# What 30 Days Tells Us About 3 Years: Identifying Early Signs of User Abandonment and Non-Adherence

**Louis Faust**
University of Notre Dame
lfaust@nd.edu

**Priscilla Jiménez-Pazmino**
University of Notre Dame
pjimenez@nd.edu

**James K. Holland**
University of Notre Dame
James.K.Holland.45@nd.edu

**Omar Lizardo**
University of California, Los Angeles
olizardo@soc.ucla.edu

**David Hachen**
University of Notre Dame
dhachen@nd.edu

**Nitesh V. Chawla**
University of Notre Dame
nchawla@nd.edu

## ABSTRACT

The ubiquity of wearable fitness trackers offers extensive opportunities for research on personal health. However, barriers specific to these trackers such as device abandonment and non-adherence often lead to substantial losses in data. As such, further research into adherence behaviors may derive the insights necessary to address these challenges and lead to more effective long-term studies. This paper serves to explore this approach: investigating the adherence behaviors of 617 college students belonging to a three-year observational study in which participants were monitored via Fitbit Charge HRs. Using this data, our objective was to assess the association between early adherence behaviors and device abandonment/long-term adherence. Adherence behavior from as early as participants' first 10 days in the study correlated with device abandonment and adherence over the next three years. Participants with unsatisfactory adherence in their first 30 days were twice as likely to abandon their devices and were, on average, 11% less adherent each month. The findings in this paper identify the stability of adherence behaviors, feasibility of their early detection and motivate the need to address non-adherent study participants early. Throughout these results, we discuss how the insights gathered from this work may shape the design of future long-term studies to minimize user attrition and promote prolonged engagement.


*PervasiveHealth'19, May 20–23, 2019, Trento, Italy*

ACM ISBN 978-1-4503-6126-2/19/05…$15.00
https://doi.org/10.1145/3329189.3329196

## CCS CONCEPTS

• **Human-centered computing → User studies**; *Empirical studies in ubiquitous and mobile computing*;

## KEYWORDS

User abandonment, adherence, wearables, fitness trackers, Fitbit, long-term study

**ACM Reference format:**
Louis Faust, Priscilla Jiménez-Pazmino, James K. Holland, Omar Lizardo, David Hachen, and Nitesh V. Chawla. 2019. What 30 Days Tells Us About 3 Years: Identifying Early Signs of User Abandonment and Non-Adherence. In *Proceedings of The 13th International Conference on Pervasive Computing Technologies for Healthcare, Trento, Italy, May 20–23, 2019 (PervasiveHealth'19),* 9 pages.
https://doi.org/10.1145/3329189.3329196

## 1 INTRODUCTION

As wearable technology continues to evolve, the ubiquity and unobtrusive nature of personal tracking devices serve as invaluable tools for the fields of health and wellness [7, 20]. Providing continuous physiological and behavioral measurements in real-time and at scale, these devices can increase the longevity and enhance the granularity of data collected through observational and interventional studies. The wealth of opportunities these studies provide are of fundamental importance to advancing our understanding of personal health and well-being.

Accompanied by these studies' benefits, however, are significant challenges that must be addressed; among them are user attrition and non-adherence. As adherence on behalf of participants often diminishes throughout studies, user abandonment leads to costly losses in data, which can reduce sample sizes and introduce bias [3, 11, 17, 22]. As personal health tracking studies lengthen from months to years to decades, investigating approaches as to how these problems may be mitigated is critical to ensuring the gathered data is complete.

Therefore, the goal of this work was to derive insights from user adherence behaviors which may be exploited to

improve study designs. Specifically, we examined the stability of adherence behaviors and how early data may serve to inform us of long-term behavior. We evaluate this stability through two research questions:

**RQ1**: Are study participants with unsatisfactory adherence in their first 30 days of a study more likely to drop out than those with satisfactory adherence?

**RQ2**: Will study participants with unsatisfactory adherence in their first 30 days of a study be less compliant throughout a study than those with satisfactory adherence?

