel KD, Mire EF, Thethi TK, Brantley PJ, et al. Promoting Successful Weight Loss in Primary Care in Louisiana (PROPEL): rationale, design and baseline characteristics. *Contemp Clin Trials.* 2018;67:1–10. doi: 10.1016/j.cct.2018.02.002
27. Katzmarzyk PT, Martin CK, Newton RL Jr, Apolzan JW, Arnold CL, Davis TC, Price-Haywood EG, Denstel KD, Mire EF, Thethi TK, et al. Weight loss in underserved patients: a cluster-randomized trial. *N Engl J Med.* 2020;383:909–918. doi: 10.1056/NEJMoa2007448
28. Diabetes Prevention Program (DPP) Research Group. The Diabetes Prevention Program (DPP): description of lifestyle intervention. *Diabetes Care.* 2002;25:2165–2171. doi:10.2337/diacare.25.12.2165
29. Look AHEAD Research Group, Wadden TA, West DS, Delahanty L, Jakicic J, Rejeski J, Williamson D, Berkowitz RI, Kelley DE, Tomchee C, Hill JO, et al. The Look AHEAD study: a description of the lifestyle intervention and the evidence supporting it. *Obesity (Silver Spring).* 2006;14:737–752. doi: 10.1038/oby.2006.84
30. Rickman AD, Williamson DA, Martin CK, Gilhooly CH, Stein RI, Bales CW, Roberts S, Das SK. The CALERIE Study: design and methods of an innovative 25% caloric restriction intervention. *Contemp Clin Trials.* 2011;32:874–881. doi: 10.1016/j.cct.2011.07.002
31. Martin CK, Miller AC, Thomas DM, Champagne CM, Han H, Church T. Efficacy of SmartLoss, a smartphone-based weight loss intervention: results from a randomized controlled trial. *Obesity (Silver Spring).* 2015;23:935–942. doi: 10.1002/oby.21063
32. Martin CK, Gilmore LA, Apolzan JW, Myers CA, Thomas DM, Redman LM. Smartloss: a personalized mobile health intervention for weight management and health promotion. *JMIR Mhealth Uhealth.* 2016;4:e18. doi: 10.2196/mhealth.5027
33. Centers for Medicare & Medicaid Services. Decision memo for intensive behavioral therapy for obesity (CAG-00423N). Accessed March 26, 2019. https://www.cms.gov/medicare-coverage-database/details/nca-decision-memo.aspx?&NcaName=Intensive%20Behavioral%20Therapy%20for%20Obesity&bc=ACAAAAAAIAAA&NCAId=253&
34. Harper CR, Jacobson TA. Using apolipoprotein B to manage dyslipidemic patients: time for a change? *Mayo Clin Proc.* 2010;85:440–445. doi: 10.4065/mcp.2009.0517
35. Lemieux I, Lamarche B, Couillard C, Pascot A, Cantin B, Bergeron J, Dagenais GR, Després JP. Total cholesterol/HDL cholesterol ratio vs LDL cholesterol/HDL cholesterol ratio as indices of ischemic heart disease risk in men: the Quebec Cardiovascular Study. *Arch Intern Med.* 2001;161:2685–2692. doi: 10.1001/archinte.161.22.2685
36. Brunner FJ, Waldeyer C, Ojeda F, Salomaa V, Kee F, Sans S, Thorand B, Giampaoli S, Brambilla P, Tunstall-Pedoe H, et al; Multinational Cardiovascular Risk Consortium. Application of non-HDL cholesterol for population-based cardiovascular risk stratification: results from the Multinational Cardiovascular Risk Consortium. *Lancet.* 2019;394:2173–2183. doi: 10.1016/S0140-6736(19)32519-X
37. Gurka MJ, Lilly CL, Oliver MN, DeBoer MD. An examination of sex and racial/ethnic differences in the metabolic syndrome among adults: a confirmatory factor analysis and a resulting continuous severity score. *Metabolism.* 2014;63:218–225. doi: 10.1016/j.metabol.2013.10.006
38. DeBoer MD, Filipp SL, Gurka MJ. Use of a metabolic syndrome severity Z Score to track risk during treatment of prediabetes: an analysis of the Diabetes Prevention Program. *Diabetes Care.* 2018;41:2421–2430. doi: 10.2337/dc18-1079
39. DeBoer MD, Gurka MJ, Morrison JA, Woo JG. Inter-relationships between the severity of metabolic syndrome, insulin and adiponectin and their relationship to future type 2 diabetes and cardiovascular disease. *Int J Obes (Lond).* 2016;40:1353–1359. doi: 10.1038/ijo.2016.81
40. Gurka MJ, Golden SH, Musani SK, Sims M, Vishnu A, Guo Y, Cardel M, Pearson TA, DeBoer MD. Independent associations between a metabolic syndrome severity score and future diabetes by sex and race: the Atherosclerosis Risk In Communities Study and Jackson Heart Study. *Diabetologia.* 2017;60:1261–1270. doi: 10.1007/s00125-017-4267-6
41. DeBoer MD, Gurka MJ, Woo JG, Morrison JA. Severity of the metabolic syndrome as a predictor of type 2 diabetes between childhood and adulthood: the Princeton Lipid Research Cohort Study. *Diabetologia.* 2015;58:2745–2752. doi: 10.1007/s00125-015-3759-5
42. DeBoer MD, Gurka MJ, Woo JG, Morrison JA. Severity of metabolic syndrome as a predictor of cardiovascular disease between childhood and adulthood: the Princeton Lipid Research Cohort Study. *J Am Coll Cardiol.* 2015;66:755–757. doi: 10.1016/j.jacc.2015.05.061
43. DeBoer MD, Gurka MJ, Golden SH, Musani SK, Sims M, Vishnu A, Guo Y, Pearson TA. Independent associations between metabolic syndrome severity and future coronary heart disease by sex and race. *J Am Coll Cardiol.* 2017;69:1204–1205. doi: 10.1016/j.jacc.2016.10.088
44. Look AHEAD Research Group, Pi-Sunyer X, Blackburn G, Brancati FL, Bray GA, Bright R, Clark JM, Curtis JM, Espeland MA, Foreyt JP, Graves K, et al. Reduction in weight and cardiovascular disease risk factors in individuals with type 2 diabetes: one-year results of the look AHEAD trial. *Diabetes Care.* 2007;30:1374–1383. doi: 10.2337/dc07-0048.
45. Blaha MJ, Martin SS. How do statins work? Changing paradigms with implications for statin allocation. *J Am Coll Cardiol.* 2013;62:2392–2394. doi: 10.1016/j.jacc.2013.08.1626
46. Chapman MJ. Are the effects of statins on HDL-cholesterol clinically relevant? *Eur Heart J Suppl.* 2004;6:C58–C63. doi: 10.1016/j.ehjsup.2004.04.002.
47. Escolà-Gil JC, Julve J, Griffin BA, Freeman D, Blanco-Vaca F. HDL and lifestyle interventions. *Handb Exp Pharmacol.* 2015;224:569–592. doi: 10.1007/978-3-319-09665-0_18
48. Rader DJ, Hovingh GK. HDL and cardiovascular disease. *Lancet.* 2014;384:618–625. doi: 10.1016/S0140-6736(14)61217-4
49. Rader DJ. Mediterranean approach to improving high-density lipoprotein function. *Circulation.* 2017;135:644–647. doi: 10.1161/CIRCULATIONAHA.117.026278
50. Shai I, Rimm EB, Hankinson SE, Curhan G, Manson JE, Rifai N, Stampfer MJ, Ma J. Multivariate assessment of lipid parameters as predictors of coronary heart disease among postmenopausal women: potential implications for clinical guidelines. *Circulation.* 2004;110:2824–2830. doi: 10.1161/01.CIR.0000146339.57154.9B
51. Millán J, Pintó X, Muñoz A, Zúñiga M, Rubiés-Prat J, Pallardo LF, Masana L, Mangas A, Hernández-Mijares A, González-Santos P, et al. Lipoprotein ratios: physiological significance and clinical usefulness in cardiovascular prevention. *Vasc Health Risk Manag.* 2009;5:757–765. doi: 10.2147/VHRM.S6269
52. Crandall J, Schade D, Ma Y, Fujimoto WY, Barrett-Connor E, Fowler S, Dagogo-Jack S, Andres R. The influence of age on the effects of lifestyle modification and metformin in prevention of diabetes. *J Gerontol A Biol Sci Med Sci.* 2006;61:1075–1081. doi: 10.1093/Gerona/61.10.1075
53. Pagoto SL, Schneider KL, Oleski JL, Luciani JM, Bodenlos JS, Whited MC. Male inclusion in randomized controlled trials of lifestyle weight loss interventions. *Obesity (Silver Spring).* 2012;20:1234–1239. doi: 10.1038/oby.2011.140

Publication 12

Exercise-induced changes in central adiposity during a RCT: Effect of exercise dose and associations with compensation.

James L Dorling [1], John W Apolzan [2], Neil M Johannsen [2,3], Diana M Thomas [4], **Christoph Höchsmann** [5], Daniel S Hsia [2], Corby K Martin [2]

[1] Human Nutrition, School of Medicine, Dentistry and Nursing, University of Glasgow, Glasgow, UK

[2] Pennington Biomedical Research Center, Baton Rouge, LA, USA

[3] School of Kinesiology, Louisiana State University, Baton Rouge, LA, USA

[4] Department of Mathematical Sciences, United States Military Academy, West Point, NY, USA

[5] TUM School of Medicine and Health, Technical University of Munich, Munich, Germany

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
Journal of Clinical Endocrinology & Metabolism, 109(3): e997–e1005	2024	10.1210/clinem/dgad696	5.8 (2022)	Q1	Scopus: 0 Google Scholar: 0

Author contribution:

Co-author; co-developed the research question and statistical model, contributed to interpreting the results, and provided critical revisions for important intellectual content.

The Journal of Clinical Endocrinology & Metabolism, 2024, **109**, e997–e1005
https://doi.org/10.1210/clinem/dgad696
Advance access publication 29 November 2023
Clinical Research Article

Exercise-induced Changes in Central Adiposity During an RCT: Effect of Exercise Dose and Associations With Compensation

James L. Dorling,[1] John W. Apolzan,[2] Neil M. Johannsen,[2,3] Diana M. Thomas,[4] Christoph Höchsmann,[5] Daniel S. Hsia,[2] and Corby K. Martin[2]

[1]Human Nutrition, School of Medicine, Dentistry and Nursing, College of Medical, Veterinary and Life of Sciences, University of Glasgow, Glasgow G31 2ER, UK
[2]Pennington Biomedical Research Center, Baton Rouge, LA 70808, USA
[3]School of Kinesiology, Louisiana State University, Baton Rouge, LA 70803, USA
[4]Department of Mathematical Sciences, United States Military Academy, West Point, NY 10996, USA
[5]TUM School of Medicine and Health, Technical University of Munich, Munich 80809, Germany

Correspondence: James L. Dorling, PhD, University of Glasgow, New Lister Building, Glasgow Royal Infirmary, 10-16 Alexandra Parade, Glasgow G31 2ER, UK. Email: James.Dorling@glasgow.ac.uk.

Abstract

Context: Exercise can decrease central adiposity, but the effect of exercise dose and the relationship between central adiposity and exercise-induced compensation is unclear.

Objective: Test the effect of exercise dose on central adiposity change and the association between central adiposity and exercise-induced weight compensation.

Methods: In this ancillary analysis of a 6-month randomized controlled trial, 170 participants with overweight or obesity (mean ± SD body mass index: 31.5 ± 4.7 kg/m^2) were randomized to a control group or exercise groups that reflected exercise recommendations for health (8 kcal/kg/week [KKW]) or weight loss and weight maintenance (20 KKW). Waist circumference was measured, and dual-energy X-ray absorptiometry assessed central adiposity. Predicted weight change was estimated and weight compensation (weight change – predicted weight change) was calculated.

Results: Between-group change in waist circumference (control: .0 cm [95% CI, −1.0 to 1.0], 8 KKW: −.7 cm [95% CI, −1.7 to .4], 20 KKW: −1.3 cm [95% CI, −2.4 to −.2]) and visceral adipose tissue (VAT; control: −.02 kg [95% CI, −.07 to .04], 8 KKW: −.01 kg [95% CI, −.07 to .04], 20 KKW: −.04 kg [95% CI, −.10 to .02]) was similar ($P \geq .23$). Most exercisers (82.6%) compensated (weight loss less than expected). Exercisers who compensated exhibited a 2.5-cm (95% CI, .8 to 4.2) and .23-kg (95% CI, .14 to .31) increase in waist circumference and VAT, respectively, vs those who did not ($P < .01$). Desire to eat predicted VAT change during exercise ($\beta = .21$; $P = .03$).

Conclusion: In the presence of significant weight compensation, exercise at doses recommended for health and weight loss and weight maintenance leads to negligible changes in central adiposity.

Key Words: physical activity, weight loss/reduction, abdominal obesity, visceral fat, energy intake, body composition

Abbreviations: ANCOVA, analysis of covariance; BMI, body mass index; DXA, dual-energy X-ray absorptiometry; E-MECHANIC, Examination of Mechanisms of Exercise-Induced Weight Compensation; ES, effect size; KKW, kcal/kg/week; VAT, visceral adipose tissue.

More than two-thirds of the US population lives with overweight or obesity, which are characterized by an elevated body mass index (BMI) (1). Overweight and obesity are key risk factors for cardiometabolic disease development (2), and an established link between BMI and cardiometabolic disease risk exists (3). However, central adiposity, primarily visceral adipose tissue (VAT), is strongly associated with metabolic disease risk factors (eg, high blood glucose (4) and dyslipidemia (5)), cardiometabolic conditions (eg, metabolic syndrome, type 2 diabetes (6), and cardiovascular disease (5)), and mortality (7). Additionally, some have shown central adiposity is more strongly related to metabolic diseases such as type 2 diabetes than fat stored in other regions (3, 6). Interventions that reduce central adiposity are therefore needed to prevent and treat metabolic disease and improve health span.

Aerobic exercise training can decrease central adiposity in individuals with overweight or obesity, regardless of age, sex, and ethnicity (8). However, the influence of aerobic exercise dose on central adiposity is equivocal, with some showing that greater doses do not decrease markers of central adiposity (9, 10) and others suggesting that reductions in VAT are improved at higher exercise doses (11, 12). Understanding the influence of exercise dose on central adiposity is crucial to help design optimal aerobic exercise regimens that enhance central adiposity outcomes. Thus, considering the conflicting

Received: 24 August 2023. **Editorial Decision:** 24 November 2023. **Corrected and Typeset:** 3 January 2024

findings, large randomized controlled trials are needed to test the effect of exercise dose on change in central adiposity markers, including VAT, in individuals with overweight and obesity.

Exercise-induced weight compensation, which is actual weight loss lower than weight loss predicted based on the energy expenditure of exercise, is common (13, 14). Weight compensation can occur because of multiple behavioral and physiological factors, although substantial weight compensation during exercise is primarily caused by elevations in energy intake (13). The relationship between compensation and central adiposity change during aerobic exercise is poorly understood. Some have demonstrated that reductions in VAT during high volumes of exercise are attenuated when compensation occurs through increased energy intake (15, 16), but the association between compensation and changes in central adiposity during exercise at doses similar to that recommended for health (700 to 1000 kcal/week) and weight loss and weight loss maintenance (~2000 kcal/week) (17) has not been thoroughly studied. Moreover, the associations between central adiposity change and the mechanisms related to compensation—namely, increased energy intake, reduced energy expenditure and physical activity, and maladaptive eating attitudes and behaviors—during exercise training are not well understood. Assessing the compensation-related predictors of central adiposity change during exercise could pinpoint factors and/or constructs that may be targeted to enhance improvements in central adiposity.

This ancillary analysis had 2 primary aims. First, we tested the effect of aerobic exercise dose on changes in central adiposity in individuals with overweight or obesity. Second, we examined the associations between central adiposity changes and compensation during exercise training at guidelines akin to those recommended for health and weight loss and weight loss maintenance. As an exploratory aim, we assessed if mechanisms related to exercise-induced compensation predict VAT change during exercise.

Methods

Study Design

The methods of the Examination of Mechanisms of Exercise-Induced Weight Compensation (E-MECHANIC) study (ClinicalTrials.gov: NCT01264406) have been detailed elsewhere (13, 18). Briefly, the study was a 6-month randomized controlled trial that took place at Pennington Biomedical Research Center after institutional review board approval. After the provision of written informed consent, participants recruited to the study were randomized (N = 198) to 1 of 3 groups: a no-exercise control group, an exercise group that aimed to expend 8 kcal/kg/week (KKW) through exercise, or an exercise group that aimed to expend 20 KKW through exercise. The 8-KKW group reflected recommendations for general health (~700 kcal/week), whereas the 20-KKW group reflected recommendations for weight loss and weight loss maintenance (~1760 kcal/week) (13, 17). A biostatistician devised a 1:1:1 randomization ratio, and sex was stratified so that an equal number of males and females were randomized to each group. Randomization was concealed in an envelope until an interventionist or the study manager opened it with the participant. The participants and interventionists supervising exercise sessions were not blinded to group allocation, but the study investigators and the assessment team were because group allocation was not disclosed by the study manager or interventionists. Recruitment and data collection occurred from November 2010 to March 2015 (first participant enrolled 2011). Recruitment finished when the target sample size was recruited (13, 18).

Participants

Sedentary (not exercising >20 minutes on ≥ 3 days/week) individuals living with overweight or obesity (body mass index [BMI], ≥ 25 kg/m^2-≤ 45 kg/m^2) who were otherwise healthy were recruited for the trial. Further details on the participant exclusion criteria have been reported (13).

Intervention

Aerobic exercise training was conducted on a treadmill at an intensity that maintained participants within a heart rate range equivalent to 65% to 85% of baseline peak oxygen uptake. Participants in the 8-KKW group performed their complete exercise dose from the start. To acclimatize participants in the 20-KKW group, participants expended 8 KKW through exercise in week 1 and 14 KKW through exercise in week 2 before completing their complete dose (20 KKW of energy expenditure through exercise) from week 3 until the cessation of the study.

Exercise training was fully supervised and monitored. Participants were weighed weekly with a Tanita scale (Tanita Corporation, Arlington Heights, IL) and selected their exercise frequency (3, 4, or 5 sessions per week) to aid compliance. The energy expenditure target of each session was calculated by dividing the prescribed exercise dose (8 KKW or 20 KKW) by the exercise frequency. To meet the energy expenditure targets, the length of the exercise sessions varied. Real-time estimations of energy expenditure were calculated based on intensity and participant weight, and energy expenditure was measured periodically via a metabolic cart. Participants' adherence to their exercise regimen was calculated as attained exercise energy expenditure divided by prescribed exercise energy expenditure.

The control group received health information (eg, stress management, benefits of healthy foods), although they were instructed to maintain their baseline physical activity.

Outcome Measures

Body weight and waist circumference were measured at baseline and follow-up. Assessments of body composition were performed by dual-energy X-ray absorptiometry (DXA) at baseline and follow-up using Lunar iDXA with Encore software version 13.60 (GE Healthcare, Madison, WI, USA). DXA and Encore software quantified fat mass and body fat percentage for the whole body, trunk, arms, and legs, as well as VAT. The trunk-fat-to-limb-fat ratio (19) and VAT-to-total-fat ratio (20) were calculated as further assessments of central adiposity.

Compensation was calculated as actual weight change minus predicted weight change. Predicted weight at the end of the intervention was estimated using a validated dynamic energy balance model, which is a differential equation based on the first law of thermodynamics and accounts for metabolic adaptation and body composition changes during aerobic exercise training, overcoming the drawbacks of traditional predictions of weight during lifestyle regimens (21, 22). Predicted

change in body weight and body composition are in response to the change in energy expenditure resulting from an increase in physical activity expenditure, as derived from the literature (22). Compensation was not included in the dynamic energy balance model. Indeed, the predicted body weight and body composition changes represent changes without compensation, and hence the difference between model body weight predictions and observed body weight reflects the degree of weight compensation.

Several measures were conducted at baseline and follow-up to assess mechanisms related to compensation (13). Energy intake was determined through doubly labeled water. In the primary outcome manuscript, change in energy intake with doubly labeled water was adjusted for resting metabolic rate, although estimates were also made without adjustments and with adjustments for body composition (13, 23). Results in the present analysis were similar with all estimates of energy intake; thus, change in energy intake with adjustment for resting metabolic rate is reported. Resting metabolic rate was examined via Max II metabolic carts (AEI Technologies), and steps per day were measured with SenseWear armbands (Body Media). The Eating Inventory assessed dietary restraint, disinhibition, and hunger (24). The Food Craving Inventory assessed intense desires to consume certain foods irrespective of hunger (only total score was used in the present analysis) (25). The Food Preference Questionnaire determined food preferences for certain food groups, as well as a fat preference score (only fat preference score was used in the current analysis) (26). Further, retrospective visual analogue scales assessed perceptions of appetite (27), the Compensatory Health Beliefs Scale measured compensatory health-related beliefs (eg, justifying eating because of exercise) (28), and the Activity Temperament Questionnaire examined participants' tendency to move (29).

Statistical Analysis

The current analysis assesses secondary endpoints of the E-MECHANIC trial. Because our secondary endpoints analysis requires follow-up data and adherence to the exercise intervention, participants assessed in the main analysis of the primary manuscript (ie, individuals with baseline and follow up data and ≥75% adherence to their exercise regimen) were considered (13). Including individuals with ≥75% adherence negated the influence of adherence as a possible confounder during between-group comparisons. In total, 171 participants satisfied the follow-up and adherence criteria, but 1 of the 171 participants did not have a baseline DXA measurement. As a result, our reference dataset for this study was restricted to 170 participants (Supplemental Figure 1 (30)).

