

A Large-Scale Observational Analysis of Social Media Data Reveals Major Public Misperception of the Attainability of Drastic Weight Loss by Dieting

Elad Yom Tov^a Irit Hochberg^{b, c}

^aFaculty of Industrial Engineering and Management, Technion Israel Institute of Technology, Haifa, Israel; ^bInstitute of Endocrinology, Diabetes and Metabolism, Rambam Health Care Campus and Faculty of Medicine, Haifa, Israel; ^cBruce Rappaport Faculty of Medicine, Technion Israel Institute of Technology, Haifa, Israel

Keywords

Weight · Diet · Internet

Abstract

Introduction: Diet forums in social media websites provide an opportunity to glimpse the experience of different weight loss diet strategies reported by tens of thousands of individuals. **Methods:** We analyzed all postings with weight information from the six major Reddit weight loss diet forums (“subreddits”) as reported by forum participants. **Results:** Data were collected from January 2011 to April 2020 from all 55,900 users posting weight information. Average start BMI was in the overweight or obese range (26–34 kg/m²), and average goal BMI was in the normal range (21.5–24.5 kg/m²) for all subreddits. There is correlation between start BMI and goal BMI ($R^2 = 0.63, p < 10^{-10}$) and between planned weight loss and reported weight loss ($R^2 = 0.56, p < 10^{-10}$). Approximately 80% of forum participants reported a weight loss that was greater than 5% of their initial body weight. Actual reported weight loss was less than half of goal weight loss. Average reported weight loss and adherence were highest in the keto and loseit subreddits. More upvotes and fewer downvotes were associated with higher reported weight

loss in five of the six subreddits. **Conclusions:** Despite the need for cautious interpretation of these data due to self-selection of users who updated weight loss and the possibility of unreliable weight reports, the study has several findings. Average goal BMI was in the normal weight range, demonstrating a highly unrealistic perception, in a very large lay-public cohort, of the plausibility of losing all excess weight. The success in weight loss and maintenance in self-selected individuals who continued reporting weight for many months may demonstrate the subjective value some individuals can obtain from forum participation.

© 2022 The Author(s).

Published by S. Karger AG, Basel

Introduction

A meta-analysis of weight loss attempts in adults reported that over 40% of the worldwide population surveyed would like to lose weight, and more than half of these individuals are dieting [1]. Comparison of clinical trials using different diet strategies for weight loss shows that the weight loss differences between individual named diets are relatively small, with average weight loss of diets

compared to no diet being in the range of 5–10 kg at 6 months and 4–7 kg at 12 months [2]. In long-term weight maintenance studies, the average dieter ended 3.15 kg lighter than their initial weight [3].

While the clinical trials reporting the effect of different weight loss diet strategies usually include a few dozen or at a most a few hundred patients [2], diet forums in social media websites provide an opportunity to glimpse the experience of different weight loss diet strategies reported by tens of thousands of individuals. In a sense, these individuals represent “real-life” dieting better than those participating in a clinical trial in a professional academic setting with tight healthcare-team follow-up.

A few previous observational studies have reported on some aspects of diet forums, most commonly inspecting the effect of group engagement, feedback, and support that individuals receive from other forum members in a single forum [4–6]. Previous studies have not compared various forums discussing different diet strategies. In this study, we aimed to compare weight loss expectations and success between different diet strategies using weight information (initial weight, goal weight, and longitudinal weight updates) from the six major Reddit weight loss diet forums (“subreddits”) as reported by forum participants.

Materials and Methods

Our data were obtained from the social media website Reddit (www.reddit.com), where users communicate in groups, called subreddits, centered on topics of interest. Preliminary analysis identified six active dieting-related groups that included information on the weight of users: fasting, intermittentfasting, keto, xxketo, loseit, and 1200isplenty. We excluded two groups (1500isplenty and weightwatchers) that had fewer than 200 participants who stated their weights. Other diet subreddits (such as Atkins, vegan, Paleo) were excluded as they include no user weight data.

