

## Singapore Management University Institutional Knowledge at Singapore Management University

---

Research Collection School Of Information Systems

School of Information Systems

---

4-2018

# Does journaling encourage healthier choices? Analyzing healthy eating behaviors of food journalers

Palakorn ACHANANUPARP

Singapore Management University, [palakorna@smu.edu.sg](mailto:palakorna@smu.edu.sg)

Ee Peng LIM


Singapore Management University, [eplim@smu.edu.sg](mailto:eplim@smu.edu.sg)

Vibhanshu ABHISHEK

Carnegie Mellon University

**DOI:** <https://doi.org/10.1145/3194658.3194663>

Follow this and additional works at: [https://ink.library.smu.edu.sg/sis\\_research](https://ink.library.smu.edu.sg/sis_research)

 Part of the [Databases and Information Systems Commons](#), [Health Information Technology Commons](#), and the [Numerical Analysis and Scientific Computing Commons](#)

---

### Citation

ACHANANUPARP, Palakorn; LIM, Ee Peng; and ABHISHEK, Vibhanshu. Does journaling encourage healthier choices? Analyzing healthy eating behaviors of food journalers. (2018). *DH '18: Proceedings of the 2018 International Conference on Digital Health, Lyon, France, April 23-26*. 35-44. Research Collection School Of Information Systems.

**Available at:** [https://ink.library.smu.edu.sg/sis\\_research/4050](https://ink.library.smu.edu.sg/sis_research/4050)

This Conference Proceeding Article is brought to you for free and open access by the School of Information Systems at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Information Systems by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email [libIR@smu.edu.sg](mailto:libIR@smu.edu.sg).

# Does Journaling Encourage Healthier Choices? Analyzing Healthy Eating Behaviors of Food Journalers

Palakorn Achananuparp  
Singapore Management University  
Singapore, Singapore  
palakorna@smu.edu.sg

Ee-Peng Lim  
Singapore Management University  
Singapore, Singapore  
eplim@smu.edu.sg

Vibhanshu Abhishek  
Carnegie Mellon University  
Pittsburgh, PA  
vibs@andrew.cmu.edu

## ABSTRACT

Past research has shown the benefits of food journaling in promoting mindful eating and healthier food choices. Consequently, it can be implied that the act of journaling can inherently lead to healthy lifestyle change. However, the links between journaling and healthy eating have not been thoroughly examined. Beyond caloric restriction, do journalers consistently and sufficiently consume healthful diets? How different are their eating habits compared to those of average consumers who tend to be less conscious about health? In this study, we analyze the healthy eating behaviors of active food journalers using data from MyFitnessPal. Surprisingly, our findings show that food journalers do not eat as healthily as they should despite their proclivity to health eating and their food choices resemble those of the general populace. Furthermore, we find that the journaling duration is only a marginal determinant of healthy eating outcomes and sociodemographic factors, such as gender and regions of residence, are much more predictive of healthy food choices.

## CCS CONCEPTS

• Applied computing → Health informatics; Consumer health;

## KEYWORDS

Health; Food; Eating Behaviors; Food Journals; Food Diaries; MyFitnessPal; Personal Informatics; Quantified Self

## ACM Reference Format:

Palakorn Achananuparp, Ee-Peng Lim, and Vibhanshu Abhishek. 2018. Does Journaling Encourage Healthier Choices? Analyzing Healthy Eating Behaviors of Food Journalers. In *DH'18: 2018 International Digital Health Conference, April 23–26, 2018, Lyon, France*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3194658.3194663>

## 1 INTRODUCTION

Recent progress in mobile and wearable technologies has provided individuals the means to routinely track data about themselves for self-knowledge and improvement. This self-tracking practice is also known as *quantified self* or *personal informatics*. In the domain of dietary self-monitoring, mobile food journal apps, such as MyFitnessPal (MFP hereafter), are one of the most popular tracking

methods widely used by millions of people. Past research has suggested that food journaling is an effective intervention in weight loss programs [9]. The act of journaling helps create increased in-the-moment awareness (mindfulness) and can encourage healthier choices [13, 14]. As the stage of the obesity epidemic continues to worsen, understanding how the journaling practice affects eating behaviors may provide useful insights for the designs of an effective population-wide intervention. While there is growing evidence supporting the critical role of food journals in improving weight loss outcomes, little empirical work has been done to investigate the broader impacts of the journaling practice on the individuals' healthy eating behaviors. Especially, more evidence is needed to: (1) quantitatively compare long-term healthy eating habits of food journalers with other behavioral baselines, such as the general public and the dietary recommendations; and (2) measure the influences of the journaling practice on the healthy eating outcomes with respect to other factors. From a methodological perspective, it is also an opportunity to further explore the use of a large-scale self-tracking data in conjunction with offline data sources to answer these questions.

To address the research gaps, our study aims to assess the healthy eating behaviors of food journalers by analyzing public food diary entries of MFP users and comparing their eating behaviors to those of the general populace reported in other studies. Although people tend *not* to perceive healthy eating the same ways as public-health experts [5, 41], a recent survey [13] suggests that a vast number of food journalers (past and present) generally agree with experts about the notion of healthy eating, e.g., most believe that they should eat more fruits and vegetables, lean meat, and balanced diets. Thus, it is reasonable to assume that food journalers are more likely to achieve evidence-based healthy eating outcomes, as defined by public-health experts, than the average consumers who may be less informed about healthy eating. Next, we expect active food journalers, who are likely to develop a mindful eating habit, to consciously make healthy food choices and be less influenced by the sociodemographic biases. Specifically, we formulate the following research questions:

**RQ1:** *Do active food journalers have healthier eating behaviors than the general populace?*

To investigate the effectiveness of journaling in encouraging healthier food choices, we aim to characterize the healthy eating behaviors of food journalers using the corresponding intakes from the dietary guidelines and the general populace as comparison data. If food journalers tend to (1) have higher intakes of healthy diets and lower intakes of unhealthy diets than the general populace and (2) consistently meet the recommended intakes per the dietary guidelines, then such findings may provide evidence supporting the notion that not only does journaling is linked to significant weight loss, but it also plays a significant role in individuals' healthier food choices.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*DH'18, April 23–26, 2018, Lyon, France*

© 2018 Association for Computing Machinery.  
ACM ISBN 978-1-4503-6493-5/18/04...\$15.00  
<https://doi.org/10.1145/3194658.3194663>

**RQ2:** *How do the eating behaviors of food journalers significantly differ across sociodemographic groups?*

*Mindless eating* describes a situation where individuals are unaware of the influences exerted on their food choices by external factors, such as the environment and gender roles. It is generally associated with unhealthy eating habits and weight gains [45]. Here, we seek to further examine the relationship between the healthier food choices effect of journaling and the healthy eating behaviors of food journalers. Specifically, we expect food journalers of different sociodemographic backgrounds (e.g., gender, age, etc.) to be equally conscious of their food choices such that their healthy eating behavior is more homogeneous than that of the general populace. For example, male and female journalers should consume a comparable amount of fruits and vegetables. In other words, the difference in fruit and vegetable intakes between male and female journalers should not be statistically significant.

**RQ3:** *To what extent does the journaling practice influence the eating behaviors of food journalers?*

Past research has shown that weight loss outcomes are proportional to the journaling practice. That is, individuals who are more active in recording their food journals tend to lose more weights than the less active journalers [9]. As such, we aim to quantify the impacts of *journaling duration* and *persistence* on the healthy eating behaviors using regression analysis. If food journaling in fact encourages healthier food choices and mindful eating, we expect the journaling factors to have a higher influence on the eating behaviors than the sociodemographic factors. Such findings, together with those from the other research questions, will help demonstrate that food journaling is an effective healthy lifestyle intervention. On the other hand, unexpected results may provide us insights into the flaws of mobile food journals in affecting health behavior changes.

The main contributions of our work are as follows. *Firstly*, we thoroughly examine the relationship between food journaling, sociodemographic factors, and a variety of healthy eating behaviors in a large population of nearly 10,000 active food journalers over a six-month period. *Secondly*, we present our data preprocessing steps to automatically generate an analysis-ready dataset including: (1) identifying foods and beverages relevant to targeted healthy eating behaviors from the annotated food diary entries; (2) extracting portion sizes of foods and beverages from the food diary entry text and the associated caloric value; and (3) normalizing the portion sizes of varying measurement units into standard nutritional units. *Lastly*, our findings suggest, in contrast to past studies, that the journaling duration only plays a minor role in determining healthy eating behaviors, whereas several other factors, such as gender, journaling persistence, and regions of residence, are much more influential in determining the healthy eating behaviors than the journaling duration.

