



Physiology

Scientific discourse in the era of open science: a response to Hall et al. regarding the Carbohydrate-Insulin Model

David S. Ludwig¹ · Paul R. Lakin² · William W. Wong³ · Cara B. Ebbeling¹

Received: 17 July 2019 / Revised: 24 September 2019 / Accepted: 24 September 2019 / Published online: 4 October 2019
© The Author(s), under exclusive licence to Springer Nature Limited 2019. This article is published with open access

Introduction

In November 2018, we presented the results of a large macronutrient feeding study at The Obesity Society's Blackburn Symposium [1] and simultaneously in BMJ, with peer review available online [2]. In that study, total energy expenditure (TEE) measured by doubly-labeled water (DLW) at post-weight-loss and at 10 and 20 weeks of weight-loss maintenance was 200–280 kcal/day greater on a low- versus high-carbohydrate diet, consistent with the Carbohydrate-Insulin Model of obesity [3]. To facilitate scientific discourse, we made the full database immediately available on Open Science Framework.

At the Blackburn Symposium, and in subsequent comments elsewhere, Kevin Hall raised a series of concerns about our study, including the possibility of baseline instability in weight, randomization failure for other reasons, and inaccuracy of DLW due to isotope sequestration by de novo lipogenesis [4, 5]. We aimed to refute those criticisms with evidence to show baseline weight stability, a lack of difference in weight change between diet groups before and after randomization, and insignificant rates of de novo lipogenesis on relevant diets [6–8].

In their current reanalysis, Hall et al. [9] reiterate several previous concerns and raise new ones. Here, we respond to three key questions related to our study. To facilitate this response, we obtained more accurate and precise data on

energy intake during the weight-loss maintenance phase of our study [10].

What is the appropriate baseline for studying metabolism during weight-loss maintenance?

In earlier versions of this reanalysis (before Speakman joined as coauthor) [11], and elsewhere [4], Hall and Guo argue that the pre- rather than post-weight-loss measurement is the more appropriate baseline for examining diet effects on TEE. However, that approach disregards important biological variation among individuals regarding how body composition and metabolism change in response to weight loss, which averaged 9.6 kg in our study. Variation in these major confounders would introduce imprecision in statistical models involving inter-individual comparisons in this parallel design. As a general rule, baseline data should be collected as close to randomization as possible to decrease error arising from any time-varying covariate (the pre-weight-loss measurement would entail a 4-month lag until randomization). For these reasons, weight-loss-maintenance studies like Diogenes typically use the post-weight-loss timepoint.

We explored this issue by comparing the pre-weight-loss (Fig. 1a) or post-weight-loss (Fig. 1b) measurement of TEE with TEE measured at 10 and 20 weeks after randomization. As expected, the post-weight-loss timepoint yielded a stronger correlation. Furthermore, the results of the ANCOVA analysis by Hall et al. [9], without use of either baseline, yielded a diet effect similar to our change models with the post-weight-loss baseline.