We extracted all postings from these groups made between January 1, 2011, and April 30, 2020, using the Pushshift API. In the diet groups analyzed, users customarily provide their demographics and weight information through the “flair” field or in the posting text. For example, the text “32m/6’2”/SW307/CW203/GW195” refers to a 32-year-old male, 6 feet and 2 inches tall (1.9 m), who started his diet at 307 lb. (139.2 kg), is aiming for a goal weight of 195 lb. (88.5 kg), and is currently 203 lb. (92.1 kg). We matched all such patterns in postings to these subreddits.

Users’ weights were estimated over the first year since their first post in one of the diet subreddits using the following procedure: we noted all posts where the current weight was reported during the first year, and removed those reports that were identical to a previous post, under the assumption that a user may have not updated their weights. We then used linear interpolation among ad-

jacent pairs of reported weights to estimate the weight at each day. If the weight reports were more than 60 days apart, we did not perform this procedure and assumed that the data for those days were missing.

We define adherence as the length of time between the first and last postings of a user on a given subreddit. We defined the weight plan as the difference between the start and goal weight and weight change as the difference between the initial weight and the last reported weight.

This study was approved by the Technion’s Behavioral Sciences Research Ethics Committee (approval number 2018-032). Informed consent was not obtained as this analysis was of open-source data uploaded by anonymous forum participants.

Multiple ANOVA was used for assessing effect of categorical variables on subreddit choice. Kruskal-Wallis test with ANOVA among pairs was used to compare weights, weight loss, and adherence between subreddits. Ranksum test was used to compare intervals and weights between weight posts. Two-sample Kolmogorov-Smirnov goodness-of-fit hypothesis test was used to compare weight loss between pairs of subreddits.

Results

Description of Subreddit Diet Groups

Data were collected from over 55,900 users. Table 1 shows the subreddits that were analyzed and the number of users and number of posts found in each of them. Across all analyzed posts, starting weight was provided in 92% of posts, goal weight in 94%, and current weight in 82% of posts. Gender was given in 44% of posts, age in 42%, and height in 79%.

The number of posts per day from 2011 to 2020 in each diet subreddit is presented in online supplementary Figure S1 (see www.karger.com/doi/10.1159/000528803 for all online suppl. material). The most popular subreddit was keto, with participation of around 30 posts a day between 2013 and 2019. The second most popular subreddit was loseit, which rose to 30 posts a day between 2016 and 2019. The popularity of both has steeply declined since 2020.

The average age of participants in the different subreddits was 28–34 years old and the percentage of females ranged from 42.9% to 88.3%. Multiple ANOVA of age, gender, and their interaction revealed all three variables were statistically significantly ($p < 10^{-10}$) associated with the choice of diet subreddit, with xxketo having the largest percentage of female users.

Start Weights and Weight Goals of Diet Group Participants

Figure 1 shows the start and goal BMI of users in the different subreddits. Average initial BMI in all subreddits

Table 1. Analyzed subreddits, their description on the website, and numbers of users with weight information (the latter refers to users who reported all 3 weights: current, start, and goal weight)

Subreddit	Description of subreddit	Number of users with all weight information	Number of posts of users with weight information	Average age, years	Female, %
fasting	Whatever form of water fasting you practice, from IF to extended fasting and everything in between...	379	2,468	33.7	48.4
intermittentfasting	Intermittent Fasting (IF) is way of eating that restricts *when* you eat, usually on a daily or weekly schedule...	1,224	3,837	29.4	47.8
keto	The Ketogenic Diet is a low carbohydrate method of eating. /r/keto is place to share thoughts, ideas, benefits, and experiences around eating within a Ketogenic lifestyle	19,861	101,669	31.2	42.9
xxketo	A subreddit dedicated to discussing a ketogenic diet from a female-identifying perspective	4,432	29,537	31.5	88.1
loseit	A place for people of all sizes to discuss healthy and sustainable methods of weight loss...	14,911	48,533	28.2	54.5
1200isplenty	A sub for recipes, memes, and support related to low-calorie diets, targeted at people who have low TDEEs	793	4,942	28.3	88.3