In what follows, we begin by reviewing related work on food journaling and data-driven approaches to health behavior assessment. Then, we introduce the healthy eating behaviors considered in the study. Next, we describe the dataset and data preprocessing steps. In the subsequent section, we define (1) the quantitative measures of the eating behaviors; and (2) the sociodemographic and journaling factors being studied. Finally, we present the findings, discuss their significance on the food journaling practice and future design implications, and conclude the paper.

## 2 RELATED WORK

### 2.1 Personal informatics and food journaling

Much research in personal informatics has focused on modeling and characterizing the use of tools and technologies to track one's own personal data for self-discovery and behavior change in a variety of domains [19, 41]. In particular, a few researchers have explored the use of mobile food journals and other online tools for dietary self-tracking in recent years [13, 14, 18]. Cordeiro et al. identified several key challenges related to the journaling tools and practices, such as unreliable data, negative nudges, and lack of social support [14]. To overcome tracking burden and fatigue in food journaling, Epstein et al. proposed a lightweight approach to food journal for promoting mindful eating [18]. Recently, Chung et al. [11] studied the practice of food tracking amongst Instagram users and the role of social support on their healthy eating pursuit.

Our work is complementary to previous research about food journaling [13, 14, 18, 19]. While many studies aimed to characterize various aspects of the self-tracking practices using qualitative methods, few studies have taken a computational approach to examine the broader impacts of journaling on the healthy eating behaviors of food journalers. In this study, we analyze more than 1 million food diary entries to quantitatively assess the behavioral impacts of journaling.

### 2.2 Using online data to assess health behaviors

Data from online social media, quantified-self, and other online sources have been used to study various aspects of health and eating behaviors. First, a few studies [1, 15] analyzed mentions of foods in the Twitter network to track the prevalence of obesity and other public health issues. Next, Park et al. [39] investigated the impacts of user profile, fitness activity, and fitness network of Twitter users on the long-term engagement of fitness app users. Mejova et al. [31] analyzed food pictures shared by Instagram users to study the prevalence of obesity. Recently, a few studies have investigated the tasks of predicting diet compliance outcomes using MFP food diary data [46] together with Twitter data [17].

Our work is highly relevant to [17, 46] in which the researchers studied diet compliance outcomes from MFP food diary data and constructed computational models to predict diet compliance success using different types of features, such as words & food types identified from the diary entries and social and linguistic attributes extracted from the users' social media messages. While their studies particularly focused on caloric balance as the primary outcome, we examine a more comprehensive set of eating behaviors by using evidence-based healthy eating outcomes as the primary measures. Additionally, we investigate the role of sociodemographic and journaling factors, derived from the user profile and food diary data, in determining the healthy eating behaviors.

## 3 HEALTHY EATING BEHAVIORS

We begin by introducing the evidence-based healthy eating outcomes categorized by the consumption of: (1) fruits and vegetables; (2) animal-based protein sources, such as red and processed meat, poultry, and fish; and (3) added sugars and sugary drinks. They are identified based on growing scientific evidence from several

randomized controlled trials and meta-analysis about their associations with health benefits and risks. Together, they constitute dietary intakes commonly recommended by most dietary guidelines [3, 20, 22, 36, 37, 47].

### 3.1 Fruits and vegetables

High intake of fruits and vegetables (abbreviated as FV) provides a variety of long-term health benefits, such as lowering the risk of cardiovascular disease [44] and cancers [30]. On the other hand, low intakes of FV are associated with the increased prevalence of obesity and diabetes [32]. Generally, a recommended daily FV intake for healthy adults is **at least 5 servings** [32] or approximately 400 - 500 grams [22, 47]. Despite numerous health benefits, the consumption of fruits and vegetables has been persistently low in the US [6] and worldwide [47] for decades. From 1994 - 2005, the average daily FV intake amongst Americans has decreased slightly from **3.43 servings** to **3.24 servings** while the percentage of people who met the recommended daily intake has remained unchanged at about **25%** [6]. In 2015, the Centers for Disease Control and Prevention (CDC) reported that **less than 15%** of Americans sufficiently met the recommendations [22].

### 3.2 Red and processed meat, poultry, and fish

Growing evidence suggests that processed meat is carcinogenic while red meat is probably carcinogenic [30]. High consumption of red and processed meat may increase mortality rates of type-2 diabetes [4], cardiovascular disease [42] and colorectal cancer [30]. Replacing red and processed meat with healthier protein sources such as white meat (e.g., poultry and fish) [2] may lower the risk of all-cause mortality. Most dietary guidelines recommend to limit a daily intake of red and processed meat to **1 serving** or approximately 65 - 75 grams [20]. Despite high associations with various health risks, red and processed meat still accounts for more than 50% of meat consumed in the United States with almost **2.31 servings** per day [16]. Moreover, the consumption of healthier white meat, such as fish, amongst Americans is low. On average, most adults consumed **0.17 servings** (17.28 grams) of fish a day [38], 70.59% lower than the recommended daily intake of **0.29 servings** (28.57 grams) [3].

### 3.3 Added sugars and sugary drinks

Added sugars are sugars not naturally occurring in foods and beverages. Asides from sweetening and adding extra calories, they provide no nutritional benefits. Sugary drinks (e.g., sodas/soft drinks, energy drinks, coffee drinks, and fruit juices) are the largest source (36%) of added sugar intake in American diets [35]. According to the recent survey by Gallup [24], **48%** of Americans drink at least 1 glass of soft drink on an average day. Ideally, sugary drink consumption should be avoided at all cost; or else an intake should be limited to **237 milliliters** a day [36]. There is strong evidence linking high consumption of added sugars and sugary drinks to increased risks for obesity and type-2 diabetes, whereas simply lowering the intake of sugary drinks can reduce weight gain and decrease prevalence of obesity [27]. Recent dietary guidelines suggest a maximum daily intake of added sugars to **25 grams** [37]. However, an American adult consumes on average **70 grams** [21] of added sugar per day, 180% higher than the recommended amount.

## 4 DATA

### 4.1 Collecting and processing MyFitnessPal data

We used a public food diary dataset<sup>1</sup> collected from 9,896 MyFitnessPal users in March 2015 by Weber and Achananuparp [46]. The dataset includes 71,715 unique food entries recorded over 1,919,024 meals from a 6-month recording period between October 2014 to March 2015. Each user recorded 59.3 days of diaries on average (S.D. = 54.6, median = 42). The majority of users were able to achieve their daily caloric goals [46]. Users who recorded at least 7 days of diary entries were treated as active users (N=8,381; 85.69% of all users in the dataset). Each food diary text was automatically annotated with categorical information describing its composition (e.g., food groups) and cooking method in the previous work [46]. In addition, basic nutritional facts (e.g., calories, protein, sugar) for each food entry were available. Next, we performed data cleaning by removing 207 outlying diary entries (0.04% of total data) whose total daily calories are: (1) greater than 6,000 kcal or 2 standard deviations away from the mean daily calories; or (2) lower than or equal to zero kcal.

For each user in the dataset, we further collected personal information from their user profile page. In total, the profile pages of 8,794 users are publicly accessible. Next, we categorized the profile attributes such as age, geographical location, and friend list into the following groups: Age group (young adults age 18-44 years old and old adults age 45 and above), social connection quartiles (Q1: 0-6 friends, Q2: 7-18 friends, Q3: 19-41 friends, and Q4: 42 or more friends), regions of residence at a global level (US and Non-US), and regions of residence within the United States as per the US census classification (Northeast, South, Midwest, and West). As we can see in Table 1, the majority of users are female (82.01%) and young adults 18 - 44 of age (79.80%). Next, more than 50% of users have less than 19 users in their friend list. Most users reside in the United States (70.88%) while the rest of the users live outside the US. Amongst US users, 33.76% live in the Southern states, whereas 18.37% live in the Northeastern states. The geographical distribution of MFP users is representative of the population distribution across US regions [8].