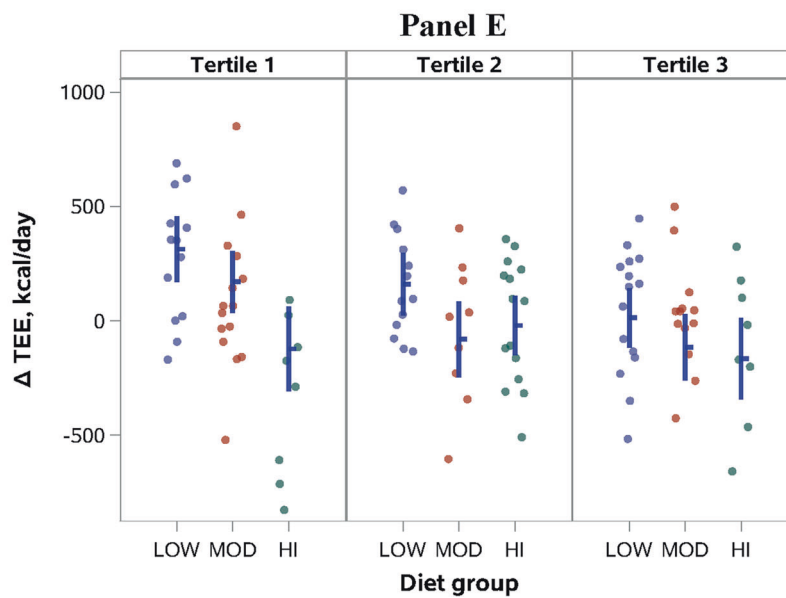
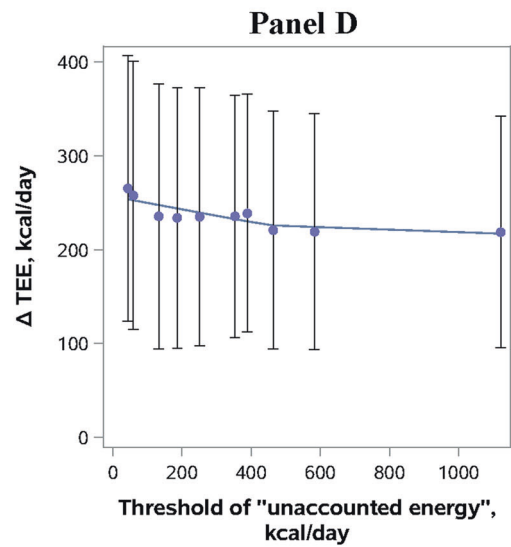
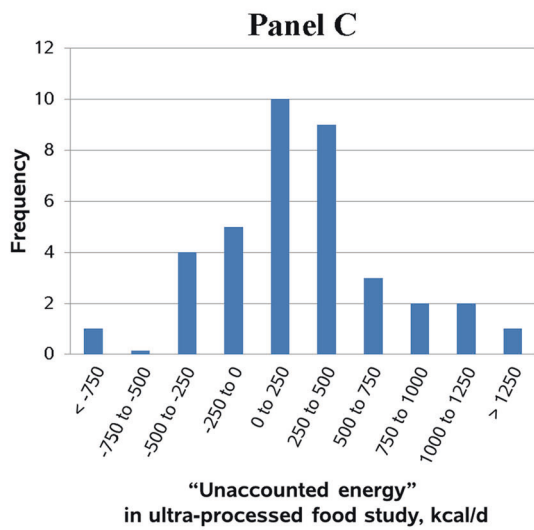
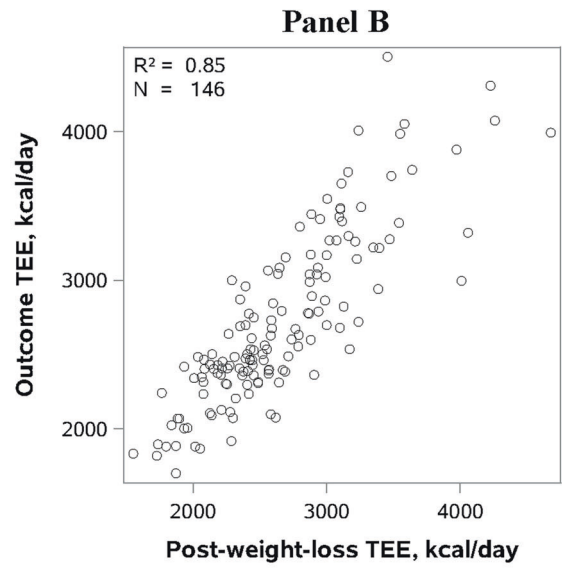
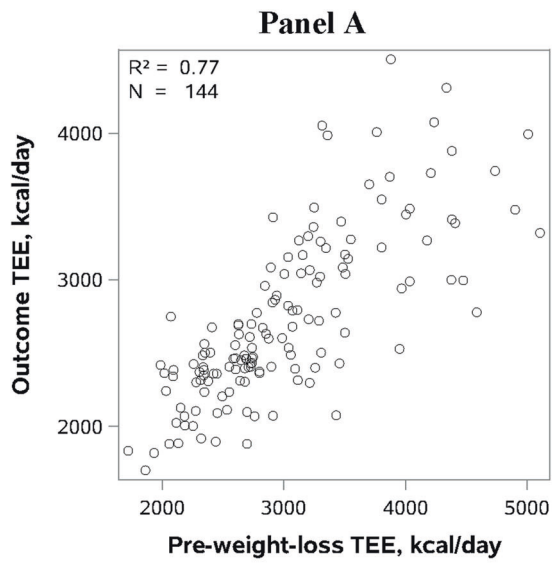
A related concern of Hall and Guo [11] is the possibility of weight instability and ongoing metabolic adaptations during the post-weight-loss baseline measurements. However, as previously considered [6], weight change during the 15-day assessment period was small, averaging 23 g/day, with no significant difference between individuals who would be assigned to the three diet groups. Moreover, the