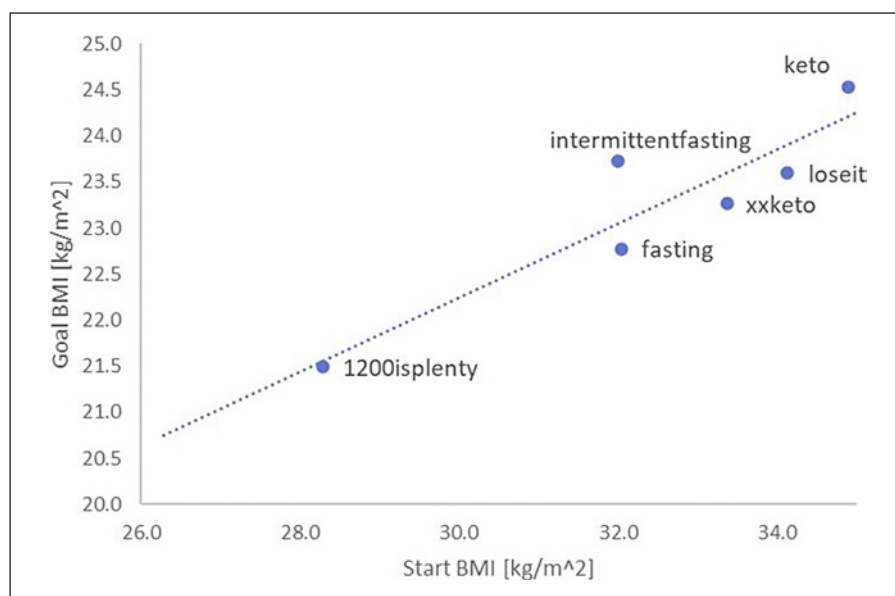


Fig. 1. Start and goal BMI for the subreddits. The dotted line is a linear regression curve with $R^2 = 0.82$.

was in the overweight to obese range (BMI 26–34 kg/m²). People in the two Keto subreddits and the loseit subreddit had the highest start BMI. The start BMIs are statistically significantly different among all subreddits, except for the fasting and intermittentfasting subreddits (Kruskal-Wallis test with ANOVA among pairs, $p < 10^{-3}$).

The goal BMI for all subreddits was in the normal weight range (21.5–24.5 kg/m²). The goal BMIs of

1200isplenty and keto are statistically significantly different (lower) from those of all other subreddits (Kruskal-Wallis test with ANOVA among pairs, $p < 10^{-3}$). The start BMI and goal BMI of users are correlated with $R^2 = 0.63$ ($p < 10^{-10}$); that is, people with a higher start weight aim for a higher goal weight and vice versa. We compared the weight plan (the difference between the start and goal weight) and weight change (the difference between the

Table 2. Weight loss, activity, and adherence on each of the diet subreddits

Subreddit	Percentage of users with multiple posts	Time span, days (standard deviation)	Starting weight, kg (STD)	Median weight loss, kg (STD)	Mean weight loss, kg (STD)	Fraction of goal weight loss attained (STD)
intermittentfasting	18	68 (82)	97.4 (30.1)	9.1 (13.4)	12.9 (13.4)	0.338 (0.363)
fasting	25	112 (168)	97.4 (35.0)	8.6 (19.5)	13.2 (19.5)	0.285 (0.550)
1200isplenty	53	171 (205)	77.0 (22.0)	9.1 (11.7)	12.3 (11.5)	0.493 (0.291)
loseit	35	221 (287)	100.6 (34.0)	11.3 (16.2)	16.2 (16.2)	0.416 (0.350)
xxketo	52	241 (353)	93.8 (25.5)	9.1 (12.4)	12.2 (12.4)	0.371 (0.448)
keto	50	291 (415)	107.4 (32.6)	12.2 (15.9)	16.3 (15.9)	0.419 (1.340)

Time span refers to the average time between the first and last post of users on the subreddit. Fraction of weight loss refers to the maximal weight loss reported by the user relative to the goal weight loss.

initial weight and the last reported weight). The correlation between the first weight plan and final weight change is 0.56 ($p < 10^{-10}$, $n = 39,124$), showing correlativity between the weight that people wanted to lose and their actual weight loss. There was no correlation or anticorrelation between adherence, defined as the length of time between the first and last postings of a user on a given subreddit, and the weight plan ($r = 0.01$, $p = 0.34$, $n = 9,737$).

