

# Computational Approaches Toward Integrating Quantified Self Sensing and Social Media

**Munmun De Choudhury**  
Georgia Institute of  
Technology  
Atlanta, GA 30332 USA  
munmund@gatech.edu

**Mrinal Kumar**  
Georgia Institute of  
Technology  
Atlanta, GA 30332 USA  
mkumar73@gatech.edu

**Ingmar Weber**  
Qatar Computing Research  
Institute  
HBKU, Doha, Qatar  
iweber@qf.org.qa

## ABSTRACT

The growing amount of data collected by quantified self tools and social media hold great potential for applications in personalized medicine. Whereas the first includes health-related physiological signals, the latter provides insights into a user's behavior. However, the two sources of data have largely been studied in isolation. We analyze public data from users who have chosen to connect their MyFitnessPal and Twitter accounts. We show that a user's diet compliance success, measured via their self-logged food diaries, can be predicted using features derived from social media: linguistic, activity, and social capital. We find that users with more positive affect and a larger social network are more successful in succeeding in their dietary goals. Using a Granger causality methodology, we also show that social media can help predict daily changes in diet compliance success or failure with an accuracy of 77%, that improves over baseline techniques by 17%. We discuss the implications of our work in the design of improved health interventions for behavior change.

## Author Keywords

diet; fitness; health; MyFitnessPal; quantified self; social media; Twitter; well-being; behavior change

## ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

## INTRODUCTION

In his State of the Union address on January 20, 2015, President Obama announced the Precision Medicine Initiative (PMI)<sup>1</sup>. The vision for precision medicine is a world where variability in the lifestyle, physiology, genes, and environmental context for each person can be accurately measured, understood, and utilized in the prevention and treatment of

<sup>1</sup><https://www.nih.gov/precision-medicine-initiative-cohort-program>

diseases. Making this vision a reality for all, in a scalable and cost-effective manner, depends on advancing the science of measurement, novel ways to harvest health related data, and methods to infer health risk [5, 11]. Consequently, Estrin [20] and Haddadi et al. [26] have recently emphasized the need for the development of computational approaches that integrate multiple forms of sensed data to improve understanding of health and wellness.

Two such forms of sensed data that have been *individually* observed to comprise valuable signals about health and well-being, include social media and quantified self sensing technologies. Specifically, in the case of the former, the continual adoption of social media sites is presenting opportunities toward pro-active and unobtrusive assessment and improvement of our health and well-being at scale. These include, observing behaviors and psychological states [16], depressive tendencies [17], fitness and diet [1, 57, 58, 56] and so on. At the same time, we are seeing a rapid increase in the adoption and use of lifelogging and self-tracking tools, popularly known as the quantified self movement [36]. A variety of applications have emerged, that allow people to use their mobile phones to track aspects of physical (e.g., step count, sleep [37]), physiological (e.g., heart rate, respiration rate [6]) and behavioral health (e.g., affect, depression, stress [32]).

However, despite the vision outlined by the PMI<sup>1</sup>, there is a dearth of work examining the relationship between these different forms of data and in combining them for positive health behavior change. Many health behaviors, e.g., addiction or obesity, have social, behavioral, cognitive and affective dimensions. Information about these constructs can be obtained from social media, but would be difficult to observe using quantified self sensing technologies alone. Similarly, quantified self sensing technologies provide rich data about fine-grained and self-reported physiological attributes, which are difficult to derive from social media alone.

Through this paper, we aim to address some of these gaps in health data integration. We explore the relationship between quantified self sensing based markers of health, and one's social behavioral, cognitive and affective context, as gleaned through social media. Specifically, we study diet compliance of individuals as self-journaled on a quantified self platform, and complementarily examine the social media activities of the same cohort. Our two specific research questions include:

**RQ 1:** *Are there behavioral measures derived from social media that are predictive of diet compliance success or failure?*

**RQ 2:** *To what extent can social media derived measures be integrated with sensed historical data about dietary practices, in order to better assess future diet compliance?*

Towards these goals, we leverage public data from a set of individuals who have chosen to connect their quantified self sensing and social media accounts. Our data comprises over 100K daily entries shared by nearly 700 individuals on a calorie and diet logging application, MyFitnessPal (MFP)<sup>2</sup>, and over 2M Twitter posts shared by the same individuals. With this data, we define six different metrics for diet compliance, based on relationships between self-reported daily calorie goals and self-reported calories consumed. We also extract a number of measures to characterize an individual's behavior on Twitter: linguistic, activity and social capital measures.

We first build statistical models to examine the relationship between social media derived measures and sensed information about diet compliance success and failure. To answer the second research question, we develop multivariate vector auto-regressive models [25] of diet compliance of individuals over time. Next, we apply a Granger causality analysis [24] based methodology to examine the gain in predictive power enabled by utilizing measurements of social media behavior, alongside using sensed data on diet alone.

Our findings reveal significant links between social media behavioral measures and diet compliance. Successful individuals tend to post more positively (and correspondingly less negatively, with less anger, anxiety and sadness), exhibit high cognitive functioning, demonstrate collective attention, and tend to be future oriented in the tone of their shared posts. These individuals also show greater access to social capital and heightened tendency for social interaction. We find that use of social media data can promisingly predict several diet compliance metrics (accuracy=77%), with improvements in accuracy up to 37% compared to baseline models. Finally, our Granger causality based auto-regressive models indicate that integrating social media data in predicting future diet compliance can improve prediction accuracies by 17%, over models that use historical data on diet compliance alone.

Our findings situate the significance of integrating quantified self sensing and social media data in predicting a specific yet important health behavior—diet compliance. We discuss the implications of our computational approaches and observational findings in defining next generation health interventions for positive behavior change.

## RELATED WORK

### Role of Technology in Health and Wellness

A rich body of work has studied the role of quantified self sensing technologies in empowering individuals with data about their health and well-being. This has included focusing on attributes that are predictive of behavior change—for instance, the role of self-efficacy [49], motivational interviewing, perceived barriers, vulnerability, social norms [47], self-monitoring, rewards and goal-setting [50, 37]. To this end, gamification of quantified self technologies have also been

proposed [59]. In a CHI 2014 workshop, Meyer et al. [36] examined notions that trade-off the cost and value of such technologies. In a more recent work, Murnane et al. [39] provided a categorization of popular mobile health applications and then examined the perceived efficacy of apps alongside the reasons behind their adoption and abandonment. Ivanov et al. [28] examined the factors that influence the sharing of health-tracking records by patients with different categories of acquaintances.

Quantified self sensing data has also been employed to assess a number of health and well-being attributes of individuals. Researchers have used various sensors to monitor changes and predict trends of physiological and physical signals [4, 30]. Wearable sensing technologies that detect galvanic skin response (GSR) signals, movement, respiration and heart rate measurements have also been found to provide valuable physiological data on challenges like stress and anxiety [32, 2]. Recently, the StudentLife Project [55, 8] leveraged passive and active sensing techniques through the use of smartphones among college students. Activity and behavioral data collected through this methodology was then correlated to academic performance. Although rudimentary Facebook profile information was obtained through the mobile phone application, this data was not analyzed.

Our work is related to the work of Weber and Achananuparp [57], who examined publicly accessible MyFitnessPal food diaries to build models that can predict diet success and caloric intake based on reported food categories. While this work highly relevant to our investigations in this paper, we examine the role of social and linguistic attributes, derived from social media, in making and helping improve inferences of diet compliance success and failure of individuals.

### Health Information Sharing on Social Media

Online communities thriving on social media have also been found to allow people afflicted by medical conditions to connect with others and find support [21, 40]. Additionally, these communities serve a range of purposes, including seeking advice [29], connecting with experts and individuals with similar experiences [21, 27], sharing questions and concerns around treatment options [21], sensemaking [33] and understanding professional diagnoses [27], enabling better management of chronic health conditions [34], and fueling discussions with healthcare providers [21].

In other works, Vickey and Breslin [53] conducted a system-level study of how fitness app data is shared on Twitter. Teodoro and Naaman [51] performed a qualitative analysis of Twitter posts, as well as conducting interviews with experienced users who post messages about exercise, diet, and weight loss activities. Their goal was to identify what motivates people to share such content on social media. Recently, Park et al. [43] studied the traits of users who share their personal health and fitness related information on Twitter via the MyFitnessPal application. They observed that persistent sharing of such health status updates on Twitter was correlated with health-related linguistic attributes, as well as presence of a fitness-oriented support network. Close to our work is also the recent work of Wang et al. [56], wherein the authors

examined weight updates shared on Twitter via a Withings internet-enabled smart scale. Akbar and Weber [3] present a similar study on sleep duration and quality. They track auto-generated tweets from a sleep-recording mobile app to connect a user’s sleep behavior to their social media feed.

This body work is valuable to our investigation, because it provides evidence linking social media use and one’s health and well-being. While most of this work focuses on using qualitative methods, in our paper, we adopt a computational approach to build models that are able to examine in what ways one’s behavior, affect, activities and language on Twitter may relate to diet success or failure.

### Assessing Health Status from Social Media

Considerable research has focused on developing approaches that can (semi-)automatically assess health and wellness states using social media. These include Twitter-based topic models to identify conditions and symptoms related to diseases [44], postpartum depression [16], depressive disorders [17], eating disorders [7], suicidal ideation [18], addictive behaviors and substance abuse [38], as well as a variety of county-level health statistics [15].

Online data also been employed to infer dietary practices, food consumption, and taste preferences of individuals and populations [54, 35]. For instance, Wagner et al. [54] analyzed data from an online recipe platform to understand the association between geographic proximity and shared dietary preferences and the extent to which temporal information helps to predict these preferences. Twitter and Facebook posts [1] as well as the content of Instagram images [23] have further been found to correlate with CDC reported prevalence of obesity in different geographical regions, as well as in helping infer caloric and nutrient consumption. Weber and Mejova [58], recently showed the feasibility of using crowd-sourcing to infer body weight categories from profile pictures shared on Twitter.

Relevant to our study is the recent work of Padrez et al. [42] wherein the authors linked social media and medical record data to determine the acceptability to patients and potential utility to researchers of a database linking patients’ social media content with their electronic medical record (EMR) data.

In our work we extend the above line of research by integrating social media with self-reported information about caloric goals and consumption on the MyFitnessPal platform. In this way, we believe our approach will be able to complement state-of-the-science health assessment tools to give rich, detailed, difficult to measure observations about context and content of social interactions, social capital, and exposure to online cues that may alleviate or mediate an individual’s success or failure toward their fitness or dietary goal.

## DATA

### Collecting MyFitnessPal and Twitter Data

We utilize self-reported information about calorie goals, intake and consumption shared via the popular smartphone application MyFitnessPal<sup>2</sup> as our source of quantified self sens-

<sup>2</sup><https://www.myfitnesspal.com/>

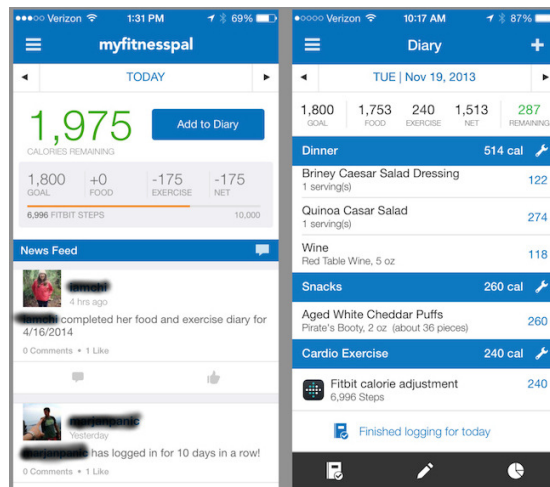


Figure 1: Screenshot of the MyFitnessPal (MFP) application, along with an MFP diary entry.

ing data. The app (henceforth MFP) allows users to track their diet and exercise in a diary form to determine optimal caloric intake and nutrients for the users’ goals. MFP also allows users to share their updates on their Facebook or Twitter profiles, thereby creating a link to their social media activity. It might seem that MFP is not a typical source of “sensed” data as outlined in the literature [20], however since we unobtrusively and non-invasively collect and utilize this data as a stream of information relating to one’s diet, it may be considered to be sensed.

