January 21, 2021

I have reproduced the text of this blog post below that I received in pdf format from Ivan Oransky at Retraction Watch on January 19. My comments to date are provided below in red font.

(Preamble: This post is timed to appear with a related post by Ethan and Sarah Ludwin-Peery, who have some pertinent questions about patterns in the data associated with the article that is discussed. *I strongly recommend that you read their analysis first*, not least because it provides a much more comprehensive introduction to the study. Here I discuss a variety of other apparent problems with the same article.)

This post looks at an article that first appeared in May 2019 describing a randomised controlled nutrition study. The authors claimed that people who were allowed to eat as much as they wished of a diet based on either "ultra-processed" or "unprocessed" food(*) consumed around 500 kcal/day more on the ultra-processed diet, and gained an average of 0.9 kg (2 lbs) in two weeks, compared to people on the unprocessed diet, who lost an average of 0.9 kg in the same period. The same 20 participants ate both diets, in a randomised order. Importantly, the amount of macronutrients (protein, fat, and carbohydrates) provided in the meals was closely matched across diets, as was the number of calories offered (logically, since calories are a function of the macronutrients). That is, the claim is that the mere fact that the food was ultra-processed, versus unprocessed, caused people to consume 500 kcal/day more and gain, rather than lose, weight in a controlled inpatient setting.

Perhaps not surprisingly, the study attracted a lot of attention. It has already been cited more than 360 times according to Google Scholar. The National Institutes of Health (NIH), which funded and conducted the study, put out an extensive news release about it, and the story was covered by both Science and Nature, as well as the BBC, the Guardian, the Washington Post, and many other major media outlets.

Here is the full reference of the article. For the first time since the appearance of the 7th edition of the APA Publication Manual (which says that we now have to list up to 20 authors' names in a reference) I'm actually going to need an ellipsis to omit some of the 25 authors:

Hall, K. D., Ayuketah, A., Brychta, R., Cai, H., Cassimatis, T., Chen, K. Y., Chung, S. T.,
Costa, E., Courville, A., Darcey, V., Fletcher, L. A., Forde, C. G., Gharib, A. M., Guo, J., Howard, R., Joseph, P. V., McGehee, S., Ouwerkerk, R., Raisinger, K., ... Zhou1, M. (2019). Ultra-processed diets cause excess calorie intake and weight gain: An inpatient randomized controlled trial of ad libitum food intake. *Cell Metabolism*, *30*(1), 67–77. https://doi.org/10.1016/j.cmet.2019.05.008

The article is published on an Open Access basis; you can find the full text here (PDF, 2 MB) or a fuller version, including the Supplemental Information, here (PDF, 23 MB). A small erratum, correcting a number of minor issues, was published on August 6, 2019; all of the issues mentioned in the erratum are already corrected in the PDF files, so you don't need to keep that to hand while reading the article.

Importantly, another erratum was published in October 2020 and is available here: https://www.cell.com/cell-metabolism/fulltext/S1550-4131(20)30427-7
The correction relates some of the issues raised below and we realize that the updated data and code were not yet deposited on the OSF website. We will do so.

This study has already been the subject of a comment on PubPeer by Edward Archer, who, I think it is fair to say, is a prolific critic of the way that much nutritional research is carried out. I am not a nutrition scientist, so this blog post will mostly concentrate on the data and statistics of the study. I do have one or two small methodological questions too, but these are based only on my 60 years of experience of consuming food and 40 or so of preparing it, rather than any understanding of how nutrition studies are run.

The study

The authors recruited 20 volunteers, 10 male and 10 female, and kept them in an in-patient environment for 28 days at the NIH Clinical Center in Bethesda, Maryland. The data show that between one and four people were in the facility at any point between the first admission on April 17, 2018 and the last recorded data collection on November 19, 2018.