Changes in Start Weights and Goal Weights over Time

Among users who were active on a single subreddit and who had more than one post, approximately 15% posted more than one start weight. The average time difference between posts that had identical start weights was 50 days and those that had different start weights was 301 days (statistically significant, ranksum test, $p < 10^{-10}$). For these 15% of users in which start weight was changed, the average difference between the maximal and minimal start weights was 8.8 kg. In 28% of cases, start weights which were posted earlier were greater than those posted later.

Among users who were active on a single subreddit and who had more than one post, approximately 19% posted more than one goal weight. The average time difference between posts which had identical goal weights was 51 days and those which had different goal weights was 215 days (statistically significant, ranksum test, $p < 10^{-10}$). The average difference between the maximal and lowest goal weights was 8.4 kg. In 61% of cases, earlier posted goal weights were higher than those posted later.

In 24% of cases where either start or goal weight was updated, both were changed. In those cases, the correlation between the change of goal weight and the change of start weight was 0.41 ($p < 10^{-10}$).

Posting a new goal weight was associated with posting a new start weight (χ^2 test, $p < 10^{-10}$). Therefore, we deduce that people posted new start weights when they viewed the later time period as a new diet.

Weight Loss Reported by Diet Group Participants

Table 2 presents the median and mean weight loss as well as the fraction of reported weight loss from the goal weight loss. Average reported weight loss was highest in the keto subreddit (16.3 kg \pm 15.9), followed by loseit (16.2 kg \pm 16.2). Median reported weight loss was also highest in these same subreddits.

Kruskal-Wallis nonparametric ANOVA shows that differences between weight loss reported in the 1200isplenty, fasting, intermittentfasting, and xxketo subreddits are not statistically significant from each other, while weight loss in the keto and loseit subreddits are significantly different from all other subreddits including each other ($p < 0.05$). The average fraction of weight loss attained (reported weight loss divided by weight plan) was between 0.28 and 0.49 (Table 2), meaning that the majority of individuals did not lose even half the amount of weight they declared they were planning to lose when entering the forum.

Adherence

We define adherence as the length of time between the first and last postings of a user on a given subreddit. Table 2 shows the average number of days between the first and last posts of users on each of the subreddits, as well as the percentage of users who had more than one post. As the table shows, users stayed longest in the two keto subreddits, followed by loseit. Users in the intermittentfasting and fasting subreddits stayed for the shortest time and were the least active.

Kruskal-Wallis nonparametric ANOVA shows that adherence in the 1200isplenty subreddit is statistically