✉ David S. Ludwig
david.ludwig@childrens.harvard.edu

¹ New Balance Foundation Obesity Prevention Center, Boston Children's Hospital and Harvard Medical School, Boston, MA, USA

² Institutional Centers for Clinical and Translational Research, Boston Children's Hospital, Boston, MA, USA

³ USDA/Agricultural Research Service, Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA



◀ **Fig. 1** All panels involve data from our weight-loss-maintenance trial [2, 10], with the exception of **c**, which comes from Hall et al. [17]. **a, b** Correlation between TEE at baseline and outcome was stronger using the post-weight-loss versus pre-weight-loss timepoint (Meng's Z-test, $p = 0.003$), demonstrating why choice of the post-weight-loss timepoint would be preferable. **c** "Unaccounted energy" in a metabolic ward study of "ultra-processed" foods [17] calculated from online data (deltabc.sas7bdat), with each participant studied on two diets for 2-week periods with DLW measurement of TEE and DXA measurement of body composition. "Unaccounted energy" (kcal/day) was calculated as energy intake (variable names: ProcAveEI, UnprocAveEI) – TEE (DLWEEProcessed, DLWEEUnproc) – change in stored energy per day (BCEBProcessed, BCEBUnproc). Only 15 of 37 observations were within ± 250 kcal/day. **d** Analysis modeled after **b** in Hall et al. [9] of "unaccounted energy," with sequential elimination of 40 (vs 81) individuals with low TEE relative to energy intake (vs low energy intake relative to TEE) in our intention-to-treat group. The observed diet effect inflates (vs deflates), demonstrating the bias introduced by postrandomization exclusion involving variables linked to the primary outcome. Because assumptions of linear regression were not satisfied (a problem that also applies to the corresponding figure in Hall et al. [9]), we used LOESS to visualize the relationship between threshold and Δ TEE. **e** Diet effect by tertiles of energy intake to TEE ratio in the Per Protocol group. Individual data indicated by circles. Mean (horizontal line) and 95% CIs (vertical bars) derived from a model adjusted for post-weight-loss TEE and other covariates (cohort, sex, age, percent weight loss, post-weight-loss weight). The difference in diet effect across the tertiles diminishes with adjustment, demonstrating the confounding present in raw data segregated by energy intake to TEE ratio. (This adjustment for baseline covariates would not remedy bias arising from postrandomization biological or measurement variation.)

experimental design protects against type I (false-positive) error due to ongoing metabolic adaptations (or any other prerandomization factor).

According to a basic principle of statistical analysis, the most powerful method available should be chosen to avoid type II (false-negative) error. Clearly, use of the post-weight-loss timepoint provides the most precise and least biased estimate.