Participants spent 14 days on each of two diets, one described as "processed" and the other as "unprocessed". The diets were presented on a 7-day rotation, so each participant ate the same meal twice, 7 days apart. Although the purpose of the study was to examine the effect of an "ultra-processed" diet, and that term tends to be used in nutrition science with a specific meaning that is different from "processed" (it's complicated), I will mostly follow the authors' use of the terms "processed" and "unprocessed" to distinguish between the two. I hope that this will avoid any confusion that might be caused by the fact that "ultra-processed" and "unprocessed" both start with the same letter. The participants were randomised to receive the processed diet first (N=10, 6 male, 4 female) or the unprocessed diet first (N=10, 4 male, 6 female); after 14 days on one diet they immediately switched to the other, as shown here.

Timeline of participants in the study. Reproduced from Figure 1 of Hall et al.'s article.

The code and data

The authors have made their data and analysis code available here. There are two datasets, named ADLDataSAScode and ADLDataSAScode1, each in its own ZIP file. The only difference between these seems to be that ADLDataSAScode1, which was uploaded on August 20, 2019 (three months after the article was first published online, which was on May 16, 2019), contains one extra data file, and the code has been extended with a few lines to produce a table from that file (more on this later). All of the analyses in this post refer to the ADLDataSAScode1 dataset.

Screenshot of the timestamps of the OSF repository for the study. A full-size version of this image is available as part of the supporting files for this post (see "Code and data", below).

The SAS code is not, as one might have hoped, a run-once script that generates all of the tables and figures from the article. Indeed, as supplied, the main script file (ADLDocumentation1.sas) produces two runtime errors at line 61 because the variables created within the SAS data file DLW at lines 42 and 43 are lost when this file is overwritten twice at lines 45 and 46. It seems that the code is best regarded as a collection of "building blocks" of code that can be run individually, possibly with minor modifications to use different subsets of the data. However, for completeness, I patched up the code so that it would run without error messages, and also to include both the original and adjusted analyses of the figures from Table 3D (see "The adjusted weight data", below), and ran it in SAS University edition. I have made the resulting code ("Nick-ADLDocumentation1.sas") and output ("(Annotated) Results_Nick-ADLDocumentation1.pdf") files available online (see "Code and data", below).

The exact length of the study

An issue that stands out immediately when one looks at any of the data files containing daily records is that there seems to be a fencepost error. Participants spent 14 days on each of two diets, with no break in between; their weight at the start of day 1 was the baseline for the first phase (processed or unprocessed diet, assigned at random), and their weight at the start of day 15 was the baseline for the second phase, when they received the other diet. It would seem, therefore, that they should have been weighed 29 times—once at the very start of the study, and then 28 more times after eating a day's worth of meals each time—but there are only 28 daily weight records for each participant. That is, we apparently do not know the effect on their weight of the last (14th) day of the second diet, because the last measurement of their weight on that second diet was the one conducted on the morning of the 14th day (their 28th in the study), *before* they proceeded to eat their food and undergo whatever other measurements were performed on that day. This seems to make little sense, from the standpoint of either study design or ethics.

Why feed your participants the controlled diet on the last day if you are not going to collect weight data from them relating to that day?

Participants were admitted the afternoon before the study began. An overnight fasted body weight measurement was collected the next morning (day 1) which served as the fiducial point for the weight change calculations during the next 14 days on the first diet. On the morning of day 15, subjects were weighed which served as the fiducial point for weight change calculations on the alternate diet that was provided after an oral glucose tolerance test (OGTT). Fasted body weight measurements were then collected each morning including day 29 when the final OGTT was performed after which the subject was discharged. Thus, there were 29 fasted body weight measurements corresponding to the fiducial markers on days 1 and 15 prior to delivery of each diet and 14 days thereafter. However, it appears that the final body weight data on day 29 was inadvertently omitted from the deposited data because it did not correspond to a diet day. Thus, the reported body weight changes correspond to days 1-14 of the first diet period and days 15-28 of the second diet period as shown in Figure 3A of the manuscript.

Which days did participants spend in the respiratory chamber?