We focus on MFP data shared by users in their food diaries, that includes detailed daily caloric and nutrient information about food items consumed, total calorie consumed, and the calorie goal for the all of the meals of the day. For instance, on a certain day the MFP food diary of a user could record that their calorie goal for the day was 1,400, and they consumed a total of 1,463 calories in all. Thus the user exceeded their goal by 63 calories. A screenshot of the app is shown in Figure 1.

**Collection Methodology.** We adopted a bootstrapping methodology to find users who have cross-posted their MFP diary updates on Twitter. Following manual inspection, we found that typically, the Twitter cross-postings of MFP diary updates have a predictable structured format containing the hashtag “#myfitnesspal” and including a shortened link, that points to the post author’s online MFP diary (e.g., “*completed his food and exercise diary for 4/22/2016 and was under his calorie goal* <bit.ly URL> #myfitnesspal”). To get this data, we used the official Streaming API of Twitter<sup>3</sup> to find publicly shared posts with the hashtag “#myfitnesspal” and containing an embedded link (e.g., “bit.ly”). From this candidate set of over 1,000 posts shared over a week long period in early 2016, we extracted the unique usernames and expanded the MFP diary links included in the posts.

Thereafter, we undertook two parallel tasks: (1) Obtaining the entire timeline (or the most recent 3,200 posts) of each

<sup>3</sup><https://dev.twitter.com/streaming/overview>

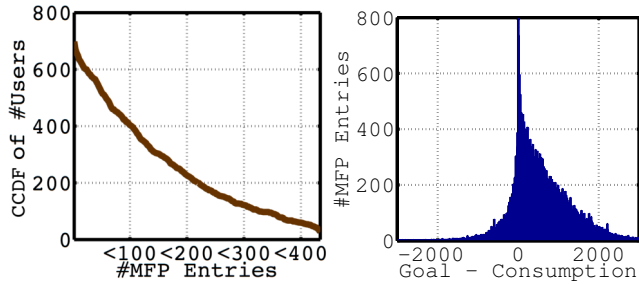


Figure 2: [Left] Complementary Cumulative Density Function (CCDF) showing the distribution of users over number of MFP entries. [Right] Distribution of the number of MFP entries over the difference between calorie goal and calorie consumption.

Twitter user using the Twitter API; and (2) Obtaining the content (or the parsed HTML page source) for each of the extracted MFP diary links, whose content was publicly accessible. We were able to obtain Twitter post timeline information for 1,038 users who still had their account active/public at the time of data collection, yielding 2,249,297 posts (mean posts/user=2,166.9; median posts/user=2,839). Separately, after data cleaning, we were able to obtain a total of 109,920 MFP diary entries of 698 users from the links. For 692 users, we had both their Twitter data and MFP diary data. We aligned the MFP and Twitter data for each of these users by considering the timestamps of their MFP entries and Twitter posts. Concretely, for a specific user, we considered only those of their MFP and Twitter posts that lay within the largest overlapping timestamp range with consecutive daily data from both data sources.

**Data Cleaning.** We further cleaned the above dataset, especially focusing on MFP food diary entries whose calorie goals or calorie consumed values were more than two standard deviations away from the user’s mean calorie goal/consumed: (1) We removed two individual MFP food diary entries since they reported unusually large calorie goals ( $\sim 1$  million). (2) We excluded MFP entries for whom the absolute difference between calorie goal and consumption was greater than 3,000 calories, since we suspected this might indicate misreporting of either goals or consumption, or specific individuals with extra-ordinary dietary/fitness regimes. (3) We additionally removed one user since they had all their calorie goals set to 0. In the rest of the paper, we will use this set of 691 users.

### Data Descriptive Statistics

We present some descriptive statistics of the data we collected above. For the 691 users, the mean and median number of MFP food diary entries per user were 157.5, and 122, indicating the presence of sufficient and temporally spread diet data. Figure 2 gives the distribution. Among all of the MFP entries, we had 88,982 entries where the calorie consumption was under the calorie goal, and 20,103 where the consumption was greater than the goal. In Figure 2 we also show the distribution of this difference over MFP food diary entries. Additionally, the distribution of the number of users over their mean calorie goal and calorie consumption across

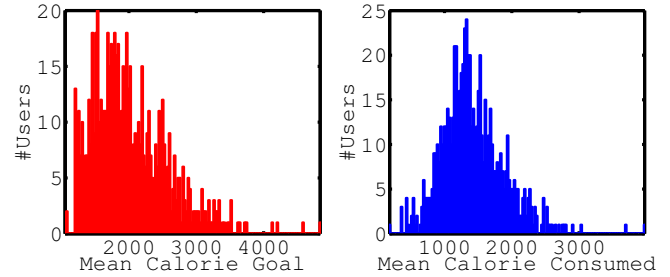


Figure 3: Distribution of number of users over their mean calorie goals and calorie consumption, as measured from the MFP food diary entries. The mean calorie goal of users is 2019 calories ( $\sigma = 566.4$ ), while the mean calorie consumption is 1418 calories ( $\sigma = 466.8$ ).

all MFP entries is given in Figure 3. Interestingly, we observe that users whose calorie consumption was, in general, under their goal, had more MFP entries ( $\mu = 129.1$ ;  $\sigma = 110.5$ ), than those for whom the mean consumption was more than their goal ( $\mu = 33.6$ ;  $\sigma = 45.4$ ). This aligns with prior findings that indicate that individuals who tend to be successful in meeting their dietary goals, tend to be persistent users of the MyFitnessPal platform [57].

## METHODS

### Defining Measures of Diet Compliance

Before proposing our methods, we first define a number of metrics of diet compliance “success” and “failure” for the users in our data, based on suitably chosen empirical thresholds that were found to distinctively demarcate the success and failure classes. Note that the six measures defined below are independent of each other, i.e., a user is assessed to be successful or not based on each individual measure. Below we provide definitions of the measures – based on a users’ MFP dairy entries until day  $t$ , their success/failure label at  $t$  is determined as:

- *Success*: if the mean caloric consumption of the user is under 2,000 calories until day  $t$  — this threshold is chosen based on the typical diet of a moderately active adult released by the USDA<sup>4</sup>; *Failure*: if the mean caloric consumption is over 2,000 calories.
- *Success*: if the mean calorie consumed by the user is under their mean calorie goal over all days until  $t$ ; *Failure*: if it is over the goal set by the user.
- *Success*: if the mean calorie consumed is 25% or more under the mean calorie goal over all days until  $t$ ; *Failure*: if it is over this threshold.
- *Success*: if the mean calorie consumed is 50% or more under the mean calorie goal over all days until  $t$ ; *Failure*: if it is over this threshold.
- *Success*: if the user’s calorie consumption was less than the corresponding calorie goal for at least 50% of the days until  $t$ ; *Failure*: if the consumption was above the goal for more than 50% of the days.

<sup>4</sup>[http://www.cnpp.usda.gov/sites/default/files/usda\\_food\\_patterns/EstimatedCalorieNeedsPerDayTable.pdf](http://www.cnpp.usda.gov/sites/default/files/usda_food_patterns/EstimatedCalorieNeedsPerDayTable.pdf)

| Metric      | “Success” #Users | “Failure” #Users |
|-------------|------------------|------------------|
| BASELINE    | 614              | 75               |
| CALDIFF     | 655              | 34               |
| CALDIFF:25  | 371              | 318              |
| CALDIFF:50  | 80               | 609              |
| PROPDAYS:50 | 657              | 32               |
| PROPDAYS:75 | 518              | 171              |

Table 1: Sizes of the user classes (“success” and “failure”) based on the different diet compliance metrics.

- *Success*: if the user’s calorie consumption was less than the corresponding calorie goal for at least 75% of the days until  $t$ ; *Failure*: if the consumption was above the goal for more than 25% of the days.

Refer to Table 1 for information on the distribution of success and failure user classes obtained by employing the above six measures in our MFP data. Here, each user’s whole overlapping range of MFP and Twitter data was used as the period of consideration. We note imbalance in class sizes for some measures, including a heightened propensity for success – the implications of these data artifacts are presented in the Discussion section.

### Social Media Behavioral Measures

We now identify various attributes to characterize the behavior manifested through the Twitter posts of the MFP users. We consider three categories of linguistic measures: (1) **affective attributes**, (2) **cognitive attributes**, and (3) **linguistic style attributes**. These measures are largely based on the psycholinguistic lexicon LIWC<sup>5</sup>, and were motivated from prior literature that examines associations between the behavioral expression of individuals and their health and well-being [9]. For our experiments we used the 2015 version of the LIWC dictionary.

(1) We consider two measures of affect derived from LIWC: positive affect (PA), and negative affect (NA), and four other measures of emotional expression: *anger*, *anxiety*, *sadness*, and *swear*.

(2) We use LIWC to define cognitive measures as well: (a) cognition, comprising *cognitive mech*, *discrepancies*, *inhibition*, *negation*, *causation*, *certainty*, and *tentativeness* words; and (b) perception, comprising a set of words in LIWC around *see*, *hear*, *feel*, *percept*, *insight*, and *relative*.

(3) Next, we consider four measures of linguistic style: (a) **Lexical Density**: consisting of words that are *verbs*, *auxiliary verbs*, *adverbs*, *prepositions*, *conjunctions*, *articles*, *inclusive*, and *exclusive*. (b) **Temporal References**: consisting of *past*, *present*, and *future* tenses. (c) **Social/Personal Concerns**: words belonging to *family*, *friends*, *social*, *work*, *health*, *humans*, *religion*, *bio*, *body*, *money*, *achievement*, *home*, *sexual*, and *death*. (d) **Interpersonal Awareness and Focus**: words that are *1st person singular*, *1st person plural*, *2nd person*, and *3rd person* pronouns.

<sup>5</sup><http://liwc.wpengine.com/compare-dictionaries/>

We additionally consider two measures of user **activity**: (1) *Interactivity*, as given by the fraction of @-replies in a user’s timeline of posts on Twitter; and (2) *Information sharing*, as given by the fraction of posts containing an external link. Our final two measures assess a user’s **social capital**, given by the number of *in-links* (followers) in their ego-centric network on Twitter, and the number of *out-links* (followees).

### Analytic Techniques

#### Models for Diet Compliance Success and Failure

To address RQ 1, we propose statistical models to predict a user’s diet compliance success or failure using the proposed social media measures. Specifically, we build six different regularized logistic regression models (we employ ridge regression): one corresponding to each diet compliance metric. The goal of these logistic regression models is to explain, model, and classify the social media attributes of MFP users who succeed or fail in diet compliance. Thus the models use the diet compliance success/failure labels (BASELINE, CALDIFF, CALDIFF:25, CALDIFF:50, PROPDAYS:50, and PROPDAYS:75 respectively) of all user over all of their MFP entries as the *response variable*. Thereafter, as *explanatory variables*, we consider the three sets of social media measures: language, activity, and social capital. Note that, we excluded any tokens in posts that contained the string “my-fitnesspal”, in order to not corrupt the explanatory variables with information about the response variable.

To evaluate the goodness of fits of our six models, we use *deviance*. Briefly put, deviance is a measure of the lack of fit to data, hence lower values are better. It is calculated by comparing a model with the saturated model—a model with a theoretically perfect fit, which we consider to be the intercept-only model and refer to as *Null*. Additionally, we perform  $k$ -fold cross validation ( $k = 5$ ) to determine the best tuning of the model parameters, and also to prevent overfitting to the dataset. It also helps us assess the performance of the models in predicting diet compliance success and failure on unseen test data. We evaluate model performance here through the metrics accuracy, precision, and recall.