These questions were assessed through a data set featuring 617 users' daily fitness tracker usage as part of a three year observational health study. To the best of our knowledge, this has been the longest continuous study monitoring personal health through wearable fitness trackers, making it the ideal data set for understanding long-term adherence behaviors. In addition to early adherence behaviors, we assess the big-five personality traits, given prior work has observed associations between these traits and adherence [1, 6, 12, 13]. As such, we adjust for and further examine this relationship. Finally, we reevaluate our research questions, redefining early adherence by participants first *10* days to assess how a shorter interval affects likelihood of dropout and long-term adherence.

Our contribution to the field of ubiquitous computing in health care is the insight into the use and abandonment of wearable fitness trackers. In particular, the identification of adherence behaviors as a stable trait and the viability of their early detection. By leveraging these findings, practitioners and researchers can make modifications to study designs to address non-compliant participants early in a study, minimizing user attrition and non-adherence, which we discuss in greater detail toward the end of this manuscript.

## 2 RELATED WORKS

Wearable fitness trackers have provided a unique opportunity for researchers to gather information on personal health. However, a large proportion of studies have reported poor adherence. A long-term observational study of 711 participants monitoring adherence found a "slow exponential decay" of use: reporting 73.9% of participants still engaged after 100 days and only 16% after 320 days [11]. Even in shorter studies, dropout rates have been reported as high as 75% within the first two months [3, 8, 17, 22]. With only a handful of participants remaining engaged in these long-term studies until the end, insights gained from the data become limited and difficult to generalize.

Beyond quantitative measures, researchers have investigated *why* users abandon their activity trackers. Commonly reported reasons for discontinued use have included technical device issues, losing the device and forgetfulness [4, 9, 14, 15, 22]. Aspects of a device's design have also been attributed to abandonment, including poor aesthetics and the device being uncomfortable or ill-fitting [3, 10, 15]. Functionality was found to play a role as well with concerns of data inaccuracies, limited insights, and the failure of devices to support evolving personal goals [2, 3, 11].

Patterns of usage was also a common theme among usability studies, finding that users who remained in a study long-term typically held to a certain type of pattern [5, 17, 18]. An analysis of these patterns resulted in five common types of user: *try-and-drop*, *slow-starter*, *experimenter*, *intermittent user*, and *power user*. The consistency of these usage patterns suggested they may be exploited by different types of interventions to boost adherence.

Such findings emphasize the importance of gathering insights about participant's adherence behaviors early in a study. To the best of our knowledge, no prior research has investigated whether early adherence behaviors can predict likelihood of abandonment and long-term adherence.

## 3 CONTEXT - LONGITUDINAL STUDY

To provide the reader with a background on the study from which our analyses were conducted, the following section outlines the study's objectives, recruitment practices and all interventions conducted on the study population.

### Objectives

The NetHealth study conducted at the University of Notre Dame features an ongoing collection of demographic, psychometric, social network and physical activity data with the intent of modeling the co-evolution of health behaviors and social networks [21]. To monitor health behaviors, participants were issued a Fitbit Charge HR. To ensure completeness in the data collected, participants were asked to wear the device as much as possible and sync their device every four to seven days.

### Recruitment of the participants

Participants included 698 individuals who entered the university as first-year students over the course of the 2015-2016 academic year. Participants' ages ranged from 17 to 19 years. This initial cohort of participants was split across three tiers based on when they entered the study (Figure 1). Three hundred and ninety-one tier 1 students were recruited via interest surveys, e-mail, and a Facebook page in June 2015. Selection was on a first come, first served basis in keeping with the overall demographic distributions of the university. Ninety-seven tier 2 students were then recruited in November and December 2015, nominated by existing participants in the study. Finally, 210 tier 3 students were recruited via email in April 2016.

Across the 698 enrolled participants, 65 participants were removed from consideration as they were not issued Fitbits,