Participants spent one day per week in a respiratory chamber to enable their energy expenditure to be studied in detail. The article states that "On the chamber days, subjects were presented with identical meals within each diet period, and those meals were not offered on non-chamber days" (p. 72), which makes sense from an experimental control point of view, in that all participants would have consumed the same food on that day. The article's Supplemental Information [PDF, 21MB] further states (on pp. 15, 16, 17, 37, 38, and 39) that the chamber day was day 5 of each weekly diet rotation, corresponding to days 5 and 12 of each participant's time on each diet.

However, the records in the data file `chamber` appear to contradict this. I looked for precise matches between the recorded energy intake on the chamber days and the records for each participant in the `dailyintake` file, and found exactly one match for each participant and chamber day. Support for the idea that these matches are not coincidental is provided by the fact that the calendar dates of each record of the matched pairs (one in `chamber` and one in `dailyintake`) are identical. The matched records imply that of the 80 chamber days (20 participants x 2 diets x 2 chamber days per diet), only 7 took place on day 5 of the weekly diet rotation (whereas 2 were on day 1, 24 on day 3, 3 on day 4, 31 on day 6, and 13 on day 7). Furthermore, of the 40 pairs of chamber days within the same diet, 15 were on different diet rotation days within the pair (e.g., for participant ADL002 on the unprocessed diet, the chamber days were 3 and 8, corresponding to the third and first days of the diet rotation, respectively), meaning that the participant would have eaten different meals on their two chamber days for a given diet in 37.5% of cases. Of course, it is possible that the participants did indeed all spend days 5 and 12 of each diet in the chamber, as reported in the article and supplement, but that would mean that 73 out of 80 records in the data file `chamber` of both the date and how much they ate on those days are inconsistent with the equivalent records for diet days 5 and 12 in `dailyintake`.

The article and supplement do not claim that "participants did indeed all spend days 5 and 12 of each diet in the chamber". Rather, the main manuscript describes that participants spent one day each week in the respiratory chambers but does not specify the days of the week. The Supplementary Materials provide information about the rotating 7-day menu of meals provided on each diet and the chamber days were listed as occurring on day 5 of each week. This was not intended to indicate that the chamber days only occurred on day 5 but rather that the meals provided during the chamber days were prespecified and did not vary between subjects on the same diet no matter what day the chamber days occurred. The clinical protocol (available on the OSF website) indicates in Appendix A that the proposed schedule (page 34) had chamber days planned for days 3 and 10 on each diet. However, the protocol also notes on pages 13-14 that "Every effort will be made to adhere to the proposed timelines, but some flexibility is required for scheduling of other studies, unanticipated equipment maintenance, etc. Scheduling variations will not be reported." Thus, while chamber days varied to accommodate such scheduling challenges, the meals provided on the chamber days were constant within each diet.

Counting the calories

The data file `dailyintake` contains information about the amount of calories and individual nutrients consumed by the participants on each day. The total number of calories consumed is reported to two decimal places, but the individual readings of calories for protein, fat, and carbohydrates that sum to that total are reported to six decimal places, which on visual inspection do not appear to contain any regular patterns (recurring decimals, etc).

Extract from `dailyintake` file, showing six digits of precision for macronutrient calorie counts. Some columns have been reduced to zero width to enable the image to fit on this web page.

It is not clear how such numbers could have been generated, however, as the process for calculating the amount of calories consumed presumably ought to have been a fairly simple multiplicative one, based on estimates of the numbers of grams of protein, fat, carbohydrates, and water in the uneaten portions of each food that was offered. (Edward Archer's comment on PubPeer mentions this issue, and suggests that a bomb calorimeter might have been a better measure.) The authors report that the diets were designed and analyzed using ProNutra software, made by Viocare of Princeton, NJ. I wrote to Viocare to ask how this software calculate calories from macronutrients—for example, whether it uses the Atwater values of 4.0 kcal/g for protein and carbohydrates and 9.0 kcal/g for fat, and whether it typically generates long mantissas in its output. Its founder and president, Rick Weiss, sent me this reply:

ProNutra's standard nutritional database is from USDA which we load into ProNutra with the resolution as USDA provides. Typically a research group using ProNutra would round off to the decimal place that they need. So I agree, seeing a value to the 6th decimal doesn't make sense. The analysis of calories from macronutrients does use Atwater values.