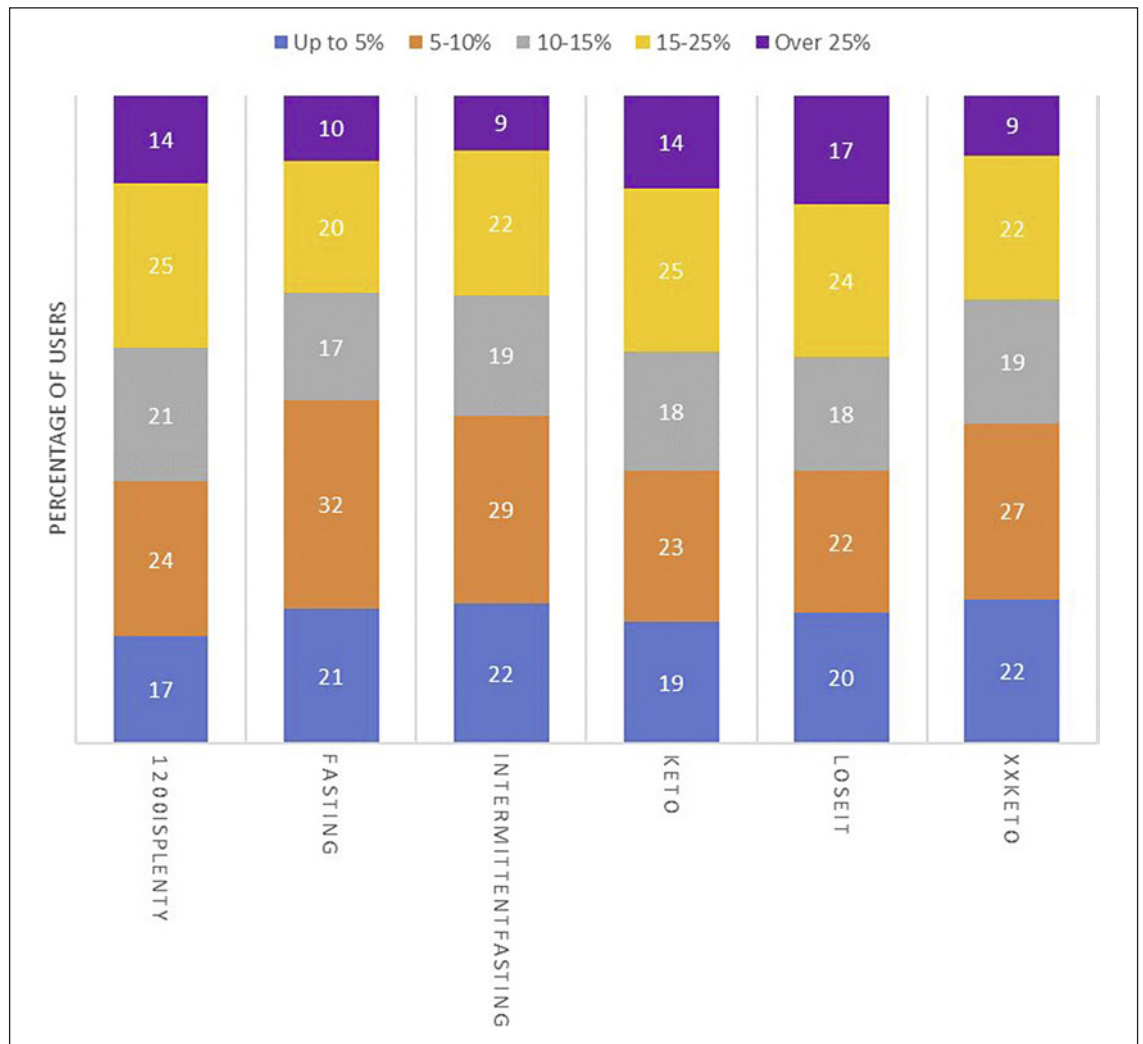


Fig. 2. Distribution of highest weight loss percentage subgroups for different diet subreddits.

significantly different (lower) from all other subreddits ($p < 0.05$). Adherence on the fasting and intermittentfasting subreddits is not statistically significantly different from each other but is significantly lower than other subreddits ($p < 0.05$). Similarly, the xxketo and loseit subreddits are not statistically significantly different from each other but are different from all other subreddits ($p < 0.05$). Adherence to the keto subreddit is statistically significantly longer than to all other subreddits ($p < 0.05$).

Weight Change

Figure 2 shows the distribution of maximum percentage of weight loss (computed as the difference between the start and minimum reported weights, divided by the start weight) reported by users in the different

diet subreddits. In all subreddits, the majority of individuals that reported longitudinally achieved a clinically meaningful weight loss, defined as at least 5% of start weight: 22–32% lost 5–10% of their start weight, 17–21% lost 10–15% of their start weight, 20–25% lost 15–25% of their start weight, and 9–17% lost over 25% of their start weight. Comparing weight loss between pairs of subreddits using a two-sample Kolmogorov-Smirnov goodness-of-fit hypothesis test found statistically significant differences between several pairs of subreddits (Fig. 2). Compared to other diets, fewer people in the fasting subreddit lost more than 10% of their weight. The largest proportion of people losing over 15% of their weight was reported in the Loseit keto and 1200isplenty subreddits.