Was the change in our analysis plan involving choice of baseline proper?

Hall et al. [9] criticize us for revising the final analysis plan to the post-weight-loss baseline, after our initial registry specified the pre-weight-loss timepoint. As previously discussed [7], the pre-weight-loss specification was an inadvertent holdover from a prior crossover study [12], in which no post-weight-loss timepoint was obtained (or needed). Unlike in parallel studies, use of a pre-weight-loss baseline would not introduce major confounding in crossover studies because of the within-individual nature of the comparisons. Biological changes during weight loss for each participant, with randomization, would affect all diet periods equally.

We identified this misspecification during preparation of our final analysis plan under supervision of our statistician, made the change before breaking the data blind, and

disclosed the registry change in the manuscript. An analysis with the pre-weight-loss baseline is available in online peer review [2], and we released the full database to facilitate alternative analyses.

In the CALERIE study, the final analyses were similarly "prespecified before initiation of analyses" [13], even though data had accumulated during the study [14]. Arguably the three largest macronutrient diet studies of the last decade—CALERIE, DIETFITS, and Diogenes—each made changes involving the primary outcome from initial registry posting to final publication. With the greater methodological heterogeneity of diet versus drug studies [15], registry changes and discrepancies are more the rule than the exception. Hall should appreciate this point. The registry for his only macronutrient RCT [16] did not specify a primary outcome, among other deficiencies that remain as of the 58th revision in February 2019.

Thus, our registry procedures are entirely consistent with standard practice in nutrition research and carry little risk of bias.

Could nonadherence to the test diets in our study explain the primary finding?

Before addressing this question, the notion of "unaccounted energy" introduced by Hall et al. warrants examination. In the reanalysis, they compared reported energy intake and expenditure in our study, identifying discrepancies as violating the "physical law of energy conservation." But this treatment disregards substantial cumulative error arising from measurement of the various components of energy balance, each with recognized imprecision and temporal variation. In the recent study by Hall et al. of ultra-processed food [17], conducted in the optimal environment of a metabolic ward, mean energy discrepancy on one diet was large (382 kcal/day) and "unaccounted energy" exceeded 250 kcal/day for most participants (Fig. 1c).

Recognizing this inherent variability and imprecision in the measurement of both energy intake and TEE, we can see why exclusions involving their difference or ratio (as in Fig. 1b of Hall et al.) would produce highly misleading results. Whereas individuals with low energy intake relative to TEE might have been nonadherent (i.e., unobserved food intake), they would also tend to be at the upper end of the natural distribution for TEE (related to true biological differences or randomly distributed measurement variation). Therefore, eliminating them would deplete the cohort of those with the greatest TEE denominator, deflating the diet effect.

We can demonstrate this phenomenon in three ways. First, we conducted the converse analysis, sequentially eliminating individuals with "unaccounted energy" arising from low TEE to energy intake (here, TEE resides in the

numerator). As illustrated in Fig. 1d, the diet effect now increases with the progressive threshold because individuals at the lower end of the TEE distribution are eliminated, leaving a residual cohort enriched for hyper-responders. However, these models, involving postrandomization variables inextricably linked to the outcome, violate a basic principle of statistical inference and should be discarded as fatally flawed.

Second, we divided the per protocol (weight stable) group into tertiles, based on the ratio of energy intake to TEE (Fig. 1e). In an unadjusted model, those in the lowest tertile (i.e., those eliminated in the analysis of Hall et al.) demonstrated a substantially larger diet effect. However, they were also more likely than those in the other two tertiles to have a baseline TEE above the median (OR 2.7 [95% CI 1.2–6.1], $p = 0.02$). With adjustment for baseline TEE and other relevant covariates, the differences between the tertiles for diet effect diminished.