All statistical analyses were performed in SPSS version 28, with the significance level set to $\alpha = .05$. Differences in change scores among the 3 study groups were examined by 1-way analysis of covariance (ANCOVA), with adjustments for sex and age. Subgroup analyses were performed to examine variations in study group differences between: (1) those with high waist circumference (≥102 cm for males, ≥ 88 cm for females) (31) and healthy waist circumference (<102 cm for males, <88 cm for females) at baseline; (2) males and females; and (3) Black participants and participants of other races. These subgroup analyses were conducted via 2-way ANCOVA adjusted for sex (except for the male vs female subgroup analysis), age, and baseline values. Adjusted post hoc comparisons (Holm-Bonferroni) were performed when ANCOVA omnibus tests were significant to ascertain where differences lay. In exercisers, percent compensation (percentage compensation = [actual weight change − predicted weight change]/predicted weight loss) was calculated (13). Differences in waist circumference change, VAT change, VAT-to-total-fat ratio change, and trunk-fat-to-limb-fat ratio change were examined between those who showed positive compensation (percent compensation > 0%) and those with zero or negative compensation (percent compensation ≤ 0%) via 1-way ANCOVA adjusted for age, sex, and baseline values. Multiple linear regression models adjusted for sex, age, and baseline values also assessed the association between change in central adiposity indices and percent compensation. Pearson correlations tested the relationship between change in VAT and change in mechanisms related to compensation, and significant variables were then entered into a multiple linear regression model along with age, sex, and VAT at baseline to assess the predictors of VAT change in exercisers. We calculated absolute Cohen's d effect size (ES) values to supplement between-group comparisons (32). Comparisons were considered negligible, small, medium, and large when ES values were <.20, .20 to .49, .50 to .79, and ≥.80, respectively, based on previous literature (32). Unless noted otherwise, values from inferential tests are estimated marginal mean (95% CI), whereas descriptive data are mean (SD).

Results

Descriptive

Descriptive characteristics of the participants included in the present analysis are shown in Table 1. Characteristics were similar when the full recruited sample (N = 198) was observed (data not shown). Most of the participants were female (N = 123; 72.4%) and White (N = 113; 66.5%). The mean age, weight, and BMI of the participants was 48.8 (±11.4) years, 88.6 (±15.4) kg, and 31.5 (±4.7) kg/m^2, respectively.

Intervention Data

The 8-KKW group completed 101.0% (±6.3%) of prescribed exercise energy expenditure and the 20-KKW group completed 98.1% (±6.0%) of prescribed exercise energy expenditure, demonstrating the high adherence in both groups. Total energy expended by the 8-KKW group and 20-KKW group during exercise was 17 114 (±3175) kcal and 38 992 (±7308) kcal, respectively. On average, 680 (±123) kcal/week were expended by the 8-KKW group and 1521 (±263) kcal/week were expended by the 20-KKW group. Additional training data are shown in Supplemental Table 1 (30).

The average percent compensation shown by the 8-KKW group and the 20-KKW group was 70.0% (±129.2%) and 58.0% (±61.7%), respectively, and overall, 90 exercise participants (82.6%) showed positive compensation (ie, lost less weight than expected). Those who displayed positive compensation and those who displayed zero or negative compensation showed similar baseline characteristics ($P \geq .17$; Supplemental Table 2 (30)).

Weight and Total Body Composition Change

A difference between groups was identified for weight and BMI change ($P = .02$), with the 20-KKW group exhibiting a

Table 1. Descriptive characteristics of participants included in the per protocol analysis at baseline

		Control (N = 61)	8 KKW (N = 59)	20 KKW (N = 50)	All (N = 170)
Age, y		49.5 (10.8)	48.3 (11.1)	48.5 (12.5)	48.8 (11.4)
Sex					
	Male	16 (26.2)	16 (27.1)	15 (30.0)	47 (27.6)
	Female	45 (73.8)	43 (72.9)	35 (70.0)	123 (72.4)
Race					
	White	38 (62.3)	39 (66.1)	36 (72.0)	113 (66.5)
	Black	21 (34.4)	20 (33.9)	12 (24.0)	53 (31.2)
	Other	2 (3.3)	0 (.0)	2 (4.0)	4 (2.4)
Income					
	<$30 000	10 (16.4)	8 (13.6)	3 (6.0)	21 (12.4)
	$30 000-$49 999	9 (14.8)	7 (11.9)	7 (14.0)	23 (13.5)
	$50 000-$79 999	15 (24.6)	13 (22.0)	15 (30.0)	43 (25.3)
	$80 000-$99 999	12 (19.7)	10 (16.9)	9 (18.0)	31 (18.2)
	≥$100 000	14 (23.0)	20 (33.9)	15 (30.0)	49 (28.8)
	Don't know or missing	1 (1.6)	1 (1.7)	1 (2.0)	3 (1.8)

Continuous data are mean (SD); categorical data are number (%).
Abbreviation: KKW, kcal/kg/week.

decrease compared with control ($P < .03$; ES ≥ .51; Table 2). Akin to results reported in the primary outcomes manuscript (13), total fat mass and body fat percent was reduced in the 20-KKW group compared with other groups ($P < .05$; ES ≥ .44), whereas no differences were seen for total lean body mass ($P = .51$; ES ≤ .22).

Regional Adiposity Change

No significant between-group difference was observed for change in waist circumference ($P = .23$), despite confidence interval data indicating a within-group reduction in the 20-KKW group (Table 2). Changes in VAT, the VAT-to-total-fat ratio, and the trunk-fat-to-limb-fat ratio were similar in all groups, with negligible ES values seen (all $P \geq .65$; all ES ≤ .17). In spite of CI data showing reductions in trunk fat mass and trunk percent fat in the 20-KKW group, between-group differences were not significantly different and small to negligible ES values were revealed (all $P \geq .06$; all ES ≤ .43). Arm percent fat, leg fat mass, and leg percent fat were statistically different between groups (all $P \leq .04$). Post hoc tests for leg fat mass and leg percent fat indicated a greater reduction in leg fat in the 20-KKW group compared with other groups ($P \leq .04$), although no significant between-group variations were observed for arm percent fat following adjustment ($P > .08$: Table 2). Results from subgroup analyses are shown in Supplemental Tables 3-5 (30). The effect of study group on changes in weight, BMI, waist circumference, and DXA endpoints was not modified by waist circumference at baseline or race (all P for interaction ≥ .07). There was a 2-way interaction between study group and sex for VAT-to-total-fat ratio change and trunk-fat-to-limb-fat ratio change (P for interaction < .05), but following adjustments for multiple comparisons, no significant between-group differences were seen when males and females were analyzed separately ($P > .05$).

Individuals with positive weight compensation displayed a 2.5-cm (95% CI, .8-4.2; ES = .76), .23-kg (95% CI, .14-.31; ES = 1.31), .0033 (95% CI, .0017-.0050; ES = 1.02), and .06 (95% CI, .01-.10; ES = .60) increase in waist circumference, VAT, the VAT-to-total-fat ratio, and the trunk-fat-to-limb-fat ratio, respectively, compared with individuals with zero or negative weight compensation (all $P \leq .02$; Fig. 1). Multiple linear regression analyses showed greater weight compensation during exercise training was associated with increases in waist circumference, VAT, the VAT-to-total-fat ratio, and the trunk-fat-to-limb-fat ratio (all $\beta \geq .24$; $P \leq .01$; Supplemental Table 6 (30)).

In Pearson correlation analyses, change in compensatory health beliefs ($r = .20$; $P = .04$) and retrospective desire to eat ($r = .23$; $P = .02$) were related with VAT change during exercise (Supplemental Table 7 (30)); hence, these variables were entered into the multiple linear regression analysis. This regression analysis revealed that retrospective desire to eat was a positive predictor of VAT change ($\beta = .21$; $P = .03$; Table 3). This model also suggested that compensatory health beliefs was not a significant predictor of VAT change, although a similar standardized β was observed ($\beta = .16$; $P = .09$).

Discussion

Overall, in this ancillary analysis of a large, 6-month randomized controlled trial in individuals with overweight and obesity, we showed negligible differences in central adiposity change between a no-exercise control group and 2 aerobic exercise groups—1 similar to guidelines recommended for health and 1 similar to guidelines recommended for weight loss and weight loss maintenance. We also showed that exercisers who displayed positive weight compensation (ie, lost less weight than predicted) showed reduced improvements in central adiposity relative to those who did not compensate. These results indicate that exercise dose has no significant impact on central adiposity, and that significant compensation is likely to negate central adiposity improvements during exercise at guidelines for health and weight loss and weight loss maintenance.

Table 2. Baseline values and 6-month change in weight, waist circumference, total adiposity, total lean mass, and regional adiposity in the control group, 8-KKW group, and 20-KKW group

		Control (N = 61)	8 KKW (N = 59)	20 KKW (N = 50)	*P*	Between-group ES		
						Control vs 8 KKW	Control vs 20 KKW	8 KKW vs 20 KKW
Weight, kg								
	Baseline	90.1 (86.2-94.0)	88.7 (84.8-92.7)	86.7 (82.4-91.0)				
	Change	−.3 (−1.0 to .4)	−.5 (−1.3 to .2)	−1.8 (−2.6 to −.9)[b]	.02[a]	.07	.51	.43
BMI, kg/m^2								
	Baseline	32.3 (31.1-33.5)	31.4 (30.2-32.6)	30.6 (29.3-31.9)				
	Change	−.1 (−.4 to .2)	−.2 (−.5 to .1)	−.6 (−.9 to −.4)[b]	.02[a]	.10	.53	.43
Waist circumference, cm								
	Baseline	101.1 (98.0-104.3)	98.5 (95.3-101.7)	97.0 (93.6-100.5)				
	Change	.0 (−1.0 to 1.0)	−.7 (−1.7 to .4)	−1.3 (−2.4 to −.2)	.23	.15	.33	.17
Total fat mass, kg								
	Baseline	38.9 (36.3-41.4)	37.0 (34.5-39.6)	36.3 (33.5-39.0)				
	Change	.0 (−.6 to .7)	−.3 (−1.0 to .4)	−1.4 (−2.2 to −.7)[c]	.01[a]	.12	.57	.44
Total lean mass, kg								
	Baseline	48.2 (45.7-50.8)	48.8 (46.2-51.4)	47.4 (44.6-50.3)				
	Change	−.4 (−.7 to −.1)	−.2 (−.6 to .1)	−.1 (−.5 to .2)	.51	.13	.22	.09
Total body fat %								
	Baseline	43.2 (41.3-45.0)	41.6 (39.7-43.5)	41.8 (39.8-43.8)				
	Change	.1 (−.4 to .6)	−.1 (−.6 to .4)	−.9 (−1.5 to −.4)[c]	.02[a]	.09	.52	.44
VAT, kg								
	Baseline	1.40 (1.17-1.63)	1.25 (1.01-1.48)	1.30 (1.04-1.55)				
	Change	−.02 (−.07 to .04)	−.01 (−.07 to .04)	−.04 (−.10 to .02)	.71	.02	.13	.15
VAT-to-total-fat ratio								
	Baseline	.0361 (.0302-.0420)	.0336 (.0276-.0396)	.0363 (.0298-.0428)				
	Change	−.0008 (−.0016 to .0001)	−.0002 (−.0011 to .0007)	−.0004 (−.0014 to .0006)	.65	.17	.11	.06
Trunk fat mass, kg								
	Baseline	21.1 (19.6-22.7)	20.0 (18.4-21.6)	19.3 (17.5-21.0)				
	Change	.0 (−.5 to .4)	−.3 (−.8 to .2)	−.8 (−1.3 to −.3)	.07	.14	.43	.29
Trunk fat %								
	Baseline	46.9 (45.0-48.7)	45.0 (43.2-46.9)	44.8 (42.7-46.8)				
	Change	−.1 (−.7 to .5)	−.2 (−.8 to .4)	−1.1 (−1.8 to −.4)	.06	.04	.42	.38
Arms fat mass, kg								
	Baseline	4.01 (3.72-4.30)	3.84 (3.54-4.14)	3.60 (3.28-3.93)				
	Change	.06 (−.02 to .15)	.06 (−.03 to .15)	−.07 (−.17 to .02)	.06	.01	.41	.39
Arms fat %								
	Baseline	39.9 (37.5-42.3)	38.5 (36.1-40.9)	38.4 (35.7-41.0)				
	Change	.7 (.2 to 1.1)	.7 (.2 to 1.1)	−.1 (−.5 to .4)	.04[a]	.01	.42	.43

(*continued*)

Table 2. Continued

		Control (N = 61)	8 KKW (N = 59)	20 KKW (N = 50)	*P*	Between-group ES		
						Control vs 8 KKW	Control vs 20 KKW	8 KKW vs 20 KKW
Legs fat mass, kg								
	Baseline	12.74 (11.70-13.77)	12.19 (11.14-13.24)	12.42 (11.28-13.57)				
	Change	−.01 (−.28 to .26)	−.05 (−.33 to .22)	−.54 (−.83 to −.24)[c]	.02[a]	.04	.50	.45
Legs fat %								
	Baseline	42.0 (39.6-44.3)	41.0 (38.6-43.3)	41.6 (39.0-44.1)				
	Change	.3 (−.2 to .7)	−.2 (−.7 to .2)	−1.0 (−1.5 to −.5)[c]	<.01[a]	.26	.72	.46
Trunk-fat-to-limb-fat ratio								
	Baseline	1.32 (1.22-1.42)	1.28 (1.18-1.38)	1.27 (1.16-1.38)				
	Change	−.03 (−.06 to .00)	−.02 (−.05 to .01)	−.02 (−.05 to .02)	.90	.06	.08	.03

Data are estimated marginal mean (95% CI). Change values are adjusted for age and sex.
Abbreviations: ANCOVA, analysis of covariance; BMI, body mass index; ES, effect size; KKW, kcal/kg/week; VAT, visceral adipose tissue.
[a]Significant ANCOVA omnibus test for study group ($P < .05$).
[b]Significantly different from control group ($P < .05$).
[c]Significantly different from the control group and 8-KKW group ($P < .05$).

Greater energy expenditure during aerobic exercise training leads to increased weight loss because of a higher energy deficit (33), but the influence of aerobic exercise dose on central adiposity is equivocal over 6 months or more. Although Recchia and colleagues demonstrated that greater energy expenditure through aerobic exercise leads to small yet significant improvements in central adiposity (12), a smaller trial showed that increasing exercise dose does not lead to improvements in VAT (9), and others have demonstrated no differences in waist circumference between individuals performing exercise at 50%, 100%, and 150% of guidelines (10). Results from our trial displayed no significant differences in central adiposity changes over 6 months between a control group and 2 groups exercising at doses resembling that recommended for health (680 kcal/week) and for weight loss and weight loss maintenance (1521 kcal/week). We did observe significant reductions in total fat and adiposity in other noncentral regions in the 20-KKW group, which could provide metabolic benefits for individuals with overweight and obesity (34). Confidence intervals also indicated that the 20-KKW group demonstrated a reduction in some central adiposity indices (eg, waist circumference, trunk fat), and between-group significance levels for trunk fat mass and trunk percent fat were close to the significance threshold. However, estimated marginal mean and effect size data show variations in central adiposity between groups are negligible or small at best. The 20-KKW group, for example, exhibited a 1.3-cm and .7-cm decrease in waist circumference compared with the control group and the 8-KKW group, respectively, and these differences are considered clinically unimportant (35, 36) based on associations between waist circumference change and metabolic disease (37) and mortality (38). Thus, on balance, though relatively small levels of physical activity can improve central adiposity (10, 39), we believe exercise doses that expended ∼700 kcal/week and ∼1500 kcal/week induced clinically trivial changes in central adiposity during this trial. Doses with even greater differences in exercise-induced energy expenditure may be required to detect clinically meaningful improvements in central adiposity.

A reason why exercise groups exhibited negligible changes in central adiposity compared with control in our study could be the compensation displayed by exercise groups. The majority (82.6%) of exercisers exhibited positive weight compensation (ie, lost less weight than expected) and these participants displayed a 2.5-cm and .23-kg increase in waist circumference and VAT, respectively, compared with those who did not show positive weight compensation. Although few have examined the link between compensation and changes in central adiposity, 2 studies in individuals with high waist circumference showed that aerobic exercise without weight loss (ie, with compensation) led to attenuated reductions in central adiposity relative to exercise with weight loss (ie, without compensation) in males (15) and females (16), supporting our findings. Nonetheless, contrary to our results, these studies still found significant central adiposity improvements in exercisers who compensated (15, 16). That these earlier studies solely recruited individuals with high waist circumference is unlikely to explain why the previous studies saw improvements in central adiposity in individuals who compensated and we did not, as we found no exercise-induced differences in central adiposity change between those with high and healthy waist circumference at baseline. Rather, the discrepancies could occur because the previous studies were only 3

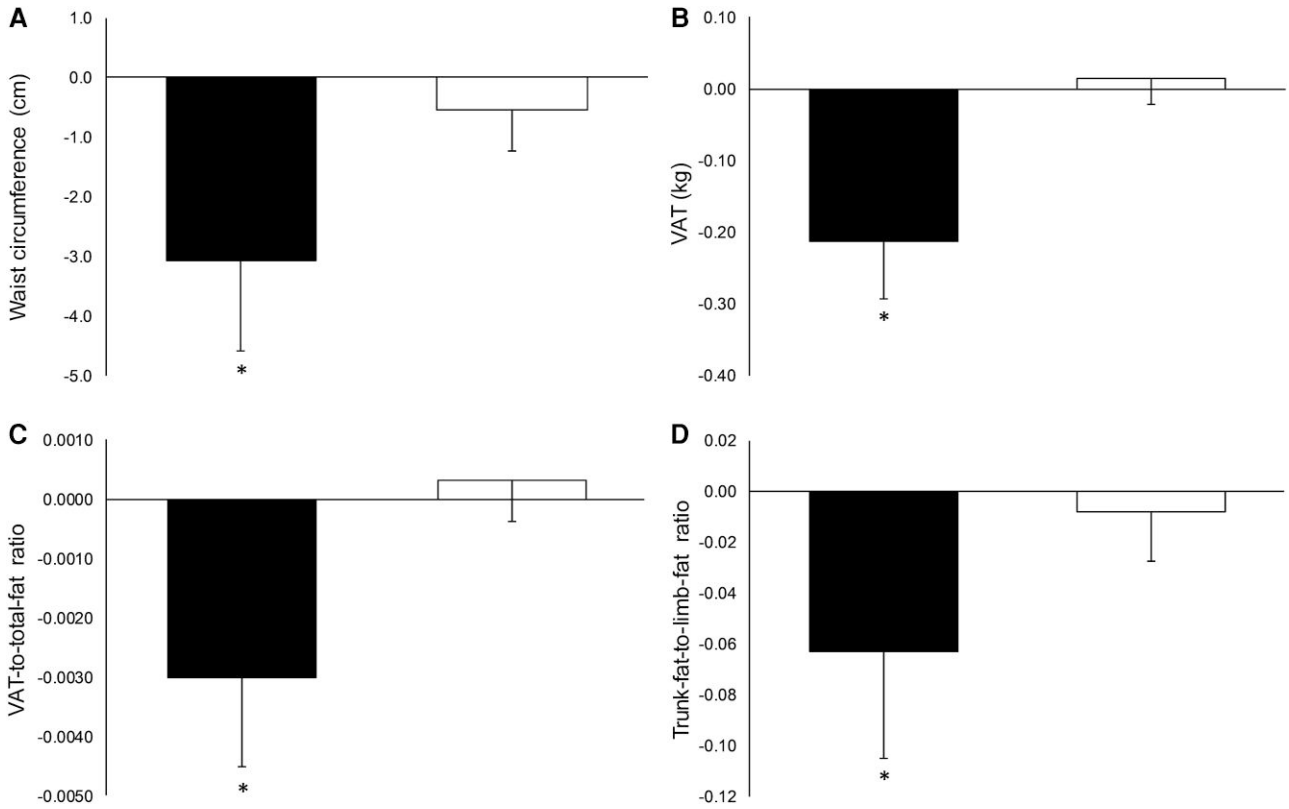


Figure 1. Change in waist circumference (A), VAT (B), VAT-to-total-fat ratio (C), and trunk-fat-to-limb-fat ratio (D) in individuals who displayed zero or negative weight compensation (ie, those who lost more or equal weight to that which was predicted) (N = 19) and individuals who displayed positive weight compensation (ie, those who lost less weight than predicted) (N = 90) during exercise. Abbreviations: ANCOVA, analysis of covariance; VAT, visceral adipose tissue. Black bars are individuals who showed zero or negative compensation; white bars are individuals who showed positive compensation. Data are estimated marginal means (95% CI) adjusted for age, sex, and baseline values. *Significant ANCOVA omnibus comparison between those who displayed zero or negative compensation and those who displayed positive compensation ($P < .05$).

months and/or they implemented far greater exercise energy expenditures of 3500 kcal/week in women (16) and 4900 kcal/week in men (15). These findings could collectively indicate that exercise in the presence of significant weight compensation improves central adiposity during short interventions in which exercise volumes are high, but not during medium to long-term regimens in which exercise performed is similar to that recommended for health and weight loss and weight loss maintenance.

By highlighting the compensatory mechanisms that predict central adiposity change during exercise, effective strategies can be developed to improve central adiposity outcomes during exercise. Several mechanisms could drive positive weight compensation during exercise training: an increase in energy intake, maladaptive changes in eating behaviors and physical activity patterns, and reductions in exercise and nonexercise energy expenditure (13). In this analysis, an increase in desire to eat positively predicted change in VAT. A similar association between change in VAT and compensatory health beliefs (eg, justifying an eating episode because of exercise) was also observed, albeit the coefficient was smaller and tended to be significant. Along with findings from the primary outcome manuscript that showed exercise-induced elevations in energy intake (13), these results may indicate that changes in eating habits attenuated reductions in central adiposity during exercise. This is in line with previous studies (15, 16) and implies that strategies targeting desire to eat and compensatory behaviors during exercise could enhance central adiposity outcomes. Akin to other regimens (40), such strategies could include behavioral sessions that help manage appetite by encouraging participants to increase consumption of foods with low energy density (41). Additional strategies and sessions focussing on meal planning and portion and stimulus control could also decrease compensatory meals and/or snacking during exercise training (40). Nevertheless, it should be acknowledged that more work is needed to elucidate the role of compensatory behaviors in modifying central adiposity changes during exercise because most compensatory behaviors (including energy intake) were not related to VAT and our models explained a small proportion of VAT variance.