Table 1: Sociodemographic distributions

		Count	%
Gender	Female	7,212	82.01%
	Male	1,582	17.99%
Age group	18-44	7,015	79.80%
	45+	1,776	20.20%
Social connection	Q1 (0-6)	2,681	30.49%
	Q2 (7-18)	2,170	24.68%
	Q3 (19-41)	2,040	23.20%
	Q4 (42+)	1,903	21.64%
Global region	US	6,233	70.88%
	Non-US	2,561	29.12%
US region	South	2,104	33.76%
	Midwest	1,593	25.56%
	West	1,391	22.32%
	Northeast	1,145	18.37%

<sup>1</sup> <http://bit.ly/2hNzRHT>

**Table 2: Amounts of food equivalent to 1 standard serving size**

	Grams	kcal
Fruit	150	75
Vegetable	75	40
Red meat	65	160
Poultry	80	160
Fish	100	160

## 4.2 Normalizing and extracting portion sizes

Each diary entry contains a free-text description of the food item (i.e., food name and portion size) and nutritional facts, such as calories, protein, sugar. When recording a diary entry, MFP users can choose to enter the amount of food and beverage in various measurement units. For example, the measurement units for *apple* may include weights (grams, ounces, pounds, and kilograms), volumes (cup), physical sizes (small, medium, and large), and nutritional units (servings). To make it possible to directly compare the dietary intakes with other sources, we apply the following steps to normalize the portion sizes. First, we extracted quantity and unit from the diary entry text using regular expressions and convert weights and volumes from other measurement systems to grams and milliliters, respectively. Then, we used the annotated categories to identify specific types of diets. First, for solid foods containing only a category of fruit, vegetable, red meat, poultry, or fish, we determined standard serving sizes using the conversions [23] in Table 2. For example, 100 grams of *apple* is equal to 0.67 serving of fruit. If a gram-equivalent weight cannot be found in the diary entry text, the amount of corresponding calories in the diary entry was used for the conversion instead. E.g., a 320-kcal *tuna* contains 2 servings of fish. Next, for composite foods typically served with grains and other types of ingredients, we subtract 240 (an average kcal for 1 regular serving of grain-based foods) from the total calories and calculate the serving sizes using the corresponding caloric values in Table 2. With these steps, we can identify 1 serving of fish from a 400-kcal *tuna sandwich*.

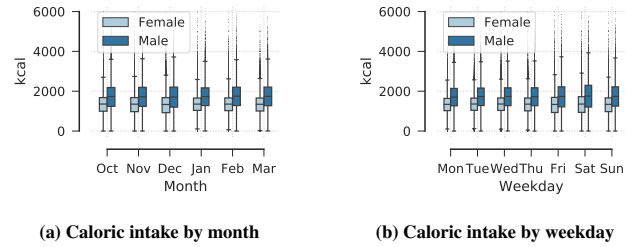
## 4.3 Caloric intake patterns

Figure 1 displays the average daily caloric intakes over time. Overall, the average daily caloric intakes, as seen in Figure 1(a) are around 1,700 kcal and 1,300 kcal for male and female users, respectively. As expected, these are much lower than the estimated calorie requirements of an average adult (2,000 - 2,200 kcal), suggesting that the users were likely dieting. Furthermore, we observe the largest interquartile range of the daily caloric intakes in December, followed by the smallest in January, a possible effect of a new year’s weight-loss resolution. Lastly, we compare the daily calorie intakes between weekdays in Figure 1(b). As can be seen, the daily caloric intakes follow the weekday-weekend lifestyle pattern, trending slightly upward from Monday before reaching the largest interquartile range on Saturday.

# 5 METHODS

## 5.1 Behavioral measures

Based on the eating behaviors introduced previously, we define the following measures to quantify the behavioral outcomes for each MFP user:

**Figure 1: Average daily caloric intakes over time (kcal)**

**Journaling behaviors:** Two journaling behaviors are defined to represent the journaling duration and persistence. First, *recording days* is defined as a number of days the user records food diary entries. Second, *normalized lapsing frequency* is defined as a fraction of days the user temporarily stops recording any diary entries with respect to her journaling lifetime (recording days + lapsing days).

**Eating behaviors:** To measure the energy intake from foods and beverages, we compute *median daily caloric intake* (in kcal). Next, for each diet type, we compute *median daily intake* (in servings for foods, grams for added sugars, and milliliters for drinks) and *normalized intake frequency* (percentages of days in which the diets were consumed). We used the following criteria to identify specific diet types from the annotated food diary entries. First, fruits and vegetables are selected from entries tagged with *fruit* and *vegetable* categories. Next, red and processed meats are chosen from entries tagged with *beef*, *pork*, *lamb*, *game* (meat from wild animals), *sausage*, and *meatball* sub-categories. Furthermore, poultry and fish are identified from entries tagged with *poultry* and *fish* sub-categories, respectively. Next, added sugars are non-zero sugar content entries tagged with the following categories and sub-categories: *beverage*, *dessert*, *snack*, *condiment*, and *dairy product*. Lastly, sugary drinks include any non-zero sugar content entries tagged with *beverage* category, whereas soft drinks included entries tagged with *soft drink* subcategory. This results in 14 diet-specific behavioral measures.

## 5.2 Sociodemographic and journaling factors

Now, we introduce the following factors known to be associated with specific healthy eating behaviors below.

**Gender and age:** Gender and age have been found as contributors of behavioral differences in many dietary behavior studies. For example, women tended to eat more healthy diets than men, e.g., consuming more fruits and fiber [6]. High consumption of meat and red meat was closely associated with being male [16, 25]. Men also consumed more sugary drinks than women [7]. Next, as people get older, they tended to eat less and changed their eating behaviors. For instance, compared to younger adults, older adults were more likely to consume more fruits and vegetables [6], less red meat as well as all meat [16, 25] but more fish [16], and less sugary drink [7]. In the mean time, younger adults tended to consume more poultry [16]. In this work, we investigate (1) the behavioral differences between genders (male and female), age groups (young and old adults) in RQ2 and (2) the influences of the gender and age factors on all behavioral measures in RQ3.

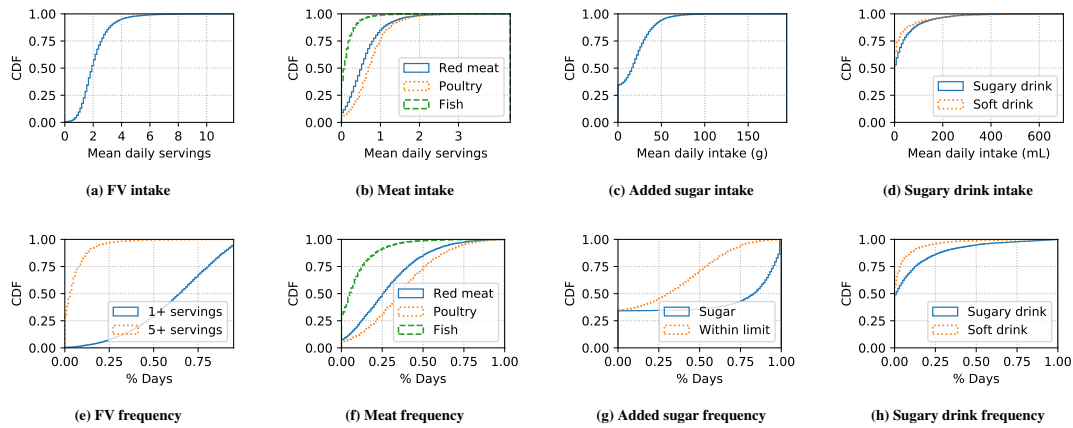


Figure 2: Average daily servings and frequencies

**Social connections:** Social ties have shown to have both the positive and negative influences on health behaviors [43]. For instance, daily consumption of fruits and vegetables [34] and overall weight loss [28] were associated with high levels of social support. On the other hand, having obese friends substantially increased one’s own risk of obesity [10]. In this study, we explore (1) the behavioral differences between social connection quartiles in RQ2 and (3) the influence of social connections on all behavioral measures in RQ3.