Hall et al. modeled CO₂ production (rCO₂) in our cohort to circumvent the need for respiratory quotient (RQ), deviating from well-established DLW methodology and introducing severe bias against the low-carbohydrate diet. Because food quotient (FQ) equals RQ during weight (and body composition) stability, as applies to our per protocol group, a third approach is to conduct a sensitivity analysis examining how varying degrees of nonadherence would affect FQ and thereby TEE. As shown in Table 1, the low-versus high-carbohydrate diet comparison remained statistically significant through 50% nonadherence. Of particular interest, the diet effect relative to carbohydrate proportion remained remarkably stable throughout the range of assumed nonadherence, and consistently above the hypothesized 50 kcal/day for every 10% decrease in the proportion of energy as carbohydrate [3]. Moreover, among participants in the lowest tertile of energy intake to TEE (for whom estimates of FQ may be least accurate), the unadjusted change in rCO₂ was itself significantly greater on the

low- versus high-carbohydrate diet (10.3 vs –47.0 L/day, $p = 0.01$). That is, the diet effect on TEE in this subgroup was so large as to require no assumptions about FQ, providing further evidence against nonadherence as an explanation for study findings.

As stated in the BMJ article, our preliminary estimates of energy intake, used by Hall et al., “would tend to selectively underestimate those with high energy expenditure” and were not intended to be definitive. With more precise and accurate data [10], we found that energy requirements for weight stability (i.e., by calorie titration) showed a similar magnitude of effect (≈ 200 – 300 kcal/day) and hierarchical order (low > moderate > high carbohydrate) among diets as TEE, as predicted by the Carbohydrate-Insulin Model. Due to imprecision involved in these (and all) methods for determining outpatient energy intake and expenditure, the magnitude of effect should be interpreted cautiously.

Other issues

Hall et al. set a high bar for the Carbohydrate-Insulin Model by stating that “[p]roponents of low-carbohydrate diets have claimed that such diets result in a substantial increase in ... [TEE] amounting to 400–600 kcal/day”. However, the original source for this assertion, Fein and Feinman [18], characterized this estimate as a “hypothesis that would need to be tested” based on extreme assumptions about gluconeogenesis, with the additional qualification that “we [do not] know the magnitude of the effect.” An estimate derived from experimental data—and one that would still hold major implications for obesity treatment if true—is in the range of 200 kcal/day [3]. At the same time, they set a low bar for themselves, citing a 6-day trial [16] (confounded by transient adaptive responses to macronutrient change [3]) and a nonrandomized pilot study [5] (confounded by weight loss [8]) as a basis for questioning DLW methodology.

Table 1 Sensitivity analysis of TEE. This analysis examines how potential nonadherence could influence the diet effect on TEE in the per protocol (weight stable) group, considering how FQ (used in DLW methodology) changes with macronutrient ratio

| | Degree of nonadherence (proportion of total) ^a | | | | | | |
|--|---|------------------|------------------|------------------|------------------|------------------|--------|
| | 0% | 10% | 20% | 30% | 40% | 50% | 60% |
| FQ low-carbohydrate diet | 0.7880 | 0.7938 | 0.7996 | 0.8054 | 0.8112 | 0.8170 | 0.8228 |
| FQ high-carbohydrate diet | 0.9040 | 0.8982 | 0.8924 | 0.8866 | 0.8808 | 0.8750 | 0.8692 |
| TEE diet effect ^b kcal/day | 280 ^c | 249 ^c | 219 ^c | 188 ^c | 158 ^c | 127 ^c | 97 |
| TEE diet effect kcal/day per 10% decrease in carbohydrate | 70 | 69 | 68 | 67 | 66 | 64 | 60 |

^aAssumes that any foods eaten off protocol for both low-carbohydrate (20% carbohydrate, 60% fat) and high-carbohydrate (60% carbohydrate, 20% fat) diet groups reflected the macronutrient composition of the moderate-carbohydrate (40% carbohydrate, 40% fat) diet group

^bDifference between low- and high-carbohydrate diet groups, minimally adjusted for cohort, as per analyses in BMJ [2]

^cRemained statistically significant at $p \leq 0.05$

Elsewhere, Hall interpreted these studies as sufficient to “falsify” the Carbohydrate-Insulin Model [19]—but they do nothing of the kind. Indeed, a recent reanalysis of that pilot study suggests an effect similar to ours (≈ 250 kcal/day) [20].