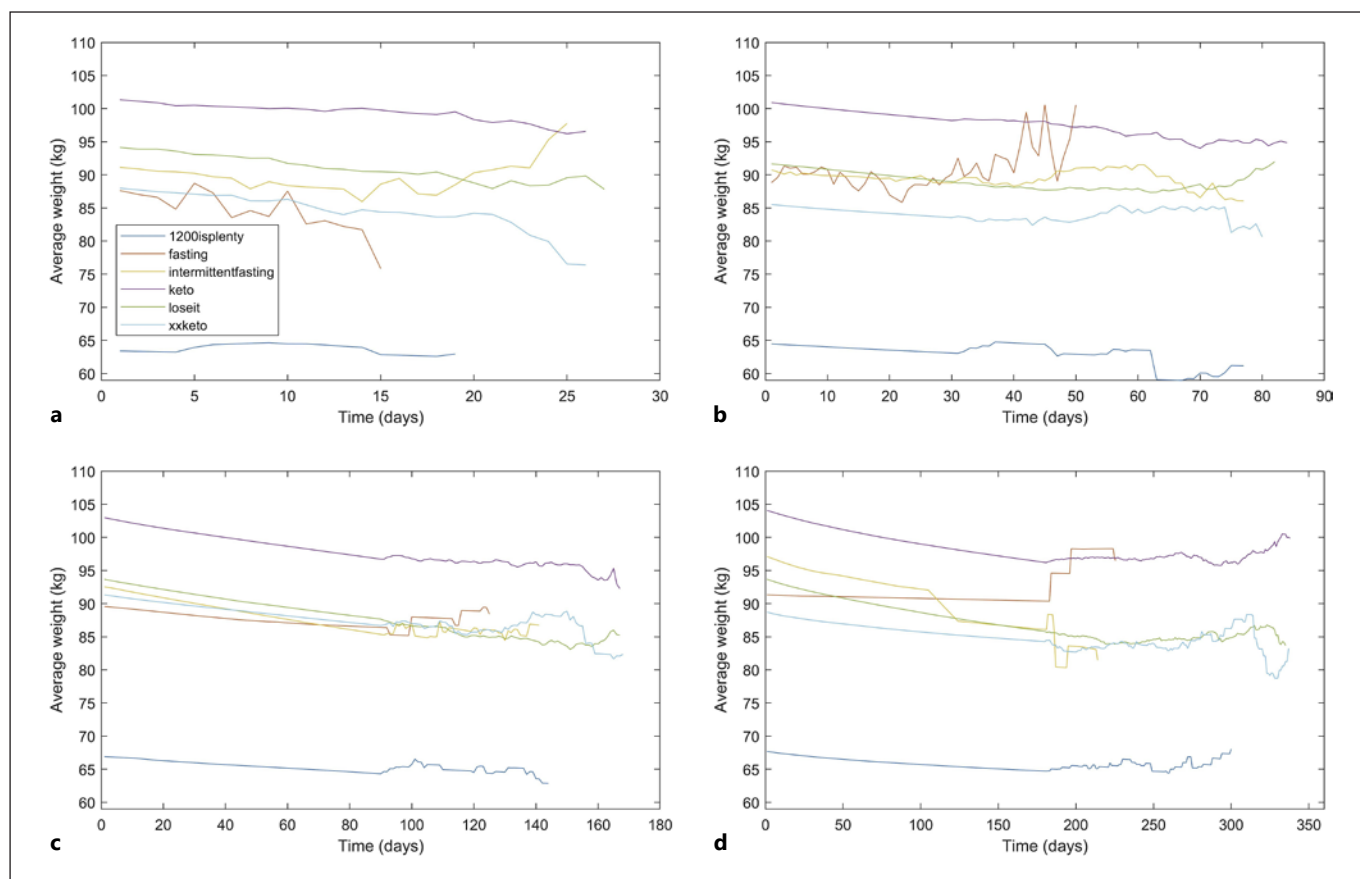


Fig. 3. Average weight of users over time, stratified by days of adherence. **a** Adherent 2–30 days. **b** Adherent 31–60 days. **c** Adherent 61–180 days. **d** Adherent 181–360 days.

Figure 3 shows the reported weights over time of the users, stratified by the adherence of users to continued participation in the diet forum. Online supplementary Figure S2 shows the same analysis with error bars. Long-term (over 180 days) adherence is associated with an average weight loss of 11.5 kg for people in the intermittentfasting subreddit, 10.5 kg in the loseit subreddit, and 10.2 kg in the keto subreddit.

A linear model of weight loss as a function of subreddit and adherence reached a low R^2 (0.03 , $p < 10^{-10}$, $n = 9,737$). Nevertheless, several statistically significant interactions between adherence, subreddit, and weight loss were found. These include an association of higher weight loss to participation in specific subreddits (keto, xxketo, and loseit). Additionally, shorter adherence in the intermittentfasting subreddit and longer adherence in the keto and xxketo subreddits were associated with lower weight loss.