A strength of the present analysis is that it comprises energy intake, energy expenditure, and physical activity outcomes assessed with gold-standard techniques, as well as questionnaires examining eating attitudes and behaviors. One limitation is that we did not use computed tomography or magnetic resonance imaging, which are considered gold standard tools for VAT assessment. Although VAT assessments via DXA are linked to computed tomography-derived measurements (42) and our primary findings were consistent among several indices related to VAT, future studies using computed tomography or magnetic resonance imaging are warranted. Additionally, because we did not assess potential physiological mediators, further studies are needed to examine the mechanisms underpinning our findings. Another limitation is that most of the sample were female and white, so it is

Table 3. Multiple linear regression analysis for association between retrospective desire to eat and compensatory health beliefs, and change in VAT

		R^2	B	95% CI	β	*P*
VAT (kg)		.115				
	Retrospective VAS, desire to eat		.0023	(.0002 to .0045)	.21	.03[a]
	Compensatory health beliefs		.0047	(−.0007 to .0101)	.16	.09
	Age		.0011	(−.0023 to .0046)	.07	.51
	Sex[b]		−.0868	(−.1939 to .0202)	−.20	.11
	VAT at baseline		−.0001	(−.0001 to .0000)	−.23	.07

Abbreviations: VAS, visual analogue scale; VAT, visceral adipose tissue.
[a]Statistically significant ($P < .05$).
[b]Male = 1, female = 2.

possible we were underpowered to detect consistent and significant interactions in our subgroup analyses. It is also noteworthy that this manuscript reports an ancillary project, though it should be acknowledged that it used data from a large, randomized control trial where exercise adherence was excellent and exercise dose was fastidiously supervised and monitored.

Taken together, the present study indicates that in the presence of significant weight compensation, there are negligible differences in central adiposity change during aerobic exercise at doses similar to that recommended for health and weight loss and weight loss maintenance in individuals with overweight or obesity. Moreover, higher weight compensation was associated with reduced improvements in central adiposity, and exercisers with increased subjective desire to eat exhibited poorer change in central adiposity. During exercise at guidelines for health and weight loss and weight loss maintenance, exercise-induced compensation should be treated and reduced in individuals with overweight or obesity to enhance central adiposity reductions, potentially through strategies that manage appetite and compensatory food behaviors.

Acknowledgments

The late Conrad P. Earnest, Texas A&M University, College Station, TX, is duly recognized by the authors for his contribution to the organization and execution of the E-MECHANIC study. The authors thank Sultan Alenezi, University of Glasgow, for feedback and comments on the results of the study; he received no financial support for his participation.

Funding

Study supported by National Institutes of Health (NIH) (grant: HL102166); Nutrition Obesity Research Center (NORC) grant P30 DK072476, titled "Nutritional Programming: Environmental and Molecular Interactions" sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases; the National Institute of General Medical Sciences grant U54 GM104940; and the American Heart Association (AHA) (grant: 20POST35210907). The sponsor had no input in the study's design, data collection, data analysis, interpretation, or publication write up.

Disclosures

All authors declare no relevant conflicts of interest.

Data Availability

Some or all datasets generated during and/or analyzed during the current study are not publicly available but are available from the corresponding author on reasonable request.

Clinical trials registration: NCT01264406.

References

1. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011-2012. *JAMA*. 2014;311(8):806-814.
2. Jensen MD, Ryan DH, Apovian CM, *et al*. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association task force on practice guidelines and the obesity society. *J Am Coll Cardiol*. 2013;63(25):2985-3023.
3. Ohlson LO, Larsson B, Svardsudd K, *et al*. The influence of body fat distribution on the incidence of diabetes mellitus. 13.5 years of follow-up of the participants in the study of men born in 1913. *Diabetes*. 1985;34(10):1055-1058.
4. Fox CS, Massaro JM, Hoffmann U, *et al*. Abdominal visceral and subcutaneous adipose tissue compartments: association with metabolic risk factors in the Framingham heart study. *Circulation*. 2007;116(1):39-48.
5. Mathieu P, Poirier P, Pibarot P, Lemieux I, Després JP. Visceral obesity: the link among inflammation, hypertension, and cardiovascular disease. *Hypertension*. 2009;53(4):577-584.
6. Neeland IJ, Ayers CR, Rohatgi AK, *et al*. Associations of visceral and abdominal subcutaneous adipose tissue with markers of cardiac and metabolic risk in obese adults. *Obesity*. 2013;21(9):E439-E447.
7. Kuk JL, Katzmarzyk PT, Nichaman MZ, Church TS, Blair SN, Ross R. Visceral fat is an independent predictor of all-cause mortality in men. *Obesity*. 2006;14(2):336-341.
8. Ross R, Soni S, Houle SA. Negative energy balance induced by exercise or diet: effects on visceral adipose tissue and liver fat. *Nutrients*. 2020;12(4):891.
9. Cowan TE, Brennan AM, Stotz PJ, Clarke J, Lamarche B, Ross R. Separate effects of exercise amount and intensity on adipose tissue and skeletal muscle mass in adults with abdominal obesity. *Obesity*. 2018;26(11):1696-1703.
10. Church TS, Martin CK, Thompson AM, Earnest CP, Mikus CR, Blair SN. Changes in weight, waist circumference and compensatory responses with different doses of exercise among sedentary, overweight postmenopausal women. *PLoS One*. 2009;4(2):e4515.
11. Slentz CA, Aiken LB, Houmard JA, *et al*. Inactivity, exercise, and visceral fat. STRRIDE: a randomized, controlled study of exercise intensity and amount. *J Appl Physiol*. 2005;99(4):1613-1618.
12. Recchia F, Leung CK, Yu AP, *et al*. Dose-response effects of exercise and caloric restriction on visceral adiposity in overweight and obese

adults: a systematic review and meta-analysis of randomised controlled trials. *Br J Sports Med*. 2023;57(16):1035-1041.
13. Martin CK, Johnson WD, Myers CA, *et al*. Effect of different doses of supervised exercise on food intake, metabolism, and non-exercise physical activity: the E-MECHANIC randomized controlled trial. *Am J Clin Nutr*. 2019;110(3):583-592.
14. Blundell JE, Stubbs RJ, Hughes DA, Whybrow S, King NA. Cross talk between physical activity and appetite control: does physical activity stimulate appetite? *Proc Nutr Soc*. 2003;62(3):651-661.
15. Ross R, Dagnone D, Jones PJH, *et al*. Reduction in obesity and related comorbid conditions after diet-induced weight loss or exercise-induced weight loss in men: a randomized, controlled trial. *Ann Intern Med*. 2000;133(2):92-103.
16. Ross R, Janssen I, Dawson J, *et al*. Exercise-induced reduction in obesity and insulin resistance in women: a randomized controlled trial. *Obesity Res*. 2004;12(5):789-798.
17. West Suitor C, Kraak V. *Institute of Medicine. Adequacy of Evidence for Physical Activity Guidelines Development: Workshop Summary*. National Academies Press; 2007.
18. Myers CA, Johnson WD, Earnest CP, *et al*. Examination of mechanisms (E-MECHANIC) of exercise-induced weight compensation: study protocol for a randomized controlled trial. *Trials*. 2014;15(1):212.
19. Imboden MT, Welch WA, Swartz AM, *et al*. Reference standards for body fat measures using GE dual energy x-ray absorptiometry in Caucasian adults. *PLoS One*. 2017;12(4):e0175110.
20. Redman LM, Heilbronn LK, Martin CK, Alfonso A, Smith SR, Ravussin E. Effect of calorie restriction with or without exercise on body composition and fat distribution. *J Clin Endocrinol Metab*. 2007;92(3):865-872.
21. Thomas DM, Martin CK, Lettieri S, *et al*. Can a weight loss of one pound a week be achieved with a 3500-kcal deficit? Commentary on a commonly accepted rule. *Int J Obesity*. 2013;37(12):1611-1613.
22. Thomas DM, Bouchard C, Church T, *et al*. Why do individuals not lose more weight from an exercise intervention at a defined dose? An energy balance analysis. *Obes Rev*. 2012;13(10):835-847.
23. Dorling JL, Höchsmann C, Fearnbach SN, *et al*. Initial weight change and long-term changes in weight and compensation during supervised exercise training. *Med Sci Sports Exerc*. 2021;53(8):1675-1684.
24. Stunkard AJ, Messick S. The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *J Psychosom Res*. 1985;29(1):71-83.
25. White MA, Whisenhunt BL, Williamson DA, Greenway FL, Netemeyer RG. Development and validation of the food-craving inventory. *Obes Res*. 2002;10(2):107-114.
26. Geiselman PJ, Anderson AM, Dowdy ML, West DB, Redmann SM, Smith SR. Reliability and validity of a macronutrient self-selection paradigm and a food preference questionnaire. *Physiol Behav*. 1998;63(5):919-928.
27. Womble LG, Wadden TA, Chandler JM, Martin AR. Agreement between weekly vs. Daily assessment of appetite. *Appetite*. 2003;40(2):131-135.
28. Knäuper B, Rabiau M, Cohen O, Patriciu N. Compensatory health beliefs: scale development and psychometric properties. *Psychol Health*. 2004;19(5):607-624.
29. Anderson SE, Bandini LG, Dietz WH, Must A. Relationship between temperament, nonresting energy expenditure, body composition, and physical activity in girls. *Int J Obesity*. 2004;28(2):300-306.
30. Dorling J. Supplemental Material: "Exercise-induced changes in central adiposity during an RCT: Effect of exercise dose and associations with compensation". *Figshare*. Online resource. https://doi.org/10.6084/m9.figshare.24648183.v2
31. Lean MEJ, Han TS, Morrison CE. Waist circumference as a measure for indicating need for weight management. *BMJ*. 1995;311(6998):158-161.
32. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates; 1988.
33. Donnelly JE, Blair SN, Jakicic JM, *et al*. Appropriate physical activity intervention strategies for weight loss and prevention of weight regain for adults. *Med Sci Sports Exerc*. 2009;41(2):459-471.
34. Agrawal S, Klarqvist MDR, Diamant N, *et al*. BMI-adjusted adipose tissue volumes exhibit depot-specific and divergent associations with cardiometabolic diseases. *Nat Commun*. 2023;14(1):266.
35. Bonafiglia JT, Swinton PA, Ross R, *et al*. Interindividual differences in trainability and moderators of cardiorespiratory fitness, waist circumference, and body mass responses: a large-scale individual participant data meta-analysis. *Sports Med*. 2022;52(12):2837-2851.
36. Hammond BP, Stotz PJ, Brennan AM, Lamarche B, Day AG, Ross R. Individual variability in waist circumference and body weight in response to exercise. *Med Sci Sports Exerc*. 2019;51(2):315-322.
37. Vazquez G, Duval S, Jacobs DR, Silventoinen K. Comparison of body mass index, waist circumference, and waist/hip ratio in predicting incident diabetes: a meta-analysis. *Epidemiol Rev*. 2007;29(1):115-128.
38. Cerhan JR, Moore SC, Jacobs EJ, *et al*. A pooled analysis of waist circumference and mortality in 650,000 adults. *Mayo Clin Proc*. 2014;89(3):335-345.
39. Dorling JL, Höchsmann C, Tudor-Locke C, Beyl RA, Martin CK. Effect of an office-based intervention on visceral adipose tissue: the WorkACTIVE-P randomized controlled trial. *Appl Physiol Nutr Metab*. 2021;46(2):117-125.
40. Rickman AD, Williamson DA, Martin CK, *et al*. The CALERIE study: design and methods of an innovative 25% caloric restriction intervention. *Contemp Clin Trials*. 2011;32(6):874-881.
41. Rolls BJ. The relationship between dietary energy density and energy intake. *Physiol Behav*. 2009;97(5):609-615.
42. Micklesfield LK, Goedecke JH, Punyanitya M, Wilson KE, Kelly TL. Dual-energy X-ray performs as well as clinical computed tomography for the measurement of visceral fat. *Obesity*. 2012;20(5):1109-1114.

Publication 13

Challenges in defining successful adherence to calorie restriction goals in humans: Results from CALERIE™ 2.

Corby K Martin [1], **Christoph Höchsmann** [1,2], James L Dorling [1,3], Manjushri Bhapkar [4], Carl F Pieper [4], Susan B Racette [5], Sai Krupa Das [6], Leanne M Redman [1], William E Kraus [4], Eric Ravussin [1] for the CALERIE™ Phase 2 Study Group

[1] Pennington Biomedical Research Center, Baton Rouge, LA, USA

[2] Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany

[3] Human Nutrition, School of Medicine, Dentistry and Nursing, University of Glasgow, Glasgow, UK

[4] Duke University School of Medicine, Durham, NC, USA

[5] Washington University School of Medicine, St. Louis, MO, USA

[6] Human Nutrition Research Center on Aging at Tufts University, Boston, MA, USA

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
Experimental Gerontology, 162: 111757	2022	10.1016/j.exger.2022.111757	3.9 (2022)	Q2	Scopus: 4 Google Scholar: 4

Author contribution:

Second author; co-developed the research question and statistical model, co-drafted the manuscript, and created tables and figures.

Experimental Gerontology 162 (2022) 111757

Contents lists available at ScienceDirect

Experimental Gerontology

journal homepage: www.elsevier.com/locate/expgero

Challenges in defining successful adherence to calorie restriction goals in humans: Results from CALERIE™ 2

Corby K. Martin [a,*], Christoph Höchsmann [a,b], James L. Dorling [a,c], Manjushri Bhapkar [d], Carl F. Pieper [d], Susan B. Racette [e], Sai Krupa Das [f], Leanne M. Redman [a], William E. Kraus [d], Eric Ravussin [a], for the CALERIE™ Phase 2 Study Group

[a] Pennington Biomedical Research Center, Baton Rouge, LA, USA
[b] Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany
[c] Human Nutrition, School of Medicine, Dentistry and Nursing, College of Medical, Veterinary and Life of Sciences, University of Glasgow, Glasgow, Scotland, UK
[d] Duke University School of Medicine, Durham, NC, USA
[e] Program in Physical Therapy and Department of Medicine, Washington University School of Medicine, St. Louis, MO, USA
[f] JM, USDA, Human Nutrition Research Center on Aging at Tufts University, Boston, MA, USA

ARTICLE INFO

Section Editor: Ricki Colman

Keywords:
Calorie restriction
Adherence
Weight graph
Weight loss
Lifestyle change
Intensive lifestyle intervention
Dietary adherence
Energy intake

ABSTRACT

Background: The Comprehensive Assessment of Long-term Effects of Reducing Intake of Energy (CALERIE™) phase 2 trial tested the effects of two years of 25% calorie restriction (CR) on aging in humans. CALERIE 2 was one of the first studies to use a graph of predicted weight loss to: 1) provide a proxy of dietary adherence, and 2) promote dietary adherence. Assuming 25% CR, each participant's weight over time was predicted, with upper and lower bounds around predicted weights. Thus, the resulting weight graph included a zone or range of body weights that reflected adherence to 25% CR, and this was named the zone of adherence. Participants were considered adherent if their weight was in this zone. It is unlikely, however, that the entire zone reflects 25% CR.
Objectives: To determine the level of CR associated with the zone of adherence and if the level of CR achieved by participants was within the zone.
Methods: Percent CR associated with the upper and lower bounds of the zone were determined via the Body Weight Planner (https://www.niddk.nih.gov/bwp) for participants in the CALERIE 2 CR group (N = 143). Percent CR achieved by participants was estimated with the intake-balance method.
Results: At month 24, the zone of adherence ranged from 10.4(0.0)% to 19.4(0.0)% CR [Mean(SEM)], and participants achieved 11.9(0.7)% CR and were in the zone.
Conclusion: The results highlight the challenges of: 1) setting a single CR goal vs. a range of acceptable values, and 2) obtaining real-time and valid measures of CR adherence to facilitate adherence.

1. Introduction

1.1. Few methods exist to accurately quantify dietary adherence in real-time, particularly over the long-term

Promoting adherence to calorie-restricted diets has been very difficult due to the challenges of accurately quantifying energy intake. Traditional self-report methods to assess energy intake (e.g., food records, dietary recall) are commonly used, though the accuracy of these methods has been questioned (Beaton et al., 1997; Tran et al., 2000; Schoeller et al., 1990; Bandini et al., 1990) and it is difficult for participants to use them over the long term. The doubly labeled water (DLW) method can be used to estimate energy intake accurately over the short term (e.g., two weeks) and long-term when changes in body

Abbreviations: AL, ad libitum; CALERIE™, Comprehensive Assessment of Long-term Effects of Reducing Intake of Energy; CTS, Computer Tracking System; CR, calorie restriction; DLW, doubly labeled water; DXA, dual energy X-ray absorptiometry; IBM SPSS, International Business Machines Statistical Package for the Social Sciences; kJ, kilojoules; NIDDK, National Institute of Diabetes and Digestive and Kidney Diseases; PAL, Physical Activity Level; RMR, resting metabolic rate; SEM, standard error of the mean; SD, standard deviation; TDEE, total daily energy expenditure.

* Corresponding author at: Pennington Biomedical Research Center, 6400 Perkins Rd., Baton Rouge, LA 70808, USA.
E-mail address: Corby.Martin@pbrc.edu (C.K. Martin).

https://doi.org/10.1016/j.exger.2022.111757
Received 26 August 2021; Received in revised form 9 February 2022; Accepted 24 February 2022
Available online 28 February 2022

composition are measured.(Racette et al., 2012) This approach requires repeated DLW and body composition measurements, as well as isotope analyses; therefore, it is not practical for many studies and cannot provide real-time estimates of calorie restriction (CR) to inform intervention delivery.

1.2. Estimating dietary adherence in real-time by using body weight as a proxy for dietary adherence

An alternative method for estimating adherence to a calorie-restricted diet is to calculate expected body weight for study participants based on the prescribed level of CR. This allows the participant's actual weights to be compared to expected weights over time and, if the participant's actual body weight reflects the expected weight, adherence to the CR goal can be inferred. If the participant's actual body weight deviates from the expected weight, then it can be inferred that the participant is not adhering to the CR goal. One challenge to developing and deploying this approach is the erroneous assumption that human participants can control their body weight precisely enough to closely mirror a single body weight at any point in time. A second challenge is inherent error in calculations of expected body weight. One way to address these limitations is to provide participants with a range of acceptable body weights that reflects CR adherence.

To develop and facilitate this approach, a mathematical model was developed by Pieper et al. (2011) that predicted the distribution of percent weight change over 12 months assuming 25% CR. The output from the model was used to create weight graphs for participants that reflect the goal weight, which is represented by the green line in Fig. 1 (the green line reflects the 50th percentile of expected weight change from the model). Upper and lower bounds around this goal weight are represented by the yellow and light blue lines in Fig. 1 (the yellow and light blue lines reflect the 80th and 10th percentiles of weight change from the model, respectively). The result is a weight graph that includes a zone or range of body weights that reflects adherence to 25% CR.

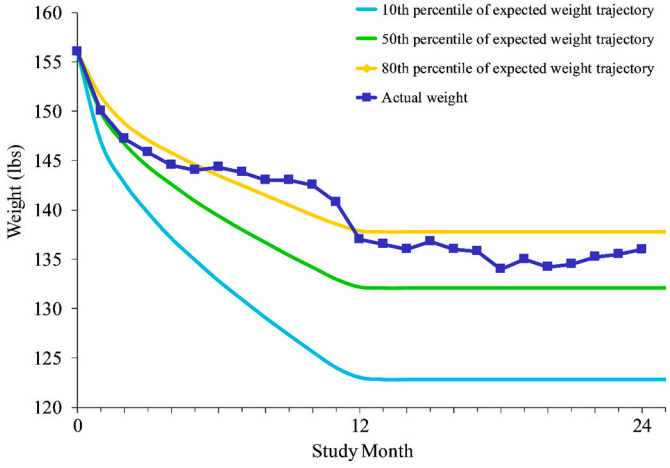


Fig. 1. A sample weight graph is displayed for a hypothetical participant in the Comprehensive Assessment of Long-term Effects of Reducing Intake of Energy (CALERIE™) phase 2 trial (CALERIE 2), which tested the effects of two years of CR on biomarkers of aging in humans. The light blue, green, and yellow lines correspond to the 10th, 50th, and 80th percentiles of expected weight trajectories, respectively. The dark blue line depicts the hypothetical participant's measured weight trajectory. The participant's starting weight was 70.7 in kilograms. From months 12 to 24, the yellow, green, and light blue lines represent 62.5, 60, and 55.7 kg, respectively. Reprinted from Contemporary Clinical Trials, Vol 32, Issue 6; Amy D. Rickman, Donald A. Williamson, Corby K. Martin, Cheryl H. Gilhooly, Richard I. Stein, Connie W. Bales, Susan Roberts, and Sai Krupa Das; The CALERIE Study: Design and methods of an innovative 25% caloric restriction intervention; Page No. 880, 2011, with permission from Elsevier.

Hence, this zone of acceptable weights is called the zone of adherence. A participant's measured body weights are plotted over time on the weight graph and the participant is considered adherent to 25% CR if his/her weight is within the zone. Because the Pieper et al. (2011) model was not designed to predict body weight beyond 12 months, the zone of adherence is flat between months 12 and 24, as depicted Fig. 1.

1.3. Using a weight graph and zone of adherence to personalize intervention delivery and promote CR adherence

The model and weight graphs from Pieper et al. (2011) were integrated into the intervention (Rickman et al., 2011) for the Comprehensive Assessment of Long-term Effects of Reducing Intake of Energy (CALERIE™) phase 2 trial (CALERIE 2), which tested the effects of two years of CR on biomarkers of aging in humans. As detailed by Rickman et al. (2011) the model and weight graph were central in directing the delivery of the intervention by assessing if a participant was adherent or nonadherent in real time and adjusting treatment delivery accordingly. Participants were weighed at each intervention session and their weights were plotted onto their weight graph. Participants were considered adherent to the 25% CR goal if their weight was within the zone of adherence. When weight was above the zone, participants were considered nonadherent to 25% CR and intervention strategies were deployed to help participants better restrict energy intake. Conversely, a weight below the zone indicated that the participant had been too restrictive and efforts were needed to increase energy intake.