**Regions of residence:** Differences in eating behaviors were observed amongst Americans in different regions, which could be attributed to *social norms* and *environmental contexts*. People in the Northeast and the West were more likely to consume more fruits and vegetables than those in the Midwest and the South [6]. Next, beef consumption was highest in Midwest and lowest in the South [25]. In addition, people in the Midwest and the Northeast consumed more sugary drinks than those in the South and the West [40]. Particularly, the consumption of soft drinks was the highest in the Northeast, compared to other regions [40]. In this work, we examine (1) the behavioral differences between regions of residence within the United States (Northeast, South, Midwest, and West) as well as global regions of residence (US vs. non-US) in RQ2 and (2) their influences on all behavioral measures in RQ3.

**Journaling:** Food journaling can promote mindful eating [13], leading to weight loss [26], and healthier behaviors [18]. Furthermore, the health benefits were greater amongst highly active journalers [9]. In this work, we explore the influence of the journaling behaviors, i.e., recording days and lapsing frequency, on all caloric and diet-specific behavioral measures in RQ3.

## 6 RESULTS

### 6.1 RQ1: Distributions of the healthy diet intakes

We first examine the average daily intake and the intake frequency of each diet type over the 6-month period. For each food type, we calculated the cumulative distribution function (CDF) of the mean daily servings and the percentage of days consumed for each user. All CDF plots are shown in Figure 2.

**Fruits and vegetables:** As shown in Figure 2(a), most food journalers do not consume a sufficient amount of fruits and vegetables on a daily basis. Especially, the average daily FV intake of food journalers is much lower than that of the general populace. In particular, 50% of users consume up to **1.97 servings** of FV per day, whereas only **2.18%** of users (N=183) manage to meet the recommended daily intake of 5 servings. In terms of the frequency of intake, shown in Figure 2(e), 50% of users consume fruits and vegetables up to 64.96% of the time. Furthermore, **less than 1%** of users (N=6) are able to meet the recommended intake at least 80% of the time. This is very surprising since fruits and vegetables should be an essential part of any healthy diets, particularly amongst people trying to lose weight and live a healthier lifestyle.

**Red and processed meat, poultry, fish:** Unlike typical American consumers, food journalers consume higher portions of healthy protein sources, especially poultry, than unhealthy red and processed meat. As shown in Figure 2(b), 50% of users consume up to **0.53 servings** of red and processed meat and **0.72 servings** of poultry per day. Moreover, the average intakes are within the recommended daily limit of 2 servings. Next, poultry is also consumed more frequently than red and processed meat. According to Figure 2(f), 50% of users consume red meat and poultry up to **26%** and **36%** of the time, respectively. However, the average daily fish intake is much lower than that of red and processed meat and poultry. Similar to average American consumers, food journalers consume much fewer servings of fish per day than the recommended daily intake. Particularly, 50% of users consume **0.09 servings** of fish a day, whereas only **15%** of users (N=1,444) meet the recommendation as shown in Figure 2(b).

**Added sugar and sugary drinks:** The average daily added sugar intake amongst food journalers is much lower than that of the general populace. According to Figure 2(c), 50% of users consume up to **16 grams** of added sugar per day. However, the top 25% of users consume higher amount of daily added sugar intake (**29.83 grams**) than the recommended limits. Furthermore, the consumption occurs very frequently. As shown in Figure 2(g), 50% of users consume added sugars up to **83%** of the time, whereas the consumption within the recommended limit only accounts for up to **30%** of the time. Next, the overall consumption of sugary drinks and soft drinks amongst

**Table 3: Behavioral differences between groups.**

Behavioral measure	Gender	Age	Friends	Region	
				US	Global
Recording days	Male**	45+**	Q4**	Northeast**	Non-US*
Lapsing frequency	Female**	18-44**	Q1**		
Caloric intake	Male**	18-44*	Q4**		Non-US**
FV intake**	Male**		Q4*	West**	Non-US**
FV frequency	Female**		Q4**	West**	Non-US**
Red meat intake	Male**	45+**		Midwest**	Non-US**
Poultry intake	Male**	18-44**	Q4*	South**	US**
Fish intake					US**
Red meat frequency	Male**	45+**	Q4*	Midwest**	Non-US**
Poultry frequency	Male**	18-44**	Q4**	South*	US**
Fish frequency	Male**	45+**			Non-US**
Added sugar intake		18-44*	Q1**		US**
Sugary drink intake				Midwest*	US*
Soft drink intake	Male**				US*
Added sugar frequency	Female**		Q1**		US**
Sugary drink frequency	Female*	18-44**	Q1**		
Soft drink frequency		18-44**	Q1**	South**	US**

Significance thresholds are \*  $p < 0.05$  and \*\*  $p < 0.01$ . Caloric intake is in kcal. Food intakes are in numbers of servings. Sugar intake is in grams. Drink intakes are in mL. Intake frequencies are in percentages of days. Cell values represent groups with the highest median. Non-statistically significant results are omitted.

food journalers is comparable to that of the general populace. As shown in Figure 2(d), **53.25%** and **60.66%** of users do not consume any sugary/soft drinks, respectively. Amongst drinkers, the average daily intakes for both sugary and soft drinks are less than 1 glass. However, on any drinking days, the daily intake exceeds the limit most of the time.

**Key insights:** Unexpectedly, we find that the healthy eating behaviors of active food journalers do not differ much from those of the general populace in several areas. Given that our user population is highly skewed toward females, who tend to be health-conscious, the numbers of healthy eating lapses observed are even more surprising. For example, a vast majority of food journalers did not eat enough fruits and vegetables and fish as per the dietary guidelines. Next, their sugary and soft drink consumption was about the same level as the general populace, which is a worrying trend. On a positive note, the consumption of red and processed meat and added sugars is lower in food journalers than the general populace.

## 6.2 RQ2: Behavioral differences

In this section, we present the differences in eating behaviors of food journalers across sociodemographic groups. Comparisons of behavioral measures, shown in Table 3, were performed by (1) Mann-Whitney U test for genders, age groups, and global regions; and (2) Kruskal-Wallis H test for social connections and US regions. Dunn's multiple comparisons, shown in Table 4, were performed for social connections and US regions for post-hoc tests.

**Journaling and caloric intakes:** First, significantly higher median recording days are observed amongst the following groups than other groups: males ( $p < 0.01$ ), older adults ( $p < 0.01$ ), largest social connections ( $p < 0.01$ ), residing in the Northeast ( $p < 0.01$ ), and residing outside the US ( $p < 0.05$ ) as displayed in Table 3; pairwise differences are significant for all social connection pairs and the

**Table 4: Dunn's multiple comparisons for different social connections and US regions.**

Behavioral measure	Friends	US regions
Recording days	12 <sup>*</sup> , 13 <sup>**</sup> , 14 <sup>**</sup> , 23 <sup>**</sup> , 24 <sup>**</sup> , 34 <sup>**</sup>	NS <sup>*</sup> , SM <sup>*</sup>
Lapsing frequency	14 <sup>**</sup> , 24 <sup>**</sup> , 34 <sup>**</sup>	
Caloric intake	14 <sup>*</sup> , 24 <sup>**</sup> , 34 <sup>**</sup>	
FV intake	24 <sup>*</sup>	SM <sup>*</sup> , MW <sup>**</sup>
FV frequency	14 <sup>**</sup> , 23 <sup>*</sup> , 24 <sup>**</sup>	NS <sup>**</sup> , NM <sup>**</sup> , SW <sup>**</sup> , MW <sup>**</sup>
Red meat intake		NS <sup>*</sup> , NM <sup>**</sup>
Poultry intake	34 <sup>**</sup>	NM <sup>*</sup> , SM <sup>*</sup>
Fish intake		
Red meat frequency	14 <sup>*</sup>	NS <sup>**</sup> , NM <sup>**</sup> , NW <sup>**</sup>
Poultry frequency	24 <sup>**</sup> , 34 <sup>**</sup>	SM <sup>*</sup>
Fish frequency		
Added sugar intake	12 <sup>**</sup> , 13 <sup>**</sup> , 14 <sup>**</sup> , 24 <sup>**</sup> , 34 <sup>**</sup>	
Sugary drink intake		
Soft drink intake		
Added sugar frequency	12 <sup>**</sup> , 13 <sup>**</sup> , 14 <sup>**</sup> , 24 <sup>**</sup> , 34 <sup>**</sup>	
Sugary drink frequency	13 <sup>**</sup> , 14 <sup>**</sup> , 24 <sup>**</sup>	
Soft drink frequency	14 <sup>**</sup>	NS <sup>**</sup> , SM <sup>*</sup> , SW <sup>**</sup>