Effect of Group and Initial Interaction on Weight Loss

We modelled weight change as a function of the parameters of the first interaction of a user in the subreddit. The interactions of other users with a users' post include the number of up and down votes that the post received as well as the number of comments made in response to the post. Table 3 shows the parameters of the model. Although the model explains only a small percentage of the variance ($R^2 = 0.016$), a positive introduction to the subreddit (more upvotes and, especially, fewer downvotes) is associated with higher reported weight loss in five of the six subreddits. The effect was the opposite in the loseit subreddit, in which positive upvotes led to lower weight loss (Table 3).

Table 3. Model of weight loss as a function of the responses to the users' first post on the subreddit

Variable	Slope
Ups	0.016*
Downs	-0.363*
Number of comments	-0.005
fasting	-0.391
intermittentfasting	-6.282
keto	-0.494
loseit	5.698
xxketo	-1.737

Asterisk marks statistically significant interactions ($p < 0.05$). $N = 4,109$, $R^2 = 0.016$.

Discussion

In this study, we collected data on weight reported in Reddit forums by over 55,900 people attempting to lose weight by dieting. The most popular diet subreddit in the study period was keto (ketogenic diet), followed by loseit (Table 1; Fig. 1).

The average weight loss of participants in the different diet groups was 12–16 kg, with 78–83% of users in the different subreddits reporting over 5% weight loss. These results are not inferior to the weight loss reported in short-term clinical studies [2], but several potential caveats make their interpretation questionable. Some of the weight reports may be unreliable for various reasons [7]. People reported an update on their current weight at a self-selected-point in time that may reflect their best period of diet adherence, unlike clinical studies where all participants are weighed at prespecified follow-up times. We can expect a significant self-selection bias where people that were successful in continuously losing weight would be more likely to stay in the forum and report additional weight measurements than people failing to lose weight. Due to these limitations, weight loss reported in this cohort cannot support or refute the effectiveness of diet forum participation or the importance of structured weight loss programs with long-term clinical follow-up. On average, participants who reported weights longitudinally in the keto and loseit subreddits lost the most weight (Table 2).

The most distinct finding in our view is that the average goal BMI in all subreddits is in the normal weight range (Table 2). This demonstrates a highly unrealistic perception, in a very large lay-public cohort, of the plausibility of losing all excess weight and achieving a drastic

metamorphosis of body habitus by dieting. While it is well established that the actual weight loss in clinical studies of different diet strategies is 4–7 kg [2] and long-term maintained weight loss is an average of 3.1 kg [3], the average Reddit user aimed to lose 24–37 kg (Fig. 1). Such unrealistic expectations for drastic weight loss have been described 2 decades ago in patients presenting to obesity outpatient clinics [8, 9]. Our study repeats these results in a current, much larger cohort and proves that, despite several decades of clear data in the scientific community on the chronic nature of obesity and modest effect of diets, the general public is greatly misinformed on the subject.

According to our results, expectations to lose all excess weight are not detrimental to short-term weight loss forum participants who continued reporting their weight. Our analysis of the effect of weight expectations showed a positive effect on weight loss, and individuals whose weight goal was lower lost more weight on average.

This result is not concordant with that of most previous studies on unrealistic weight loss expectations that have reported that such expectations are associated with failure to adhere to weight loss interventions [10]. We might speculate that our different result has to do with the strong self-selection of the group who continued reporting their weight in the forum, and possibly some of the people who stopped updating their weight may be those whose weight loss is much lower than goal weight loss.

Continued forum participation, which we defined as adherence, was longest in the keto and xxketo subreddits, an average of approximately 10 and 8 months, respectively (Table 2). Data on weight maintenance are quite sparse, but when looking at the average reported weight in each subreddit stratified by adherence (Fig. 3), we see that on average, the subgroup of forum participants that continued participating and reporting weight were apparently successful in achieving and maintaining a reduced weight. Participants who continued reporting their weight over 6 months maintained, on average, over 10 kg of weight loss compared to their initial reported weight. These self-selected individuals demonstrate the subjective value some individuals can obtain from forum participation and the support they receive or provide to other users. Approximately half of the participants in half of the forums wrote more than one post (Table 2). Receiving positive feedback and not negative feedback in response to the first post is correlated with weight loss in most forums (Table 3), a finding that supports the value of online group participation for some individuals aiming to achieve lifestyle changes.