#### Granger Causality Analysis

Next, per RQ 2, we are concerned with the question whether social media data correlates with temporal changes in diet compliance, defined by our various metrics of success and failure<sup>6</sup>. To answer this question, we adopt the econometric technique of Granger causality analysis [24] to the daily time series produced by the Twitter specific measures, versus success or failure in diet compliance. Granger causality analysis rests on the assumption that if a variable  $X$  causes  $Y$  then changes in  $X$  will systematically occur before changes in  $Y$ . We will thus find that the lagged values of  $X$  will exhibit a statistically significant correlation with  $Y$ . We note that here, correlation however does not prove causation; we are merely testing whether one time series has predictive information about the other.

<sup>6</sup>Since here we are studying the temporal changes in diet compliance and its relationship to social media, we compute daily diet compliance metrics of each user, instead of over the entire period.

For our task, we propose two different lagged multivariate vector auto-regressive (VAR) models for each compliance metric [25]. Our first model is meant to forecast success (or failure) on day  $t_k$  based on (i) success/failure in compliance over  $n$  previous days ( $t_{k-1}, t_{k-2}, \dots, t_{k-n}$ ), as well as (ii) the number of calories consumed over  $t_{k-1}, t_{k-2}, \dots, t_{k-n}$ . The second model would forecast the same, but using  $n$  lagged values of both success/failure and calories consumed, along with all of the three categories of social media derived measures (language, activity, social capital) over  $t_{k-1}, t_{k-2}, \dots, t_{k-n}$ . Using the log likelihood metric of the models, we propose to assess the role of social media—the second model would enable better forecast of diet compliance if the log likelihood of the model is higher than the first.

Next, we statistically establish whether the second model, that uses social media data, indeed improves our diet compliance forecasting ability over the one that uses MFP data alone. For this purpose, we utilize the Granger causality test [24]. The Granger causality test is a statistical hypothesis test to determine whether a time series  $X$  (the predictor variable, or MFP or linguistic, activity, social capital measures) is useful in forecasting another time series  $Y$  (the predicted variable, or diet compliance success/failure) by attempting to reject the null hypothesis that  $X$  does not help predict, i.e., Granger-cause,  $Y$ . The alternative hypothesis is that adding  $X$  does help predict  $Y$ .

However we note in our case, we cannot apply the traditional Granger causality test, as the time series we intend to forecast, say,  $Y$  is not a vector over time, but a *set* of such vectors spanning multiple MFP users. Hence, we adopt methods for generalizing Granger causality to sets of time series, where  $m$  time series ( $X$ ) Granger-cause  $m$  other time series ( $Y$ ) [24]. Specifically, we use canonical-correlation analysis (CCA) as a multivariate statistical technique that has been applied to Granger causality analysis to infer information flow between sets of time series [41]. It finds linear combinations of the predictor ( $X$ ) and predicted variables ( $Y$ ) which have maximum correlation with each other.

After applying the CCA, we propose to examine the null hypothesis stated above; in other words, if *any* of the different predictor variables  $X$  is less than a chosen level of significance  $\alpha$ . For this purpose, we will use the  $F$ -statistic and the Wilks' statistic [48]. If there is at least one such variable, one would conclude that Granger causality is present, i.e., there is a predictor variable (MFP or Twitter measure) that leads to the predicted outcome (diet compliance success/failure).

## RESULTS

### RQ 1: What Predicts Diet Compliance?

In Table 2 we report a summary of the different model fits. For each of the six diet compliance metrics, we build three models using the various Twitter derived measures—(i) language measures only, (ii) language and activity measures, and (iii) all of the language, activity and social measures. Due to the randomness introduced by cross-validation, we report the results corresponding to the lowest deviances that we obtained in any of the runs.

|                       | Deviance | df | $\chi^2$ | LL      | $p$         |
|-----------------------|----------|----|----------|---------|-------------|
| <b>BASELINE</b>       |          |    |          |         |             |
| Null                  | 1037.6   | 0  |          |         |             |
| Language              | 585.96   | 49 | 451.67   | -106.3  | $< 10^{-5}$ |
| Language + Activity   | 418.24   | 51 | 619.39   | -138.5  | $< 10^{-5}$ |
| Lang + Activ + Social | 388.40   | 53 | 649.23   | -194.2  | $< 10^{-6}$ |
| <b>CALDIFF</b>        |          |    |          |         |             |
| Null                  | 633.12   | 0  |          |         |             |
| Language              | 354.43   | 49 | 279.8    | -41.29  | $< 10^{-3}$ |
| Language + Activity   | 348.80   | 51 | 284.32   | -74.60  | $< 10^{-4}$ |
| Lang + Activ + Social | 315.25   | 53 | 317.87   | -96.99  | $< 10^{-4}$ |
| <b>CALDIFF:25</b>     |          |    |          |         |             |
| Null                  | 1164.7   | 0  |          |         |             |
| Language              | 814.73   | 49 | 349.9    | -129.37 | $< 10^{-5}$ |
| Language + Activity   | 703.45   | 51 | 461.25   | -218.43 | $< 10^{-6}$ |
| Lang + Activ + Social | 586.84   | 53 | 577.95   | -293.42 | $< 10^{-6}$ |
| <b>CALDIFF:50</b>     |          |    |          |         |             |
| Null                  | 847.93   | 0  |          |         |             |
| Language              | 202.37   | 49 | 645.56   | -71.18  | $< 10^{-3}$ |
| Language + Activity   | 186.49   | 51 | 661.44   | -90.24  | $< 10^{-4}$ |
| Lang + Activ + Social | 149.86   | 53 | 698.07   | -136.3  | $< 10^{-4}$ |
| <b>PROPDAYS:50</b>    |          |    |          |         |             |
| Null                  | 641.29   | 0  |          |         |             |
| Language              | 197.27   | 49 | 444.02   | -56.81  | $< 10^{-3}$ |
| Language + Activity   | 134.86   | 51 | 506.42   | -89.06  | $< 10^{-3}$ |
| Lang + Activ + Social | 114.94   | 53 | 526.34   | -124.8  | $< 10^{-4}$ |
| <b>PROPDAYS:75</b>    |          |    |          |         |             |
| Null                  | 1378.3   | 0  |          |         |             |
| Language              | 962.54   | 49 | 415.79   | -223.5  | $< 10^{-6}$ |
| Language + Activity   | 872.14   | 51 | 506.15   | -291.4  | $< 10^{-9}$ |
| Lang + Activ + Social | 759.08   | 53 | 619.24   | -379.5  | $< 10^{-9}$ |

Table 2: Summary of different model fits. Null is the intercept-only model. For each model, we also report the log likelihood (LL) and the  $p$ -value of statistical significance.

**Examining Model Fits.** Compared to the Null models, all of our models provide explanatory power with significant improvements in deviances in predicting the six diet compliance metrics: BASELINE, CALDIFF, CALDIFF:25, CALDIFF:50, PROPDAYS:50, and PROPDAYS:75. The difference between the deviance of a Null model and the deviances of each of the other models approximately follows a  $\chi^2$  distribution, with degrees of freedom equal to the number of additional variables in the more comprehensive model. E.g., comparing the deviance of the model that uses Language + Activity measures in predicting CALDIFF:25 with that of the Null model, we see that the information provided by these measures has significant explanatory power:  $\chi^2(51, N = 613) = 1164.7 - 703.4 = 461.3, p < 10^{-6}$ .

We observe similar deviance results when we compare the set of three models (Language, Language + Activity, Language + Activity + Social) corresponding to each diet compliance metric: the latter models give better fits. This is further apparent in the measures of the log likelihood given by the three models corresponding to each diet compliance metric. For instance, the log likelihood of the Language, Language + Activity, Language + Activity + Social model in predicting CALDIFF:50 is 91.3% higher than that given by the model that uses language measures only. This indicates that the three categories of measures derived from Twitter, together provide



|                                 | Model 1 | Model 3 | Model 4 | Model 6 |
|---------------------------------|---------|---------|---------|---------|
| <b>Affective attributes</b>     |         |         |         |         |
| PA                              | 8.699   | 10.489  | 7.612   | 10.352  |
| NA                              | -11.680 | -13.776 | -6.095  | -8.888  |
| anger                           | -2.535  | -5.496  | -6.398  | -4.169  |
| anxiety                         | -6.389  | -4.073  | -2.402  | -3.864  |
| sadness                         | -4.700  | -7.180  | -3.218  | -7.106  |
| swear                           | -2.081  | -0.848  | -1.631  | -0.978  |
| <b>Cognitive attributes</b>     |         |         |         |         |
| <b>Cognition</b>                |         |         |         |         |
| cognitive_mech                  | 5.085   | 8.064   | 5.949   | 8.480   |
| discrepancies                   | 3.379   | 1.468   | 1.332   | 0.971   |
| negation                        | -1.784  | -2.793  | -2.564  | -2.292  |
| inhibition                      | -2.197  | -2.576  | -3.130  | -2.068  |
| causation                       | 1.077   | 0.838   | 1.248   | 1.084   |
| certainty                       | 1.627   | 1.428   | 0.998   | 1.324   |
| tentativeness                   | 1.756   | 0.632   | 0.652   | 1.230   |
| <b>Perception</b>               |         |         |         |         |
| see                             | 2.421   | 2.007   | 2.472   | 1.741   |
| hear                            | 5.809   | 1.881   | 3.201   | 1.695   |
| feel                            | 1.879   | 3.480   | 4.417   | 2.414   |
| percept                         | 3.513   | 1.502   | 3.890   | 3.115   |
| insight                         | 1.445   | 1.365   | 2.056   | 3.603   |
| relative                        | 1.247   | 1.079   | 1.877   | 2.089   |
| <b>Lexical Density</b>          |         |         |         |         |
| verbs                           | 1.388   | 0.847   | 1.259   | 0.315   |
| auxiliary verbs                 | 1.684   | 0.293   | 1.043   | 0.367   |
| adverbs                         | 0.782   | 0.564   | 0.934   | 0.419   |
| prepositions                    | 0.610   | 0.594   | 0.769   | 0.520   |
| conjunctions                    | 1.168   | 0.624   | 0.431   | 0.656   |
| articles                        | 1.132   | 1.999   | 1.692   | 0.395   |
| inclusive                       | 0.577   | 0.563   | 0.422   | 0.504   |
| exclusive                       | 1.251   | 0.596   | 0.853   | 0.758   |
| <b>Temporal references</b>      |         |         |         |         |
| past tense                      | -4.126  | -4.459  | -5.866  | -3.681  |
| present tense                   | 3.422   | 4.666   | 3.429   | 2.351   |
| future tense                    | 4.585   | 3.880   | 3.849   | 2.387   |
| <b>Social/Personal Concerns</b> |         |         |         |         |
| family                          | 7.265   | 7.251   | 7.710   | 6.070   |
| friends                         | 6.782   | 6.558   | 5.227   | 6.977   |
| social                          | 6.829   | 8.062   | 5.045   | 9.115   |
| work                            | -1.741  | -1.112  | -1.198  | -1.508  |
| health                          | 11.381  | 10.270  | 7.135   | 6.988   |
| humans                          | -0.809  | -0.236  | -1.620  | -0.742  |
| religion                        | -0.616  | -0.306  | -0.629  | -0.356  |
| bio                             | 1.716   | 2.147   | 2.971   | 2.348   |
| body                            | 6.402   | 5.611   | 4.809   | 6.251   |
| money                           | -1.928  | -1.460  | -1.132  | -1.327  |
| achievement                     | 1.938   | 2.187   | 2.085   | 1.832   |
| home                            | 2.261   | 1.154   | 1.532   | 1.671   |
| sexual                          | -0.874  | -0.421  | -0.773  | -0.134  |
| death                           | -0.590  | -1.146  | -0.851  | -1.430  |
| <b>Interpersonal awareness</b>  |         |         |         |         |
| 1st p. singular                 | -4.228  | -3.304  | -4.117  | -2.014  |
| 1st p. plural                   | 3.825   | 1.144   | 1.304   | 2.046   |
| 2nd p.                          | 5.379   | 2.045   | 2.508   | 3.621   |
| 3rd p.                          | 1.763   | 1.599   | 2.078   | 1.042   |
| <b>Activity</b>                 |         |         |         |         |
| Interactivity                   | 8.512   | 4.277   | 5.370   | 9.669   |
| Information sharing             | 4.539   | 3.706   | 5.524   | 3.974   |
| <b>Social capital</b>           |         |         |         |         |
| # inlinks                       | 3.718   | 11.083  | 10.688  | 8.977   |
| # outlinks                      | 7.176   | 4.935   | 4.940   | 5.538   |

Table 3: Predictor variables and their coefficients ( $\beta$ ) obtained from four logistic regression models, with the highest  $\chi^2$ -statistic in Table 2. (1) Model 1 predicts: BASELINE ; (2) Model 3 predicts: CALDIFF:25 ; (3) Model 4 predicts: CALDIFF:50 ; and (4) Model 6 predicts: PROPDAYS:75 .

improved explanatory power for diet compliance, compared to the categories alone or their sub-combinations.