1.4. Current objectives

The use of the model and weight graphs to foster adherence during CALERIE 2 was novel and provided a much-needed real time metric of adherence. The approach also provided a framework to personalize intervention delivery and to guide deployment of treatment strategies (Rickman et al., 2011). Nonetheless, the utility of the weight graphs and the success of the CALERIE 2 intervention require further analysis. To that end, the objectives of this analysis were twofold. First, determine the level of CR associated with the zone of adherence by utilizing a validated weight loss calculator that was not used during CALERIE 2 (Hall et al., 2011; Hall and Chow, 2011). Second, determine if participants' actual level of CR was within the zone by using the intake-balance method, which is considered accurate (Ravussin et al., 2015), but cannot provide data in real time and thus necessitates a post-hoc analysis. It was hypothesized that the upper bound of the zone at months 12 and 24 would be less than 25% CR. It was further hypothesized that the level of CR achieved by participants would be within the zone at months 12 and 24.

2. Methods

The CALERIE™ phase 2 randomized controlled trial was a multi-site study conducted at Pennington Biomedical (Baton Rouge, LA, USA), Washington University School of Medicine (St. Louis, MO, USA), and Tufts University (Boston, MA, USA). The coordinating center was Duke Clinical Research Institute (Durham, NC, USA). The clinicaltrials.gov registration number is NCT00427193. All sites received Institutional Review Board approval and all participants provided written informed consent. The CALERIE 2 study aimed to test the effects of two years of 25% CR on aging biomarkers in comparison to an ad libitum (AL) control group. The study design (Rochon et al., 2011), screening and recruitment procedures (Stewart et al., 2013), and intervention (Pieper et al., 2011; Rickman et al., 2011) have been described extensively.

2.1. Participants and randomization

CALERIE 2 recruited participants who were 20–50 years old (men) or 20–47 years old (women) and had body mass index $\geq$22.0 and $<$ 28.0

kg/m^2. Exclusion criteria included significant medical conditions (e.g., cardiovascular disease, diabetes, hypertension), psychological disorders, high levels of physical activity (≥ 30 mins ≥5 days/week), and women who were pregnant or planning to become pregnant during the trial.

Participants were randomized into the CR or AL group in a 2:1 ratio favoring CR. A permuted block randomization approach was used to stratify by study site, sex, and BMI category (normal weight: BMI 22.0–24.9 kg/m^2 and overweight: BMI 25.0–27.9 kg/m^2). 145 participants were randomized to the CR group and 75 participants were randomized to the AL group. The AL group was asked to continue eating their habitual diet and did not receive any intervention; they are not included in the analyses reported herein.

2.2. Description of the CR intervention

The goal of the CR intervention was to promote 25% CR for two years. As described in Rickman et al. (2011) the CR participants received an intensive lifestyle intervention to foster adherence, including individual sessions with an interventionist once per week for the first month, twice per month from month 2 through 12, and once per month from month 13 through 24. Additional sessions were scheduled as needed. Finally, participants attended group sessions twice per month from month 1 through 6, and once per month from month 7 through 24.

As noted earlier, the model and weight graphs developed by Pieper et al. (2011) were central to guiding intervention delivery throughout the two-year intervention, which was deployed via an Internet-based Computer Tracking System (CTS) that was created for the project (Rickman et al., 2011). Briefly, the CTS facilitated intervention fidelity and provided structure to how the intervention was deployed over time, across interventionists, and across participants. A central feature of the CTS was tracking weight as a proxy of CR adherence. Participants' demographic information was entered into the CTS, as well as their starting body weight and their energy intake target, which reflected 25% CR. A personalized weight graph was then generated for each participant based on the Pieper et al. (2011) model. Participants were weighed at each session and the interventionist entered the measured body weight into the CTS, which plotted the participant's weight onto his/her graph. Adherence was considered acceptable if the participant's weight was within the zone. A sample weight graph is provided in Fig. 1 and illustrates that this hypothetical participant was in the zone and adherent in the early period of the intervention. The participant's weight was above the zone, however, from around month 6 to month 11, indicating non-adherence to the CR prescription. During this period, the CTS would automatically suggest toolbox options or specific intervention strategies (e.g., use of portion-controlled foods) to support the participant in achieving their prescribed energy intake level and re-entering the zone. This also helped standardize the delivery of treatment options among participants when they presented with similar challenges (e.g., difficulty adhering their prescribed energy intake level, weight being above or below the zone, poor attendance to sessions, etc.). As indicated in Fig. 1, this hypothetical participant re-entered the zone around month 12 and maintained adherence throughout the rest of the trial.

2.3. Percent CR calculations

The purpose of the analyses reported in this paper were to: 1) determine the level of CR associated with the zone of adherence in CALERIE 2, and 2) examine the level of CR achieved by participants in relation to the percent CR values from the zone of adherence.

2.3.1. Percent CR associated with the zone of adherence

To calculate the percent CR associated with the zone of adherence, a model was needed that was both valid and different from the model that was used in CALERIE 2 (i.e., the Pieper et al. (2011) model). The NIDDK Body Weight Planner (Hall et al., 2011; Hall and Chow, 2011) was selected since the models used in the planner have been validated (Hall et al., 2011; Hall and Chow, 2011) and the models were found to accurately quantify change in energy intake over two years in the CALERIE 2 study when compared to the intake-balance method (Sanghvi et al., 2015). Additionally, the NIDDK Body Weight Planner provides the ability to adjust each participant's physical activity level (PAL) to match participants' baseline energy requirements with the energy requirement measured in CALERIE 2. Thus, the NIDDK Body Weight Planner provided: 1) a valid method to quantify the percent CR associated with the zone of adherence, 2) a model that was not used to generate the zone of adherence during CALERIE 2, and 3) the ability to adapt PAL such that energy requirements were most accurate for each individual participant.

The percent CR associated with the zone of adherence, specifically, the upper bound of the zone (80th percentile), the lower bound of the zone (10th percentile), and the 50th percentile line, was calculated with the NIDDK Body Weight Planner. To do so, the following procedures were followed, and the example provided is to determine the percent CR associated with the 80th percentile or the upper bound of the zone at month 12. First, each participant's weight, sex, age, height, and baseline weight from CALERIE 2 were entered into the planner. The physical activity level (PAL) was adjusted in the planner until each participant's baseline energy requirement in the planner matched the energy requirement measured in CALERIE 2 (each participant's PAL was also measured during CALERIE 2, and agreement between this measure and the value entered into the planner was evaluated). Second, each participant's predicted weight at the 80th percentile from the weight graph at month 12 was entered into the planner as the goal weight, and the duration to achieve the goal was set to 12 months. Third, the planner then produced the energy intake value needed to achieve this goal. Fourth, this energy intake value was used in conjunction with the baseline energy requirements to calculate the percent CR reflective of the 80th percentile at month 12. This process was repeated for the 80th percentile at month 24, and for the 10th and 50th percentiles at months 12 and 24.

Once percent CR was calculated for the zone of adherence for each participant, the mean (and standard error of the mean or SEM) percent CR values for the zone of adherence were calculated across all of the CR participants.

2.3.2. Percent CR achieved by participants in CALERIE 2

The second purpose of the analyses reported herein was to examine the level of CR achieved by participants in relation to the percent CR values from the zone of adherence. This process determined if the participants were adherent to the CR goal, as defined by the zone of adherence, even if the level of CR that they achieved failed to reach 25%.

The previous section outlined the methods to calculate the percent CR associated with the zone of adherence, and these calculations relied on the NIDDK Body Weight Planner. As detailed in the following paragraphs, determining each participant's percent CR required different methods, namely, the intake-balance method (Ravussin et al., 2015), which relied on state-of-the art measures that were collected during CALERIE 2.

The intake-balance method (Ravussin et al., 2015) relies on measures of total daily energy expenditure (TDEE) and, if weight is not stable during the TDEE assessment, a measure of change in body energy stores, which can be determined by measuring change in body composition during the TDEE assessment. During energy balance or weight stability, energy intake is equal to TDEE. Hence, measured TDEE is equal to energy intake during weight stability. If weight is not stable, then TDEE is not equal to energy intake. In this case, TDEE must be corrected for the energy cost of the change in body composition during the period of TDEE assessment. Hence, energy intake is calculated as the difference between energy expenditure (TDEE) and the energy cost of changes in body composition.

During CALERIE 2, TDEE and body composition were assessed at several time points, allowing our team to calculate the percent CR that each participant achieved for different intervals in the study (e.g., from baseline to month 12, and baseline to month 24). Specifically, participants' TDEE was measured with doubly labeled water over four weeks at baseline to establish baseline energy requirements. Doubly labeled water was also used to measure TDEE for two weeks at months 6, 12, 18, and 24. To quantify change in body composition during the TDEE assessments, fat mass and fat-free mass were measured with dual energy X-ray absorptiometry (DXA; Hologic QDR 4500A; Hologic, Bedford, MA) before and after each TDEE assessment.

The TDEE and body composition measures outlined above allowed the mean energy intake of each participant to be estimated with the intake-balance method. Specifically, each participant's mean daily energy intake from baseline to month 12, and baseline to month 24, was calculated. The calculation for energy intake was: mean TDEE for each interval plus changes in body energy stores. Change in energy stores was calculated assuming 9300 kcal/kg (38,893 kJ/kg) for the energy content of fat mass change and 1100 kcal/kg (4602 kJ/kg) for fat-free mass change (Racette et al., 2012). The mean daily energy intake values were then used to calculate percent CR in relation to each participant's baseline energy requirements.

Once percent CR at months 12 and 24 were calculated for each participant, mean (and SEM) percent CR at months 12 and 24 were calculated across all CR participants.

2.4. Physical activity level (PAL)

Physical activity level was calculated as TDEE from DLW divided by resting metabolic rate (RMR). RMR was measured via indirect calorimetry using a Vista-MX metabolic measurement system (Vacumed, Ventura, CA).

2.5. Data analytic plan

Measured PAL from CALERIE 2 and the PAL used in the Body Weight Planner were compared with Pearson correlation coefficients. As noted earlier, mean percent CR across all CR participants was calculated for the 80th, 50th, and 10th percentiles at months 12 and 24. These values were graphed and participants' actual percent CR was plotted in relation to these values. Independent sample *t*-tests and analysis of variance (ANOVA) were used to determine if participants' actual percent CR, and the percent CR values for the 80th, 50th, and 10th percentiles at months 12 and 24, differed by sex, BMI category, or race. Alpha was set at 0.05. All analyses were conducted using IBM SPSS, Version 27 (Armonk, NY, IBL Corp).

3. Results

3.1. CALERIE 2 results

The CALERIE 2 results have been reported extensively but, in brief, indicated that two years of CR was safe, resulted in significantly improved aging and longevity biomarkers, and reduced risk factors for age-related diseases (Ravussin et al., 2015; Kraus et al., 2019; Dorling et al., 2021; Romashkan et al., 2016; Kebbe et al., 2021). Additionally, CR was found to have no detrimental, and some positive effects, on health-related quality of life (Martin et al., 2016a).

3.2. Participant characteristics

The descriptive characteristics of the sample are provided in Table 1. The sample was predominantly female (69.2%) with a slightly higher proportion of participants in the overweight (52.4%) vs. normal weight (47.6%) BMI stratum. The sample was comprised of 143 CR participants who started the intervention, as reflected in Table 1, though data were available for 128 participants at month 12, the first time point of interest for this analysis. Table 2 includes the sample sizes at each time point in total and by grouping variable (i.e., sex, race, and BMI stratum).

Table 1
Baseline characteristics of participants in the calorie restriction group ($N = 143$).

Sex, n (%)		
Male	44	(30.8)
Female	99	(69.2)
Race, n (%)		
White	111	(77.6)
African American	15	(10.5)
Asian	11	(7.7)
Other	6	(4.2)
Age (years), mean (SD)	38.2	(7.3)
Weight (kg), mean (SD)	73.7	(9.9)
BMI (kg/m^2), mean (SD)	25.8	(1.9)
BMI Category, n (%)		
Normal weight (22.0–24.9 kg/m^2)	68	(47.6)
Overweight (25.0–27.9 kg/m^2)	75	(52.4)

Abbreviations: BMI, body mass index; kg, kilogram; SD, standard deviation.

3.3. Physical activity level (PAL)

Overall, PAL entered into the Body Weight Planner [1.66 (0.02)] correlated significantly with measured PAL [1.75 (0.02)] from CALERIE 2 ($n = 127$, $r = 0.60$, $p < 0.001$).

3.4. Percent CR associated with the zone of adherence

As hypothesized, the upper bound of the zone of adherence (the 80th percentile) reflected less than 25% CR (Fig. 2, Table 2). At months 12 and 24, the mean CR levels for the upper bound of the zones were approximately half (13.7% CR) and less than half (10.4% CR) of the 25% CR goal, respectively. The lower bound of the zone (the 10th percentile) essentially reflected 25% CR (24.9% CR) at month 12 only, with the CR value decreasing to 19.4% at month 24. Moreover, the 50th percentile, which many participants considered their body weight target, reflected 17.8% and 13.6% CR at months 12 and 24, respectively.

The percent CR associated with the zone of adherence was greater for women and participants in the overweight BMI stratum; only the sex effect for the 10th percentile at month 24 was non-significant (Table 2 includes the sex and BMI effects; Fig. 3 illustrates the BMI effect). Race effects for percent CR associated with the zone of adherence were present at month 12 only, with African Americans have greater percent CR than Whites and Asians at the 80th percentile (Table 2). African Americans and Whites had greater percent CR values than Asians at the 50th and 10th percentiles at month 12.

3.5. Percent CR achieved by participants

As hypothesized, the actual level of CR achieved by participants, assessed with the intake-balance method, was within the zone of adherence at both month 12 (15.2% CR) and month 24 (11.9%) (Table 2 and Fig. 2). Percent CR did not differ by sex or race at month 12 or 24, though the race effect at month 24 had a *p*-value of 0.057. Inspection of the means suggests that participants who identified as Asian and Other had lower percent CR, although the sample sizes in these cells are small (Table 2). Participants in the overweight BMI stratum achieved higher percent CR at month 12; this effect was not statically significant at month 24 ($p = 0.056$) (Table 2).

4. Discussion

The hypotheses of the study were supported. First, the upper bound of the zone of adherence reflected a percent CR that was well below the 25% CR goal at months 12 and 24. Second, the average level of CR

Table 2
Percent CR at the 80th, 50th, and 10th percentile, and the percent CR achieved by participants at months 12 and 24 during CALERIE 2.

	All participants*	
	Mean	(SEM)
% CR at 80th percentile		
M12	13.7	(0.1)
M24	10.4	(0.0)
% CR at 50th percentile		
M12	17.8	(0.1)
M24	13.6	(0.0)
% CR at 10th percentile		
M12	24.9	(0.1)
M24	19.4	(0.0)
Actual % CR		
M12	15.2	(0.7)
M24	11.9	(0.7)

	Men†		Women†		*t*	df	*p*
	Mean	(SEM)	Mean	(SEM)			
% CR at 80th percentile							
M12	13.3	(0.1)	13.8	(0.1)	−2.9	126	**0.004**
M24	10.3	(0.0)	10.5	(0.0)	−3.1	116	**0.002**
% CR at 50th percentile							
M12	17.4	(0.2)	18.0	(0.1)	−3.2	126	**0.002**
M24	13.4	(0.0)	13.6	(0.0)	−2.9	116	**0.004**
% CR at 10th percentile							
M12	24.3	(0.2)	25.2	(0.1)	−3.5	126	**<0.001**
M24	19.3	(0.1)	19.4	(0.1)	−1.4	116	0.156
Actual % CR							
M12	15.4	(1.2)	15.2	(0.8)	0.1	123	0.904
M24	11.7	(1.2)	12.0	(0.8)	−0.2	113	0.845

	Normal Weight‡		Overweight‡		*t*	df	*p*
	Mean	(SEM)	Mean	(SEM)			
% CR at 80th percentile							
M12	13.3	(0.1)	14.0	(0.1)	−4.3	126	**<0.001**
M24	10.3	(0.0)	10.5	(0.0)	−2.7	116	**0.008**
% CR at 50th percentile							
M12	17.4	(0.1)	18.1	(0.1)	−4.2	126	**<0.001**
M24	13.5	(0.0)	13.6	(0.0)	−2.7	116	**0.009**
% CR at 10th percentile							
M12	24.4	(0.2)	25.4	(0.2)	−4.2	126	**<0.001**
M24	19.3	(0.1)	19.5	(0.1)	−2.7	116	**0.008**
Actual % CR							
M12	13.8	(1.0)	16.5	(0.8)	−2.1	123	**0.036**
M24	10.5	(1.0)	13.1	(0.9)	−1.9	113	0.056

	White§		African American§		Asian§		Other§		*F*	df	*p*
	Mean	(SEM)	Mean	(SEM)	Mean	(SEM)	Mean	(SEM)			
% CR at 80th percentile											
M12	13.7	(0.1) [a]	14.2	(0.4) [b]	13.1	(0.2) [a]	13.3	(0.3) [a,b]	3.5	3	**0.018**
M24	10.4	(0.0)	10.4	(0.1)	10.3	(0.1)	10.4	(0.1)	1.2	3	0.315
% CR at 50th percentile											
M12	17.8	(0.1) [a]	18.3	(0.3) [a]	17.0	(0.3) [b]	17.3	(0.3) [a,b]	3.4	3	**0.020**
M24	13.6	(0.0)	13.5	(0.1)	13.3	(0.1)	13.5	(0.1)	1.7	3	0.167
% CR at 10th percentile											
M12	25.0	(0.1) [a]	25.5	(0.4) [a]	24.0	(0.4) [b]	24.4	(0.4) [a,b]	2.7	3	**0.047**
M24	19.4	(0.1)	19.3	(0.2)	19.1	(0.2)	19.3	(0.2)	1.4	3	0.238
Actual % CR											
M12	15.7	(0.7)	16.0	(2.1)	11.8	(2.3)	10.6	(2.3)	1.6	3	0.186
M24	12.5	(0.8)	12.7	(1.9)	6.3	(1.9)	6.8	(2.7)	2.6	3	0.057

Data are mean (SEM). Superscripts that differ from each other within a row indicate significant differences between subgroups ($P < 0.05$).
Abbreviations: BMI, body mass index; CR, calorie restriction; df, degrees of freedom; *F*, F-value; M, month; SEM, standard error of mean; *t*, t-value
Significant p-values are represented in bold text.

* Percent CR at the 80th, 50th, and 10th percentiles are available for 128 (M12) and 118 (M24) participants. Actual percent CR is available for 125 (M12) and 115 (M24) participants.

† For men, percent CR at the 80th, 50th, and 10th percentile are available for 39 (M12) and 35 (M24) participants, and actual percent CR is available for 38 (M12) and 35 (M24) participants. For women, percent CR at the 80th, 50th, and 10th percentile is available for 89 (M12) and 83 (M24) participants, and actual percent CR is available for 87 (M12) and 80 (M24) participants.

‡ For the low BMI category, percent CR at the 80th, 50th, and 10th percentile is available for 61 (M12) and 57 (M24) participants, and actual percent CR is available for 57 (M12) and 54 (M24) participants. For the high BMI category, percent CR at the 80th, 50th, and 10th percentile is available for 67 (M12) and 61 (M24) participants, and actual percent CR is available for 68 (M12) and 61 (M24) participants.

§ For Whites, percent CR at the 80th, 50th, and 10th percentile is available for 99 (M12) and 92 (M24) participants, and actual percent CR is available for 97 (M12) and 90 (M24) participants. For African Americans, percent CR at the 80th, 50th, and 10th percentile is available for 14 (M12 and M24) participants, and actual percent CR is available for 13 (M12 and M24) participants. For Asians, percent CR at the 80th, 50th, and 10th percentile and actual percent CR are available for 10 (M12) and 7 (M24) participants. For other races, percent CR at the 80th, 50th, and 10th percentile and actual percent CR are available for 5 (M12 and M24) participants.

achieved by participants was within the zone at months 12 and 24. The lower bound of the zone nearly reflected 25% CR only at month 12 and, by month 24, the lower bound of the zone reflected ~19% CR. This highlights a problem that CALERIE 2 faced when using a model designed to predict weight loss over 12 months in a 24-month trial. The predicted weight loss trajectory was flat between months 12 and 24 because the model was not designed to predict weight loss past 12 months. This is problematic since different levels of CR are required to produce the same amount of weight loss over two different periods of time and body weight was used as a proxy measure of CR. Specifically, more severe CR is necessary to produce the same level of weight loss over a shorter duration, resulting in different levels of CR for the same goal weights at months 12 and 24.

Based on the weight graph and the definition of adherence used during the CALERIE 2 trial to inform intervention delivery, participants were, on average, adherent. Moreover, participants would need to have achieved a weight loss below the lower bound of the zone to achieve 25% CR between months 12 and 24. While it cannot be confirmed if 25% CR is feasible for most participants, the interpretation that the 25% CR intervention was a relative failure, and that participants could only achieve 12% CR on average over the two years, is confounded by the accuracy of the tool used to guide participants toward the prescribed goals. Indeed, the present analyses uncovered a discrepancy between the adherence metric that was obtained in real-time to guide intervention delivery with adherence calculations computed post hoc from state-of-the-art techniques, such as the intake-balance method. This highlights challenges with quantifying the success of a study or an intervention. When the intake-balance method is used to estimate participants' percent CR, it is noted that the level of achieved CR is below the 25% CR target; hence, the CALERIE 2 intervention is interpreted as failing to achieve its goal. Conversely, when a zone of adherence is used to determine adherence, as it was during delivery of the CALERIE 2 intervention, participants were considered adherent, on average, by virtue of their weights being in the zone of adherence. This discrepancy is noteworthy since a measure of percent CR from the intake-balance method cannot be obtained in real time to modify intervention delivery. Moreover, determining adherence with the intake-balance method creates a conundrum since any deviation from 25% CR technically reflects non-adherence, unless there is an a priori decision to consider a range of percent CR (e.g., 22% to 28% CR) as adherent.