Finally, we agree with Hall et al. that the component(s) of energy expenditure that might underlie our findings remain unknown. The diet effects on resting energy expenditure (REE) and moderate- to vigorous-intensity physical activity were of borderline significance in the hypothesized direction ($p = 0.06$ – 0.09) for pair-wise comparisons. Because individual components of energy expenditure might each contribute <100 kcal/day to the diet effect, our study lacked power to examine these secondary outcomes. In our crossover study [12], we found a significantly greater REE on a very-low-carbohydrate versus low-fat (high-carbohydrate) diet in the fasting state, when the thermic effect of food would have dissipated.

Conclusion

We aim to show that the latest series of criticisms by Hall et al., like those previously addressed, have little merit. Contrary to their claims, the data from our BMJ study, together with new data on energy intake, provide substantial support for the Carbohydrate-Insulin Model. Nevertheless, we recognize that these results require replication and that the relative advantages of dietary carbohydrate- versus fat-restriction on a population basis have not been established.

Of course, debate and criticism lie at the heart of science and should be encouraged, including with public posting of full databases. However, in the new era of open science, with widespread availability of raw data, the inevitable deficiencies in trials will be on full display. To minimize distraction and promote constructive discourse, it will be critical to distinguish inconsequential discrepancies and omissions from flaws that pose high risk of bias. Thus, we must all resist the admittedly natural tendency to hold a double standard for studies that support versus oppose our own views.

Data availability

The protocol and data set for the original trial findings study are available at Open Science Framework (<https://osf.io/rvbuy/>). New primary data on energy intake and body composition will be posted upon publication of the related findings in a peer-review journal.

Acknowledgements We thank Arne Astrup, Henry Feldman, Steven Heymsfield, and Walter Willett for reviewing the manuscript.

Funding The trial considered here was funded by Nutrition Science Initiative (made possible by gifts from the Laura and John Arnold Foundation and Robert Lloyd Corkin Charitable Foundation), New Balance Foundation, Many Voices Foundation, and Blue Cross Blue Shield. DSL was supported by a mid-career mentoring award from the National Institute of Diabetes and Digestive and Kidney Diseases (K24DK082730). The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the study sponsors.

Compliance with ethical standards

Conflict of interest DSL and CBE have conducted research studies examining the Carbohydrate-Insulin Model of obesity funded by the National Institutes of Health and philanthropic organizations unaffiliated with the food industry; DSL received royalties for books on obesity and nutrition that recommend a low-glycemic load diet. The other authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- George L. Blackburn Symposium to feature weight loss studies. News Release, The Obesity Society; 14 November 2018. <https://www.globenewswire.com/news-release/2018/11/14/1651612/0/en/George-L-Blackburn-Symposium-to-Feature-Weight-Loss-Studies.html>. Accessed 9 Jul 2019.
- Ebbeling CB, Feldman HA, Klein GL, Wong JMW, Bielak L, Steltz SK, et al. Effects of a low carbohydrate diet on energy expenditure during weight loss maintenance: randomized trial. *BMJ*. 2018;363:k4583. <https://doi.org/10.1136/bmj.k4583>.
- Ludwig DS, Ebbeling CB. The carbohydrate-insulin model of obesity: beyond “calories in, calories out”. *JAMA Int Med*. 2018;178:1098–103.
- Hall KD, Guo J. No significant effect of dietary carbohydrate versus fat on the reduction in total energy expenditure during maintenance of lost weight. *BMJ Rapid Resp*. 2018. <https://www.bmj.com/content/363/bmj.k4583/rr-16>.
- Hall KD, Guo J, Chen KY, Leibel RL, Reitman ML, Rosenbaum M, et al. Methodologic considerations for measuring energy expenditure differences between diets varying in carbohydrate using the doubly labeled water method. *Am J Clin Nutr*. 2019;109:1328–34.
- Ludwig DS, Ebbeling CB, Feldman HA. Choice of baseline for primary endpoint. *BMJ Rapid Resp*. 2018. <https://www.bmj.com/content/363/bmj.k4583/rr-11>.