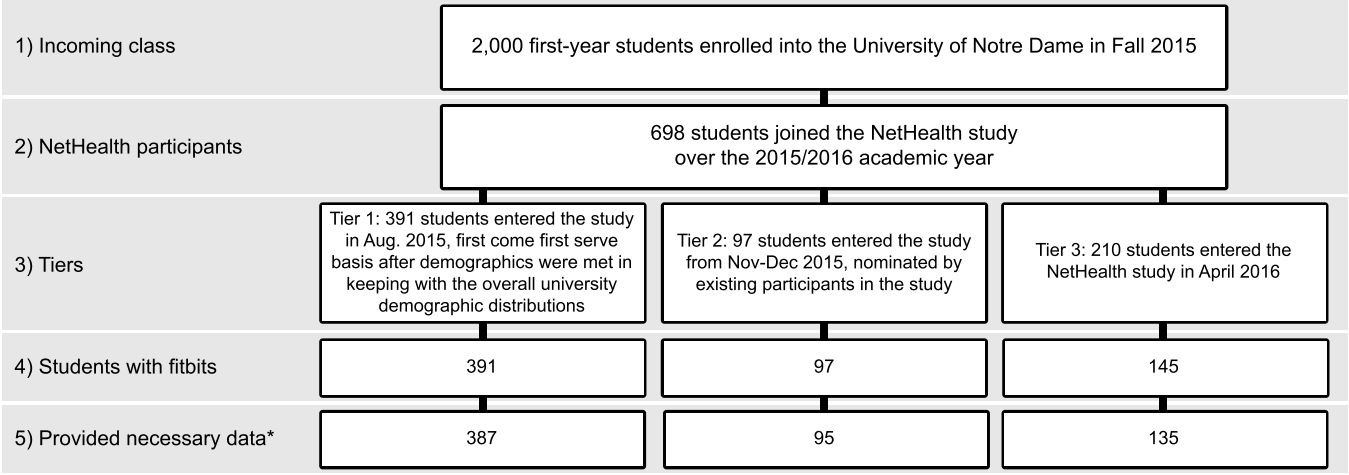


**Figure 1: Consort diagram of NetHealth recruitment.**

**Table 1: Demographic overview of study cohort.**

| | n (%) | | n (%) |
|---|---|---|---|
| **Sex** | 617 (100) | **Race** | 617 (100) |
| Male | 303 (49) | Foreign* | 39 (6) |
| Female | 314 (51) | Latino | 78 (12) |
| | | Black | 37 (6) |
| | | Asian | 58 (9) |
| | | White | 405 (66) |

*Foreign race refers to international students, participants that were not born or raised in the USA.

with reasons ranging from participants declining them to dropping the study before the device could be issued (Figure 1, level 4). Among the participants who received Fitbits, 16 were further removed, with 11 providing no Fitbit data and 5 providing no sex or race data (Figure 1, level 5). To provide an overview of the study cohort, Table 1 outlines basic demographic statistics.

### Participant support

For participants who encountered technical issues or had any questions regarding the study, office hours were held twice a week each semester. An email address was also provided which participants could contact with any questions or to setup a meeting if they were unable to make office hours. During the seasonal breaks, participants could only reach support via email unless still on campus, meaning device issues in these time periods often went unresolved until the participant returned to campus.

### Interventions for maintaining participation

Across the three year study period, in an effort to maintain adherence and retain user participation, the following interventions were conducted on the study population:

- *Provide incentives:* Monetary incentives were provided on a monthly basis during the study. For the first two months of the study, a uniform compensation strategy was implemented with participants paid $20 each month. Starting in the third month (November 2015), an compliance-based strategy shown to improve participant's adherence was implemented [19]. Participants were paid $20 for having at least 80% adherence, $15 for 50-80%, $10 for 30-50% and no payments were received for less than 30% adherence.
  One year into the study (September 2016) students were given the option to participate in a one week step challenge where they would be given monetary incentives for completing daily step goals. Participants could choose between two incentive schemes: *standard* and *risky*. Each student began the week with $10. The *standard* scheme compensated an additional $1 each day the student met their step goal with a payout range of $10-$17. The *risky* scheme compensated an additional $2 each day participants met their step goal but lost $1 every day they did not meet it, with a payout range of $3-$24.
  Finally, after a year into the study (December 2016), three workshops were scheduled to check-in with students to ensure their device was working properly, offering compensation of $5 for attendance.
- *Implement a reminder system:* A monthly email (e-newsletter) was sent to participants with study updates including reminders to sync their devices, office hours, information about technical support, and payments. To help participants remember to sync their Fitbit, an automated SMS

service was introduced in the sixth month of the study (February 2016) which sent monthly reminders to participants.
- *Provide adherence feedback:* Participants had access to their adherence scores by logging into the study adherence portal which provided their current adherence score for the month and daily scores for the past 60 days.
- *Additional strategies:* Throughout February and March of the second year, to boost adherence, phone calls were made to non-adherent and unresponsive participants to set up times to meet and fix any issues with devices.