Further, we observe differences in the deviances (and the corresponding  $\chi^2$ -statistics and log likelihoods) given by the models predicting the different diet compliance metrics. The  $\chi^2$ -statistics for the three models predicting BASELINE , CALDIFF:50 , and PROPDAYS:75 metrics tend to be larger than the ones predicting CALDIFF , CALDIFF:25 , and PROPDAYS:50 metrics. For instance, the model with the best fit, that uses language, activity and social measures in predicting CALDIFF:50 has considerably low deviance compared to the Null model; this difference follows a  $\chi^2$  distribution:  $\chi^2(53, N = 613) = 847.93 - 149.86 = 698.07, p < 10^{-4}$ .

**Measures with High Predictive Power.** Next, we examine: what measures provide the most explanatory power in predicting the different diet compliance metrics. For this purpose, in Table 3, we report the  $\beta$  coefficients given by the logistic regression models predicting a subset of the diet compliance metrics; we report on those models that use all of the three categories of social media measures, language, activity, and social capital, and include the four models for which we obtained the largest  $\chi^2$ -statistic values in Table 2. These models are the ones that predict the following diet compliance metrics: BASELINE , CALDIFF:25 , CALDIFF:50 , and PROPDAYS:75 .

We now present a discussion of the measures with the highest (absolute values) of the  $\beta$  coefficients, that is, ones that depict the most predictive association (positive or negative) with diet compliance success/failure, across all of the models:

**Observation 1:** *Users who succeed in diet compliance express higher positive affect; those who do not succeed express more negative affect, anger, anxiety, and sadness in their Twitter posts.* That is, likelihood of diet compliance success is higher in users whose content exhibits a more pronounced hedonic focus on positive emotions and a positive outlook towards life (PA:  $\beta = 7.6$  to 10.4).

“Train smart like a trainer. eat clean like a nutritionist. sleep like well bathed baby. win like 6 time champion” (↑ positive affect)

“Kids enjoying what they love best! festival fun! #blessed #enjoythemoment #festivalfun” (↑ positive affect)

Users who do not succeed express more negative emotions in their posts (NA:  $\beta = -6.1$  to  $-13.7$ ), for instance, frustrations in not being able to meet fitness goals, or sharing less pleasing experiences.

“I have had the shittest year, I have not done what I needed to do and I have gained 30 pounds. how shit, how shit (↑ negative affect)

“Silence is my go to when I’m upset, sad, mad or emotional. It’s obvious something’s wrong cuz I’m nearly silent (↑ negative affect)

Users who fail in complying with their dietary goals also express higher levels of anxiety (anxiety:  $\beta = -2.4$  to  $-6.3$ ). They tend to be anxious because of the lack of emotional control as well as due to certain activities and events of daily life.

“I’m pretty sure I’m going to lose my mind. Completely lose what little is left. I cry at the thought of stupid things..” (↑ anxiety)

“Feel rough as old boots this morning :/ Ankle hurts, shin hurts, chest hurts, head hurts” (↑ anxiety)

**Observation 2:** *Users who succeed in complying with their diet plan exhibit high cognitive functioning and perception; conversely, those who fail show signs of cognitive impairment.* Specifically, we observe that the former cohort speaks more insightfully in their posts, shows a self-reflective shift and connection to their own cognition through use of cognitive\_mech ( $\beta = 5.1$  to  $8.4$ ), inhibition ( $\beta = -2.1$  to  $-3.1$ ), and insight words ( $\beta = 1.4$  to  $3.6$ ).

“If we never stumble we never fall. If we never fall we never fail, and if we never fail we never grow! #noexcuses” (↑ cognitive\_mech)

“If you could find something that would heal hurt, forgive guilt, calm fear, inspire hope, reveal purpose, and give life would that be good news?” (↑ cognitive\_mech)

**Observation 3:** *Successful diet compliance users engage in greater Twitter discourse around health issues and show heightened interest in topics relating to body and physical health.* Specifically these health ( $\beta = 6.9$  to  $11.3$ ) and body ( $\beta = 4.8$  to  $6.4$ ) related discourses include sharing tips toward continuing to maintain fitness progress, sharing advice and information around improved eating habits and nutrition, and disclosures of fitness parameters like weight and body fat.

“Ever felt sick and tired of being sick and tired? I have been exercising away chronic fatigue” (↑ health)

“Here are some awesome nutritional advice for health and wellbeing ::URL:: #food #recipes #health #healthy” (↑ health)

“Week 1, down 7.2 lbs, 1” off the waist. chest, hips, neck :) body fat remain unchanged. #letsdothis” (↑ body)

“When you wake up, stretch your arms, legs to get your blood circulating in your body. this helps the body wake up faster” (↑ body)

**Observation 4:** *Users who are successful in diet compliance express greater sense of achievement in their Twitter content.* We find that these users engage in motivational and inspirational content sharing around improved health and well-being, as well as objective goals and outlook toward meeting their desired fitness and diet (achievement:  $\beta = 1.8$  to  $2.1$ ).

“If your not failing your not trying hard enough. try, fail, try, fail and repeat for success. #success” (↑ achievement)

“Just set behavior goals instead of outcome goals:instead of a goal to lose 20 pounds, set a goal of strength training 3 days” (↑ achievement)

**Observation 5:** *Successful diet compliance users tend to share content around family and friend relations, and topics relating to their social life and events.* Sharing of such content (family:  $\beta = 6.1$  to  $7.7$ ; friends:  $\beta = 5.2$  to  $6.9$ ; social:  $\beta = 5.0$  to  $9.1$ ) indicates the presence of a supportive social structure for the successful users, ability to bond and engage with it from time to time—aspects that are positively linked to improved health and well-being [56].

“Enjoying family. My uncle’s with their niece’s and nephews. Minus a few nieces and nephews” (↑ family)

“Best friends since 5th grade; future best friends! blessed by the wonderful life long friendships ::URL::” (↑ friends)

**Observation 6:** *Users successful in complying with their diet plans show a temporal orientation that focuses on the here and now in their social media content, and demonstrate a future temporal reference.* This observation also suggests that the successful users are goal-oriented and bear a positive outlook towards the diet compliance process in the near future (present tense:  $\beta = 2.3$  to  $4.6$ ; future tense:  $\beta = 2.3$  to  $4.5$ ). Users who fail in complying with their dietary goals tend to share posts that are ruminative of the past, and can be regretful or nostalgic of past events.

“Transforming your body will have a domino effect. Your body will change, your mind will change, your life will change. Eric Eisenberg” (↑ future tense)

“I will stick to my fitness plan. It will be difficult. It will take time. It is going to require sacrifice. But it will be worth it.” (↑ future tense)

**Observation 7:** *Users who succeed in complying with their diet goals show higher collective attentional focus in their social media writing (1st person plural pronouns), as well as more social involvement and attention to people and objects, manifested in the use of 2nd and 3rd person pronouns respectively.* Unsuccessful users, on the other hand, show high pre-occupation with their own selves in their Twitter posts, as measured through the use of 1st person singular pronouns (1st pp. singular:  $\beta = -2.0$  to  $-4.2$ ).

**Observation 8:** *Users successful in diet compliance tend to be more socially interactive in their social media activity (as measured by the sharing of @-replies), and tend to disseminate more external information to their networks via links.* We conjecture this might indicate the desire for these users to seek and reach out to others, as well as openness to consuming and sharing novel information (interactivity:  $\beta = 4.2$  to  $9.6$ ; information sharing:  $\beta = 3.7$  to  $5.5$ ).

**Observation 9:** *Finally, diet complying users exhibit greater availability of social capital in their networks, as measured by the number of inlinks and outlinks in their social media network.* Together, these measures (inlinks:  $\beta = 3.7$  to  $11.1$ ; outlinks:  $\beta = 4.9$  to  $7.1$ ) indicate the presence of a strong support system for users who succeed in their dietary goals.

**Cross Validation and Performance Evaluation.** Next, we assess the performance of the six logistic regression models that give the lowest deviances in Table 2, in predicting diet compliance success/failures labels of users in a test set. Note that these models use all of the language, activity and social capital measures.

In Figure 4 we present the mean accuracy, precision, recall, and receiver operating characteristic (ROC) curves for each model, averaged across the five folds of cross validation performed. Our best performing model is the one that predicts success/failure for the CALDIFF:25 diet compliance metric: the “chance” accuracy for this model is 53.8%, and our model improves over it by 37.2%. Close second



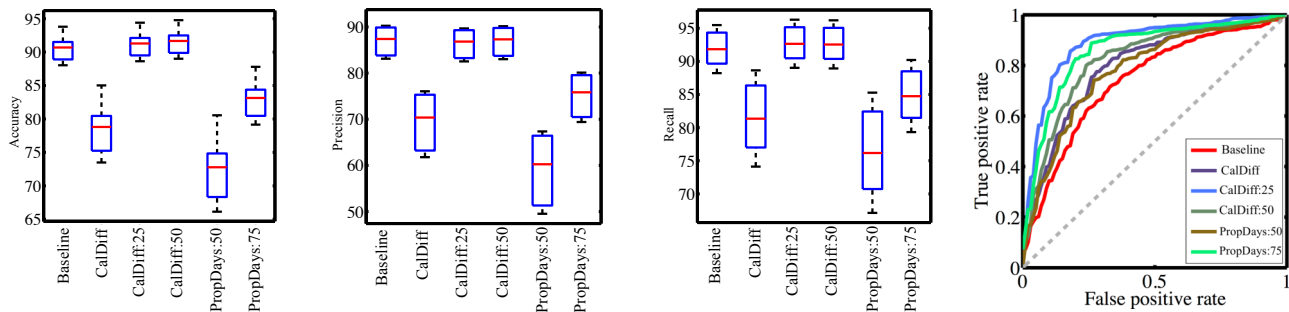


Figure 4: Performance of different logistic regression models (accuracy, precision, recall, and receiver operating characteristic—ROC curve) in classifying success and failure corresponding to the six different diet compliance metrics: BASELINE , CALDIFF , CALDIFF:25 , CALDIFF:50 , PROPDAYS:50 , and PROPDAYS:75 .

are the models that predict diet compliance success/failure for BASELINE , CALDIFF:25 , and PROPDAYS:75 , however their improvements over chance accuracies are modest: 1.4%, 3.1% and 7.7% respectively. We attribute this to the large class imbalance in both the success/failure data for the three diet compliance measures. Models corresponding to CALDIFF and PROPDAYS:50 , that have a very high imbalance class sizes (95.1% and 95.4% chance accuracies respectively), yield 78.4% and 72.3% accuracy. These models thus fail to improve over chance accuracy by 16.6% and 23.1% respectively. Excluding them, our models that leverage Twitter measures to predict diet compliance success and failure improve over baseline chance models by 1.3–37.2%.