The results of the study also indicate that the percent CR associated with the zone of adherence varied by sex, BMI stratum, and race. Specifically, the zone of adherence resulted in greater percent CR for women and for participants in the overweight BMI stratum. Nonetheless, due to the low variability of these measures, relatively small differences in percent CR were significant. The percent CR achieved did not differ by sex, but it did differ by BMI stratum. Specifically, the participants in the overweight BMI stratum achieved higher percent CR compared to those in the normal weight BMI stratum at month 12. Finally, Asians had lower percent CR associated with the zone of adherence compared to African Americans and sometimes Whites, though the percent CR achieved did not vary by race. These results highlight the need to: 1) build and validate models on representative samples of participants, and 2) build and validate models that better model the effects of sex and body mass on energy balance, which has been the focus of recent efforts (Hall et al., 2011; Thomas et al., 2011). The effects of race likely require further investigation, as body composition (Wagner and Heyward, 2000) and metabolism may differ among races (DeLany et al., 2005), even after adjusting for fat-free mass. Lastly, the results indicate that different groups of participants inadvertently may be held to different standards of adherence, which will affect the delivery of their intervention. This is an important area of study, particularly given the challenges of applying models and techniques to individual participants when they were validated at the group level.

The primary aim of the CALERIE 2 trial was to determine if CR favorably slowed biomarkers of aging, as it does in animal models, among human participants without obesity, including normal weight participants. A lower BMI limit of 22.0 kg/m^2 was established to allow investigation of the anti-aging effects of CR among participants of normal weight, specifically avoiding a study of obesity treatment, which has been the focus of many prior studies (Rochon et al., 2011). A rigorous safety plan was established that included monitoring bone mineral density and BMI, and CR was discontinued temporarily or permanently if participants' values went below predefined limits (e.g., BMI < 18.5 kg/m^2) (Romashkan et al., 2016). The level of CR achieved in CALERIE 2 was found to be generally safe and well-tolerated, with no significant differences in adverse events between the CR and control (ad libitum) group (Romashkan et al., 2016). Within the CR group, however, participants of normal weight had a significantly higher incidence of nervous system, musculoskeletal, and reproductive system adverse events compare to the CR participants in the overweight BMI stratum at baseline (Romashkan et al., 2016). Bone mineral density decreases with weight loss, and the CR group experienced expected levels of decreased bone mineral density, though the increase in predicted osteoporotic fracture risk over 10 years was minimal (0.2%). Loss of fat free mass also occurs during weight loss and, as expected, this was the case in CALERIE 2. Nonetheless, the CR participants, compared to the ad lib control, experienced an increase in the percent of body weight that was fat free mass, and a decrease in the percent of body weight that was fat mass (Das et al., 2017), and CR did not negatively affect aerobic capacity (Racette et al., 2017).

The findings from CALERIE 2 indicate that CR is feasible and generally safe in adults without obesity. The lower level of CR achieved by participants in the normal weight BMI stratum compared to those in the overweight stratum at baseline suggests that leaner individuals may have experienced more difficulty adhering to CR, though this conclusion is confounded by the fact that the zone of adherence resulted in a greater percent CR for participants in the overweight BMI stratum. Further research is needed to evaluate the influence of weight status and BMI on adherence to a CR regimen.

This study has many strengths, including frequent TDEE and body composition assessments, which were necessary to estimate percent CR using the intake-balance method. An additional strength was the use of individualized weight graphs and a mathematical model to guide intervention delivery by estimating adherence to the 25% CR goal in real time throughout the two-year trial. The study also has limitations, including the inherent limitations in estimating percent CR with both the intake-balance method and a mathematical model and weight graph. Regarding estimation of percent CR with the intake-balance method, the method requires an accurate estimate of TDEE and changes in energy

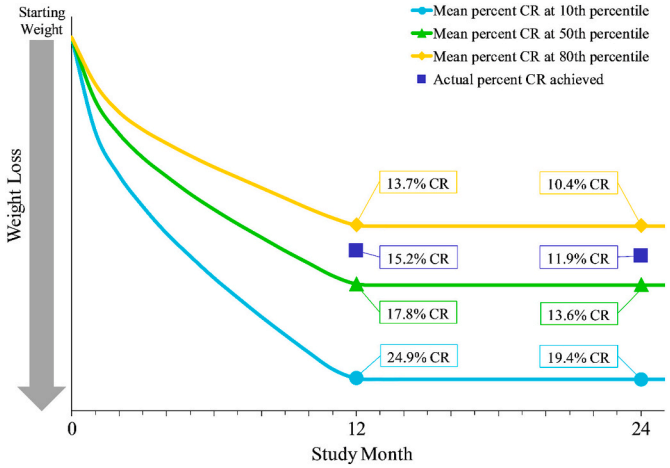


Fig. 2. Percent CR, determined by the Body Weight Planner, at months 12 and 24 for the upper (80th percentile; yellow line) and lower (10th percentile; blue line) bounds of the adherence zone, as well as the 50th percentile (green line). Actual percent CR achieved by participants at months 12 and 24 is depicted by the dark blue squares and was measured with the intake-balance method.

stores throughout the period of interest. It is not possible to obtain accurate estimates of TDEE throughout the intervention without frequent DLW measurements, which is impractical in most trials. Rather, mean TDEE was determined from DLW assessments at baseline and months 6, 12, 18, and 24, with the assumption that changes in TDEE were linear over time. Linear change in TDEE is unlikely since change in body weight, which is tightly associated with TDEE, is curvilinear, and changes in physical activity between DLW assessments will not be detected. Similarly, change in body composition requires repeated assessments with DXA or other techniques, and these measurements include inherent error, in addition to the error associated with the estimated energy costs of changes in fat mass and fat-free mass (Racette et al., 2012). A final limitation is the application of mathematical models of energy balance, as well as other techniques, to individual participants since the models are typically validated at the group level. Importantly, the mathematical model used in this study (Hall et al., 2011; Hall and Chow, 2011) provides valid estimates of energy intake (Sanghvi et al., 2015) and was different from the model used to direct intervention delivery in CALERIE 2 (Pieper et al., 2011).

5. Conclusions

The mathematical model and zone of adherence used in CALERIE 2 were novel and represent a pragmatic approach for estimating and promoting adherence to CR goals in real time. The clinical significance of the approach is exemplified by its integration into adaptive interventions that can be deployed remotely via mobile devices, such as smartphones and tablets (Martin et al., 2016b). Such interventions have been found to promote clinically significant weight loss of 9.4% among healthy adults when delivered remotely (Martin et al., 2015) and to decrease the proportion of pregnant women who exceed gestational weight gain guidelines (Redman et al., 2017). The zone of adherence in CALERIE 2, however, considered CR far less than the 25% goal as being adherent. This must be considered in designing CR interventions and strategies to promote adherence. For example, by structuring adherence zones that are lower, which would result in higher levels of CR being achieved when participants' weights were in the zone of adherence. The results also demonstrate the need to better understand the effects of sex, BMI, and race on zones of adherence, as well as intervention delivery. Specifically, research is needed to determine if the widths of adherence zones are sufficient to account for error in the models and to not hold some participants to a more stringent (or lenient) adherence metric.

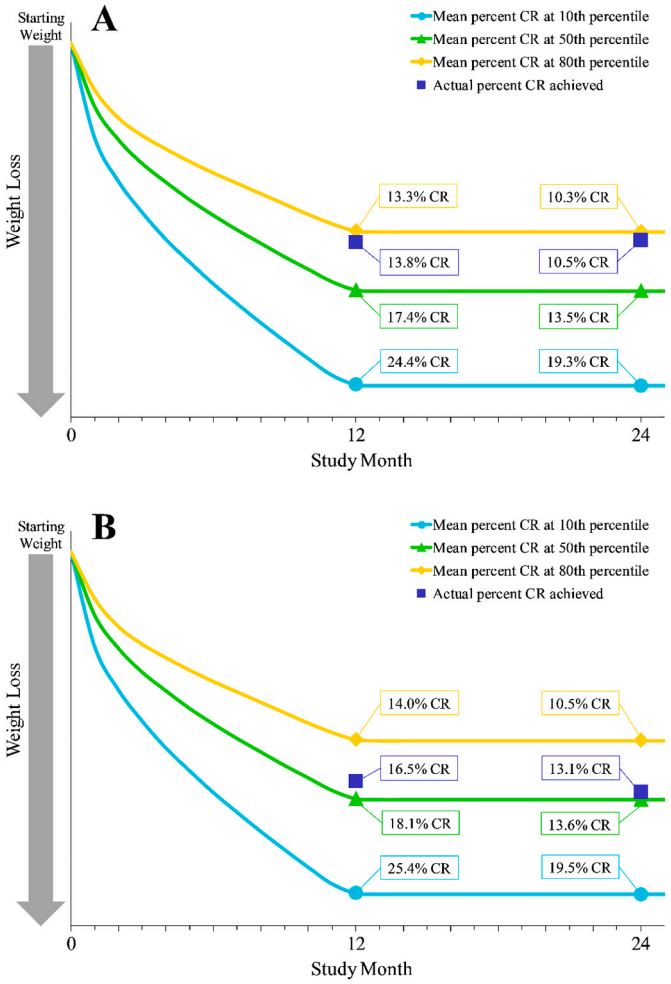


Fig. 3. Percent CR, determined by the Body Weight Planner, at months 12 and 24 for the normal weight BMI category (22.0–24.9 kg/m2, **Panel A**) and the overweight BMI category (25.0–27.9 kg/m2, **Panel B**). The yellow line depicts the upper (80th percentile) and the blue line the lower (10th percentile) bounds of the adherence zone. The green line depicts the 50th percentile. Actual percent CR achieved by participants at months 12 and 24 is depicted by the dark blue squares and was measured with the intake-balance method.

Finally, the way in which intervention success is evaluated after a trial requires further exploration since even state-of-the-art techniques, including the intake-balance method, have limitations and will not always align with measures of adherence used during intervention delivery.

Funding

The research was supported by the National Institute on Aging and the National Institute of Diabetes and Digestive and Kidney Diseases (grants U01AG022132, U01AG020478, U01AG020487, and U01AG020480), as well as the Nutrition Obesity Research Center (grant P30 DK072476), sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases; and the National Institute of General Medical Sciences of the National Institutes of Health, which funds the Louisiana Clinical and Translational Science Center (grant U54 GM104940). Author JLD was supported by American Heart Association grant 20POST35210907) and author CH was supported by a National Institutes of Health National Research Service Award (T32 DK064584).

Declaration of competing interest

The authors report no conflicts of interest.

Acknowledgments

We thank the entire study team and the study participants, who dedicated considerable time and effort over two years to volunteer in the project.

Appendix A. Investigators and Staff Participating in CALERIE (Comprehensive Assessment of the Long-term Effects of Reducing Intake of Energy)

The following is a list of the principal investigators (PIs), Co-investigators (CIs), site intervention leaders (SILs), intervention counselors (ICs), study managers (SMs), project leaders (PLs), study coordinators (SCs), and other staff (OS) participating in the CALERIE study.

Pennington Biomedical Research Center (clinical site)—PI: Eric Ravussin, PhD; CI: Catherine Champagne, RD, PhD; Alok Gupta, MD; Corby Martin, PhD; Leanne Redman, PhD; Steven Smith, MD; Donald Williamson, PhD; SIL: Corby Martin, PhD; Tiffany Stewart, PhD; IC: Michelle Begnaud, RD; Barbara Cerniauskas, RD; Allison Davis, MS; Jeanne Gabrielle, PhD; Heather Walden, MS; SM: Natalie Currier, RD; Mandy Shipp, RD; SC: Sarah Masters; Melody McNicoll; OS: Shelly Prince, MS, RD; Courtney Brock, RD; Renee Puyau, RD; Conrad Earnest, PhD; Jennifer Rood, PhD; Tiffany Stewart, PhD; Lillian Levitan, PhD; Crystal Traylor, WHNP; Susan Thomas, WHNP; Valerie Toups, LPN; Karen Jones, RN; Stephanie Tatum, RN; Celeste Waguespack, RN; Kimberly Crotwell, LPN; Lisa Dalfrey, LPN; Amy Braymer, LPN; Rhonda Hilliard, LPN; Onolee Thomas, RN; Jennifer Arceneaux, RN; Stacie LaPrarie, RN; Allison Strate, RN; Jana Ihrig, RN; Susan Mancuso, RN; Christy Beard, RN; Alicia Hymel; Desti Shepard; John Correa; Denise Jarreau; Brenda Dahmer; Grace Bella; Elizabeth Soroe; Bridget Conner; Paige McCown; Stephanie Anaya; Melissa Lupo.

Tufts University (clinical site)—PI: Susan B. Roberts, PhD; CI: Sai Krupa Das, PhD; Simin Meydani, PhD; Roger Fielding, MD; Isaac Greenberg, PhD; Anastassios Pittas, MD; Edward Saltzman, MD; Tammy Scott, PhD; SIL: Cheryl Gilhooly, RD, PhD; IC: Kimberly Gerber, PhD; Isaac Greenberg, PhD; Marjory Kaplan, PhD; Christy Karabetian, MA; Russell Kennedy, PhD; Lisa Robinson, RD; OS: Paul Fuss, Assefa, Senait; Verona Bembridge; Maria Berlis; Scarlett Buer; Robert Carabello; Cherie Campbell; Lauren Collins, RN; Marybeth Doherty, RN; Alicia Freed, RD; Chervonte Hernandez; Gyna Jean-Baptiste, RN; Mary Krasinski, RN; Marie Lim-Lucas, MPH, RD; Ekaterina Maslova; Barbara Maxwell, RN; Jean McShea, RN; Ann Muchowski, RN; Margaret Mulkerrin; Kerry Murphy; Carol Nelsen, RN; Megan O'Neill; Helen Rasmussen, RD, PhD; Brenda Roche; Eneida Roman; Gregory Sproull; Marie St. Victor, RN; Susan Storer, RN; Katherine Strissel, PhD; Stephanie Valliere; Margaret Vilme, RN; Justin Wheeler; Jill Wiley, RN; Fania Yangarber.

Washington University (clinical site)—PI: John O. Holloszy, MD; CI: Luigi Fontana, MD; Sam Klein, MD; Charles Lambert, PhD; B. Selma Mohammed, MD, PhD; Susan Racette, PhD; Dennis Villareal, MD; SIL: Rick Stein, PhD; IC: Karen Cotton, Psy D; Margaret Hof, MS, RD, LD; Cherie Massmann, MA, LPC, NCC; Kathleen Obert, MS, RD, LD; Marni Pearlman, MA, PLPC; Tina M Reising, Psy D; Laura Weber, MSEd, RD, LD; SM: Mary Uhrich, MS; SC: Morgan Schram, MS; OS: Mel Meyer, RN, BSN, CRC; Chelsea Carlen, BS; Lisa Kee, DTR; Barbara Larson, DTR; Mary McFerson, BS, DTR; Rebecca Sabatino, BS; Bridgett Toennies, RRT.

Duke Clinical Research Institute (coordinating center)—PI: James Rochon, PhD; CI: Connie W. Bales, PhD; Carl F. Pieper, DrPH; William Kraus, MD; PL: Katherine M. Galan, RN; OS: Richard Adrian, BS; Eleanor Law Allen, BA; William Blasko, BS; Manjushri Bhapkar, MS; Nikka Brown, BSN; Maria Butts, RN, BSN; Elaina K. Cossin, BS; Jennifer Curry, AAS; Jamie Daniel, BS, MS; Kathleen S. Diemer, RN; Lee Greiner, BS, MS; Darryl Johnson, BS; Cassandra Jones, BSEE; Lauren Lindblad, MS; Luanne McAdams, RN, MSN; Marty Mansfield, BA, PhD; Senthil Murugesan, MS; Lucy Piner, MS, ACSM CES; Christopher Plummer, BS; Mike Revoir, BS; Pamela Smith, RN, BSN; Monica Spaulding, MPH; James Topping, MS.

Baylor College of Medicine (doubly labeled water laboratory)—PI: William W. Wong, PhD; OS: Lucinda L. Clarke, AA; Chun W. Liu, BS; J. Kennard Fraley, MPH.

University of California at San Francisco (dual-energy x-ray absorptiometry reading center)—PI: Ann V. Schwartz, PhD; CI: John Shepherd, PhD; OS: Lisa Palermo, MS; Susan Ewing, MS; Michaela Rahorst; Caroline Navy.

University of Vermont (biochemistry laboratory)—PI: Michael Lewis, MD, MBA; CI: Russell P. Tracy, PhD; OS: Rebekah Boyle, BS, MS; Elaine Cornell, BS; Patrick Daunais, BS; Dean Draayer, PhD; Melissa Floersch, BS; Nicole Gagne, BA; Florence Keating, BS; Angela Patnoad, BS.

University of Cincinnati (nutrition reading center)—PI: Marcia Schmidt, MS, RD, LD; OS: Marcia Gavin BS, RD, LD; Frida Wiener MS, RD, LD; Ashley Hughes, DTR; Laura Benken.

University of Pittsburgh (intervention counseling curriculum)—PI: Amy Otto, PhD.

Data and safety monitoring board—Jeffrey Halter, MD (chair); David M. Buchner, MD, MPH; Patricia Elmer, PhD; Mark Espeland, PhD; Steven B. Heymsfield, MD; Xavier Pi-Sunyer, MD; Thomas Prohaska, MD; Sue Shapses, PhD; John Speakman, DSc; Richard Weindruch, PhD.

National Institute on Aging (primary funding agency)—Evan C. Hadley, MD; Judy Hannah, PhD; Sergei Romashkan, MD.

National Institute of Diabetes and Digestive and Kidney Diseases (cosponsor)—Mary Evans, PhD.

References

Bandini, L.G., Schoeller, D.A., Cyr, H.N., Dietz, W.H., 1990. Validity of reported energy intake in obese and nonobese adolescents. Am. J. Clin. Nutr. 52 (3), 421–425.
Beaton, G.H., Burema, J., Ritenbaugh, C., 1997. Errors in the interpretation of dietary assessments. Am. J. Clin. Nutr. 65 (4 Suppl), 1100S–1107S.
Das, S.K., Roberts, S.B., Bhapkar, M.V., et al., 2017. Body-composition changes in the comprehensive assessment of long-term effects of reducing intake of energy (CALERIE)-2 study: a 2-y randomized controlled trial of calorie restriction in nonobese humans. Am. J. Clin. Nutr. 105 (4), 913–927.
DeLany, J.P., Jakicic, J.M., Lowery, J.B., Hames, K.C., Kelley, D.E., Goodpaster, B.H., 2005. African american women exhibit similar adherence to intervention but lose less weight due to lower energy requirements. Int. J. Obes. 38 (9), 1147–1152, 2014.
Dorling, J.L., van Vliet, S., Huffman, K.M., et al., 2021. Effects of caloric restriction on human physiological, psychological, and behavioral outcomes: highlights from CALERIE phase 2. Nutr. Rev. 79 (1), 98–113.
Hall, K.D., Chow, C.C., 2011. Estimating changes in free-living energy intake and its confidence interval. Am. J. Clin. Nutr. 94 (1), 66–74.
Hall, K.D., Sacks, G., Chandramohan, D., et al., 2011. Quantification of the effect of energy imbalance on bodyweight. Lancet 378 (9793), 826–837.
Kebbe, M., Sparks, J.R., Flanagan, E.W., Redman, L.M., 2021. Beyond weight loss: current perspectives on the impact of calorie restriction on healthspan and lifespan. Expert. Rev. Endocrinol. Metab. 16 (3), 95–108.
Kraus, W.E., Bhapkar, M., Huffman, K.M., et al., 2019. 2 years of calorie restriction and cardiometabolic risk (CALERIE): exploratory outcomes of a multicentre, phase 2, randomised controlled trial. Lancet Diabetes Endocrinol. 7 (9), 673–683.
Martin, C.K., Miller, A.C., Thomas, D.M., Champagne, C.M., Han, H., Church, T., 2015. Efficacy of SmartLoss, a smartphone-based weight loss intervention: results from a randomized controlled trial. Obesity (Silver Spring) 23 (5), 935–942.
Martin, C.K., Bhapkar, M., Pittas, A.G., et al., 2016. Effect of calorie restriction on mood, quality of life, sleep, and sexual function in healthy nonobese adults: the CALERIE 2 randomized clinical trial. JAMA Intern. Med. 176 (6), 743–752.
Martin, C.K., Gilmore, L.A., Apolzan, J.W., Myers, C.A., Thomas, D.M., Redman, L.M., 2016. Smartloss: a personalized Mobile health intervention for weight management and health promotion. JMIR Mhealth Uhealth 4 (1), e18.
Pieper, C., Redman, L.M., Bapkar, M., et al., 2011. Development of adherence metrics for caloric restriction interventions. Clin. Trials 8 (2), 155–164.
Racette, S.B., Krupa Das, S., Bhapkar, M., et al., 2012. Approaches for quantifying energy intake and %calorie restriction during calorie restriction interventions in humans: the multicenter CALERIE study. Am. J. Physiol. Endocrinol. Metab. 201, E441–E448.
Racette, S.B., Rochon, J., Uhrich, M.L., et al., 2017. Effects of two years of calorie restriction on aerobic capacity and muscle strength. Med. Sci. Sports Exerc. 49 (11), 2240–2249.