## 4 METHODS

We begin this section providing details on the data chosen to address the research questions driving this manuscript. Following this, we outline the analyses necessary to appropriately address each question.

### Data

Among the data gathered throughout the NetHealth study, we selected daily adherence scores, big-five personality traits, and demographics for our analysis.

*Adherence behaviors.* Adherence data was derived from participants Fitbits. As Fitbits' recorded participants' heart rates every minute the device was worn, daily adherence scores could be calculated through the sum of minutes in a day a heart rate was detected, divided by the total number of minutes in a day (1440).

Throughout previous NetHealth study analyses, satisfactory adherence was defined by wearing the Fitbit 19 out of 24 hours (80%) each day as this threshold has been shown to provide a good indication of activity and sleep [21].

Given the goal of this research was to address whether early adherence behaviors are indicative of long-term adherence and dropout, we defined early adherence as participants' first thirty days in the study, labeling each participant as *satisfactory* or *unsatisfactory* based on this 80% adherence threshold. Specifically, a participant was labeled *satisfactory* if they had at least 80% adherence on average across their first thirty days and labeled *unsatisfactory* if they fell below this 80% average. A distribution of early adherence scores is provided in Figure 2, with a median early adherence score of 81% ($\sigma$ = 22%).

*Big-five Personality Traits.* As part of the NetHealth study, participants completed surveys before arriving on campus for their first year and then once per semester.

From these surveys, we utilized the big-five personality traits: extroversion, agreeableness, conscientiousness, neuroticism, and openness, as prior work had linked these traits to variations in adherence [1, 6, 12, 13]. Personality assessments resulted in a score between 1-5 for each trait, with higher scores indicating the participant was more representative of that trait. As personality was assessed in participants' entrance survey, first semester survey and sixth semester survey, personality traits were treated as time-varying covariates; updated after each consecutive assessment to capture any changes in personality. For participants with missing personality assessments, fill-forward/backward imputation was utilized, allowing these imputed values to most closely reflect participants true scores at the time.

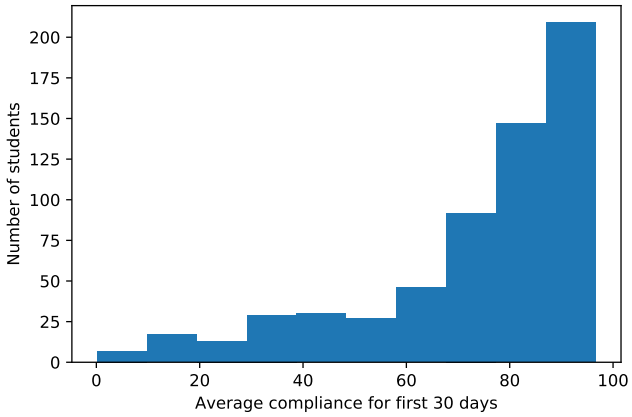


**Figure 2: Distribution of participants average adherence for their first 30 days in the study.**

*Demographics.* Finally, we included sex and race with both variables taken from each participant's first survey.

### Modeling time to dropout (RQ1)

The first research question asked was whether participants with unsatisfactory adherence in their first 30 days of the study would be more likely to dropout (RQ1). Participants were deemed "dropped" if they ceased providing Fitbit data before the study ended. A participant's first month of 0% adherence was designated as the month in which they dropped the study, provided all consecutive months were also 0%.

A discrete-time survival analysis approach was taken to modeling participants time to dropout, with three separate models constructed. The first model considered early adherence, defined as a binary feature based on participants adherence habits throughout their first 30 days in the study. A total of 322 participants were labeled as *satisfactory* early adherence and 295 as *unsatisfactory*. The second model examined the big-five personality traits and the third model considered early adherence and big-five while also adjusting for sex and race. The big-five personality traits were treated as time-varying covariates, while early adherence, sex, and race were static.

### Modeling monthly adherence (RQ2)

The second research question we addressed was whether participants with unsatisfactory adherence in their first 30 days would be less compliant throughout the study compared to participants with satisfactory adherence in their first 30 days (RQ2).