Significant thresholds are \*  $p < 0.05$  and \*\*  $p < 0.01$ . Social connection groups are abbreviated as 1: Q1, 2: Q2, 3: Q3, and 4: Q4. Region groups are abbreviated as N: Northeast, S: South, M: Midwest, and W: West. Comparisons are denoted as a pair of letters, e.g., 12 represents a comparison of two social connection groups Q1 and Q2. Groups with a higher rank sum are in bold. Non-statistically significant results are omitted.

median recording days of the Northeast and the Midwest are significantly higher than that of the South as displayed in Table 4. Second, median lapsing frequency are significantly higher amongst users in the following groups: females ( $p < 0.01$ ), younger adults (0.01), and small connections (0.01); pairwise differences are significant between smaller social connections (Q1, Q2, and Q3) vs. the largest (Q4). Next, significantly higher median caloric intakes are observed amongst the following groups than other groups: males ( $p < 0.01$ ), younger adults ( $p < 0.05$ ), largest social connections ( $p < 0.01$ ), and residing outside the US ( $p < 0.01$ ); pairwise differences are significant between smaller social connections (Q1, Q2, and Q3) vs. the largest (Q4).

**Fruits and vegetables:** Compared to the general populace, the differences in FV consumption of food journalers are associated with similar sociodemographic groups (gender, social connections, and regions of residence). First, females consume fruits and vegetables more frequently than males ( $p < 0.01$ ). Though, unlike the general populace, males have a significantly higher median FV intake ( $p < 0.01$ ). Next, the differences in FV intake and intake frequency are significant between groups with different social connections ( $p < 0.05$  and  $p < 0.01$ , respectively); pairwise differences are significant between (Q2 vs. Q4) for FV intake and (Q1 vs. Q4), (Q2 vs. Q3), and (Q2 vs. Q4) for intake frequency. Moreover, significantly higher median FV intake and frequency are observed amongst those in the West ( $p < 0.01$ ); pairwise intake differences are significant between (South vs. West) and (Midwest vs. West), whereas pairwise frequency differences are significant between all region pairs except (Northeast vs. West) and (South vs. Midwest). In contrast to the general populace, there are no behavioral differences between age groups for FV consumption.

**Red and processed meat, poultry, fish:** Similar to the general populace, the differences in meat consumption amongst the MFP populace are associated with gender, age, and regions of residence.

**Table 5: Coefficients ( $\beta$ ) of 17 OLS regression models.**

Predicted variable	Male	Age	Log (Friends)	Northeast	South	West	Non-US	Recording days	Lapsing frequency	Adjusted R <sup>2</sup>
Recording days	9.470**	0.653**	7.725**	2.837	-4.292*	-2.939	2.291			0.055
Lapsing frequency	-0.047**	-0.004**	-0.019**	-0.003	0.011	0.011	-0.007			0.026
Caloric intake	343.71**	-3.458**	7.498*	-14.235	-19.207	19.598	84.805**	1.163**	-312.951**	0.219
FV intake	0.203**	-0.002	0.011	0.071	0.04	0.107*	0.198**	-0.001	-0.268**	0.012
FV frequency	-0.072**	0.001*	0.004	0.032**	0.002	0.036**	0.104**	0.0002*	-0.107**	0.069
Red meat intake	0.228**	0.004**	0.012*	-0.083**	-0.041*	-0.031	0.016	-0.0004*	-0.014	0.040
Poultry intake	0.235**	-0.006**	0.010*	0.051*	0.055**	0.039*	-0.121**	0.0001	0.0056	0.051
Fish intake	0.006	0.0003	-0.0003	0.012	0.005	0.005	0.055**	-0.0002**	-0.0193	0.011
Red meat frequency	0.057**	0.002**	0.006**	-0.037**	-0.003	-0.008	0.018*	0.0002**	-0.0561**	0.062
Poultry frequency	0.055**	-0.001**	0.004*	0.007	0.024**	0.007	-0.043**	0.0003**	-0.0718**	0.052
Fish frequency	0.009*	0.0003*	-0.001	0.006	0.007	0.004	0.047**	-0.0004	-0.0309**	0.029
Added sugar intake	0.291	-0.133**	-1.454**	-0.407	-0.238	-0.439	-3.701**	-0.015*	-7.142**	0.028
Sugary drink intake	-3.066	0.117	-0.821	5.548*	5.045*	1.151	7.229**	-0.04	-6.96	0.002
Soft drink intake	6.714**	0.143	0.595	-3.613	-1.326	-6.412*	0.23	-0.042	-1.209	0.002
Added sugar frequency	-0.056**	-0.001*	-0.044**	0.008	0.013	0.021	-0.049**	-0.0008**	-0.169**	0.032
Sugary drink frequency	-0.014*	-0.001*	-0.006**	0.012	0.012	0.007	0.022**	-0.0001	-0.0293**	0.005
Soft drink frequency	0.009**	-0.0003*	0.0001	-0.003	0.009*	-0.006	0.017**	-0.0003	-0.013*	0.008

Caloric intake is in kcal. Food intakes are in numbers of servings. Sugar intake is in grams. Drink intakes are in mL. Intake frequencies are in percentages of days. Intercept terms are omitted. Significance thresholds are \*  $p < 0.05$  and \*\*  $p < 0.01$ .

First, males have a significantly higher median intake ( $p < 0.01$ ) and frequency ( $p < 0.01$ ) than females, for all protein sources except fish. Second, older adults have a significantly higher fish intake frequency ( $p < 0.01$ ). In addition, poultry consumption (median intake and frequency) in younger adults is significantly higher ( $p < 0.01$ ). Next, differences in red and processed meat intake frequency are significant amongst regions of residence ( $p < 0.01$ ). Users in the Midwest generally have the highest intake and intake frequency; pairwise intake differences are significant between (Northeast vs. South) and (Northeast vs. Midwest), whereas pairwise frequency differences are significant between the followings: (Northeast vs. South), (Northeast vs. Midwest), and (Northeast vs. West). As opposed to the general populace, older adults have a significantly higher median red and processed meat intake ( $p < 0.01$ ) and frequency ( $p < 0.01$ ) than younger adults.

**Added sugars and sugary drinks:** Lastly, the differences in added sugar and sugary drink consumption of food journalers are associated with gender, age, and regions of residence – reflecting the overall differences in the general populace. First, males have a significantly higher median intake of soft drink ( $p < 0.01$ ) than females; however, in contrast to the general populace, females have significantly higher intake frequencies of added sugars ( $p < 0.01$ ) and sugary drink ( $p < 0.05$ ). Second, the median intake of added sugars is significantly higher in younger adults ( $p < 0.05$ ). Furthermore, younger adults have significantly higher intake frequencies of sugary drinks ( $p < 0.01$ ) and soft drinks ( $p < 0.01$ ), than older adults. In contrast to the general populace, we find that food journalers in the Midwest and the South tend to consume more soft drinks than those in the Northeast and the West. Specifically, users in the Midwest have a significantly higher median soft drink intake ( $p < 0.05$ ), whereas those in the South have a significantly higher soft drink intake frequency ( $p < 0.01$ ) than

other regions; pairwise differences are significant for the South vs. other regions for the soft drink intake frequency, whereas there are no pairwise differences between regions for the median soft drink intake.

**Key insights:** Overall, food journalers who are male, 45 years or older, and have the largest social network tend to have significantly longer journaling duration and more persistent in recording food journals than others. Next, the healthy eating behaviors of food journalers within the sociodemographic groups are not as homogeneous as we initially expected. As can be seen in Table 3, 58 of 85 (68%) behavioral differences are statistically significant. Furthermore, the differences in eating behaviors across sociodemographic groups are, in many cases, fairly similar to those naturally observable in the general populace. Specifically, both positive and negative eating behaviors occurred within the expected sociodemographic groups, e.g., high intake frequency of fruits and vegetables in females, high added sugar intake in younger adults, and more healthy eating behaviors in users with larger social connections, etc. The results are quite interesting as they may suggest that food journalers were not being as mindful of their healthier food choices as they should have been. We will further investigate the influences of sociodemographic and journaling factors in the next section.