On average, all of our models produce satisfactory performance in predicting diet compliance success/failure with mean accuracy 84.4%, precision 77.3% and recall 86.7%. From the ROC curve, we further observe that the models yield large area under curve measures; that is, all of them produce high true positive rates in predicting diet compliance, compromising a manageably small false positive rate.

## RQ 2: Assessing Role of Social Media in Diet Compliance

### Results of Auto-Regressive Models

Presenting the findings of RQ 2, we first discuss the results of fitting two different VAR models toward forecasting each diet compliance metric. Recall, our first VAR model forecasts success (or failure) on day  $t$  based on (i) success/failure in compliance over  $n$  previous days, as well as (ii) the number of calories consumed. The second VAR model forecasts the same, but uses  $n$  lagged daily values of both previous success/failure and calories consumed, along with all of the three categories of social media derived measures (language, activity, social capital), aggregated per day. Since both of these models use time series data, we needed to focus on users who had sufficiently long durations of MFP entries as well as corresponding Twitter data. We consider all users who had at least 100 days of MFP entries; if a user had more than two weeks of gap between two consecutive MFP entries, we did not include them in our analysis. Additionally, in order to determine the optimal size of the lag parameter  $n$  in the two VAR models, we consider historical (Twitter/MFP) measures in increments of 10 days, from the day on which prediction is sought:  $n = 10, 20, 30, \dots, 90$ .

Next we tested whether our Twitter measure inclusive time series predicts diet compliance success and failure. For the purpose, we assessed the additional predictive capability of Twitter data over MFP data, by employing five-fold cross validation in each model, and then reporting accuracy in the test sets. Figure 5 gives these performance measures for the two auto-regressive models. We show one plot corresponding to each diet compliance metric; in each plot we show the accuracy across different values of  $n$ . On an average, the Twitter inclusive model improves accuracy by 17% compared to the model that does not include this data.

Finally, for the VAR model that combines both MFP and Twitter measures (Figure 5), we note that across the different values of lag  $n$  (duration of historical data used in prediction), log likelihood of the model is maximum, or correspondingly, peak accuracy is observed at  $n = 41$  days for all of the diet compliance metrics. This also indicates that values of  $n$  smaller than this may not encapsulate sufficient information for forecasting purposes, whereas when  $n$  is longer, the historical data may include diet behaviors that have evolved over time, and are thus less predictive.

### Granger Causality Tests

As a last result, utilizing the optimal lag parameter  $n$  for the individual metrics of diet compliance, we perform Granger causality statistical tests. The tests examine the individual role each of the MFP and Twitter measures play in forecasting future diet compliance success and failure.

We begin by examining whether the pairs of time series we consider for the Granger causality test: one corresponding to diet compliance success or failure of a user, and the other corresponding to a particular linguistic, activity or social capital measure, satisfy the Dickey Fuller tests for stationary time series [19]. This is an important step to ensure the Granger causality tests are applicable to the different pairs of time series we consider here. We find that the pairs satisfy the Dickey Fuller test ( $p < .05$ ).

Following this test, in Table 4, we report the Wilks'  $\lambda$  (likelihood ratio) statistic, Rao's approximate  $F$  statistic, and the  $p$  value of CCA between the two sets of time series, for each pair of (MFP/Twitter based) predictor and predicted variables (diet compliance success/failure) for all users. For a large

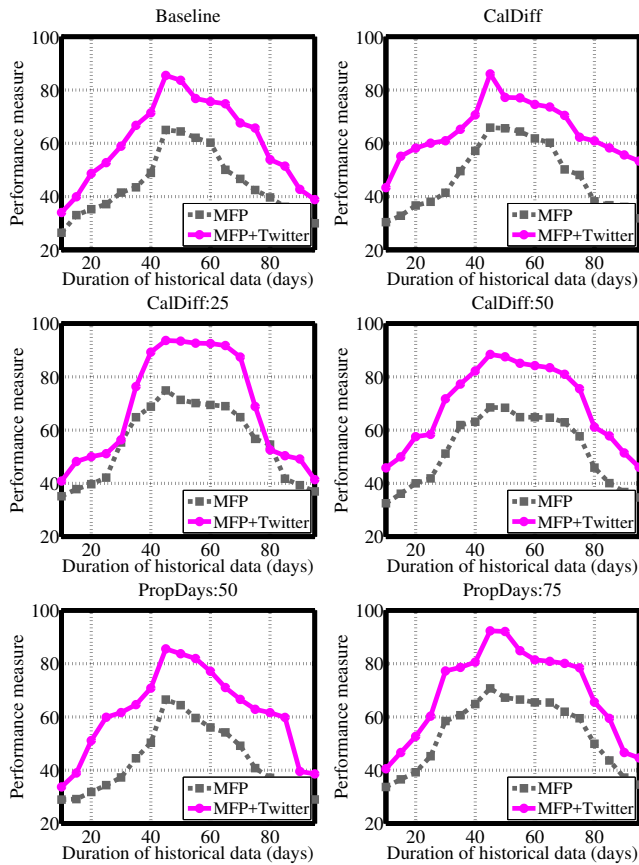


Figure 5: Performance measure (accuracy) of several lagged VAR models using MFP data alone, and using both MFP and Twitter data.

number of the Twitter derived predictor measures we consider, there is statistical significance in Granger causation. This is indicated by lower values of the Wilks’ statistic and higher values of the  $F^2$ -statistic. The measures include: affective attributes like PA, anger, anxiety and sadness; cognitive attributes like cognitive\_mech, inhibition, certainty, as well as perception words like see, hear, feel; temporal references like past, present and future tense use; and interpersonal awareness quantified by the use of 1st person singular and 2nd person pronoun words.

All measures belonging to the activity and social capital categories also indicate high Granger significance: interactivity, number of inlinks, outlinks. While the MFP based predictor variables, like past calorie consumed or proportion of successful diet compliance days in the past, on their own, exhibit significant Granger causation, the significance of the other Twitter measures reveals the additional gain given by their inclusion in the forecasting task. This, in turn, bolsters our earlier observation about the utility of integrating social media with quantified self sensing data on diet compliance.

## DISCUSSION

### Theoretical and Practical Implications

Although our motivation and computational advances align with the vision outlined by the PMI<sup>1</sup>, we note that, in the

scope of this paper, we did not include all possible forms of sensed data such as genetics or biology, typically also included in PMI’s agenda. Instead, we explored how two important elements of health related to each other: one’s physiological attributes (diet compliance) and one’s behavioral attributes (social media). We believe that the integration approach proposed in this paper will enable us to better understand the factors that affect the health of individuals as measured in terms of their diet, the spatio-temporal trends of physical fitness as observed in quantified self sensing and social media platforms integratively, or early signs that may indicate a forthcoming risk to diet compliance. Our work can also enable the development of novel interventions that can proactively monitor risk to diet compliance, and bring appropriate clinical and psychosocial help to those in need. For example, there is a growing body of work in the field of clinical nutrition to build models that can predict a patient’s weight loss [45, 52, 22]. This research is motivated by the desire “to make a decision about the benefit, risk, or cost-efficacy of continued intervention” [22]. These predictive models take into account both weight-specific variables, e.g., the starting weight or the weight lost within an initial certain period, and demographic factors, e.g., age and gender. Incorporating social media data, as shown in our work, could potentially improve the predictive power of such interventions.

Additionally, consideration of two complementary sources of data allows us to validate, for the first time, the efficacy of social media and quantified self sensing measures in revealing risk to diet compliance. The Granger analysis methodology enables us to identify the additional gain in predictive power made possible by examining social media data, over approaches that use an individual’s historical diet compliance data alone. In the future, our technique may also be improved, such as with stratification or propensity score matching [18], to delineate causal pathways associated with diet compliance by aligning multi-level, multi-modal, comprehensive assessments as gleaned from social media and quantified self sensing together. For instance, by temporally aligning social media, quantified self sensing and self-reported attributes, statistical models may be able to explore dynamics of events around *when* or *how soon* an individual is likely to fail in meeting their dietary goal.

### Implications for Design

Integration of diverse health data can lay groundwork for the design of appropriate, adaptive, and privacy-honoring health interventions that may be delivered via a variety of channels: social media, or mobile phone applications. Such tools could be personalized using the quantified self and social media data collected on each individual. They can also “machine learn” patterns over time as more data becomes available. We describe design considerations of some such tools below:

**Persuasion Strategies.** First, social media technologies can apply a variety of automated or semi-automated persuasion strategies that are typically used by people to influence others towards positive behavior change goals—positive feedback, modeling target behaviors or attitudes, and influencing normative rules and social dynamics [13]. Should our methods

discover someone to be at risk of diet compliance failure in the future, motivational messages could be issued as pop-ups in one's MFP application or social media.

Gamification is also possible wherein daily diet compliance success/failure can translate to confidence boosting concrete metrics like rewards, points or badges [59]. Our ability to estimate risk of failure can enable timely and tailored deployment of gamification approaches. For instance, on forecasting that an individual may be at risk of diet compliance failure, subtle nudges (e.g., access to help resources) and alerts could be provided such that individuals can make plans or arrangements that can help them accrue specific incentives or rewards in the future.

**Social Support.** Additionally, mechanisms could be created on social media platforms to provide socially supportive interventions to individuals at risk of diet compliance failure (e.g., those expressing high NA or high cognitive impairment in their Twitter posts). Such interventions could suggest them to engage with users observed to be successful at diet compliance, for instance, through recommendations surfaced on one's social media or quantified self sensing application. Individuals can thus explore and learn about what to expect while trying to meet specific dietary goals, tips and tricks on how to combat urges or situations that make them vulnerable to failure, or for general positive reinforcement to meet their desired dietary goal. Alternatively, at-risk users identified by our methods, such as those who exhibit lower access to social capital, could be recommended to bond as a virtual community based on hashtags or specific social media accounts. This can allow individuals provide peer support to each other, calibrate and compare their own diet compliance success/failure trend with that of others, and motivate each other in helping meet dietary goals.

**Self-Reflection.** While quantified self sensing applications like MFP are great at allowing individuals to log their diet and desired goals [12], our methods could be used to enhance their abilities in promoting improved diet compliance. Trends of diet compliance success/failure may be provided to users of such applications, including forecasts of their likelihood to succeed. Social, activity and other linguistic attributes obtained from social media that relate to these forecasts could also be overlaid with the trends, so that individuals can explore these relationships in an intuitive manner. Broadly, these tools hold the potential to enable better self-reflection and self-awareness of one's dietary practices, so that they could adapt appropriately to forecasted changes in diet compliance. Forecasting abilities can also help self-experimentation. For instance, if an individual finds out they are unlikely to meet their current dietary goals, they could employ changes in their day-to-day activities or alter specific strategies currently in use.

Certainly, designing and deploying the right kind of interventions that fall in any of the above groups needs to carefully consider an individual's choices, lifestyle, and context. These are active directions of future research. What we describe here is the potential of our methods in informing some of

the decisions that underlie next generation personalized and adaptive intervention tools.

### Privacy and Ethical Considerations

Integration of multiple forms of data to promote positive health behaviors is promising, but such approaches are particularly vulnerable to privacy and ethical lapses. We note that individuals may be unaware of the implications of sharing content on social media or of the data being collected via quantified self sensing technologies. Importantly, they might be ignorant of the latent relationships that may be discovered by integrating these diverse forms of health, social, linguistic and activity data, as enabled by our approach. Even though a lot of the health inferences found in prior research are derived from implicit patterns in activity on social media [31], the ability to automatically derive any information about a person's health state may have serious repercussions (e.g., higher insurance rates). Developing tools that remind users of these risks through interactive ways of interpreting their own data, is an important area of future exploration.

Nevertheless, introducing transparency of data analysis to individuals can be challenging. This is because many machine learning approaches typically applied for health inference, like the ones we developed in this paper, consist of non-intuitive, and complex workflows. Promoting actual understanding of such systems and their methods to a non-expert would require developing methods that can abstract operational details but at the same time allow intuitive characterization. Importantly, to promote transparency and still be able to tackle the challenges it poses, novel intervention systems need to revisit the regulations around access, precision and adaptive rights of those individuals who are algorithmically inferred to be at heightened risk of a certain health challenge.