But if the calories per gram are always integers, the presence of six decimal places of precision in the macronutrient information of every meal would seem to imply that the authors calculated the amount of food that was (a) served and (b) remained uneaten to the nearest microgram, which seems like it would require a lot of effort.

The data for the total energy consumed and the percentage from each macronutrient were calculated to 2 decimal places. For example, 15.68% of energy consumed as protein and a total energy intake of 2003.47 kcal. Therefore, the kcal provided from each macronutrient was calculated to six decimal places in the data file as follows: 2003.47*0.1568 = 314.144096 kcal from protein.

I also wonder what was done in the case of processed snacks, where one would expect the authors to have simply used the nutrition information provided by the manufacturers. For example, on four days of the processed diet, three participants (ADL006 on days 3 and 4, ADL007 on day 8, and ADL015 on day 9) are in the data file `intakebymeal` as having consumed 403.14 kcal in snacks, with 42.007956, 202.218222, and 158.933010 kcal coming from protein, fat, and carbohydrates respectively (these amounts are precisely identical on all four days). The chances that three people left exactly the same amount of snack food unfinished on a total of four occasions would seem to be negligible, so this duplication presumably corresponds to these participants having completely finished the contents of the same combination of snack packages on each day. But the nutrition information for each of these packaged snacks reports the amount of macronutrients with a precision of 1 g, so the calories from each of these macronutrients ought also to be an integer (a multiple of 4 or 8), unless the authors perhaps contacted the manufacturers and obtained analyses down to the microgram level.

Three different participants, four different days, identical snack consumption
A further problem here is that these records show that the three participants in question consumed more calories in the form of fat than carbohydrates from their snacking on these four days, but substantially fewer calories from protein than from carbohydrates. The only processed snack in the image on p. 24 of the Supplemental Information that has more calories from fat than from carbohydrates is the 28 g package of Planters salted peanuts (see my file `snacks.xls`), but this also has more calories from protein than from carbohydrates. I have not been able to identify any combination of packaged snacks that would get even close to the proportions of calories from protein, fat, and carbohydrates that is reported for these four participants, especially given the presumed constraint of counting only entire packages.

Nutrition information for Planters salted peanuts snack package (source), showing total grams of protein, fat, and carbohydrate. The corresponding calorie amounts would be protein, 7 x 4 = 28 kcal; fat, 14 x 9 = 126 kcal; carbohydrates, 5 x 4 = 20 kcal.

We are in the process of investigating these questions about the intake measurements and this will require consultation with the dietitians and staff at the NIH Clinical Center Nutrition department. We will fully investigate this issue, but have not been given sufficient time to respond at this time.

The participants

Participants are identified in the data by sequentially numbered labels from ADL001 through ADL021. That represents a span of 21 unique values, but there are no records with the label ADL011. Whether this is due to an error in assigning a label or a participant dropping out is not clear; however, there is no mention in the article of anyone dropping out of the study.

ADL011 declined to participate in the study after their successful screening visit when they were assigned their subject number. No participants dropped out or were withdrawn from the study after admission.

Participation in this study seems to have been what many people would probably regard as a major undertaking. The subjects spent 28 days in a highly controlled environment. The study was invasive, with subcutaneous sensors to monitor glucose levels as well as multiple finger stick operations daily. Yet no mention is made of how participants were recruited, how they were compensated, and what resources were available to look after their mental and physical wellbeing.

We encourage the blogger to read the clinical protocol (available on the OSF website) that was approved by the NIDDK Institutional Review Board which describes recruitment, risks, care, and compensation of our research volunteers.