### 6.3 RQ3: Factors influencing eating behaviors

Ordinary least squares (OLS) regression was used to assess (1) the influences of sociodemographic factors on the journaling behaviors and (2) the influences of sociodemographic and journaling factors on the caloric and diet-based behavioral measures. To that end, we built 17 regression models in which each predicted variable corresponds to each behavioral measure. For predictor variables, we included gender (dummy coded 1 for male and 0 for female), age, social



connections (logarithmic scale), US region (dummy coded into 3 variables using Midwest as a reference category), global region (dummy coded 1 for non-US and 0 for US), recording days, and lapsing frequency.

Table 5 displays the predictor variables and their coefficients ( $\beta$ ) from 17 OLS regression models. The values of adjusted  $R^2$  of all the models vary from 0.002 - 0.219, which are fairly low. As the predicted variables are derived from a long-term (up to 6 months) consumption data, we expect the predictors of the regression models to modestly explain a small portion of variance of the predicted variables as they do not take into account the temporal variability of the behaviors. This also points to the fact that there are many other factors which could potentially influence these behaviors. In what follows, we summarize the influences of different factors on all behaviors in the order of importance (by absolute coefficient values and total number of behavioral measures influenced).

**Gender:** Overall, gender appears to be the most important factor in influencing most of the behaviors in the study. Being male will substantially change the value of most behavioral measures compared to being female after controlling for other variables. Specifically, gender significantly influences 14 of 17 different behavioral measures. Amongst all predictor variables, gender has the highest predictive power on 7 behavioral measures, such as recording days, lapsing frequency, and caloric intake, and the relatively high predictive power (the top-3 highest coefficients) on 7 other measures, such as FV intake, FV intake frequency, and poultry intake frequency.

**Lapsing frequency:** Next, lapsing frequency is the second most influential factor of eating behaviors after gender. Specifically, it has a relatively high predictive power on 10 behavior measures, such as caloric intake, FV intake, and FV intake frequency. Interestingly, it has adverse relationships with most eating behaviors as indicated by negative coefficients  $\beta$ . In some cases, the negative relationships seem somewhat counter-intuitive, e.g., an increase in lapsing frequency decreases median caloric intake as well as intakes of unhealthy diets. This could be partially explained by the fact that some food journalers may be more inclined to record less and less diary entries before they temporarily stop journaling [46]. For some, this may positively indicate that they have already achieved their self-tracking goals – an example of *successful abandonment* [12]. In a recent study, De Choudhury et al. [17] concluded that this phenomenon should not be common and long-term food journalers are more likely to complete their diary entries when they choose to record a journal.

**Regions of residence:** Regions of residence have a substantial influence on journaling and eating behaviors, particularly global regions (US vs. Non-US). After adjusting for other factors, being outside the US will considerably affect most behavioral measures compared to being in the US. 13 of 17 behavioral measures are significantly influenced by global regions. Furthermore, it has a relatively high predictive power on 12 behavioral measures, such as caloric intake, FV intake, and FV intake frequency. More importantly, it is one of the two significant predictors of fish intake. Next, US regions of residence (Northeast, South, and West) significantly influence 10 of 17 behavioral measures. Amongst the behavioral measures, it is a relatively high predictor of red and processed meat intake (Northeast), red and processed meat intake frequency (Northeast), sugary drink intake (Northeast), recording days (South), red and

processed meat intake (South), poultry intake (South), sugar drink intake (South), soft drink intake frequency (South), and soft drink intake (West).

**Age:** Age is not as predictive of the healthy eating and journaling behaviors as gender, lapsing frequency, and regions of residence. Even though 13 of 17 behaviors are influenced by age, its effects ( $\beta$ ) on these behaviors are fairly modest compared to many predictor variables. For example, a one year increase in age will increase the median red and processed meat intake by 0.004 servings after adjusting for other variables.

**Recording days:** Recording days is one of the least predictive factors of most dietary behaviors. A one day increase in recording days will marginally change most behavioral measures after controlling for other factors. Only half of the behavioral measures (8 of 16) are significantly influenced by recording days. In addition, it is the second worst predictor of 7 influenced measures, such as caloric intake, FV intake frequency, and red and processed meat intake. Interestingly, it is one of the two significant predictors of fish intake.

**Social connections:** Interestingly, social connection seems to be the least influential factor of most behavioral measures. That is, it has the least predictive power on 10 influenced behavior measures, such as recording days, lapsing frequency, and caloric intake. For instance, a 1% increase in social connections will only increase the number of recording days by 0.007725 days after adjusting for other factors.

**Key insights:** Between the two journaling factors considered in this study, journaling persistence is more predictive of the healthy eating behaviors than journaling duration. Moreover, the journaling duration is one of the least influential factors of behaviors compared to other sociodemographic factors. More importantly, many sociodemographic factors, especially gender and regions of residence, still play a more critical role in determining the healthy eating behaviors of food journalers than the journaling factors. Contrary to its contribution in the weight loss outcomes, the results show that the journaling duration is a marginal determinant of the healthy eating behaviors.

## 7 DISCUSSION

Millions of people use mobile food journals as a tool for tracking their caloric intake in order to achieve specific health goals, such as losing weight and living a healthy lifestyle. Prior findings have shown that food journals who actively record what they ate tend to meet their caloric goals [17, 46] and lose more weights [9]. However, beyond managing caloric intake and weight, we found several lapses in the food journalers' healthy eating behaviors, e.g., low consumption of fruits and vegetables and fish, and high consumption of sugary drinks and soft drinks, preventing them from fully achieving their healthy lifestyle goal. Compared to the sociodemographic factors, the journaling duration has the least amount of influence on most eating behaviors. Most journalers' healthier food choices are significantly influenced by their gender roles and environmental contexts, suggesting that their food choices are still largely subconscious [45]. Overall, our study helps to further investigate the claim about the journaling effect on healthier food choices [13, 14, 18] using a large-scale data. From the public health perspective, we find

that their effectiveness in helping people achieve healthy eating behaviors as defined by evidence-based outcomes are marginal at best. Based on our findings, it appears that the calorie counting aspect of food journals is not sufficient and potentially unhelpful in facilitating sustained health behavior change. By exclusively focusing on caloric and weight control, food journals may unintentionally mislead individuals into pursuing caloric management as the only health goal and disregarding the physiological and metabolic effects of different types of diets on health and well-being. Our study also confirms findings from past study [13], suggesting that there exists a mismatch between food journalers' preconceived notions of health eating and the energy-centric design of mobile food journals, inadvertently leading to negative nudges.

Our findings have several implications to the designs of mobile food journals, such as MyFitnessPal, or other forms of *mHealth* (mobile health) interventions that can better facilitate healthy eating behaviors. First, from the goal setting perspective [33], a variety of *behavioral goals* could be suggested or posed as daily food challenges [18] to individuals by learning from their past food journal data. The aim is to: (1) supplement the existing caloric and nutrient intake goals; (2) provide a well-defined and quantifiable steps to help people achieve the healthy lifestyle goals, such as high daily FV intake; and (3) improve behavioral compliance and self-efficacy. Next, the goal-setting mechanism could incorporate individuals' backgrounds and experiences such that the behavioral goals could be dynamically and incrementally adjusted to suit them. Next, the goal-setting mechanism could be designed to focus more on targeting whole food consumption (e.g., *consuming at least 5 servings of fruits and vegetables a day*) than isolated nutrients and constituents (e.g., *consuming at least 30 grams of fiber per day*). This could effectively help educate individuals about the importance of food synergy and promote the idea of dietary variety [29]. Lastly, behavioral interventions tailored to individuals' sociodemographic backgrounds could be introduced to mobile food journals. The aim is to identify individuals who are highly susceptible to certain eating behavior lapses and provide them with additional guidance and actions relevant to their current goals. For instance, male users, who are relatively prone to infrequent fruit and vegetable consumption, could receive more targeted notification messages designed to remind and persuade them to meet their FV goals. Additionally, other complementary sources of data, e.g., social media [17], could be included to improve the adaptability and effectiveness of the interventions.