For this analysis, all participants who had dropped out of the study were removed in order to examine adherence behaviors outside of abandonment. This removed 316 participants, with the remaining 301 participants then partitioned by the 80% adherence threshold. Doing so labeled 198 students as having *satisfactory* early adherence and 103 with *unsatisfactory* early adherence.

Linear mixed effects models were employed to appropriately model the time-varying dependent variable of monthly adherence. Three models were again constructed using the same three sets of variables from the survival analysis. We note that because a participant's first 30 days in the study were used to calculate the early adherence variable, each participant's first month of adherence was omitted from this analysis to avoid violating assumptions of the mixed effects model.

Finally, for the survival analysis and mixed effects models, statistical significance was defined as $P < .05$. All analyses were conducted using Python 3.6 and R 3.4.

## 5 RESULTS

In this section, we first provide an overview of participant dropout rates and then address our findings from the two research questions motivating this manuscript. We then end with results from a supplementary analysis in which we modify the time window for determining early adherence.

### Dropout

Over the three year study period, 316 participants dropped from the initial cohort of 617 participants (Figure 3). Participants either left voluntarily as they were no longer interested in participating in the study or deemed "dropped" after several months of non-adherence. Among the 387 tier 1 participants, 209 dropped over the course of three years. For the tiers who entered the study later, 56 of the 95 tier 2 participants dropped and 51 of the 135 tier 3 participants dropped.

### Time to dropout (RQ1)

The findings from the discrete-time survival analysis are presented in Table 2. We note odds ratios (ORs) above 1.0 suggest an increased likelihood of dropping out and ORs below 1.0 suggest a decreased likelihood of dropping out. In the unadjusted and adjusted models, unsatisfactory early

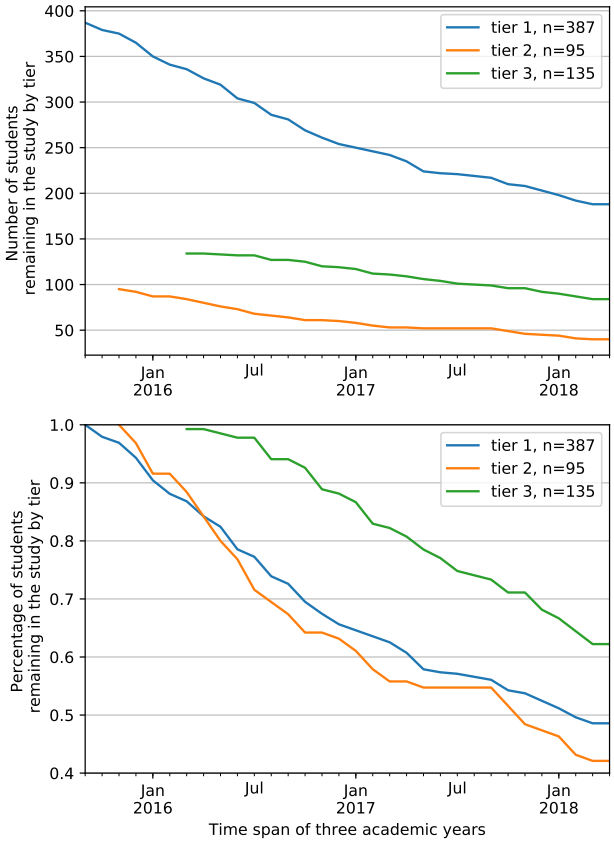


**Figure 3: Participants who withdrew from the study across three years stratified by tier, visualized by number (top) and percentage (bottom).**

adherence was associated with increased likelihood of dropping out of the study (aOR: 2.05; 95% CI: 1.61, 2.63) compared to satisfactory early adherence. Regarding the big-five personality traits, higher levels of *extraversion* were associated with increased likelihood of dropping out (aOR: 1.30; 95% CI: 1.10, 1.54). For *agreeableness*, *conscientiousness*, *neuroticism*, *openness*, *race* and *sex*, no significant associations were observed for likelihood of dropping out.