### Limitations and Future Directions

**Truthfulness of MFP data.** We note that it is challenging to accurately estimate how diligent MFP users are in logging their meals. Cordeiro et al. [14] present a qualitative study based on surveys and online forums that analyzes the difficulties faced by users of food journals. Further, a previous analysis of MFP food diaries [57] found that, just before stopping to use the food logging capability in MFP, users would log fewer and fewer calories. Though this could be a positive sign of success, more likely it indicates a drop in "logging morale". However, the same study also found evidence that users prefer to not log at all rather than enter incomplete data. Concretely, users were least likely to log anything on Saturdays and Sundays but, those users that did, had the highest probability across all week days to be above their calorie goal. So whereas some users might not log their caloric intake on a "cheat day" at all, those who do, publicly admit to fail in their goal for these days. This gives us confidence that, except for potential "end of user lifetime" effects, long-term users can largely be assumed to truthfully report their meals, if they choose to report at all.

Still, as we cannot be sure of the food consumed by the users in our study, our findings ultimately only apply to *self-reported* diet compliance, rather than the (unobservable) ac-

tual diet compliance. In the future, technical solutions aimed at facilitating and automating food logging might overcome this limitation [60].

**Choice of diet compliance measures.** Recall, we considered six different measures of diet compliance in this paper. These measures are not meant to be the only possible ways to assess diet compliance. By the same token, we do not suggest employing these specific measures in the interventions we propose above. Instead, the choice of the diet compliance measures was meant to enable comprehensive assessment of the role of social media data in improving predictions of health behavior, albeit a specific one relating to diet.

In fact, having demonstrated the role of social media, appropriately chosen diet compliance measures may be used in specific application contexts. We tested this claim by choosing two intuitive and simple alternative diet compliance metrics and re-running the analysis in RQ 1: 1) the difference between calorie goal and consumption of an individual; and 2) the proportion of days an individual’s calorie consumption is less than their goal. We built and evaluated two OLS regression models to examine if social media can be adequate in predicting these two diet compliance measures. We found satisfactory performance for both, with  $R^2$  values .24 and .13 respectively ( $p < .05$ ). This indicates that social media information is valuable in not just classifying diet compliance success and failure in individuals defined by our six measures, but can also help predict the extent of one’s success or failure in contexts that use alternative diet compliance measures.

**Bias toward diet compliance success.** Relatedly, for the set of six diet compliance measures we considered in this paper, our MFP data appears to be inherently biased toward those who finally succeed (see Table 1). This is evident in the notable class imbalance observed for BASELINE , CALDIFF , and PROPDAYS:50 . But this is not surprising: the quantified selfers who use an application like MFP are inherently more motivated to watch their calories, than those who do not use these technologies. Therefore, in building our predictive models, we evaluated performance of our models over what can be given by a simple chance model that labels everyone in the test set to succeed. Nevertheless, it is possible for the chosen success measures to be correlated with the predictors; and therefore we suggest caution in interpreting our findings.

**Representativeness of Twitter data.** Concerning the use of public Twitter posts, there are also a number of limitations to consider, in addition to the privacy and ethical considerations discussed above. These issues include self-selection bias concerning individuals who choose to share MFP diary updates on Twitter, and self-censorship and impression management related issues. This includes conscious decisions to filter and to not post specific bits of information that may portray one’s less pleasant health or diet related experiences.

Although the individuals who share MFP updates on Twitter might not be similar to Twitter users who use MFP without sharing that information on the social media platform, we do observe the users in our dataset to use Twitter for reasons beyond health information sharing—many of them had Twitter

accounts that date prior to the start of MFP logging. Nevertheless, our methods can easily be extended to contexts where diet compliance information is collected via a survey or another self-report based method.

We also suspect the influence of self-censorship on our findings to be minimal. This is because the capability of sharing MFP updates on Twitter are typically set up during installation of the MFP application. Thus while theoretically possible, it is practically less likely that individuals would regularly curate their Twitter posting behaviors, per the information shared by the MFP application at a given point in time.

**Generalizability of temporal predictions.** Apart from data quality concerns, we also note the limitations of our temporal prediction task based on Granger causality analysis. Recall, we utilized individuals who had at least 100 days of MFP log data, and well as Twitter data spanning the same period. This filtering was important to ensure that we have sufficient length of temporal data available to train and test our models, since smaller timeframes if considered, can greatly aggravate the influence of noise. However we also note that, many quantified self applications have high abandonment rates [10]. Estimates of 30-day *install retention* rates for health and fitness apps range from 40%<sup>7</sup> to 75%<sup>8</sup>. Interestingly, of those users who do use the app beyond the first day, more than half continue to use it for at least 30 days.

Hence we acknowledge that our findings do not generalize to those users who had quit using MFP sooner than 100 days, or those who did not have Twitter data matching the period of MFP data. Such missing data is a known limitation of traditional supervised learning, and can be addressed by employing methods like the survival analysis that take into account “censoring” of missing data. Nevertheless, our approach and findings provide insights into the ability to predict these individual-level trends of diet compliance by combining quantified self sensing data with social media data.

**Peer effects.** Further, we do not analyze social influence effects on general health and well-being, or more specifically on diet compliance. Prior work indicates a significant link between one’s own health and friends with similar attributes [46]. We believe many of these effects are captured via one’s posting behavior and social network structural attributes we considered here. However it is possible that there are latent peer effect variables, e.g., feedback from specific contacts, or the accrued social status in one’s network, that were not accounted for in our models. An interesting future direction could examine these influence effects more thoroughly, perhaps by collecting data for clustered communities of individuals who use social media and who are attempting to succeed in their dietary goals over time.

**Causality.** We suggest caution in inferring causality from our findings. Although Granger causality is widely adopted in econometrics to obtain causal links between two different

<sup>7</sup><http://flurrymobile.tumblr.com/post/144245637325/appmatrix>

<sup>8</sup>[https://www.similarweb.com/corp/wp-content/uploads/2016/04/086-6\\_Uninstalls-Report.pdf](https://www.similarweb.com/corp/wp-content/uploads/2016/04/086-6_Uninstalls-Report.pdf)

time series distributions, we refrain from generalized conclusions regarding causality in our specific setting of forecasting an individual's future diet compliance success or failure. While it is possible that experiences and events on social media can trigger one's success or failure, there could be many latent factors, not observable via either MFP or Twitter data, that might influence these outcomes. Future work could expand the types of data that could be linked to diet compliance as a way to account for as many potential ways to capture these latent factors.

**Intent.** Finally, while we do demonstrate that valuable signals may be gleaned from social media to forecast diet compliance success or failure, as with any quantitative method, our findings do not explain *why* these signals are useful. Neither do they provide insights into the motivations and intent of individuals who choose to share diet and health related information on a public social media platform or *why* they had chosen the specific dietary goal, though some prior qualitative work provides helpful insights [51]. These are questions that are ripe to be investigated in the future through mixed methods approaches.

## CONCLUSION

Previous work had shown the promise that both quantified self tools and social media hold for modeling an individual's health. However, quantified self data and social media data have largely been studied in isolation. In this paper, we demonstrated the relationship that exists between the two, using public data from the MyFitnessPal application and Twitter. We used MyFitnessPal food diaries to evaluate an individual's adherence to their self-defined calorie goals. We then linked this information to a number of linguistic, activity, and social capital features derived from their Twitter feed. In doing so, we showed that (i) there are behavioral measures derived from social media that are predictive of diet compliance success of individuals (RQ 1), and that (ii) social media derived measures can be integrated with sensed historical data about diet to better predict future diet compliance (RQ 2). Looking into the factors linked to diet compliance success, we observed that more successful users: (i) express higher positive affect and a greater sense of achievement, (ii) exhibit high cognitive functioning and perception and a stronger future orientation, and (iii) are generally more social, both in terms of topics discussed, interaction patterns and in terms of the size of their social network.

We hope that our work contributes to achieving a more holistic view of an individual's health, including general behavioral and lifestyle information from social media and self-logged physiological data, going beyond what is recorded in traditional health records.

## ACKNOWLEDGMENTS

De Choudhury was partly supported through NIH grant #1R01GM11269701.

## REFERENCES

1. Sofiane Abbar, Yelena Mejova, and Ingmar Weber. 2015. You Tweet What You Eat: Studying Food

Consumption Through Twitter. In *Conference on Human Factors in Computing Systems (CHI)*. 3197–3206.

2. Phil Adams, Mashfiqui Rabbi, Tauhidur Rahman, Mark Matthews, Amy Volda, Geri Gay, Tanzeem Choudhury, and Stephen Volda. 2014. Towards personal stress informatics: comparing minimally invasive techniques for measuring daily stress in the wild. In *PervasiveHealth*. 72–79.
3. Fatema Akbar and Ingmar Weber. 2016. #Sleep\_as\_Android: Feasibility of Using Sleep Logs on Twitter for Sleep Studies. In *IEEE ICHI*. <http://arxiv.org/abs/1607.06359>.
4. M Al'Absi and DK Arnett. 2000. Adrenocortical responses to psychological stress and risk for hypertension. *Biomedicine & pharmacotherapy* 54, 5 (2000), 234–244.
5. Euan A Ashley. 2015. The precision medicine initiative: a new national effort. *JAMA* 313, 21 (2015), 2119–2120.
6. Søren Brage, Niels Brage, PW Franks, U Ekelund, and NJ Wareham. 2005. Reliability and validity of the combined heart rate and movement sensor Actiheart. *European journal of clinical nutrition* 59, 4 (2005), 561–570.
7. Stevie Chancellor, Zhiyuan (Jerry) Lin, Erica Goodman, Stephanie Zerwas, and Munmun De Choudhury. 2016. Quantifying and Predicting Mental Illness Severity in Online Pro-Eating Disorder Communities. In *CSCW*. 1171–1184.
8. Fanglin Chen, Rui Wang, Xia Zhou, and Andrew T Campbell. 2014. My smartphone knows i am hungry. In *Proceedings of the 2014 workshop on physical analytics*. 9–14.
9. Cindy Chung and James W Pennebaker. 2007. The psychological functions of function words. *Social communication* (2007), 343–359.
10. 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 *UbiComp*. 647–658.
11. Francis S Collins and Harold Varmus. 2015. A new initiative on precision medicine. *New England Journal of Medicine* 372, 9 (2015), 793–795.
12. Sunny Consolvo, Katherine Everitt, Ian Smith, and James A Landay. 2006. Design requirements for technologies that encourage physical activity. In *CHI*. 457–466.
13. Sunny Consolvo, Predrag Klasnja, David W McDonald, and James A Landay. 2009. Goal-setting considerations for persuasive technologies that encourage physical activity. In *Proceedings of the 4th international Conference on Persuasive Technology*. ACM, 8.