The study has several limitations: the weight data are self-reported by individuals participating in the diet forums and cannot be validated; the timespans between weight measurements were inconsistent; we do not have information on use of weight loss medications or exercise; and most participants stopped recording their weight after weeks or months. Another point we should emphasize it that we could only analyze weight data from diet forums in which participants actually recorded their weights, and people choosing this type of forum may differ from those participating in other diet forums where weight is not reported. Confounding factors that impede comparison of weight loss in different diet strategies include different lengths of participation and absence of long-term information on individuals that stopped reporting weight.

The main strengths of this study are the large sample size and the supposedly credible information on initial weight and target weight that was consistently provided by all users analyzed. Even if they are inaccurate for some of the users, they can provide valid insights into common perceptions.

In future work, we plan to analyze differences in adherence according to the demographic factors available in the data, such as gender. This may shed more granular information on adherence among different populations.

Conclusion

Our results imply that there is widespread misinformation on the magnitude of weight loss possible to achieve by dieting. This may harm individuals who could benefit from modest weight loss, possibly leading to diet

discontinuation and weight regain when unrealistic goals are unmet. How to best correct this misinformation for the benefit of the public is worth a serious discussion.

Statement of Ethics

The study was approved by the Technion's Behavioral Sciences Research Ethics Committee, approval number 2018-032. Patient consent was not required as this study was based on publicly available data.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

The research relevant to the study had no external funding and was done as part of the authors' academic hours. Service of an English editor was self-funded by the authors.

Author Contributions

Eyal Yom Tov: conceptualization, data curation, formal analysis; investigation, methodology, software, visualization, writing – original draft, and writing – review and editing. Irit Hochberg: conceptualization, investigation, methodology, project administration, writing – original draft, and writing – review and editing.

Data Availability Statement

The data are open online on Reddit; further inquiries can be directed to the corresponding author.

References

- Santos I, Sniehotta FF, Marques MM, Carraça EV, Teixeira PJ. Prevalence of personal weight control attempts in adults: a systematic review and meta-analysis. *Obes Rev*. 2017 Jan;18(1):32–50.
- Johnston BC, Kanters S, Bandayrel K, Wu P, Naji F, Siemieniuk RA, et al. Comparison of weight loss among named diet programs in overweight and obese adults: a meta-analysis. *JAMA*. 2014 Sep 3;312(9):923–33.
- Anderson JW, Konz EC, Frederich RC, Wood CL. Long-term weight-loss maintenance: a meta-analysis of US studies. *Am J Clin Nutr*. 2001 Nov 1;74(5):579–84.
- Hekler EB, Dubey G, McDonald DW, Poole ES, Li V, Eikey E. Exploring the relationship between changes in weight and utterances in an online weight loss forum: a content and correlational analysis study. *J Med Internet Res*. 2014 Dec 8;16(12):e254.
- Pappa GL, Cunha TO, Bicalho PV, Ribeiro A, Couto Silva AP, Meira W Jr, et al. Factors associated with weight change in online weight management communities: a case study in the LoseIt Reddit community. *J Med Internet Res*. 2017 Jan 16;19(1):e17.
- Liu Y, Yin Z. Understanding weight loss via online discussions: content analysis of Reddit posts using topic modeling and word clustering techniques. *J Med Internet Res*. 2020 Jun 8;22(6):e13745.
- Drouin M, Miller D, Wehle SMJ, Hernandez E. Why do people lie online? "Because every-one lies on the internet. *Comput Hum Behav*. 2016 Nov;64:134–42.
- Foster GD, Wadden TA, Phelan S, Sarwer DB, Sanderson RS. Obese patients' perceptions of treatment outcomes and the factors that influence them. *Arch Intern Med*. 2001 Sep 24; 161(17):2133.
- Dalle Grave R, Calugi S, Molinari E, Petroni ML, Bondi M, Compare A, et al. Weight loss expectations in obese patients and treatment attrition: an observational multicenter study. *Obes Res*. 2005 Nov;13(11):1961–9.
- Moroshko I, Brennan L, O'Brien P. Predictors of dropout in weight loss interventions: a systematic review of the literature. *Obes Rev*. 2011 Nov;12(11):912–34.