This last point seems important, since some of the participants would appear to have been potentially rather vulnerable:

- Participant ADL006 (male) had a baseline BMI of 18.050 kg/m², which is below the minimum specified in the inclusion criteria on pp. e1–e2 of the article (18.5 kg/m²). That is, on the authors' own terms it seems that he ought to have been excluded from the study.

This participant met inclusion criteria at their screening visit, but their starting BMI was lower once admitted for the study.

- Participant ADL005 (female) had a baseline BMI of 42.459 kg/m². One might perhaps question the ethics of including a morbidly obese 40-year-old in a study where participants are encouraged to eat as much as they want.

Participants were instructed to eat as much *or as little* as they desired. We did not encourage overeating by instructing participants "to eat as much as they want". One might perhaps question the ethics of excluding volunteers with obesity in a study designed to investigate determinants of ad libitum calorie intake.

Participants ADL019 (female, BMI 38.762 kg/m²) and ADL008 (male, BMI 36.404 kg/m², who consumed an average of 5,228 kcal/day and gained 2.64 kg on the processed diet) were also severely obese.

While the blogger describes the weight gain of these participants during the ultra-processed diet, these participants gained an average of only ~1 kg during the entire study which amounted to < 1% of their body weight.

- Participant ADL020 (a 32 year old female) had a baseline BMI of 26.853. During her 14 days on the unprocessed diet she consumed an average of just 836 kcal/day and lost a total of 4.3 kg (9.4 lbs) in weight, accounting on her own for nearly a quarter (23.7%) of the total weight loss of the sample on the unprocessed diet. On day 12 of the same diet she obtained 22% of her calories (128 kcal out of 578 kcal total) from carbohydrates, which was the lowest daily percentage of anybody on any day on either diet in the entire study, whereas on the next day, day 13, she obtained 62% of her energy intake (602 kcal out of 962 kcal total) calories from carbohydrates, which was the *highest* daily percentage of anybody on any day on either diet in the entire study. This combination of extraordinary weight loss, very low levels of energy intake, and highly variable eating patterns make me wonder how much we can generalise from this participant to a broader understanding of the effects of different types of diet on the wider population.

The limitations of our study regarding generalizability were discussed in the manuscript. It is well-known in human nutrition research that individual subjects have large day-to-day diet variability and that there is large individual variability in weight loss.

Errors in the data for individual participants

ADL002

The data file `intakebymeal` contains one record for every meal consumed by participants during the study (breakfast, lunch, dinner, and one record for all of the snacks that they took) containing an assortment of nutritional information about that meal, including the type of diet that the participant was following on that day (and, hence, at each meal). For participant ADL002, however, something strange seems to have happened. The three meals (but not the snacks) that he consumed on days when he was on the processed diet are marked with the "unprocessed" diet flag, and vice versa, for all 14 days of each diet.

Extract from data file `intakebymeal` showing that participant ADL002 apparently consumed unprocessed meals and processed snacks on the same day. Some columns have been reduced to zero width to enable the image to fit on this web page.

It is not at all clear how this could have happened, because one would expect the data to have been recorded directly at the end of the day in question (in a spreadsheet or directly into the ProNutra software) such that the type of diet would either have been completed automatically by the system, or obvious based on the records from the preceding day. Certainly one would expect the snacks for any given day to have the same diet code as the three meals. (I believe that the three meals have the wrong diet code and the snacks have the right one, rather than the reverse, based on the fact that the `dailybw` and `dailyintake` files both show ADL002 being on the processed diet for the first 14 days of the study and the unprocessed diet for the last 14 days, whereas `intakebymeal` shows "unprocessed" as the diet for the breakfast, lunch, and dinner records for the first 14 days, and "processed" for the last 14 days.)

This error was previously discovered and an erratum was published in October of 2020 that corrected this error and is available here: https://www.cell.com/cell-metabolism/fulltext/S1550-4131(20)30427-7

We realize that we have yet to update the files in the OSF website to correct this previously identified error and apologize for the delay.