7. Ludwig DS, Ebbeling CB. Author response to Hall and Guo regarding data reanalysis and other criticisms. *BMJ Rapid Resp*. 2018. <https://www.bmj.com/content/363/bmj.k4583/rr-17>.
8. Ludwig DS, Ebbeling CB, Wong JMW, Wolfe RR, Wong WW. Methodological error in measurement of energy expenditure by the doubly labeled water method: much ado about nothing? *Am J Clin Nutr*. 2019. (In press).
9. Hall KD, Guo J, Speakman JR. Do low-carbohydrate diets increase energy expenditure? *Int J Obes*. 2019. <https://www.nature.com/articles/s41366-019-0456-3>.
10. Ebbeling CV, Bielak L, Lakin PR, Klein GL, Wong JMW, Luoto PK, et al. Higher energy requirement during weight-loss maintenance on a low- versus high-carbohydrate diet: secondary analyses from a randomized controlled feeding study. *medRxiv*. 2019. <https://doi.org/10.1101/19001248>. [preprint].
11. Hall KD, Guo J. Carbs versus fat: does it really matter for maintaining lost weight? Version 5. *bioRxiv*. 2019. <https://doi.org/10.1101/476655>. [preprint].
12. Ebbeling CB, Swain JF, Feldman HA, Wong WW, Hachey DL, Garcia-Lago E, et al. Effects of dietary composition on energy expenditure during weight-loss maintenance. *JAMA*. 2012;307:2627–34.
13. Ravussin E, Redman LM, Rochon J, Das SK, Fontana L, Kraus WE, et al. A 2-year randomized controlled trial of human caloric restriction: feasibility and effects on predictors of health span and longevity. *J Gerontol A Biol Sci Med Sci*. 2015;70:1097–104.
14. Doubly-Labeled Water (DLW) Procedures for the CALERIE study. 1–13. https://calerie.duke.edu/sites/calerie.duke.edu/files/14.0_double-labeled_water_dlw_procedures.pdf. Accessed 9 Jul 2019.
15. Ludwig DS, Ebbeling CB, Heymsfield SB. Improving the quality of dietary research. *JAMA*. 2019. <https://doi.org/10.1001/jama.2019.11169>.
16. Hall KD, Bemis T, Brychta R, Chen KY, Courville A, Crayner EJ, et al. Calorie for calorie, dietary fat restriction results in more body fat loss than carbohydrate restriction in people with obesity. *Cell Metab*. 2015;22:427–36.
17. Hall KD, Ayuketah A, Brychta R, Cai H, Cassimatis T, Chen KY, et al. Ultra-processed diets cause excess calorie intake and weight gain: an inpatient randomized controlled trial of ad libitum food intake. *Cell Metab*. 2019;30:67–77.
18. Fine EJ, Feinman RD. Thermodynamics of weight loss diets. *Nutr Metab*. 2004;1:15.
19. Hall KD. A review of the carbohydrate-insulin model of obesity. *Eur J Clin Nutr*. 2017;71:323–6.
20. Friedman MI, Appel S. Energy expenditure and body composition changes after an isocaloric ketogenic diet in overweight and obese men: a secondary analysis of energy expenditure and physical activity. *bioRxiv*, Version 5. 2019. <https://www.biorxiv.org/content/10.1101/383752v5>.