## 7.1 Limitations

Due to the self-reported and user-contributed natures of data, it is difficult to verify the accuracy of the MFP data [14, 17, 46]. These issues could bias the estimations of portion sizes and nutrient intakes for various diet types, especially high-sugar diets which can be deliberately omitted from the journals due to guilt [14]. Next, external data about the dietary consumption of the general populace are generally collected through traditional dietary assessment instruments, such as a Food Frequency Questionnaire (FFQ), which are more susceptible to recall bias than daily food journals. Higher prices of healthier food choices, such as fish, may also affect the healthy eating patterns. Moreover, since it is not possible to control for food journalers' personal belief and perception about food and nutrition

from the MFP data, the effects (or lack thereof) of the journaling practice may be confounded by the differences in healthy eating perception. Although most journalers generally hold the views of healthy eating which are consistent with evidence-based recommendations [13], some may choose to follow a specific dietary regimen, e.g., vegetarian diet, ketogenic diet, etc., for various reasons. Nevertheless, by allowing some noise in the data, our findings can be fairly compared to epidemiological studies'. Lastly, since our findings are obtained from online observational data, therefore they are limited in determining the causal associations between sociodemographic and journaling factors and healthy eating behaviors.

## 8 CONCLUSION

In this study, we investigated the healthy eating behaviors of food journalers using data from MyFitnessPal, a popular online calorie counter and diet journal app. Despite the claim about the benefit of journaling in promoting healthier choices, we found that most food journalers did not eat more healthful diets than the general public. First, much of their dietary consumption did not meet the daily recommended intakes of healthy and unhealthy food sources. Next, their dietary patterns were not as uniform as we initially expected and the distinct patterns mostly resembled those of the general populace who may be less health conscious. Moreover, journaling duration, which was previously shown to be associated with improved weight loss outcomes, appeared to have a marginal influence on the healthy eating behaviors, whereas gender, lapsing frequency, and regions of residence are much more predictive of the healthy eating outcomes.

## 9 ACKNOWLEDGEMENT

This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its International Research Centres in Singapore Funding Initiative.

## REFERENCES

- [1] Sofiane Abbar, Yelena Mejova, and Ingmar Weber. 2015. You Tweet What You Eat: Studying Food Consumption Through Twitter. In *Proceedings of the 33rd international conference on Human factors in computing systems - CHI '15*. ACM Press, New York, New York, USA, 3197–3206. <https://doi.org/10.1145/2702123.2702153>
- [2] Itziar Abete, Dora Romaguera, Ana Rita Vieira, Adolfo Lopez de Munain, and Teresa Norat. 2014. Association between total, processed, red and white meat consumption and all-cause, CVD and IHD mortality: a meta-analysis of cohort studies. *British Journal of Nutrition* 112, 05 (sep 2014), 762–775. <https://doi.org/10.1017/S000711451400124X>
- [3] American Heart Association. 2016. Fish and Omega-3 Fatty Acids. (2016). Retrieved from [http://www.heart.org/HEARTORG/HealthyLiving/HealthyEating/HealthyDietGoals/Fish-and-Omega-3-Fatty-Acids\\_UCM\\_303248\\_Article.jsp](http://www.heart.org/HEARTORG/HealthyLiving/HealthyEating/HealthyDietGoals/Fish-and-Omega-3-Fatty-Acids_UCM_303248_Article.jsp)
- [4] D. Aune, G. Ursin, and M. B. Veierød. 2009. Meat consumption and the risk of type 2 diabetes: a systematic review and meta-analysis of cohort studies. *Diabetologia* 52, 11 (nov 2009), 2277–2287. <https://doi.org/10.1007/s00125-009-1481-x>
- [5] Carole A Bisogni, Margaret Jastran, Marc Seligson, Alyssa Thompson, E.J. Lengerich, J. Atwell, and A.R. Kristal. 2009. How people interpret healthy eating: contributions of qualitative research. *Journal of nutrition education and behavior* 44, 4 (jul 2009), 282–301. <https://doi.org/10.1016/j.jneb.2011.11.009>
- [6] Heidi Michels Blanck, Cathleen Gillespie, Joel E Kimmons, Jennifer D Seymour, and Mary K Serdula. 2008. Trends in fruit and vegetable consumption among U.S. men and women, 1994-2005. *Preventing chronic disease* 5, 2 (apr 2008), A35. <http://www.ncbi.nlm.nih.gov/pubmed/18341771http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2396974>
- [7] S. N Bleich, Y C. Wang, Y. Wang, and S. L Gortmaker. 2008. Increasing consumption of sugar-sweetened beverages among US adults: 1988-1994 to 1999-2004. *American Journal of Clinical Nutrition* 89, 1 (dec 2008), 372–381. <https://doi.org/10.3945/ajcn.2008.26883>