Ravussin, E., Redman, L.M., Rochon, J., et al., 2015. A two-year randomized controlled trial of human caloric restriction: feasibility and effects on predictors of health span and longevity. J. Gerontol. 70 (9), 1097–1104.
Redman, L.M., Gilmore, L.A., Breaux, J., et al., 2017. Effectiveness of SmartMoms, a novel eHealth intervention for management of gestational weight gain: randomized controlled pilot trial. JMIR Mhealth Uhealth 5 (9), e133.
Rickman, A.D., Williamson, D.A., Martin, C.K., et al., 2011. The CALERIE study: design and methods of an innovative 25% caloric restriction intervention. Contemp. Clin. Trials 32, 874–881.
Rochon, J., Bales, C.W., Ravussin, E., et al., 2011. Design and conduct of the CALERIE study: comprehensive assessment of the long-term effects of reducing intake of energy. J. Gerontol. A Biol. Sci. Med. Sci. 66A (1), 97–108.
Romashkan, S.V., Das, S.K., Villareal, D.T., et al., 2016. Safety of two-year caloric restriction in non-obese healthy individuals. Oncotarget 7 (15), 19124–19133.
Sanghvi, A., Redman, L.M., Martin, C.K., Ravussin, E., Hall, K.D., 2015. Validation of an inexpensive and accurate mathematical method to measure long-term changes in free-living energy intake. Am. J. Clin. Nutr. 102 (2), 353–358.
Schoeller, D.A., Bandini, L.G., Dietz, W.H., 1990. Inaccuracies in self-reported intake identified by comparison with the doubly labelled water method. Can. J. Physiol. Pharmacol. 68 (7), 941–949.
Stewart, T.M., Bhapkar, M., Das, S., et al., 2013. Comprehensive assessment of long-term effects of reducing intake of energy phase 2 (CALERIE phase 2) screening and recruitment: methods and results. Contemp. Clin. Trials 34, 10–20.
Thomas, D.M., Martin, C.K., Heymsfield, S., Redman, L.M., Schoeller, D.A., Levine, J.A., 2011. A simple model predicting individual weight change in humans. J. Biol. Dyn. 5 (6), 579–599.
Tran, K.M., Johnson, R.K., Soultanakis, R.P., Matthews, D.E., 2000. In-person vs telephone-administered multiple-pass 24-hour recalls in women: validation with doubly labeled water. J. Am. Diet. Assoc. 100 (7), 777–783.
Wagner, D.R., Heyward, V.H., 2000. Measures of body composition in blacks and whites: a comparative review. Am. J. Clin. Nutr. 71 (6), 1392–1402.

Publication 14

Initial evidence of the acute effect of electronic nicotine delivery system use on energy intake.

Gabrielle T Maldonado [1,2], **Christoph Höchsmann** [3,4], Akansha Anbil [2], Karissa Neubig [4], Rabia Imran [2], Bernard F Fuemmeler [5], Thokozeni Lipato [6], Vineela Rachagiri [2], Andrew J Barnes [2,7], Corby K Martin [4], Caroline O Cobb [1,2]

[1] Department of Psychology, Virginia Commonwealth University, Richmond, VA, USA
[2] Center for the Study of Tobacco Products, Virginia Commonwealth University, Richmond, VA, USA
[3] TUM School of Medicine and Health, Technical University of Munich, Munich, Germany
[4] Pennington Biomedical Research Center, Baton Rouge, LA, USA
[5] Department of Family Medicine, Virginia Commonwealth University, Richmond, VA, USA
[6] Department of Internal Medicine, Virginia Commonwealth University, Richmond, VA, USA
[7] Department of Health Behavior and Policy, Virginia Commonwealth University, Richmond, VA, USA

Journal	Year	DOI	Impact Factor	SCImago	Citations to date
Experimental and Clinical Psycho-pharmacology, *ahead of print*	2024	10.1037/pha0000710	3.5 (2022)	Q2	Scopus: 0 Google Scholar: 1

Author contribution:

Second author; obtained funding, developed the research question, conceptualized the project including its methodology, supported manuscript writing, and provided critical revisions for important intellectual content.

AMERICAN PSYCHOLOGICAL ASSOCIATION

Experimental and Clinical Psychopharmacology

ISSN: 1064-1297

https://doi.org/10.1037/pha0000710

Initial Evidence of the Acute Effect of Electronic Nicotine Delivery System Use on Energy Intake

Gabrielle T. Maldonado[1, 2], Christoph Höchsmann[3, 4], Akansha Anbil[2], Karissa Neubig[3], Rabia Imran[2], Bernard F. Fuemmeler[5], Thokozeni Lipato[6], Vineela Rachagiri[2], Andrew J. Barnes[2, 7], Corby K. Martin[5], and Caroline O. Cobb[1, 2]

[1] Department of Psychology, Virginia Commonwealth University
[2] Center for the Study of Tobacco Products, Virginia Commonwealth University
[3] Ingestive Behavior, Weight Management and Health Promotion Laboratory, Pennington Biomedical Research Center, Louisiana State University System
[4] Department of Health and Sport Sciences, School of Medicine and Health, Technical University of Munich
[5] Department of Family Medicine and Population Health, Virginia Commonwealth University School of Medicine
[6] Department of Internal Medicine, Virginia Commonwealth University School of Medicine
[7] Department of Health Behavior and Policy, Virginia Commonwealth University School of Medicine

Previous work has aimed to disentangle the acute effects of nicotine and smoking on appetite with mixed findings. Electronic nicotine delivery systems (ENDS) have yet to be examined in this regard despite evidence of use for weight control. The present study tested the influence of an ENDS on acute energy intake and associated subjective effects. Participants ($n = 34$; 18–65 years) with current ENDS use completed two randomly ordered clinical lab sessions after overnight abstinence from tobacco/nicotine/food/drinks (other than water). Sessions differed by the product administered over 20 min: active (20 puffs of a JUUL ENDS device; 5% nicotine tobacco-flavored pod) or control (access to an uncharged JUUL with an empty pod). About 40 min after product administration, participants were provided an ad lib buffet-style meal with 21 food/drink items. Subjective ratings were assessed at baseline, after product use, and before/after the meal. Energy intake (kcal) was calculated using pre–post buffet item weights. Repeated measures analyses of variance and pairwise comparisons were used to detect differences by condition and time ($\alpha < .05$). Mean ± standard error of the mean energy intake did not differ significantly between active (1011.9 ± 98.8 kcal) and control (939.8 ± 88.4 kcal; $p = .108$) conditions. Nicotine abstinence symptoms significantly decreased after the active condition, while satiety significantly increased. Following the control condition, satiety remained constant while hunger significantly increased relative to baseline. Findings indicate that acute ENDS use did not significantly impact energy intake, but there was an ENDS-associated subjective increase in satiety and relative decrease in hunger. Results support further investigation of ENDS on appetite.

Public Health Significance

Perceptions that electronic nicotine delivery systems (ENDS) assist with appetite suppression and weight loss have been linked to initiation and may present challenges to ENDS cessation. The present results suggest that acute ENDS use did not significantly affect energy intake relative to the control condition. ENDS use was associated with increased satiety and lower subjective hunger. These findings can be used to develop prevention/intervention messaging related to ENDS use for weight control.

Keywords: electronic nicotine delivery systems, weight control, hunger, satiety, nicotine

Supplemental materials: https://doi.org/10.1037/pha0000710.supp

Gabrielle T. Maldonado https://orcid.org/0009-0006-9124-8969
Caroline O. Cobb https://orcid.org/0000-0003-3258-252X

Research reported in this publication was supported by the National Institute on Drug Abuse and U.S. Food and Drug Administration Center for Tobacco Products under Grant U54DA036105, the National Institute of Diabetes and Digestive and Kidney Diseases National Research Service Award T32DK064584 awarded to Christoph Höchsmann, a Nutrition Obesity Research Centers grant P30DK072476 entitled "Nutrition and Metabolic Health Through the Lifespan" sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases awarded to Corby K. Martin, and the National Institute of General Medical Sciences Grant U54GM104940, which funds the Louisiana Clinical and Translational Science Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Food and Drug Administration.

Corby K. Martin and Bernard F. Fuemmeler are named on a patent for a smartphone app that determines electronic cigarette device and liquid characteristics. Corby K. Martin is also named on a second patent application for a smoking cessation intervention.

Results from this project were presented at the 17th Annual Women's Day Health Conference on April 26, 2023; Virginia Commonwealth

continued

Weight control is a common motive for initiation and continued smoking, and fear of postcessation weight gain is cited as an important barrier to smoking cessation and/or relapse (Beebe & Bush, 2015; Pinto et al., 1999; White, 2012). The relations among smoking, appetite, and body weight are complex and incompletely understood, but evidence supports the hypothesis that nicotine delivered alone or via cigarettes reduces appetite, and ultimately body weight, by acting upon brain and hormonal mechanisms (Audrain-McGovern & Benowitz, 2011). Whether these effects extend to other tobacco products such as electronic nicotine delivery systems (ENDS or e-cigarettes) is unknown. Nonetheless, marketing for ENDS has included weight control messages (Lyu et al., 2022), and ENDS are being used for this purpose among people who smoke cigarettes and use ENDS (Jackson et al., 2019; Piñeiro et al., 2016; Strong et al., 2015). Better understanding of the effects of ENDS on appetite and weight-related factors could inform prevention and intervention efforts and considerations for people who use ENDS, including those who use ENDS as a harm reduction or smoking cessation aid (Hartmann-Boyce et al., 2021; National Academies of Sciences, Enginering, and Medicine, 2018).

Nicotine use for appetite suppression is complex, and there are multiple mechansims by which nicotine likely exerts this effect (Audrain-McGovern & Benowitz, 2011). For example, nicotine delivered via cigarettes or other methods is thought to act on the hormone leptin found in adipose tissue by increasing its quantity and regulating feelings of hunger and inducing satiety (Jo et al., 2002). Another potential hypothesis is that nicotine triggers a response in the propiomelanocortin cells in the hypothalamus, which play a critical role in food inhibition and increased energy expenditure (Picciotto & Mineur, 2013). Although much previous research has tried to disentangle these effects including preclinical rodent models (e.g., Grunberg et al., 1984; Rupprecht et al., 2016), acute clinical designs represent one means to better understand how nicotine use in various forms immediately impacts appetite, energy intake, and related subjective effects.

One of the earliest clinical studies on the effects of cigarette smoking on food intake found that cigarette smoking reduced consumption of and preference for sweet-tasting foods while abstinence from cigarette smoking increased sweet consumption and preference for sweets. Consumption of salty and bland foods was not significantly affected (Grunberg, 1982). Subsequent clinical lab work has revealed mixed effects of acute nicotine administration on appetite with some evidence of nicotine-associated supression of food/energy intake under certain conditions (Bulik et al., 1991; Perkins et al., 1991; Yannakoulia et al., 2018) and no effect or the opposite effect in others (Perkins et al., 1992, 1994). For example, in one examination, among 20 males (10 who smoked and 10 who did not) who abstained from food/smoking overnight, the hunger-reducing effects of nicotine only occurred following a simulated breakfast but not after water consumption. Nicotine also reduced energy intake during a test meal, with no effect of the type of food (sweet vs. high fat; Perkins et al., 1991). During a more recent study involving 14 males who smoked cigarettes, a significant reduction in energy intake was observed following a seven-item ad lib test meal when preceded by cigarette smoking (two cigarettes over 15 min) compared to a sham condition (holding an unlit cigarette), but no differences between conditions were observed for appetite-related subjective effects or hormones (Yannakoulia et al., 2018). This prior work highlights the potential influence of nicotine delivered via ENDS to impact acute energy intake and related subjective effects, and other forms of evidence reinforce this premise among people who use ENDS.

Among adolescents, ENDS use has been associated with intentions to lose weight (Mantey et al., 2020; Sanchez et al., 2021), and among young adults, higher weight concerns were associated with greater ENDS use frequency (Bennett & Pokhrel, 2018). The idea that ENDS can prevent weight gain after quitting smoking and their use for other weight control purposes has been endorsed by adults who use tobacco (Jackson et al., 2019). There is a subset of people who use ENDS who endorse use for weight management (Morean & Wedel, 2017).

Given previous mixed literature on the acute effects of nicotine on appetite and current trends in motivation for ENDS use and behavior, a well-controlled clinical laboratory study to assess the impact of ENDS on energy intake and associated subjective effects is warranted. Here, we tested the effect of an ENDS capable of a cigarette-like level of nicotine delivery (JUUL; Prochaska et al., 2022) as compared to a control condition (no nicotine or aerosol exposure) on energy intake during a buffet meal and associated subjective effects. Building on previous literature, we hypothesized there would be less energy intake following ENDS use compared to control. Further, we hypothesized that subjective feelings of hunger/food craving would decrease and satiety would increase following ENDS use compared to control.

Method

Overview

This study employed a two-condition randomized cross-over design with an active condition (20 puffs of a 5% nicotine Virginia Tobacco flavor JUUL pod) and a control condition (access to an

University Undergraduate Research Symposium on April 26, 2023, and at the International Conference for Eating Disorders in Washington, D.C. on June 1–2, 2023. This study was preregistered at https://ClinicalTrials.gov (No. NCT04219189).

All authors have contributed to the article in a significant way, and all have read and approved the article.

Gabrielle T. Maldonado played a lead role in data curation, formal analysis, investigation, methodology, project administration, validation, and writing–original draft. Christoph Höchsmann played a lead role in conceptualization and methodology and a supporting role in writing–review and editing. Akansha Anbil played a supporting role in project administration. Karissa Neubig played a supporting role in conceptualization and methodology. Rabia Imran played a supporting role in project administration and writing–review and editing. Bernard F. Fuemmeler played a supporting role in conceptualization and writing–review and editing. Thokozeni Lipato played a supporting role in project administration. Vineela Rachagiri played a supporting role in formal analysis and writing–review and editing. Andrew J. Barnes played a supporting role in formal analysis, methodology, supervision, and writing–review and editing. Corby K. Martin played a supporting role in conceptualization, methodology, supervision, and writing–review and editing. Caroline O. Cobb played a lead role in supervision and a supporting role in conceptualization, formal analysis, investigation, project administration, writing–original draft, and writing–review and editing.

Correspondence concerning this article should be addressed to Caroline O. Cobb, Department of Psychology, Virginia Commonwealth University, 100 West Franklin Street, Room 234, Richmond, VA 23220, United States. Email: cobbco@vcu.edu

empty JUUL-compatible pod and uncharged JUUL battery with the option to take up to 20 puffs). The criterion of 20 puffs (nicotine exposure estimated to be equivalent to ~1.5 cigarettes; Digard et al., 2013; Goniewicz et al., 2019) was chosen to increase the likelihood of nicotine-related effects on energy intake (similar to some prior research; Yannakoulia et al., 2018). Following screening for eligibility and obtainment of informed consent, participants took part in two sessions, about 1 week apart, in which they were randomly assigned the control or active condition followed by access to an ad lib test meal. No condition blinding was used, but participants were not informed of their condition order/assignment prior to sessions.

Participants

Thirty-four adult participants who reported either everyday ENDS use with a liquid concentration of at least 0.3% (3 mg/mL) nicotine or ENDS use at least three times a week at a liquid concentration of at least 3% (30 mg/mL) nicotine for the past 30 days completed both study sessions. Prior to participation, participants were screened for medical and psychological conditions that may have interfered with participant safety and their ability to successfully complete the study. Participants were also screened for any food allergies that could limit them from consuming any items provided in the buffet meal portion of the session. Participants also could not report alcohol use >25 days, cannabis use >20 days, or any illicit drug use in the past 30 days. Participants who were pregnant or breastfeeding were excluded using self-report and urinalysis. Participants who used progestin intrauterine devices, birth control injections (Depo-Provera, etc.), or had received a hysterectomy and still had ovaries were not eligible to participate; these criteria were used to ensure that cycle phase did not interfere with energy intake, as progesterone only intrauterine devices/birth control pills do not mimic a natural cycle. Informed consent from all participants was obtained prior to enrollment and full institutional review board approval of the study was approved prior to initiation of participant recruitment (IRB No. HM20018382).

Power Analysis

The target sample of 34 study completers was determined utilizing data from Yannakoulia et al. (2018) in which mean 152 (standard deviation [*SD*] = 190) kcal lower food intake was observed following the active cigarette condition versus the sham condition 45 min after the 15-min condition administration. Here, we assumed a more conservative mean difference of food intake between the active and control condition of 100 (*SD* = 200) kcal (Cohen's *d* effect size of 0.50) with power at 80% and a two-sided α level of 0.05.

Procedures and Materials

Prior to each session, participants abstained from any nicotine products, beverages (other than water), and food for 12 hr. All sessions occurred before 11:00 a.m. to help control for time of day, and for biologically female participants, sessions were scheduled during the luteal phase of their cycle (days 16–28 for an average cycle) to help control for hormonal effects on appetite. Presession abstinence was confirmed by obtaining participant exhaled carbon monoxide (CO) levels (must be half or lower of baseline CO level observed) and having participants sign an attestment form to affirm abstinence. Participants' blood pressure (BP) and heart rate were also obtained prior to the beginning of the session to ensure that they were within the appropriate safety limits (BP < 140/90). Participants also completed a presession questionnaire which inquired about any changes to their medical history since their last visit. Following the administration of the presession questionnaire, participants were asked to sit in the study room for an hour, during which they had access to researcher-approved reading items and puzzles to control for content that may influence appetite. This waiting period was used to ensure at least some abstinence from nicotine-containing products, as CO only measures exposure to combustible tobacco products.

After the waiting period, participants were provided the condition-specific product (active or control) and completed a subjective effect questionnaire at four time points during the session. Participants also completed the Food Craving Questionnaire–State (FCQ-state) at one time point during the session. Following a 35–40-min lapse after condition administration, participants partook in an up to 30-min ad lib meal.

Test Meal

The ad lib buffet meal consisted of a total of 21 food items in which each portion was set up identically across sessions and participants. Standard plate sizes (10.5"; ~23–26.2 g) and bowls (~15.6–16.2 g) were used for each session. Food items consisted of commercially available and ready-to-eat items. A mixture of salty (popcorn, nuts), savory (chicken tenders), sweet (M&Ms, swiss rolls), and fatty (cheese, cheese dip). All food items were provided at once with the average total calories provided ~4,228 kcal; which is in excess of a standard 2,000 kcal diet. Food items during Yannakoulia et al. (2018) were fewer in number (i.e., seven items vs. 21 items). Food items chosen were high in fat and in sugar content in order to promote consumption and were sized appropriately (0.5–1.5 cups) as to allow for repeated servings and to ensure that participants did not remember the exact consumption amount they consumed between sessions. Participants were instructed to eat until they felt comfortably full. Of note, previous work has shown that emulating a normal eating environment using a test buffet meal design is a suitable means for measuring intervention effects (Allirot et al., 2012). During the test meal, participants were in a room by themselves with unobtrusive monitoring via Zoom to ensure protocol compliance.

Baseline Measures

Participants' height and weight were measured at baseline and participants completed self-report items including gender, sex assigned at birth, race and ethnicity, current health conditions, and other measures not described here. The E-cigarette Dependence Scale (Morean et al., 2019) and Patient-Reported Outcomes Measurement Information System scale among participants who reported smoking cigarettes in the past 30 days (Shadel et al., 2014) were used to assess nicotine dependence. The Power of Food Scale was used to assess the psychological influence of existing in a food abundant environment (Lowe et al., 2009). Feelings toward appetite were assessed utilizing the Eating Inventory which was scored using a standardized method resulting in three domains: cognitive restraint, disinhibition, and hunger (Stunkard & Messick, 1985). To assess participant endorsement of ENDS use for body image concerns and as a way to control appetite an adapted version of the

Smoking-related Weight and Eating Episode Test was administered (adapted SWEET; Adams et al., 2011).

Energy and Macronutrient Intake

Energy (kcal) and macronutrient intake (total fat, saturated fat, cholesterol, etc.) were the primary outcome measures. These outcomes were measured by directly weighing the food provision and waste for each food item. Intake was calculated by taking the difference of pre and post-plate weights. Macronutrient details per gram for each food item were abstracted from food packaging and/or publicly available nutritional information found online.

Subjective Measures

Subjective questionnaires included questions that were targeted toward nicotine abstinence symptoms and related side effects using the adapted Minnesota Nicotine Withdrawal Scale (Hughes & Hatsukami, 1986; nine items) and the Direct Effects of Nicotine Scale (Evans et al., 2006; Perkins et al., 1993; 10 items). All items used a 100-point visual analog scale ranging from 0 = *not at all* to 100 = *extremely*.

Eleven appetite domains were rated via Visual Analog Scale items using the same 100-point line and the same or similar anchors as above (Flint et al., 2000; Parker et al., 2004; Stubbs et al., 2000). To measure the overall intensity of food craving, the FCQ-State was administered once, about 30 min after the condition administration and 5–10 min before the buffet meal. The scale consisted of 15 multiple choice items on a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree)* inquiring around the intensity of cravings for food. Questions included but were not limited to asking about intensity of desire to eat a certain food item, satiety around eating a specific food item, craving one or more specific foods, and so forth. Scoring of the FCQ–State involved utilizing the five-factor technique which is outlined by Cepeda-Benito et al. (2000): an intense desire to eat; anticipation of positive reinforcement that may result from eating; anticipation of relief from negative states and feelings as a result of eating; lack of control over eating; and craving as a physiological state (Cepeda-Benito et al., 2000; Meule, 2020).

Statistical Analysis

We characterized the sample descriptively using baseline measures. Then relevant coding and/or scale-specific calculations were performed. Statistical assumptions were reviewed and tested with appropriate transformations as needed. There were no missing data for energy intake. Missing data across subjective items ranged from 0% to 2.9% and on average was 1.5%. Due to the low proportion of missing data, we used mean replacement to correct relevant subjective items to ensure their inclusion in subsequent statistical analyses. We initially performed a mixed analysis of variance (ANOVA) to examine the effect of condition (active, control; within-subjects) and condition order (active then control, control then active; between-subjects) on energy intake. The main effect of condition order and the interaction of condition order and condition were not statistically significant for energy intake: condition order, $df(1, 32) = 0.795$, $p = .373$; Condition × Condition Order, $df(1, 32) = 0.411$, $p = .526$); subsequent analyses excluded condition order as a between-subjects factor for parsimony and power-related reasons, resulting in a repeated measures ANOVA.