ADL010

Participant ADL010 has a baseline weight of 91.97 kg in the data file `deltabw` but 93.17 kg in the data file `baseline`. This affects the results shown in Table S1. If 91.97 kg is the correct weight then the Total mean for weight is correct but the Male mean (79.2 reported, 79.0 actual) and Male SE (6.6 reported, 6.5 actual) are not. If 93.17 kg is correct then the Male mean and SE are correct, but the Total mean (78.2 reported, 78.3 actual) isn't.

This problem might appear to be trivial, but it is not at all clear how it could have arisen. ADL010's weight on day 2 is recorded as 93.17 kg in `deltabw`, so one possibility is that for this participant only, the copying process (which one might hope would have been automated) that generated the `baseline` table somehow picked up the day 2 value rather than the day 1 value. Interestingly, according to that same file, this participant's weight fell back again to exactly 91.97 kg on day 3, which seems like quite a strong yo-yo effect.

Weight of participant ADL010 in the data files `baseline` (top; I renamed the variable from "Bw" to "Day1BW" for my own clarity while performing the analyses in Excel) and `deltabw` (bottom).

The baseline information in Table S1 contains body composition measurements obtained by DXA. All of the subjects except ADL001 and ADL010 had their first DXA measurement on day 1, but ADL001 and ADL010 were measured on day 2. For ADL001, their body weight measurements were the same on days 1 and 2, but ADL010 had different weights on these days. Therefore, the body weight measurement on day 2 for ADL010 was included in the baseline information to correctly correspond to the day of the DXA measurement.

Other oddities in the data

As mentioned above, data file `intakebymeal` contains a record for each meal (plus snacks), with information such as macronutrient and total calories, free water consumption, the total mass of the food consumed, etc. Meanwhile, data file `dailyintake` has a record for each day's consumption for each participant, broken down similarly. One would therefore expect the values in the four records in `intakebymeal` to sum to the values in the corresponding record in `dailyintake`. Curiously, however, this is not the case. Indeed, while the energy intake (EI) field in `dailyintake` matches the sum of the per-meal EI values in `intakebymeal` to within 0.05 kcal in every case (once the diet code error for participant ADL002, discussed above, has been corrected), the calories for protein, fat, and carbohydrates from the four meal records each day almost never sum to anything even close to the equivalent values in the daily record. Per-meal (top, with sum for all four meals under "Total") and per-day intake for participant ADL001 on the first day of the processed diet. Note that while the total energy intake ("EI") from the meals is identical to within 0.01 kcal, the total for each of the macronutrients (protein, fat, and carbohydrates) is different by between 9 and 26 kcal. Some columns have been reduced to zero width to enable the image to fit on this web page.

A related problem is that, within `intakebymeal`, the three macronutrient calorie observations for a meal frequently do not sum to the overall energy intake from that same meal. A spectacular example of this is the dinner of participant ADL005 on day 3 of her unprocessed diet, which provided 393.523334 kcal from protein, 311.910799 kcal from fat, and 530.105951 kcal from carbohydrates, for a total of 1235.540084 kcal, but whose total energy content is shown as 1720.21 kcal—a net discrepancy of about 484.67 kcal. A total of 639 of the 2,240 participant x day x meal records in `intakebymeal` suffer from this problem, whereas none of the records in `dailyintake` do. Put simply, a large number of the per-meal macronutrient values in the `intakebymeal` data file seem to bear little relation to the necessities of arithmetic. (Interestingly, all of these discrepancies are on the positive side—that is, when the overall energy intake differs substantially from the total of the energy intake from the macronutrients, the overall energy intake is always larger— suggesting that whatever process is responsible for these discrepancies might not be entirely random.)

We are in the process of investigating the issues with the individual meal data and this will require consultation with the dietitians and staff at the NIH Clinical Center Nutrition department. We will fully investigate this issue, but have not been given sufficient time to respond at this time.