- [8] U.S. Census Bureau. 2017. United States Population Growth by Region. (2017). Retrieved from [https://www.census.gov/popclock/data\\_tables.php?component=growth](https://www.census.gov/popclock/data_tables.php?component=growth).
- [9] Lora E. Burke, Mindi A. Styn, Susan M. Sereika, Molly B. Conroy, Lei Ye, Karen Glanz, Mary Ann Sevick, and Linda J. Ewing. 2012. Using mHealth Technology to Enhance Self-Monitoring for Weight Loss. *American Journal of Preventive Medicine* 43, 1 (jul 2012), 20–26. <https://doi.org/10.1016/j.amepre.2012.03.016>
- [10] Nicholas A. Christakis and James H. Fowler. 2007. The Spread of Obesity in a Large Social Network over 32 Years. *New England Journal of Medicine* 357, 4 (jul 2007), 370–379. <https://doi.org/10.1056/NEJMsa066082>
- [11] Chia-Fang Chung, Elena Agapie, Jessica Schroeder, Sonali Mishra, James Fogarty, and Sean A. Munson. 2017. When Personal Tracking Becomes Social: Examining the Use of Instagram for Healthy Eating. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. ACM Press, New York, New York, USA, 1674–1687. <https://doi.org/10.1145/3025453.3025747>
- [12] James Clawson, Jessica A. Pater, Andrew D. Miller, Elizabeth D. Mynatt, and Lena Mamykina. 2015. No Longer Wearing: Investigating the Abandonment of Personal Health-Tracking Technologies on Craigslist. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*. ACM, 647–658. <https://doi.org/10.1145/2750858.2807554>
- [13] Felicia Cordeiro, Elizabeth Bales, Erin Cherry, and James Fogarty. 2015. Rethinking the Mobile Food Journal. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press, New York, New York, USA, 3207–3216. <https://doi.org/10.1145/2702123.2702154>
- [14] Felicia Cordeiro, Daniel A. Epstein, Edison Thomaz, Elizabeth Bales, Arvind K. Jagannathan, Gregory D. Abowd, and James Fogarty. 2015. Barriers and Negative Nudges: Exploring Challenges in Food Journaling. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press, New York, New York, USA, 1159–1162. <https://doi.org/10.1145/2702123.2702155>
- [15] Aaron Culotta. 2014. Estimating County Health Statistics with Twitter. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM Press, New York, New York, USA, 1335–1344. <https://doi.org/10.1145/2556288.2557139>
- [16] Carrie R Daniel, Amanda J Cross, Corinna Koebeck, and Rashmi Sinha. 2011. Trends in meat consumption in the USA. *Public Health Nutrition* 14, 04 (apr 2011), 575–583. <https://doi.org/10.1017/S1368980010002077>
- [17] Munmun De Choudhury, Mrinal Kumar, and Ingmar Weber. 2017. Computational Approaches Toward Integrating Quantified Self Sensing and Social Media. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17*. ACM Press, New York, New York, USA, 1334–1349. <https://doi.org/10.1145/2998181.2998219>
- [18] Daniel A. Epstein, Felicia Cordeiro, James Fogarty, Gary Hsieh, and Sean A. Munson. 2016. Crumbs: Lightweight Daily Food Challenges to Promote Engagement and Mindfulness. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, New York, New York, USA, 5632–5644. <https://doi.org/10.1145/2858036.2858044>
- [19] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. 2015. A lived informatics model of personal informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*. ACM Press, New York, New York, USA, 731–742. <https://doi.org/10.1145/2750858.2804250>
- [20] American Institute for Cancer Research. 2017. Lower Red Meat Consumption to Prevent Cancer. (2017). Retrieved from [http://www.aicr.org/reduce-your-cancer-risk/recommendations-for-cancer-prevention/recommendations\\_05\\_red\\_meat.html](http://www.aicr.org/reduce-your-cancer-risk/recommendations-for-cancer-prevention/recommendations_05_red_meat.html).
- [21] Centers for Disease Control and Prevention. 2013. Consumption of Added Sugars Among U.S. Adults, 2005–2010. (2013). Retrieved from <https://www.cdc.gov/nchs/data/databriefs/db122.pdf>.
- [22] Centers for Disease Control and Prevention. 2015. Adults Meeting Fruit and Vegetable Intake Recommendations — United States, 2013. (2015). Retrieved from <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6426a1.htm>.
- [23] Eat for Health. 2015. Serve sizes. (2015). Retrieved from <https://www.eatforhealth.gov.au/food-essentials/how-much-do-we-need-each-day/serve-sizes>.
- [24] Gallup. 2012. Nearly Half of Americans Drink Soda Daily. (2012). Retrieved from <http://www.gallup.com/poll/156116/Nearly-Half-Americans-Drink-Soda-Daily.aspx>.
- [25] Marcia Hill Gossard and Richard York. 2003. Social Structural Influences on Meat Consumption. *Human Ecology Review* 10 (2003), 1–9. <https://doi.org/10.2307/24707082>
- [26] Jack F Hollis, Christina M Gullion, Victor J Stevens, Phillip J Brantley, Lawrence J Appel, Jany D Ard, Catherine M Champagne, Arlene Dalcin, Thomas P Erlinger, Kristine Funk, Daniel Laferriere, Pao-Hwa Lin, Catherine M Loria, Carmen Samuel-Hodge, William M Vollmer, and Laura P Svetkey. 2008. Weight loss during the intensive intervention phase of the weight-loss maintenance trial. *American journal of preventive medicine* 35, 2 (Aug. 2008), 118–26.
- [27] Frank B. Hu. 2013. Resolved: there is sufficient scientific evidence that decreasing sugar-sweetened beverage consumption will reduce the prevalence of obesity and obesity-related diseases. *Obesity Reviews* 14, 8 (aug 2013), 606–619. <https://doi.org/10.1111/obr.12040>
- [28] Kevin O. Hwang, Allison J. Ottenbacher, Angela P. Green, M. Roseann Cannon-Diehl, Oneka Richardson, Elmer V. Bernstam, and Eric J. Thomas. 2010. Social support in an Internet weight loss community. *International Journal of Medical Informatics* 79, 1 (jan 2010), 5–13. <https://doi.org/10.1016/j.ijmedinf.2009.10.003>
- [29] D. R. Jacobs, M. D. Gross, and L. C. Tapsell. 2009. Food synergy: an operational concept for understanding nutrition. *American Journal of Clinical Nutrition* 89, 5 (may 2009), 1543S–1548S. <https://doi.org/10.3945/ajcn.2009.26736B>
- [30] Paule Latino-Martel, Vanessa Cottet, Nathalie Druesse-Pecollo, Fabrice H F Pierre, Marina Touillard, Mathilde Touvier, Marie-Paule Vasson, Mélanie Deschasaux, Julie Le Merdy, Emilie Barrandon, and Raphaëlle Ancellin. 2016. Alcoholic beverages, obesity, physical activity and other nutritional factors, and cancer risk: A review of the evidence. *Critical reviews in oncology/hematology* (jan 2016). <https://doi.org/10.1016/j.critrevonc.2016.01.002>
- [31] Yelena Mejova, Hamed Haddadi, Anastasios Noulas, and Ingmar Weber. 2015. #FoodPorn: Obesity Patterns in Culinary Interactions. In *Proceedings of the 5th International Conference on Digital Health 2015 - DH '15*. ACM Press, New York, New York, USA, 51–58. <https://doi.org/10.1145/2750511.2750524>
- [32] Ali H. Mokdad, Barbara A. Bowman, Earl S. Ford, Frank Vinicor, James S. Marks, and Jeffrey P. Koplan. 2001. The Continuing Epidemics of Obesity and Diabetes in the United States. *JAMA* 286, 10 (sep 2001), 1195. <https://doi.org/10.1001/jama.286.10.1195>
- [33] Sean Munson and Sunny Consolvo. 2012. Exploring Goal-setting, Rewards, Self-monitoring, and Sharing to Motivate Physical Activity. In *Proceedings of the 6th International Conference on Pervasive Computing Technologies for Healthcare*. IEEE. <https://doi.org/10.4108/icst.pervasivehealth.2012.248691>
- [34] Tarja Nieminen, Ritva Prättälä, Tuija Martelin, Tommi Härkönen, Markku T Hyypää, Erkki Alanen, and Seppo Koskinen. 2013. Social capital, health behaviours and health: a population-based associational study. *BMC public health* 13 (jun 2013), 613. <https://doi.org/10.1186/1471-2458-13-613>
- [35] U.S. Department of Health and Human Services. 2010. Dietary Guidelines for Americans 2010. (2010). Retrieved from <https://health.gov/dietaryguidelines/dga2010/DietaryGuidelines2010.pdf>.
- [36] Harvard T.H. Chan School of Public Health. 2015. Healthy Beverage Guidelines. (2015). Retrieved from <https://www.hsph.harvard.edu/nutritionsource/healthy-drinks-full-story/>.
- [37] World Health Organization. 2015. WHO calls on countries to reduce sugars intake among adults and children. (2015). Retrieved from <http://www.who.int/mediacentre/news/releases/2015/sugar-guideline/en/>.
- [38] Yanni Papanikolaou, James Brooks, Carrol Reider, and Victor L Fulgoni. 2014. U.S. adults are not meeting recommended levels for fish and omega-3 fatty acid intake: results of an analysis using observational data from NHANES 2003–2008. *Nutrition Journal* 13, 1 (dec 2014), 31. <https://doi.org/10.1186/1475-2891-13-31>
- [39] Kunwoo Park, Ingmar Weber, Meeyoung Cha, and Chul Lee. 2016. Persistent Sharing of Fitness App Status on Twitter. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*. ACM Press, New York, New York, USA, 183–193. <https://doi.org/10.1145/2818048.2819921>
- [40] Sohyun Park, Lisa C. McGuire, and Deborah A. Galuska. 2015. Regional Differences in Sugar-Sweetened Beverage Intake among US Adults. *Journal of the Academy of Nutrition and Dietetics* 115, 12 (dec 2015), 1996–2002. <https://doi.org/10.1016/j.jand.2015.06.010>
- [41] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2014. Personal Tracking as Lived Informatics. In *Proceedings of the 32nd international conference on Human factors in computing systems - CHI '14*. ACM Press, New York, New York, USA, 1163–1172. <https://doi.org/10.1145/2556288.2557039>
- [42] Mingyang Song, Teresa T. Fung, Frank B. Hu, Walter C. Willett, Valter D. Longo, Andrew T. Chan, and Edward L. Giovannucci. 2016. Association of Animal and Plant Protein Intake With All-Cause and Cause-Specific Mortality. *JAMA Internal Medicine* 176, 10 (oct 2016), 1453. <https://doi.org/10.1001/jamainternmed.2016.4182>
- [43] Debra Umberson, Robert Crosnoe, and Corinne Reczek. 2010. Social Relationships and Health Behavior Across the Life Course. *Annual Review of Sociology* 36, 1 (jun 2010), 139–157. <https://doi.org/10.1146/annurev-soc-070308-120011>
- [44] Xia Wang, Yingying Ouyang, Jun Liu, Minmin Zhu, Gang Zhao, Wei Bao, and Frank B. Hu. 2014. Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: systematic review and dose-response meta-analysis of prospective cohort studies. *BMJ* 349, jul29 3 (jul 2014), g4490–g4490. <https://doi.org/10.1136/bmj.g4490>
- [45] Brian Wansink. 2006. *Mindless Eating : Why We Eat More Than We Think*. Bantam Books. 276 pages. <https://books.google.com.sg/books/about/Mindless>
- [46] Ingmar Weber and Palakorn Achananuparp. 2016. Insights from machine-learned diet success prediction. In *Proceedings of Pacific Symposium on Biocomputing (PSB)*.
- [47] World Health Organization. 2003. Diet, nutrition and the prevention of chronic diseases: report of a Joint WHO/FAO Expert Consultation. *WHO Technical Report Series* 916 (2003).