For energy and macronutrient intake indices and the FCQ-state, a repeated measures ANOVA was used to detect differences by condition (active, control). For all other subjective measures, a two-way repeated measures ANOVA was used to test differences by condition (active, control) and time (1, 2, 3, 4). For all ANOVAs, adjustments for sphericity violations were assessed, and Huynh–Feldt correction values were reported. Post hoc testing with a Bonferroni corrected repeated measures *t* test was used to evaluate significant model results (McHugh, 2011). All data, study materials, and analysis code are available by request. We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Results

Participant Characteristics

A total of 52 participants provided informed consent, and of those, 34 successfully completed the study. The remaining 18 were determined ineligible due to failure to meet specific study criteria (i.e., food allergies, hormonal therapy/birth control that would interfere with food intake, and specific medical criteria), or self-withdrew due to scheduling conflicts ($n = 6$).

Detailed demographic and psychosocial characteristics of the 34 completers can be found in Table 1. In terms of sex assigned at birth, 58.8% were male and 41.2% were female, with gender distributions being similar in proportion of sex assigned at birth with the exception of one individual listed as nonbinary and two individuals identifying as men who were female at birth. Our sample consisted primarily of young adults with the average age of 25.7 years ($SD = 8.4$). The sample was relatively diverse in race/ethnicity with 14.7% identifying as Black or African American, 11.8% identifying as Middle Eastern, 44.1% identifying as White, and 8.8% identifying as more than one race. Additionally, 14.7% of participants identified as Hispanic/Latino or of Spanish origin. The majority of the sample (61.8%) reported a household income of equal to or greater than $50,000 and reported some college or higher education (82.4%) and were either currently employed (38.2%) or a student (44.1%).

The Supplemental Table S1 describes alcohol, cannabis, and tobacco use characteristics in further detail. Relevant to our study, a little over 1/3 of the sample preferred JUUL (35.3%) and most preferred nontobacco flavored ENDS (91.2%). Other baseline characteristics related to eating behavior and physical characteristics are displayed in the Supplemental Table S2. The majority of participants (67.6%) were not concerned about weight gain if they were to stop using ENDS. Of those that did express a concern ($n = 11$), 63.6% reported that weight gain concerns influenced their decision to quit ENDS. The Power of Food scale demonstrated that there was only a minimal to moderate effect of the influence of food on appetite despite living in a food abundant environment ($M = 2.7$, $SD = 0.8$). The adapted SWEET test scale indicated that there were minimal concerns around ENDS use for weight control purposes ($M = 1.1$, $SD = 0.8$). The Pearson Eating Inventory indicated that on average participants scored within the low to average range for cognitive restraint, disinhibition, and hunger.

Energy and Macronutrient Intake

Analysis of energy intake (kcal) revealed no significant main effect of condition ($p = .108$) and a partial η-squared effect size of medium ($\eta_p^2 = 0.076$; see Table 2). Descriptive examination by

Table 1
Sample Demographics

Characteristic	Sample size $n = 34$
Age (years), *M* (*SD*)	25.7 (8.4)
Gender identity, *n* (%)	
Female/woman/she/her	11 (32.4)
Male/man/he/him	21 (61.8)
Nonbinary	1 (2.9)
Missing	1 (2.9)
Sex assigned at birth, *n* (%)	
Male	20 (58.8)
Female	14 (41.2)
Race, *n* (%)	
Asian	5 (14.7)
Black or African American	5 (14.7)
Middle Eastern	4 (11.8)
More than one race	3 (8.8)
White	15 (44.1)
Preferred to self-describe	2 (5.9)
Hispanic/Latino/of Spanish origin, *n* (%)	
No	29 (85.3)
Yes	5 (14.7)
Ancestry *n* (%) $n = 5$	
Cuban/Cuban American	2 (40.0)
Mexican/Mexican American	1 (20.0)
Central/South American	2 (40.0)
Annual Income, *n* (%)	
Below $50,000	9 (26.5)
$50,000 or greater	22 (61.8)
Do not know	4 (11.8)
Education, *n* (%)	
High school graduate/GED	6 (17.6)
Some college or higher	28 (82.4)
Employment status, *n* (%)	
Working now	13 (38.2)
Not working	6 (17.6)
Student	15 (44.1)

Note. For employment status, working (includes full-time, part-time, and military), not working (includes only temporarily laid off/sick leave, nonworking disabled permanent or temporary, looking for work/unemployed, keeping house, retired, and nonworking student, full-time student is an additional category). For education, GED corresponds to General Education Development diploma.

condition indicated a slightly higher mean (±standard error of the mean) energy intake during the active condition (1011.9 ± 98.8 kcal) compared to the control condition (939.8 ± 88.4 kcal). Relatedly, descriptive examination of overall food intake in grams was similar to between active (741.9 ± 61.1 g) and control (737.9 ± 56.7 g) conditions. No significant main effects of condition were observed for any macronutrient examined ($F < 3.6$, $p > .069$) with effect sizes ranging from small to medium (see Table 2).

Exploratory analyses evaluated the influence of sex assigned at birth and past 30-day cigarette use on energy intake. There were no significant main effects or interactions involving either factor (all $p > .13$).

Subjective Effects

Tobacco/Nicotine Abstinence Symptoms and Nicotine-Related Effects

Across the adapted Minnesota Nicotine Withdrawal Scale and Direct Effects of Nicotine Scale there were 12 items that had a significant condition by time interaction ($F > 2.9$, $p < .05$; see Table 3): urges to use an e-cigarette, irritability/frustration/anger, difficulty concentrating, restlessness, impatient, craving an e-cigarette, urges to smoke a cigarette, craving a cigarette, nauseous, dizzy, lightheaded, and heart pounding. Effect size estimates for these items were medium to large. These items were examined between conditions at each time point and by time point within condition (changes relative to time 0 or baseline) using paired sample *t* tests with a Bonferroni correction (10 total comparisons: $p < .005$).

Urges to use an e-cigarette was the item with the largest *F*-value for the interaction ($F = 30.9$; see Figure 1A). Relative to baseline (65.2 ± 4.8), mean urges decreased significantly following condition administration (29.4 ± 4.0; $p < .001$) and remained significantly decreased for the remainder of the session ($p < .001$). No significant changes relative to baseline were observed for the control condition. Between conditions, urges to use an e-cigarette were significantly lower for active compared to control immediately following condition administration (29.4 ± 4.0 vs. 73.3 ± 4.0; $p < .001$) and for the remainder of the session. Similar patterns to urges to use an e-cigarette were observed for the craving an e-cigarette item.

For irritability/frustration/anger in the active condition, relative to baseline (25.9 ± 4.2), mean scores decreased significantly following the condition administration *n* (10.6 ± 2.3; $p < .001$) and after the buffet meal (9.4 ± 3.4; $p < .001$; see Figure 1B). No significant changes relative to baseline were observed in the control condition. At every time point following condition administration, scores for the active condition were significantly lower than control (all $p < .001$) with the largest difference at 20 min (control 35.0 ± 5.3 vs. active 10.6 ± 2.3). A relatively similar pattern of responding was observed for difficulty concentrating, restlessness, impatient, urges to smoke a cigarette, and craving a cigarette with lower scores for the active condition relative to control following condition administration.

Dizziness had the largest *F*-value for the interaction ($F = 6.3$) among the Direct Effects of Nicotine Scale items, and the only significant difference observed was directly following the condition administration, with higher mean values for the active condition compared to control (20.5 ± 3.9 vs. 8.0 ± 2.1; $p = .002$); no other significant differences between conditions were observed. A

Table 2
Statistical Analysis Results for Energy Intake

	Condition		
Outcome	*F*	*p*	η_p^2
Kilocalories	2.7	.108	0.076
Total fat	1.4	.251	0.040
Saturated fat	1.3	.260	0.038
Cholesterol	0.2	.632	0.007
Sodium	1.9	.180	0.054
Carbohydrate	1.3	.259	0.038
Fiber	0.5	.506	0.014
Sugars	2.4	.134	0.067
Added sugar	2.7	.107	0.077
Protein	3.5	.070	0.096
Vitamin D	<0.1	.950	<0.001
Iron	0.4	.556	0.011
Calcium	<0.1	.858	0.001
Potassium	<0.1	.998	<0.001

Note. *df*(1, 33). Used Huynh–Feldt Correction.

Table 3
Statistical Analysis Results for Subjective Measures

Outcome	Condition (*C*)			Time (*T*)			*C* × *T*		
	F	*p*	η_p^2	*F*	*p*	η_p^2	*F*	*p*	η_p^2
Adapted Minnesota Nicotine Withdrawal Scale[a]									
Urges to use an e-cigarette	63.6	**<.001**	0.658	8.6	**<.001**	0.206	30.9	**<.001**	0.483
Irritability/frustration/anger	21.7	**<.001**	0.397	10.5	**<.001**	0.241	7.5	**<.001**	0.185
Difficulty concentrating	4.9	**.034**	0.129	21.0	**<.001**	0.389	5.3	**.002**	0.138
Restlessness	3.2	.085	0.087	12.4	**<.001**	0.273	3.1	**.036**	0.086
Impatient	13.1	**<.001**	0.285	14.4	**<.001**	0.304	5.6	**.002**	0.145
Craving an e-cigarette	71.3	**<.001**	0.684	8.5	**<.001**	0.204	29.3	**<.001**	0.470
Drowsiness	1.6	.212	0.047	25.1	**<.001**	0.432	2.5	.077	0.071
Urges to smoke a cigarette	11.7	**.002**	0.261	6.9	**<.001**	0.172	3.3	**.036**	0.090
Craving a cigarette	11.4	**.002**	0.257	10.2	**<.001**	0.237	4.6	**.011**	0.122
The Direct Effects of Nicotine Scale[a]									
Nauseous	0.2	.630	0.007	4.3	**.008**	0.114	3.0	**.045**	0.084
Dizzy	2.4	.133	0.067	11.6	**<.001**	0.261	6.3	**.002**	0.160
Lightheaded	0.0	.885	0.001	14.6	**<.001**	0.307	4.2	**.008**	0.113
Nervous	1.7	.196	0.050	2.7	.076	0.075	0.2	.871	0.007
Sweaty	2.2	.145	0.063	3.2	.055	0.087	2.1	.119	0.060
Headache	4.1	.051	0.110	10.9	**<.001**	0.248	2.4	.086	0.067
Excessive salvation	1.0	.317	0.030	0.8	.449	0.023	1.8	.151	0.053
Heart pounding	0.1	.810	0.002	5.5	**.003**	0.143	4.8	**.004**	0.126
Confused	0.1	.743	0.003	0.9	.418	0.027	0.7	.500	0.021
Weak	0.3	.575	0.010	15.2	**<.001**	0.315	0.2	.792	0.007
Pennington Visual Analog Scale									
How sad do you feel at the moment?[a]	1.1	.293	0.034	7.4	**<.001**	0.184	1.0	.374	0.030
How happy do you feel at the moment?[a]	2.9	.099	0.080	14.6	**<.001**	0.307	0.4	.698	0.013
How anxious do you feel at the moment?[a]	0.0	.975	<0.001	12.7	**<.001**	0.278	5.6	**.001**	0.145
How hungry do you feel at moment?[a]	3.8	.060	0.103	125.7	**<.001**	0.792	7.4	**<.001**	0.184
How full do you feel at the moment?[a]	2.0	.170	0.056	162.2	**<.001**	0.831	0.8	.482	0.024
How satisfied do you feel at the moment?[a]	25.6	**<.001**	0.437	90.5	**<.001**	0.733	4.3	**.011**	0.115
How much do you think you can eat right now?[a]	2.8	.106	0.077	95.2	**<.001**	0.743	2.4	.081	0.067
Would you like to eat something sweet?[a]	2.9	.097	0.081	36.6	**<.001**	0.526	1.3	.277	0.038
Would you like to eat something salty? ($n = 30$)[b]	1.2	.279	0.040	59.8	**<.001**	0.674	2.8	**.045**	0.088
Would you like to eat something savory? ($n = 30$)[b]	1.0	.333	0.032	76.1	**<.001**	0.724	1.4	.264	0.045
Would you like to eat something fatty? ($n = 30$)[b]	0.9	.352	0.030	45.3	**<.001**	0.610	1.5	.238	0.048

Note. Bold indicates $p < .05$.
[a] Condition (*C*) *df*(1, 33), Time (*T*) *df*(3, 31), *C* × *T* *df*(3, 31). [b] *C* *df*(1, 29), *T* *df*(3, 27), *C* × *T* *df*(3, 27); Used Huynh–Feldt Correction.

relatively similar pattern was observed for nausea, lightheadedness, and heart pounding.

Hunger/Satiety/Appetite-Related Effects

Of the 11 Pennington Visual Analog Scale items, four items had a significant condition by time interaction with a medium to large effect size: "How anxious do you feel at the moment?," "How hungry do you feel at the moment?," "How satisfied do you feel at the moment?," and "Would you like to eat something salty?" The item assessing hunger had the largest *F*-value for the interaction ($F = 7.4$; see Figure 1C). Regarding feelings of hunger in the control condition, ratings significantly increased relative to baseline (56.9 ± 4.3) following condition administration (64.8 ± 4.3) and remained high before the buffet meal (71.0 ± 3.8; all $p < .001$); there was less change in the active condition during this same time period. Hunger ratings for both conditions significantly decreased relative to baseline following the buffet meal ($p < .001$), and there were no significant between-condition differences at any time point.

For feelings of satiety, relative to baseline, there was a significant increase directly after active condition administration (26.6 ± 3.4–39.3 ± 3.4; $p < .001$) but not for the control condition (see Figure 1D). After administration of the buffet meal, there was a significant increase in satiety relative to baseline for both conditions to a similar extent ($p < .001$). Between conditions, the active condition resulted in significantly higher satiety ratings than the control, immediately following condition administration (39.3 ± 3.4 vs. 21.2 ± 3.4; $p < .001$) and directly before the buffet meal (33.8 ± 3.6 vs. 19.5 ± 3.6, $p < .001$). A relatively similar pattern of responding was observed for feelings of anxiety at the current moment. Relative to baseline, there was a significant decrease in cravings of something salty (n = 30) in the active and control conditions directly after the buffet meal administration ($p < .001$). There were no other significant differences relative to baseline or between conditions.

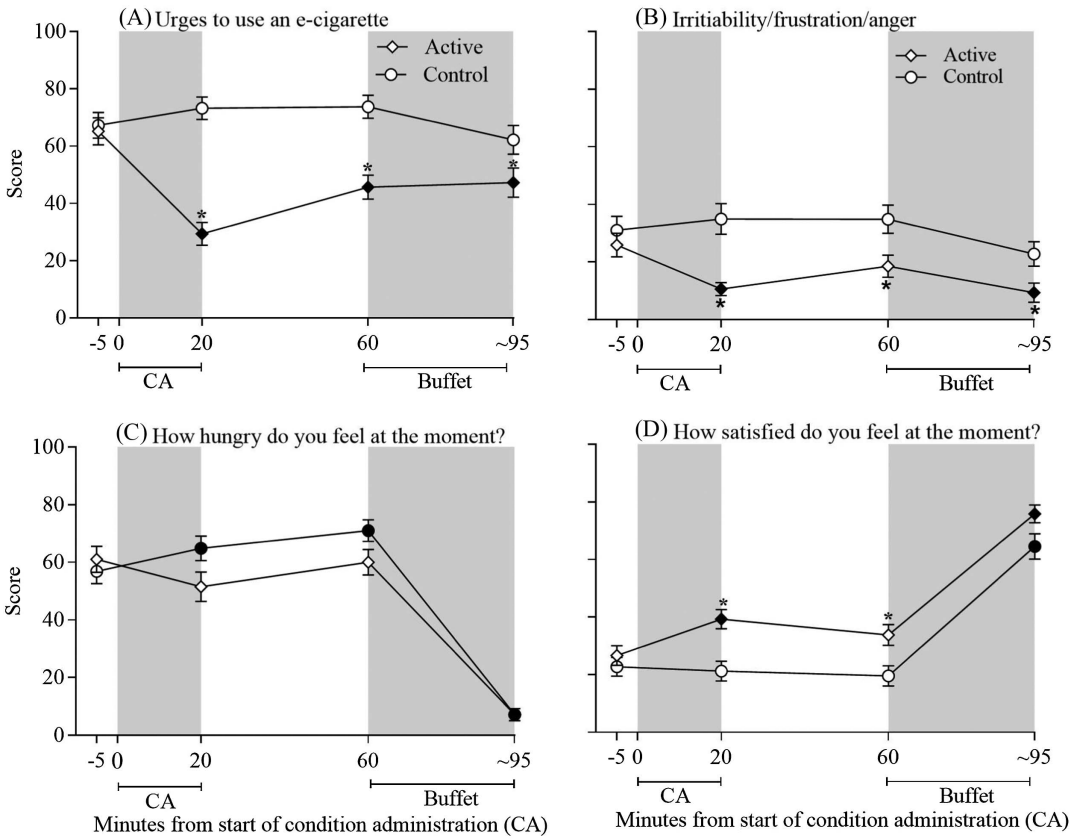
For the FCQ-state overall score and five individual factors, there were no significant main effects of condition ($F < 2.7$, $p > .112$) and small to medium effect sizes were observed (see Table 4).

Discussion

This study examined the acute effect of an ENDS capable of cigarette levels of nicotine delivery when used by people experienced with ENDS use (Prochaska et al., 2022) on energy intake during an ad lib buffet meal and associated subjective effects

Figure 1

Mean ± SEM for the Subjective Items Urges to Use an e-Cigarette (Panel A), Irritability/Frustration/Anger (Panel B), Hunger (Panel C), and Satiety (Panel D)



Note. Filled symbols represent a significant difference relative to −5 min, and asterisks represent a significant difference between conditions at that time point (all $p < .005$). *SEM* = standard error of the mean.

using a randomized cross-over design. To our knowledge, this is the first clinical lab study examining this effect with acute ENDS use; and there are mixed findings from similar work using cigarettes and nicotine nasal spray (Bulik et al., 1991; Grunberg, 1982; Perkins et al., 1991, 1992, 1994; Yannakoulia et al., 2018).

Our hypothesis that relative to the control condition, the active nicotine-containing ENDS condition would result in a significant decrease in energy and macronutrient intake during the ad lib meal was not supported; there was not a significant difference in energy or macronutrient intake. The lack of difference is consistent with results from some previous literature performed with cigarettes and nicotine nasal spray under various conditions (Grunberg, 1982; Perkins et al., 1992, 1994). However, these findings were inconsistent with smoking-related energy intake suppression observed in two previous studies examining the acute effects of smoking two own brand cigarettes in 15 min (Yannakoulia et al., 2018) and smoking

Table 4

Statistical Analysis Results and Descriptives for Food Craving Questionnaire–State

Food Craving Questionnaire (FCQ)–State	Condition *F*	*p*	η_p^2	Active *M* ± *SEM*	Control *M* ± *SEM*
FCQ–state total score	<0.1	.861	0.001	49.0 ± 1.6	49.5 ± 2.1
An intense desire to eat	1.1	.308	0.032	10.4 ± 0.5	9.6 ± 0.6
Anticipation of positive reinforcement that may result from eating	0.5	.472	0.016	10.7 ± 0.4	10.2 ± 0.5
Anticipation of relief from negative states and feelings as a result of eating	2.6	.113	0.074	10.2 ± 0.4	11.0 ± 0.5
Lack of control over eating	0.8	.369	0.024	6.9 ± 0.5	7.2 ± 0.4
Craving as a physiological state	1.4	.239	0.042	10.9 ± 0.4	11.4 ± 0.4

Note. *df*(1,33). *SEM* = standard error of the mean.

eight own brand cigarettes over a 4-hr period (Bulik et al., 1991). Compared to the more recent report (Yannakoulia et al., 2018), we observed slightly higher levels of energy intake (900–1,000 kcal vs. 700–900 kcal) and more variability (as indexed by *SD*; 500–600 kcal vs. 200–300 kcal) between participants. Based on our estimated effect size (0.5) from Yannakoulia et al. (2018), the condition-related effect observed in this study was lower than expected (0.28). Of note, our sample was 2.4 times larger and diverse in race/ethnicity and sex (e.g., 38% nonmale vs. 0% nonmale) which likely increased variability. Although JUUL was among the most preferred ENDS brands among our participants, more than 60% preferred other brands, and more than 90% preferred ENDS flavors other than tobacco. Condition instructions attempted to ensure equivalent ENDS use and associated nicotine delivery between participants, but it is possible that participants altered their puff topography (as in Hiler et al., 2017) and/or found the active condition aversive. Future work should consider the use of an own brand ENDS condition. Even considering these limitations, results of the present study suggest acute nicotine-containing ENDS use does not suppress energy intake relative to no ENDS use, and these effects may not be comparable to those observed with cigarette smoking under similar conditions.

In contrast to the energy intake effects observed in this study, consistent with our hypothesis, there was a significant increase in feelings of hunger following the control condition administration but not the active condition and a significant increase in satiety following the active condition administration but not the control condition. Interestingly, food-related cravings (indexed by the FCQ-state) measured immediately prior to the buffet meal indicated no significant differences between conditions across multiple factors. The FCQ-state focuses specifically on the intensity of food craving, especially around certain food deprivations, cues to certain foods, and food intake (Meule, 2020). Higher scores on the FCQ-state have been correlated with increased caloric consumption (Ng & Davis, 2013), but this measure also has been shown to be influenced by environmental factors that contribute to cravings of certain foods, such as deprivation of a particular food item (i.e., sweet or savory item; Meule et al., 2014). Some consider specific food cravings (i.e., feelings directed toward a certain food item, flavor, or texture) to be differentiated from physiological feelings of hunger encompassing stomach growling, irritability, and dizziness associated with not eating (Meule, 2020). This idea may help explain the discrepancy between subjective ratings of hunger and food-related cravings in this study.