The adjusted weight data

I noted earlier that the OSF repository for the project contains two ZIP files. The second of these included an extra data file named `deltabcadj14`, and the SAS code has been extended with a few lines that analyse it. This code seems to be quite important as it claims to generate the results for figure 3D of the article, which presents what are arguably the headline findings of the study: a mean weight gain of 0.9 kg per participant on the processed diet and a mean weight loss of 0.9 kg on the unprocessed diet. The code file contains this comment:

Update1: Body composition changes presented in Figure 3D are adjusted for 14 days because the body compositions were not measured exactly 14 days apart. In the previous version of SAS code and data, such adjustment was not provided. Here we have updated the SAS code at the section "data for figure 3D" and added a dataset "DeltaBCadj14";
It is not clear what adjustments were performed to make this new data file. The provided code merely re-runs the comparisons of before/after weight, fat mass, and fat-free mass for the two types of diet, using the adjusted data. When the new code is run, it produces results for the mean weight loss/gain that are around 20% different from the originals, and would presumably have been reported as a mean gain of 0.8 kg on the processed diet and a mean loss of 1.1 kg on the unprocessed diet.
Comparison of pre- and post-study weights (first two lines of each panel, for the processed and unprocessed diets, respectively) and fat/non-fat mass, using the original (top) and adjusted (bottom) data. The output from the original data file contains a descriptive label for each line, which I have removed here to allow the figures in the tables to appear in the same size font for both images.
Interestingly, the sample size for fat mass and fat-free mass on the unprocessed diet is higher with the adjusted data than the original data. The data file `deltabc` is missing these values for participant ADL002, whereas `deltabcadj14` is not. Thus, whatever the adjustment process was, it seems to have extrapolated or interpolated in some way whatever data relating to fat mass might have been missing for this participant, such that he could now be included. (I assume that fat-free mass is calculated as weight minus fat mass, so that only one missing value needs to have been inferred in this way.)
I wonder if this adjustment might be an attempt to compensate for the issue that I raised earlier under the heading "The exact length of the study". But if that is the case, it is not clear why it would be necessary to adjust the values for both diets for each participant. After all, the start of day 15 of the study—the day on which the participants changed to the other diet—ought to correspond to exactly 14 days after they were weighed on day 1.

The article states that participants were weighed at 6am every day. If it turns out that they were weighed substantially later on day 1 (or earlier on the last day), the question then arises of whether they skipped one or more meals on that day, although there are records for every scheduled meal in `intakebymeal`. On the other hand, if they were weighed only an hour or so late, the adjustment hardly seems necessary, especially since the Welch Allyn Scale-Tronix 5702 weighing scale that was used for the study has a precision of only 0.1 kg (a fact that I confirmed by e-mail correspondence with the manufacturer; see also Ethan and

Sarah's post, which explores the consequences of this constraint in more detail). The adjusted values are reported to 10 or more decimal places, which—assuming that the adjustment was indeed a function of the difference between the actual elapsed time from the first to last measurement, and exactly 14 days—suggests that the time at which participants' weight and fat mass was measured must have been recorded to a very high degree of precision indeed.

The question about the precision of the body weight measurements is addressed in our response to the blog post by Ethan and Sarah Ludwin-Peery.

Two questions arise from this operation:

- First, it would be interesting to know what the adjustment process was. It seems to have been quite powerful, because some of the differences between the original and adjusted values are substantial. For example, for participant ADL014, the loss in weight on the unprocessed diet has been adjusted from 0.10 kg to 0.95 kg, and for ADL005 the equivalent loss has gone from 0.26 kg to 1.79 kg; participant ADL019's gain of 0.30 kg on the unprocessed diet has been adjusted to a loss of 0.24 kg, while participant ADL021's loss of 0.30 kg on the processed diet has been adjusted to a gain of 0.16 kg. These changes appear to affect principally the fat-free mass rather than the fat mass, which in numerous cases (8 out of 20 on the processed diet, 2 out of 19 on the unprocessed diet) is identical to two decimal places after adjustment. For example, participant ADL010's original weight gain of 3.60 kg on the processed diet becomes 2.69 kg in the adjusted file, but his fat mass did not change at all.
- Second, if the authors believe that these adjusted figures provide a better estimate of the effects of the diets, one might wonder why they have not submitted a correction, updating the claims about weight loss that featured in the abstract of their article, rather than allowing this important new information to languish in an OSF repository. Otherwise it is not clear what the point of performing these "adjusted" analyses was.