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 *CHI*. 1159–1162.
15. Aron Culotta. 2014. Estimating county health statistics with Twitter. In *CHI*. 1335–1344.
16. Munmun De Choudhury, Scott Counts, Eric Horvitz, and Aaron Hoff. 2014. Characterizing and Predicting Postpartum Depression from Facebook Data. In *CSCW*. 626–638.
17. Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. In *ICWSM*.
18. Munmun De Choudhury, Emre Kiciman, Mark Dredze, Glen Coppersmith, and Mrinal Kumar. 2016. Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media. In *CHI*. 2098–2110.
19. David A Dickey and Wayne A Fuller. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society* (1981), 1057–1072.
20. Deborah Estrin. 2014. Small data, where n= me. *Commun. ACM* 57, 4 (2014), 32–34.
21. Gunther Eysenbach, John Powell, Marina Englesakis, Carlos Rizo, and Anita Stern. 2004. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *Bmj* 328, 7449 (2004), 1166.
22. Nicholas Finer. 2015. Predicting therapeutic weight loss. *American Journal of Clinical Nutrition* 101, 3 (2015), 419–420.
23. Venkata Rama Kiran Garimella, Abdulrahman Alfayad, and Ingmar Weber. Social Media Image Analysis for Public Health. In *Conference on Human Factors in Computing Systems (CHI)*. 5543–5547.
24. John F Geweke. 1984. Measures of conditional linear dependence and feedback between time series. *J. Amer. Statist. Assoc.* 79, 388 (1984), 907–915.
25. Rainer Goebel, Alard Roebroek, Dae-Shik Kim, and Elia Formisano. 2003. Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magnetic resonance imaging* 21, 10 (2003), 1251–1261.
26. Hamed Haddadi, Ferda Offi, Yelena Mejova, Ingmar Weber, and Jaideep Srivastava. 2015. 360-degree Quantified Self. In *International Conference on Healthcare Informatics (ICHI)*. IEEE, 587–592.
27. Jina Huh and Mark S Ackerman. 2012. Collaborative Help in Chronic Disease Management: Supporting Individualized Problems. In *CSCW*.
28. Anton Ivanov, Raj Sharman, and H. Raghav Rao. 2015. Exploring factors impacting sharing health-tracking records. *Health Policy and Technology* 4 (2015), 263–276. Issue 3.
29. Grace J Johnson and Paul J Ambrose. 2006. Neo-tribes: The power and potential of online communities in health care. *Commun. ACM* 49, 1 (2006), 107–113.
30. Rafal Kocielnik, Natalia Sidorova, Fabrizio M Maggi, Martin Ouwerkerk, and Joyce HDM Westerink. 2013. Smart technologies for long-term stress monitoring at work. In *Computer-Based Medical Systems (CBMS)*. 53–58.
31. Patty Kostkova. 2015. Public Health. In *Twitter: A Digital Socioscope*, Yelena Mejova, Ingmar Weber, and Michael Macy (Eds.). Cambridge University Press, 111–130.
32. Hong Lu, Denise Frauendorfer, Mashfiqui Rabbi, Marianne Schmid Mast, Gokul T Chittaranjan, Andrew T Campbell, Daniel Gatica-Perez, and Tanzeem Choudhury. 2012. Stresssense: Detecting stress in unconstrained acoustic environments using smartphones. In *Conference on Ubiquitous Computing (UbiComp)*. 351–360.
33. Lena Mamykina, Drashko Nakikj, and Noemie Elhadad. 2015. Collective Sensemaking in Online Health Forums. In *CHI*. 3217–3226.
34. Jennifer Mankoff, Kateryna Kuksenok, Sara Kiesler, Jennifer A Rode, and Kelly Waldman. 2011. Competing online viewpoints and models of chronic illness. In *CHI*. 589–598.
35. Yelena Mejova, Sofiane Abbar, and Hamed Haddadi. 2016. Fetishizing Food in Digital Age: #foodporn Around the World. In *ICWSM*.
36. Jochen Meyer, Steven Simske, Katie A. Siek, Cathal G. Gurrin, and Hermie Hermens. 2014. Beyond Quantified Self: Data for Wellbeing. In *CHI*. 95–98.
37. Sean A Munson and Sunny Consolvo. 2012. Exploring goal-setting, rewards, self-monitoring, and sharing to motivate physical activity. In *Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*. 25–32.
38. Elizabeth L Murnane and Scott Counts. 2014. Unraveling Abstinence and Relapse: Smoking Cessation Reflected in Social Media. In *CHI*. 1345–1354.
39. Elizabeth L Murnane, David Huffaker, and Gueorgi Kossinets. 2015. Mobile health apps: adoption, adherence, and abandonment. In *Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*. 261–264.
40. Mark W Newman, Debra Lauterbach, Sean A Munson, Paul Resnick, and Margaret E Morris. 2011. It's not that I don't have problems, I'm just not putting them on Facebook: challenges and opportunities in using online social networks for health. In *CSCW*. 341–350.



41. Pieter W Otter. 1991. On Wiener-Granger causality, information and canonical correlation. *Economics Letters* 35, 2 (1991), 187–191.
42. Kevin A Padrez, Lyle Ungar, Hansen Andrew Schwartz, Robert J Smith, Shawndra Hill, Tadas Antanavicius, Dana M Brown, Patrick Crutchley, David A Asch, and Raina M Merchant. 2015. Linking social media and medical record data: a study of adults presenting to an academic, urban emergency department. *BMJ quality & safety* (2015), bmjqs–2015.
43. Kunwoo Park, Ingmar Weber, Meeyoung Cha, and Chul Lee. 2016. Persistent sharing of fitness app status on twitter. In *CSCW*.
44. Michael J Paul and Mark Dredze. 2011. You Are What You Tweet: Analyzing Twitter for Public Health.. In *ICWSM*.
45. Patrick Ritz, Robert Caiazzo, Guillaume Becouarn, Laurent Arnalsteen, Sandrine Andrieu, Philippe Topart, and Franois Pattou. 2013. Early prediction of failure to lose weight after obesity surgery. *Surgery for Obesity and Related Diseases* 9 (2013), 118–121. Issue 1.
46. James F Sallis, Neville Owen, and Edwin B Fisher. 2008. Ecological models of health behavior. *Health behavior and health education: Theory, research, and practice* 4 (2008), 465–486.
47. Ralf Schwarzer. 2008. Modeling health behavior change: How to predict and modify the adoption and maintenance of health behaviors. *Applied Psychology* 57, 1 (2008), 1–29.
48. Samuel S Shapiro, Martin B Wilk, and Hwei J Chen. 1968. A comparative study of various tests for normality. *J. Amer. Statist. Assoc.* 63, 324 (1968), 1343–1372.
49. Victor J Strecher, Brenda McEvoy DeVellis, Marshall H Becker, and Irwin M Rosenstock. 1986. The role of self-efficacy in achieving health behavior change. *Health Education & Behavior* 13, 1 (1986), 73–92.
50. Victor J Strecher, Gerard H Seijts, Gerjo J Kok, Gary P Latham, Russell Glasgow, Brenda DeVellis, Ree M Meertens, and David W Bulger. 1995. Goal setting as a strategy for health behavior change. *Health Education & Behavior* 22, 2 (1995), 190–200.
51. Rannie Teodoro and Mor Naaman. 2013. Fitter with Twitter: Understanding Personal Health and Fitness Activity in Social Media. In *ICWSM*. 611–620.
52. Diana M Thomas, Andrada E Ivanescu, Corby K Martin, Steven B Heymsfield, Kaitlyn Marshall, Victoria E Bodrato, Donald A Williamson, Stephen D Anton, Frank M Sacks, Donna Ryan, and George A Bray. 2015. Predicting successful long-term weight loss from short-term weight-loss outcomes: new insights from a dynamic energy balance model (the POUNDS Lost study). *American Journal of Clinical Nutrition* 101, 3 (2015), 449–454.
53. Theodore A Vickey and John G Breslin. 2012. A Study on Twitter Usage for Fitness Self-Reporting via Mobile Apps. In *AAAI Spring Symposium: Self-Tracking and Collective Intelligence for Personal Wellness*. 65–70.
54. Claudia Wagner, Philipp Singer, and Markus Strohmaier. 2014. Spatial and Temporal Patterns of Online Food Preferences. In *World Wide Web Conference (WWW)*. 553–554.
55. Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. 2014. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*. 3–14.
56. Yafei Wang, Ingmar Weber, and Prasenjit Mitra. 2016. Quantified Self Meets Social Media: Sharing of Weight Updates on Twitter. In *ACM Digital Health (DH)*. 93–97.
57. Ingmar Weber and Palakorn Achananuparp. 2016. Insights from machine-learned diet success prediction. In *Pacific Symposium on Biocomputing (PSB)*. 540–551.
58. Ingmar Weber and Yelena Mejova. 2016. Crowdsourcing Health Labels: Inferring Body Weight from Profile Pictures. In *ACM Digital Health (DH)*. 105–109.
59. Jennifer R Whitson. 2013. Gaming the quantified self. *Surveillance & Society* 11, 1/2 (2013), 163.
60. Xu Ye, Guanling Chen, Yang Gao, Honghao Wang, and Yu Cao. 2016. Assisting Food Journaling with Automatic Eating Detection. In *CHI*. 3255–3262.