Compared to six prior studies in this area (Bulik et al., 1991, Grunberg, 1982; Perkins et al., 1991, 1992; Perkins et al., 1994; Yannakoulia et al., 2018), our subjective assessments are among the most comprehensive to date with the inclusion of eating-related items as well as nicotine abstinence symptomology and nicotine-related side effects. Among the four studies that included a subjective measure of hunger, there was only one where nicotine-related suppression was observed (Perkins et al., 1991) as in the present study. In the three previous studies that measured condition-related effects on subjective satiety, there were no condition-related differences, unlike the present study. Taken together, this work highlights a dearth of prior evidence indicating acute nicotine administration suppresses subjective feelings of hunger and food craving and increases satiety.

Other findings more consistent with prior work indicated that the active condition was effective in reducing nicotine abstinence-related symptoms including urges and craving, irritability, and difficulty concentrating. The scope of this suppression is consistent with studies evaluating the acute effects of JUUL and other ENDS following acute administration (Hiler et al., 2017; Maloney et al., 2020). Similar findings in cigarette-related craving relief has been observed in four previous acute studies following the administration of nicotine-containing products (Bulik et al., 1991; Perkins et al., 1992, 1994; Yannakoulia et al., 2018). Consistent with a previous acute evaluation of nicotine delivered via nasal spray (Perkins et al., 1993) but not two prior acute ENDS evaluations (Hiler et al., 2017; Yingst et al., 2019), small increases in subjective ratings for several nicotine-related side effects were observed following the active condition administration in this study. These effects are likely due to the stimulant properties associated with nicotine (Benowitz, 2009) and might have been enhanced due to overnight nicotine/tobacco and food abstinence.

In summary, the active condition suppressed subjective feelings of hunger and nicotine abstinence symptoms, and increased satiety and nicotine-associated side effects. As seen in this study, subjective effects may not be directly driving energy intake, but these effects may belie other self-reports of ENDS use for weight control purposes (Ganson & Nagata, 2021; Morean et al., 2020; Morean & Wedel, 2017). These findings could lead to actionable approaches toward enhancing ENDS cessation efforts, with a particular focus on addressing perceptions of ENDS-related appetite control and/or weight management. Further research is needed to identify what drives these perceptions and how to best implement strategies to counteract them.

Limitations

In terms of limitations, we recruited a convenience sample of individuals who currently use ENDS from the Greater Richmond Area; thus, our findings may not generalize to different populations and regions. In addition, only 1/3 of participants expressed fear regarding weight gain post-ENDS cessation. Findings may differ for these individuals or those who use ENDS for appetite/weight control. Participants were asked to maintain overnight abstinence from food and drinks (other than water) and nicotine-containing products which was confirmed by imperfect measures. Thus, it is possible that participants may not have adhered to protocol instructions. A study that involves an in-person and/or monitored abstinence period may be a more effective way to deal with this concern. The present study attempted to provide some control in this respect by using a 1-hr rest period at the start of each session.

Another limitation in our study were the conditions utilized. Due to the nature of the conditions, blinding was not possible and condition-related expectancies may have influenced responding. In addition, although ~35% of participants in our study preferred JUUL, the remainder preferred another ENDS brand and may have lacked experience with JUUL. These product characteristics may have influenced participant use behavior and the amount of nicotine that was absorbed. Given that we did not collect puff topography or blood nicotine levels, the extent of this limitation is difficult to conclude.

Another limitation common to controlled eating paradigms was that participants received the buffet meal in a lab environment while being observed by the researcher. While we took steps to minimize this effect (i.e., unobtrusive monitoring), the disclosure of monitoring and the lab setting may have influenced participants' eating behavior. Future work could consider assessing participant

perceptions of the buffet environment and the influence of the setting on their behavior as well as the impact of varying meal sizes (i.e., modest snack). Meal size and timing may be critical in understanding nicotine's ability to potentially suppress appetite (see Perkins et al., 1991).

Conclusions

This clinical lab study examined the acute effects of nicotine-containing ENDS use on energy intake and associated subjective effects. Findings indicate that acute ENDS use following overnight abstinence did not significantly impact energy intake relative to a control condition, but there was an ENDS-associated decrease in feelings of hunger and increase in feelings of satiety. Other subjective effects suggested ENDS use was effective in reducing nicotine abstinence symptoms and produced mild nicotine-related side effects.

Building from this work, future research is needed to explore perceptions of and reasons for ENDS use as an appetite and/or weight control method as well as examine other features of ENDS including liquid nicotine concentration and flavor and their impact on appetite and hunger. Considering recent increases in ENDS use among younger populations, lack of tailored ENDS cessation programs, and rapidly changing regulatory environment, this information can help guide tobacco prevention/intervention efforts and tobacco policy aimed to decrease negative health consequences associated with ENDS use.

References

Adams, C. E., Baillie, L. E., & Copeland, A. L. (2011). The Smoking-Related Weight and Eating Episodes Test (SWEET): Development and preliminary validation. *Nicotine & Tobacco Research*, *13*(11), 1123–1131. https://doi.org/10.1093/ntr/ntr162

Allirot, X., Saulais, L., Disse, E., Roth, H., Cazal, C., & Laville, M. (2012). Validation of a buffet meal design in an experimental restaurant. *Appetite*, *58*(3), 889–897. https://doi.org/10.1016/j.appet.2012.02.011

Audrain-McGovern, J., & Benowitz, N. L. (2011). Cigarette smoking, nicotine, and body weight. *Clinical Pharmacology and Therapeutics*, *90*(1), 164–168. https://doi.org/10.1038/clpt.2011.105

Beebe, L. A., & Bush, T. (2015). Post-cessation weight concerns among women calling a state tobacco quitline. *American Journal of Preventive Medicine*, *48*(Suppl. 1), S61–S64. https://doi.org/10.1016/j.amepre.2014.09.004

Bennett, B. L., & Pokhrel, P. (2018). Weight concerns and use of cigarettes and e-cigarettes among young adults. *International Journal of Environmental Research and Public Health*, *15*(6), Article 1084. https://doi.org/10.3390/ijerph15061084

Benowitz, N. L. (2009). Pharmacology of nicotine: Addiction, smoking-induced disease, and therapeutics. *Annual Review of Pharmacology and Toxicology*, *49*(1), 57–71. https://doi.org/10.1146/annurev.pharmtox.48.113006.094742

Bulik, C. M., Dahl, R. E., Epstein, L. H., & Kaye, W. H. (1991). The effects of smoking deprivation on caloric intake in women with bulimia nervosa. *International Journal of Eating Disorders*, *10*(4), 451–459. https://doi.org/10.1002/1098-108X(199107)10:4%3C451::AID-EAT2260100410%3E3.0.CO;2-%23

Cepeda-Benito, A., Gleaves, D. H., Fernández, M. C., Vila, J., Williams, T. L., & Reynoso, J. (2000). The development and validation of Spanish versions of the State and Trait Food Cravings Questionnaires. *Behaviour Research and Therapy*, *38*(11), 1125–1138. https://doi.org/10.1016/S0005-7967(99)00141-2

Digard, H., Proctor, C., Kulasekaran, A., Malmqvist, U., & Richter, A. (2013). Determination of nicotine absorption from multiple tobacco products and nicotine gum. *Nicotine & Tobacco Research*, *15*(1), 255–261. https://doi.org/10.1093/ntr/nts123

Evans, S. E., Blank, M., Sams, C., Weaver, M. F., & Eissenberg, T. (2006). Transdermal nicotine-induced tobacco abstinence symptom suppression: Nicotine dose and smokers' gender. *Experimental and Clinical Psychopharmacology*, *14*(2), 121–135. https://doi.org/10.1037/1064-1297.14.2.121

Flint, A., Raben, A., Blundell, J. E., & Astrup, A. (2000). Reproducibility, power and validity of visual analogue scales in assessment of appetite sensations in single test meal studies. *International Journal of Obesity*, *24*(1), 38–48. https://doi.org/10.1038/sj.ijo.0801083

Ganson, K. T., & Nagata, J. M. (2021). Associations between vaping and eating disorder diagnosis and risk among college students. *Eating Behaviors*, *43*, Article 101566. https://doi.org/10.1016/j.eatbeh.2021.101566

Goniewicz, M. L., Boykan, R., Messina, C. R., Eliscu, A., & Tolentino, J. (2019). High exposure to nicotine among adolescents who use Juul and other vape pod systems ('pods'). *Tobacco Control*, *28*(6), 676–677. https://doi.org/10.1136/tobaccocontrol-2018-054565

Grunberg, N. E. (1982). The effects of nicotine and cigarette smoking on food consumption and taste preferences. *Addict Behaviors*, *7*(4), 317–331. https://doi.org/10.1016/0306-4603(82)90001-6

Grunberg, N. E., Bowen, D. J., & Morse, D. E. (1984). Effects of nicotine on body weight and food consumption in rats. *Psychopharmacology*, *83*(1), 93–98. https://doi.org/10.1007/BF00427430

Hartmann-Boyce, J., McRobbie, H., Butler, A. R., Lindson, N., Bullen, C., Begh, R., Theodoulou, A., Notley, C., Rigotti, N. A., Turner, T., Fanshawe, T. R., & Hajek, P. (2021). Electronic cigarettes for smoking cessation. *Cochrane Database of Systematic Reviews*, *9*(9), Article CD010216. https://doi.org/10.1002/14651858.CD010216.pub6

Hiler, M., Breland, A., Spindle, T., Maloney, S., Lipato, T., Karaoghlanian, N., Shihadeh, A., Lopez, A., Ramôa, C., & Eissenberg, T. (2017). Electronic cigarette user plasma nicotine concentration, puff topography, heart rate, and subjective effects: Influence of liquid nicotine concentration and user experience. *Experimental and Clinical Psychopharmacology*, *25*(5), 380–392. https://doi.org/10.1037/pha0000140

Hughes, J. R., & Hatsukami, D. (1986). Signs and symptoms of tobacco withdrawal. *Archives of General Psychiatry*, *43*(3), 289–294. https://doi.org/10.1001/archpsyc.1986.01800030107013

Jackson, S. E., Brown, J., Aveyard, P., Dobbie, F., Uny, I., West, R., & Bauld, L. (2019). Vaping for weight control: A cross-sectional population study in England. *Addictive Behaviors*, *95*, 211–219. https://doi.org/10.1016/j.addbeh.2019.04.007

Jo, Y. H., Talmage, D. A., & Role, L. W. (2002). Nicotinic receptor-mediated effects on appetite and food intake. *Journal of Neurobiology*, *53*(4), 618–632. https://doi.org/10.1002/neu.10147

Lowe, M. R., Butryn, M. L., Didie, E. R., Annunziato, R. A., Thomas, J. G., Crerand, C. E., Ochner, C. N., Coletta, M. C., Bellace, D., Wallaert, M., & Halford, J. (2009). The Power of Food Scale. A new measure of the psychological influence of the food environment. *Appetite*, *53*(1), 114–118. https://doi.org/10.1016/j.appet.2009.05.016

Lyu, J. C., Huang, P., Jiang, N., & Ling, P. M. (2022). A systematic review of e-cigarette marketing communication: Messages, communication channels, and strategies. *International Journal of Environmental Research and Public Health*, *19*(15), Article 9263. https://doi.org/10.3390/ijerph19159263

Maloney, S., Eversole, A., Crabtree, M., Soule, E., Eissenberg, T., & Breland, A. (2020). Acute effects of JUUL and IQOS in cigarette smokers. *Tobacco Control*, *30*(4), 449–452. https://doi.org/10.1136/tobaccocontrol-2019-055475

Mantey, D. S., Omega-Njemnobi, O., & Kelder, S. H. (2020). E-cigarette use is associated with intentions to lose weight among high school students.

Nicotine & Tobacco Research, *22*(5), 838–842. https://doi.org/10.1093/ntr/nty245

McHugh, M. L. (2011). Multiple comparison analysis testing in ANOVA. *Biochemia Medica*, *21*(3), 203–209. https://doi.org/10.11613/BM.2011.029

Meule, A. (2020). The psychology of food cravings: The role of food deprivation. *Current Nutrition Reports*, *9*(3), 251–257. https://doi.org/10.1007/s13668-020-00326-0

Meule, A., Hermann, T., & Kübler, A. (2014). A short version of the Food Cravings Questionnaire-Trait: The FCQ-T-reduced. *Frontiers in Psychology*, *5*, Article 190. https://doi.org/10.3389/fpsyg.2014.00190

Morean, M. E., Bold, K. W., Kong, G., Camenga, D. R., Simon, P., Jackson, A., Cavallo, D. A., & Krishnan-Sarin, S. (2020). High school students' use of flavored e-cigarette e-liquids for appetite control and weight loss. *Addictive Behaviors*, *102*, Article 106139. https://doi.org/10.1016/j.addbeh.2019.106139

Morean, M. E., Krishnan-Sarin, S., Sussman, S., Foulds, J., Fishbein, H., Grana, R., & O'Malley, S. S. (2019). Psychometric Evaluation of the E-cigarette Dependence Scale. *Nicotine & Tobacco Research*, *21*(11), 1556–1564. https://doi.org/10.1093/ntr/ntx271

Morean, M. E., & Wedel, A. V. (2017). Vaping to lose weight: Predictors of adult e-cigarette use for weight loss or control. *Addictive Behaviors*, *66*, 55–59. https://doi.org/10.1016/j.addbeh.2016.10.022

National Academies of Sciences, Enginering, and Medicine. (2018). *Summary of the 2018 National Academies of Sciences, Engineering, and Medicine (NASEM) report, The public health consequences of e-cigarettes*. Retrieved November 21, 2022, from https://www.tobaccofreekids.org/assets/factsheets/0397.pdf

Ng, L., & Davis, C. (2013). Cravings and food consumption in Binge Eating Disorder. *Eating Behaviors*, *14*(4), 472–475. https://doi.org/10.1016/j.eatbeh.2013.08.011

Parker, B. A., Sturm, K., MacIntosh, C. G., Feinle, C., Horowitz, M., & Chapman, I. M. (2004). Relation between food intake and visual analogue scale ratings of appetite and other sensations in healthy older and young subjects. *European Journal of Clinical Nutrition*, *58*(2), 212–218. https://doi.org/10.1038/sj.ejcn.1601768

Perkins, K. A., Epstein, L. H., Sexton, J. E., Solberg-Kassel, R., Stiller, R. L., & Jacob, R. G. (1992). Effects of nicotine on hunger and eating in male and female smokers. *Psychopharmacology*, *106*(1), 53–59. https://doi.org/10.1007/BF02253588

Perkins, K. A., Epstein, L. H., Stiller, R. L., Fernstrom, M. H., Sexton, J. E., Jacob, R. G., & Solberg, R. (1991). Acute effects of nicotine on hunger and caloric intake in smokers and nonsmokers. *Psychopharmacology*, *103*(1), 103–109. https://doi.org/10.1007/BF02244083

Perkins, K. A., Grobe, J. E., Epstein, L. H., Caggiula, A., Stiller, R. L., & Jacob, R. G. (1993). Chronic and acute tolerance to subjective effects of nicotine. *Pharmacology, Biochemistry, and Behavior*, *45*(2), 375–381. https://doi.org/10.1016/0091-3057(93)90254-Q

Perkins, K. A., Sexton, J. E., DiMarco, A., & Fonte, C. (1994). Acute effects of tobacco smoking on hunger and eating in male and female smokers. *Appetite*, *22*(2), 149–158. https://doi.org/10.1006/appe.1994.1014

Picciotto, M. R., & Mineur, Y. S. (2013). Nicotine, food intake, and activation of POMC neurons. *Neuropsychopharmacology*, *38*(1), Article 245. https://doi.org/10.1038/npp.2012.163

Piñeiro, B., Correa, J. B., Simmons, V. N., Harrell, P. T., Menzie, N. S., Unrod, M., Meltzer, L. R., & Brandon, T. H. (2016). Gender differences in use and expectancies of e-cigarettes: Online survey results. *Addictive Behaviors*, *52*, 91–97. https://doi.org/10.1016/j.addbeh.2015.09.006

Pinto, B. M., Borrelli, B., King, T. K., Bock, B. C., Clark, M. M., Roberts, M., & Marcus, B. H. (1999). Weight control smoking among sedentary women. *Addictive Behaviors*, *24*(1), 75–86. https://doi.org/10.1016/S0306-4603(98)00034-3

Prochaska, J. J., Vogel, E. A., & Benowitz, N. (2022). Nicotine delivery and cigarette equivalents from vaping a JUULpod. *Tobacco Control*, *31*(e1), e88–e93. https://doi.org/10.1136/tobaccocontrol-2020-056367

Rupprecht, L. E., Smith, T. T., Donny, E. C., & Sved, A. F. (2016). Self-administered nicotine suppresses body weight gain independent of food intake in male rats. *Nicotine & Tobacco Research*, *18*(9), 1869–1876. https://doi.org/10.1093/ntr/ntw113

Sanchez, R., Ranjit, N., Kelder, S. H., Gill, M., & Hoelscher, D. M. (2021). Intention to lose weight and use of electronic cigarettes among adolescents. *Preventive Medicine Reports*, *23*, Article 101406. https://doi.org/10.1016/j.pmedr.2021.101406

Shadel, W. G., Edelen, M. O., Tucker, J. S., Stucky, B. D., Hansen, M., & Cai, L. (2014). Development of the PROMIS nicotine dependence item banks. *Nicotine & Tobacco Research*, *16*(Suppl. 3), S190–S201. https://doi.org/10.1093/ntr/ntu032

Strong, D. R., Myers, M., Linke, S., Leas, E., Hofstetter, R., Edland, S., & Al-Delaimy, W. K. (2015). Gender differences influence overweight smokers' experimentation with electronic nicotine delivery systems. *Addictive Behaviors*, *49*, 20–25. https://doi.org/10.1016/j.addbeh.2015.05.003

Stubbs, R. J., Hughes, D. A., Johnstone, A. M., Rowley, E., Reid, C., Elia, M., Stratton, R., Delargy, H., King, N., & Blundell, J. E. (2000). The use of visual analogue scales to assess motivation to eat in human subjects: A review of their reliability and validity with an evaluation of new hand-held computerized systems for temporal tracking of appetite ratings. *British Journal of Nutrition*, *84*(4), 405–415. https://doi.org/10.1017/S0007114500001719

Stunkard, A. J., & Messick, S. (1985). The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *Journal of Psychosomatic Research*, *29*(1), 71–83. https://doi.org/10.1016/0022-3999(85)90010-8

White, M. A. (2012). Smoking for weight control and its associations with eating disorder symptomatology. *Comprehensive Psychiatry*, *53*(4), 403–407. https://doi.org/10.1016/j.comppsych.2011.05.007

Yannakoulia, M., Anastasiou, C. A., Zachari, K., Sidiropoulou, M., Katsaounou, P., & Tenta, R. (2018). Acute effect of smoking and smoking abstinence on energy intake and appetite-related hormones blood concentrations. *Physiology & Behavior*, *184*, 78–82. https://doi.org/10.1016/j.physbeh.2017.11.007

Yingst, J. M., Foulds, J., Veldheer, S., Hrabovsky, S., Trushin, N., Eissenberg, T. T., Williams, J., Richie, J. P., Nichols, T. T., Wilson, S. J., & Hobkirk, A. L. (2019). Nicotine absorption during electronic cigarette use among regular users. *PLOS ONE*, *14*(7), Article e0220300. https://doi.org/10.1371/journal.pone.0220300

Received September 29, 2023
Revision received December 26, 2023
Accepted January 3, 2024 ■

CONCLUDING REMARKS

The overall focus of my habilitation research was on factors that can improve weight and health outcomes in exercise- and diet-based lifestyle interventions and our findings can be used to inform future behavioral weight loss interventions. For example, our research showed that behavioral factors such as habitual exercise behavior predict acute energy compensation after a single bout of exercise as well as weight loss and compensation during an exercise-based weight loss intervention. We also showed that fasting appetite hormones are strong predictors of post-exercise energy intake after acute exercise in men (but not women), and these effects should be further examined more longitudinally. In general, baseline predictors are particularly attractive, as they identify individuals who may struggle with weight loss at the earliest possible time (i.e., even before these individuals engage in a behavioral weight loss program). However, our research also showed that more dynamic, intervention-specific factors such as initial weight loss and adherence to calorie targets can substantially improve the prediction beyond these baseline variables and should be additionally incorporated in future weight loss interventions to identify when treatment modifications are necessary. Using a personalized weight graph that plots daily weights in relation to the expected weight loss trajectory when adherent to the prescribed calorie targets (“zone of adherence”) has proven very effective in many lifestyle interventions. It allows real-time feedback to patients and interventionists, puts the patient in control to improve their weight loss-related self-efficacy and motivation, and predicts long-term weight loss success, as shown in our research. Our research also addressed critical challenges when defining adherence to calorie restriction goals via the weight-graph approach, namely that the zone of adherence often considers calorie restriction far less than the prescribed goal as being adherent. We suggest lower adherence zones in future studies, which would result in higher levels of calorie restriction being achieved when participants' weights are in the zone of adherence. In addition to behavioral and endocrine factors, we also considered specific genotypes (individuals more likely to lose more weight on a high-carbohydrate or high-fat diet based on their genetic profile) and their effect on weight loss via diets with different macronutrient compositions (high-carbohydrate vs. high-fat diet). Our findings of no differences in weight loss or secondary endpoints such as body fat and blood pressure between genotype-concordant and genotype-discordant diets represent an important contribution to the field, demonstrating that with the current ability to genotype individuals, there is no evidence for a genotype-diet interaction on weight loss. Finally, my habilitation highlighted the importance of accurate methods for food intake assessment to determine diet-health interactions and examine and monitor the effects of dietary changes on obesity treatment. To date, image-based approaches with human rater-based analysis remain the state-of-the-art approach for objective food intake assessment in many settings, though mathematical modeling and the integration of multi-sensor data (e.g., CGM) have the potential to further improve the accuracy and allow automated in-the-moment feedback to patients and researchers in the near future.