The results in the published manuscript correspond to the unadjusted data and code that was originally deposited on the OSF website. The adjustments in the second file on the OSF website were performed to address the fact that the DXA body composition measurements were not performed on exactly at the same time points for all subjects and subject ADL002 did not have a DXA measurement during the unprocessed diet period. The adjusted data attempt to estimate the mean changes in fat mass that would have occurred had the DXA measurements been aligned at the same time points. To do this, we calculated the slope of the best fit regression line to the fat mass measurements over each diet period to estimate the fat mass change that would have resulted if the DXA measurements were aligned. The DXA measurement at the end of the first diet period was also used as the fiducial measurement for the start of the second diet period and subject ADL002 contributed only 2 fat mass measurements during the unprocessed diet period.

The corresponding body weight measurements on those days were used to calculate the fat-free mass estimates by subtracting the estimated fat masses on those aligned days. This explains the minor differences between mean results reported in the original file deposited in OSF (which appear in the manuscript) and the first updated file. The mean results are not materially different, and the adjusted data merely address the potential criticism that the DXA measurements were not all conducted on the same days in all subjects. The reported data in the manuscript are not in error.

Conclusion
Hall et al.'s article seems to have had a substantial impact on the field of nutrition research. However, both Ethan & Sarah's post and this one raise a number of concerning questions about the reliability of this study. There seem to be problems with the design, the data collection process, and the analyses. I only looked at about half of the 23 data files, so there may be other problems lurking. I hope that the authors and the editors of *Cell Metabolism* will take another look at this study and perhaps consider issuing a correction of some kind.

As previously noted, a correction was issued in October of 2020 regarding an error that we independently discovered. Many of the other questions raised above are the result of misinterpretations of the data and we hope that we have now clarified these issues. Several remaining questions have been raised that we have yet to be able to address due to the limited amount of time provided because the authors of the blogs wanted to publish without allowing us the opportunity to fully respond.

Code and data
I have made my R analysis code, which reproduces most of the results reported above, here. Some of my results can probably best be checked by examining the data files in a spreadsheet, so my code also includes a loop (which you need to enable, following what I hope are clear instructions) that will export the original SAS data files to CSV format. Also included at the same location is a spreadsheet file named `snacks.xls` which summarises the nutrition information for the snacks that were served on the processed diet, plus the OSF screenshot and the SAS code and results files mentioned earlier.

Acknowledgements
Thanks to Andrew Althouse and James Heathers for help with the analyses, and to Ethan and Sarah Ludwin-Peery for sharing their discoveries about the Hall et al. article and some very interesting discussions about what it all might mean.

Footnotes
(*) I have put these terms in quote marks to emphasise that they have a specific technical meaning. I don't know if that it a good idea, though; perhaps it looks like I am putting Dr Evil-style air quotes around them. That isn't my intention.
(**) I asked James Heathers, my go-to person for all things physiology, to explain these numbers. He replied: "No-one did too much of anything... even slow walking to the shops is 2.5 METs. An hour's proper exercise including walking to and from the gym, even if you do nothing else all day, will push your daily MET up to a hair under 2".

While it is unclear what this second (**) footnote refers to, free-living people have an average physical activity level (defined as the total energy expenditure divided by resting energy expenditure) of ~1.6-1.8. Despite being admitted as inpatients when subjects often become very sedentary, our subjects had physical activity levels in the free-living range as was intended by mandating 60 minutes of daily cycle ergometry exercise at fixed wattage.