|                                 | BASELINE  |           |     | CALDIFF   |           |     | CALDIFF:25 |           |     | CALDIFF:50 |           |     | PROP_DAYS:50 |           |     | PROP_DAYS:75 |           |     |
|---------------------------------|-----------|-----------|-----|-----------|-----------|-----|------------|-----------|-----|------------|-----------|-----|--------------|-----------|-----|--------------|-----------|-----|
|                                 | $\lambda$ | $F$ -stat | $p$ | $\lambda$ | $F$ -stat | $p$ | $\lambda$  | $F$ -stat | $p$ | $\lambda$  | $F$ -stat | $p$ | $\lambda$    | $F$ -stat | $p$ | $\lambda$    | $F$ -stat | $p$ |
| <b>Diet compliance measures</b> |           |           |     |           |           |     |            |           |     |            |           |     |              |           |     |              |           |     |
| past success                    | 0.49      | 4.44      | *** | 0.27      | 4.58      | *** | 0.32       | 5.92      | *** | 0.33       | 4.30      | *** | 0.33         | 4.15      | *** | 0.26         | 3.72      | *** |
| calorie consumption             | 0.24      | 5.64      | *** | 0.51      | 3.87      | *** | 0.41       | 4.94      | *** | 0.23       | 6.18      | *** | 0.58         | 4.45      | *** | 0.44         | 5.60      | *** |
| <b>Affective attributes</b>     |           |           |     |           |           |     |            |           |     |            |           |     |              |           |     |              |           |     |
| PA                              | 0.05      | 5.65      | *** | 0.51      | 5         | *** | 0.33       | 6.27      | *** | 0.78       | 3.87      | **  | 0.51         | 5.88      | *** | 0.2          | 5.13      | *** |
| NA                              | 0.65      | 3.01      | **  | 0.16      | 3.86      | **  | 0.1        | 4.56      | *** | 0.31       | 3.39      | **  | 0.41         | 6.82      | *** | 0.25         | 4.52      | *** |
| anger                           | 0.58      | 5.69      | *** | 0.24      | 3.79      | **  | 0.13       | 5.08      | *** | 0.77       | 4.99      | *** | 0.16         | 3.08      | **  | 0.66         | 6.4       | *** |
| anxiety                         | 0.06      | 5.31      | *** | 0.23      | 5.87      | *** | 0.06       | 3.36      | **  | 0.77       | 6.49      | *** | 0.54         | 5.36      | *** | 0.5          | 3.39      | **  |
| sadness                         | 0.07      | 5.02      | *** | 0.79      | 6.14      | *** | 0.64       | 6.65      | *** | 0.29       | 5.78      | *** | 0.36         | 6.34      | *** | 0.51         | 3.16      | **  |
| swear                           | 0.75      | 0.73      | -   | 0.75      | 2.01      | **  | 0.9        | 1.57      | *   | 0.98       | 1.7       | *   | 0.79         | 1.88      | *   | 0.91         | 0.66      | -   |
| <b>Cognitive attributes</b>     |           |           |     |           |           |     |            |           |     |            |           |     |              |           |     |              |           |     |
| <b>Cognition</b>                |           |           |     |           |           |     |            |           |     |            |           |     |              |           |     |              |           |     |
| cognitive_mech                  | 0.73      | 4.25      | *** | 0.08      | 5.24      | *** | 0.44       | 4.27      | *** | 0.52       | 3.23      | **  | 0.21         | 6.24      | *** | 0.77         | 6.39      | *** |
| discrepancies                   | 0.84      | 2.27      | **  | 0.74      | 0.8       | -   | 0.76       | 0.56      | -   | 0.89       | 1.99      | *   | 0.87         | 1.1       | -   | 0.82         | 2.88      | **  |
| negation                        | 0.92      | 0.23      | -   | 0.84      | 0.69      | -   | 0.8        | 0.9       | -   | 0.91       | 0.5       | -   | 0.91         | 1.76      | *   | 0.71         | 0.92      | -   |
| inhibition                      | 0.44      | 4.96      | *** | 0.48      | 5.9       | *** | 0.52       | 4.27      | *** | 0.78       | 5.15      | *** | 0.31         | 6.96      | *** | 0.32         | 5.39      | *** |
| causation                       | 0.86      | 0.97      | -   | 0.76      | 2.99      | **  | 0.93       | 0.13      | -   | 0.75       | 1.6       | *   | 0.83         | 0.65      | -   | 0.77         | 1.35      | *   |
| certainty                       | 0.55      | 6.52      | *** | 0.36      | 5.87      | *** | 0.4        | 4.71      | *** | 0.14       | 4.4       | *** | 0.14         | 5.08      | *** | 0.8          | 3.52      | **  |
| tentativeness                   | 0.81      | 2.15      | **  | 0.97      | 1.21      | -   | 0.84       | 0.62      | -   | 0.89       | 0.36      | -   | 0.82         | 0.1       | -   | 0.72         | 1.2       | *   |
| <b>Perception</b>               |           |           |     |           |           |     |            |           |     |            |           |     |              |           |     |              |           |     |
| see                             | 0.32      | 5.5       | *** | 0.1       | 3.51      | **  | 0.79       | 3.08      | **  | 0.76       | 4.11      | *** | 0.68         | 6.59      | *** | 0.43         | 6.96      | *** |
| hear                            | 0.15      | 5.41      | *** | 0.06      | 3.62      | **  | 0.36       | 5.65      | *** | 0.59       | 6.91      | *** | 0.06         | 5.13      | *** | 0.25         | 6.94      | *** |
| feel                            | 0.09      | 4.86      | *** | 0.33      | 6.69      | *** | 0.64       | 5.53      | *** | 0.59       | 4.67      | *** | 0.08         | 6.27      | *** | 0.27         | 4         | **  |
| percept                         | 0.33      | 5.4       | *** | 0.05      | 5.19      | *** | 0.38       | 5.78      | *** | 0.18       | 3.54      | **  | 0.55         | 6.79      | *** | 0.18         | 3.12      | **  |
| insight                         | 0.74      | 2.12      | **  | 0.89      | 1.9       | *   | 0.77       | 0.09      | -   | 0.89       | 1.37      | *   | 0.92         | 2.44      | **  | 0.77         | 2.11      | **  |
| relative                        | 0.72      | 1.55      | *   | 0.87      | 1.39      | *   | 0.9        | 2.46      | **  | 0.76       | 1.75      | *   | 0.99         | 0.87      | -   | 0.7          | 2.83      | **  |
| <b>Lexical Density</b>          |           |           |     |           |           |     |            |           |     |            |           |     |              |           |     |              |           |     |
| verbs                           | 0.24      | 6.63      | *** | 0.35      | 6.63      | *** | 0.3        | 6.44      | *** | 0.26       | 3.15      | **  | 0.71         | 6.83      | *** | 0.02         | 5.23      | *** |
| auxiliary verbs                 | 0.26      | 3.29      | **  | 0.6       | 3.83      | **  | 0.6        | 5.72      | *** | 0.07       | 5.46      | *** | 0.12         | 3.58      | **  | 0.19         | 6.01      | *** |
| adverbs                         | 0.81      | 2.84      | **  | 0.89      | 2.55      | **  | 0.83       | 1.47      | *   | 0.86       | 2.31      | **  | 0.89         | 2.83      | **  | 0.78         | 1.7       | *   |
| prepositions                    | 0.96      | 2.54      | **  | 0.79      | 2.71      | **  | 0.91       | 0.61      | -   | 0.98       | 1.56      | *   | 0.78         | 2.85      | **  | 0.86         | 1.06      | -   |
| conjunctions                    | 0.99      | 0.85      | -   | 0.81      | 0.36      | -   | 0.83       | 1.45      | *   | 0.74       | 1.88      | *   | 0.98         | 1         | -   | 0.82         | 1.61      | *   |
| articles                        | 0.72      | 4.84      | *** | 0.52      | 6.6       | *** | 0.32       | 4.27      | *** | 0.8        | 5.62      | *** | 0.42         | 4.37      | *** | 0.7          | 5.35      | *** |
| inclusive                       | 0.91      | 0.57      | -   | 0.94      | 2.9       | **  | 0.91       | 2.2       | **  | 0.95       | 2.1       | **  | 0.73         | 1.61      | *   | 0.95         | 2.54      | **  |
| exclusive                       | 0.79      | 2.37      | **  | 0.75      | 0.52      | -   | 0.79       | 2.24      | **  | 0.78       | 2.68      | **  | 0.85         | 0.5       | -   | 0.87         | 2.82      | **  |
| <b>Temporal references</b>      |           |           |     |           |           |     |            |           |     |            |           |     |              |           |     |              |           |     |
| past tense                      | 0.09      | 3.5       | **  | 0.01      | 5.58      | *** | 0.26       | 4.89      | *** | 0.02       | 3.2       | **  | 0.73         | 3.43      | **  | 0.25         | 3.29      | **  |
| present tense                   | 0.2       | 4.69      | *** | 0.48      | 6.24      | *** | 0.64       | 5.13      | *** | 0.58       | 3.64      | **  | 0.74         | 5.14      | *** | 0.07         | 6.11      | *** |
| future tense                    | 0.68      | 5.1       | *** | 0.74      | 6.41      | *** | 0.31       | 5.8       | *** | 0.39       | 3.32      | **  | 0.46         | 5.9       | *** | 0.09         | 6.36      | *** |
| <b>Social/Personal Concerns</b> |           |           |     |           |           |     |            |           |     |            |           |     |              |           |     |              |           |     |
| family                          | 0.68      | 6.48      | *** | 0.69      | 5.27      | *** | 0.15       | 6.97      | *** | 0.27       | 5.66      | *** | 0.68         | 6.55      | *** | 0.43         | 6.54      | *** |
| friends                         | 0.71      | 6.73      | *** | 0.23      | 3.19      | **  | 0.13       | 5.21      | *** | 0.6        | 5.04      | *** | 0.63         | 4.39      | *** | 0.55         | 5.33      | *** |
| social                          | 0.56      | 6.86      | *** | 0.37      | 4.22      | **  | 0.5        | 6.55      | *** | 0.75       | 3.46      | **  | 0.32         | 4.42      | *** | 0.35         | 5.14      | *** |
| work                            | 0.83      | 1.28      | *   | 1         | 0.6       | -   | 0.92       | 1.42      | *   | 0.88       | 2.02      | **  | 0.74         | 2.15      | **  | 0.95         | 1.52      | *   |
| health                          | 0.79      | 3.19      | **  | 0.55      | 6.23      | *** | 0.67       | 6.18      | *** | 0.64       | 6.91      | *** | 0.53         | 5.98      | *** | 0.04         | 3.53      | **  |
| humans                          | 0.91      | 1.85      | *   | 0.8       | 2.62      | **  | 0.88       | 0.01      | -   | 0.94       | 1.27      | *   | 0.78         | 2.68      | **  | 0.93         | 2.7       | **  |
| religion                        | 0.92      | 2.92      | **  | 0.97      | 1.97      | *   | 0.76       | 2.06      | **  | 0.92       | 2.68      | **  | 0.98         | 2.38      | **  | 0.91         | 0.34      | -   |
| bio                             | 0.86      | 1.69      | *   | 0.75      | 0.33      | -   | 0.98       | 1.9       | *   | 0.74       | 2.83      | **  | 0.73         | 1.44      | *   | 0.77         | 2.76      | **  |
| body                            | 0.04      | 5.84      | *** | 0.43      | 3.21      | **  | 0.06       | 3.68      | -   | 0.34       | 4.02      | *** | 0.28         | 4.37      | *** | 0.07         | 3.98      | **  |
| money                           | 0.94      | 1.8       | *   | 0.86      | 0.31      | -   | 0.74       | 0.61      | -   | 0.85       | 0.57      | -   | 0.95         | 0.63      | -   | 0.82         | 0.17      | -   |
| achievement                     | 0.25      | 3.45      | **  | 0.07      | 4.66      | *** | 0.41       | 6.64      | *** | 0.75       | 5.68      | *** | 0.6          | 4.14      | *** | 0.12         | 3.82      | **  |
| home                            | 0.83      | 2.83      | **  | 0.78      | 0.89      | -   | 0.82       | 0.08      | -   | 0.73       | 0.01      | -   | 0.78         | 0.79      | -   | 0.71         | 0.88      | -   |
| sexual                          | 0.99      | 1.01      | -   | 0.84      | 0.65      | -   | 0.71       | 2.81      | **  | 0.85       | 1.94      | *   | 0.86         | 0.54      | -   | 0.82         | 1.28      | *   |
| death                           | 0.74      | 2.88      | **  | 0.7       | 0.57      | -   | 0.77       | 1.82      | *   | 0.77       | 1.97      | *   | 0.71         | 1.63      | -   | 0.82         | 1.07      | -   |
| <b>Interpersonal awareness</b>  |           |           |     |           |           |     |            |           |     |            |           |     |              |           |     |              |           |     |
| 1st p. singular                 | 0.48      | 5.79      | *** | 0.33      | 3.04      | **  | 0.16       | 6.63      | *** | 0.75       | 3.84      | **  | 0.56         | 5.63      | *** | 0.27         | 6.36      | *** |
| 1st p. plural                   | 0.12      | 3.69      | **  | 0.25      | 4.76      | *** | 0.7        | 4.72      | *** | 0.64       | 4.14      | *** | 0.25         | 5.38      | *** | 0.7          | 3.19      | **  |
| 2nd p.                          | 0.69      | 4.06      | *** | 0.48      | 4.28      | *** | 0.63       | 5.58      | *** | 0.35       | 6.14      | *** | 0.39         | 6.98      | *** | 0.07         | 6.06      | *** |
| 3rd p.                          | 0.9       | 2.57      | **  | 0.71      | 0.88      | -   | 0.77       | 0.46      | -   | 0.94       | 2.83      | **  | 0.81         | 2.5       | **  | 0.96         | 2.2       | **  |
| <b>Activity</b>                 |           |           |     |           |           |     |            |           |     |            |           |     |              |           |     |              |           |     |
| Interactivity                   | 0.67      | 6.05      | *** | 0.53      | 6.9       | *** | 0.73       | 3.89      | **  | 0.03       | 3.03      | **  | 0.13         | 5.89      | *** | 0.05         | 4.67      | *** |
| Information sharing             | 0.33      | 6.44      | *** | 0.41      | 5.38      | *** | 0.59       | 5.28      | *** | 0.53       | 5.73      | *** | 0.22         | 6.28      | *** | 0.34         | 5.05      | *** |
| <b>Social capital</b>           |           |           |     |           |           |     |            |           |     |            |           |     |              |           |     |              |           |     |
| # inlinks                       | 0.8       | 4.4       | *** | 0.03      | 3.3       | **  | 0.55       | 5.43      | *** | 0.04       | 3.85      | **  | 0.61         | 6.02      | *** | 0.6          | 4.74      | *** |
| # outlinks                      | 0.04      | 5.74      | *** | 0.5       | 4.69      | *** | 0.61       | 6.51      | *** | 0.55       | 3.04      | **  | 0.56         | 5.56      | *** | 0.11         | 6.77      | *** |

Table 4: Results of multivariate Granger causality tests for different MFP and Twitter measures in forecasting future diet compliance success/failure. For each measure, we report the Wilks'  $\lambda$  (likelihood ratio) statistic, Rao's approximate  $F$  statistic, and the  $p$  value (-:  $p = .05$ ; \*:  $p = .01$ ; \*\*:  $p = .001$ ; and \*\*\*:  $p = .0001$ ). Results are shown for all six diet compliance metrics. Grey rows indicate measures for which none of the diet compliance predictions yielded Granger significance at  $p = .0001$  level.