### Monthly adherence (RQ2)

Having addressed factors associated with likelihood of dropout, a subsequent analysis was conducted for the 301 participants who did not drop the study to examine adherence behaviors outside of abandonment. Results from the linear mixed effects models are outlined in Table 3 with coefficients representing change in average adherence for a given month. Unsatisfactory early adherence was associated with lower monthly adherence (aCoef: -11.08; 95% CI: -15.10, -7.06) compared to satisfactory early adherence, visualized in Figure 4. Regarding the big-five personality traits, increased

**Table 2: Results from the discrete-time survival analysis (RQ1)**

| | Unadjusted* (n = 617) | Unadjusted† (n = 572) | Adjusted‡ (n = 572) |
|---|---|---|---|
| | OR (95% CI) | OR (95% CI) | aOR (95% CI) |
| **First 30 days adherence** | | | |
| Satisfactory (≥80%) | reference | | reference |
| Unsatisfactory (<80%) | 2.13 (1.69, 2.68) | | 2.05 (1.61, 2.63) |
| **Big-five Personality Traits** | | | |
| Extraversion | | 1.33 (1.12, 1.58) | 1.30 (1.10, 1.54) |
| Agreeableness | | 0.81 (0.64, 1.02) | 0.85 (0.67, 1.07) |
| Conscientiousness | | 0.80 (0.63, 1.00) | 0.80 (0.63, 1.00) |
| Neuroticism | | 1.04 (0.86, 1.26) | 1.09 (0.89, 1.34) |
| Openness | | 1.04 (0.83, 1.30) | 1.00 (0.80, 1.24) |

*includes only early adherence group, †includes only big-five personality traits
‡includes all variables from the unadjusted models along with sex and race

**Table 3: Results from the linear mixed effects model (RQ2)**

| | Unadjusted* (n = 301) | Unadjusted† (n = 293) | Adjusted‡ (n = 293) |
|---|---|---|---|
| | Coef (95% CI) | Coef (95% CI) | aCoef (95% CI) |
| **First 30 days adherence** | | | |
| Satisfactory (≥80%) | reference | | reference |
| Unsatisfactory (<80%) | -13.98 (-18.10, -9.85) | | -11.08 (-15.10, -7.06) |
| **Big-five Personality Traits** | | | |
| Extraversion | | -3.46 (-5.30, -1.63) | -3.20 (-4.99, -1.41) |
| Agreeableness | | 2.79 (0.61, 4.97) | 2.73 (0.58, 4.89) |
| Conscientiousness | | 2.79 (0.66, 4.93) | 2.75 (0.64, 4.87) |
| Neuroticism | | -1.03 (-2.87, 0.80) | -0.88 (-2.73, 0.97) |
| Openness | | -5.04 (-7.12, -2.96) | -4.54 (-6.59, -2.48) |

*includes only early adherence group, †includes only big-five personality traits
‡includes all variables from the unadjusted models along with sex and race

**Table 4: Results from the 10 Day analysis**

| | Discrete-time survival analysis | | Linear mixed effects | |
|---|---|---|---|---|
| | Unadjusted* (n = 617) | Adjusted† (n = 572) | Unadjusted* (n = 301) | Adjusted† (n = 293) |
| **First 10 days adherence** | OR (95% CI) | aOR (95% CI) | Coef (95% CI) | aCoef (95% CI) |
| Satisfactory (≥80%) | reference | reference | reference | reference |
| Unsatisfactory (<80%) | 1.82 (1.45, 2.28) | 1.71 (1.34, 2.20) | -7.23 (-11.59, -2.86) | -7.19 (-11.26, -3.11) |

*includes only early adherence group variable, †adjusted for big-five personality traits, sex and race

*extraversion* (aCoef -3.20; 95% CI: -4.99, -1.41) and *openness* (aCoef -4.54; 95% CI -6.59, -2.48) were also associated with lower monthly adherence. We observed associations between higher monthly adherence and *agreeableness* (aCoef 2.73; 95% CI: 0.58, 4.89) and *conscientiousness* (aCoef 2.75; 95% CI: 0.64, 4.87), respectively. Finally, no significant association was observed between *neuroticism*, *sex*, or *race* and changes in monthly adherence.

### Can we predict adherence in a shorter time? - A 10 day analysis

Given the significant findings from our two research questions, we wanted to determine whether a shorter window

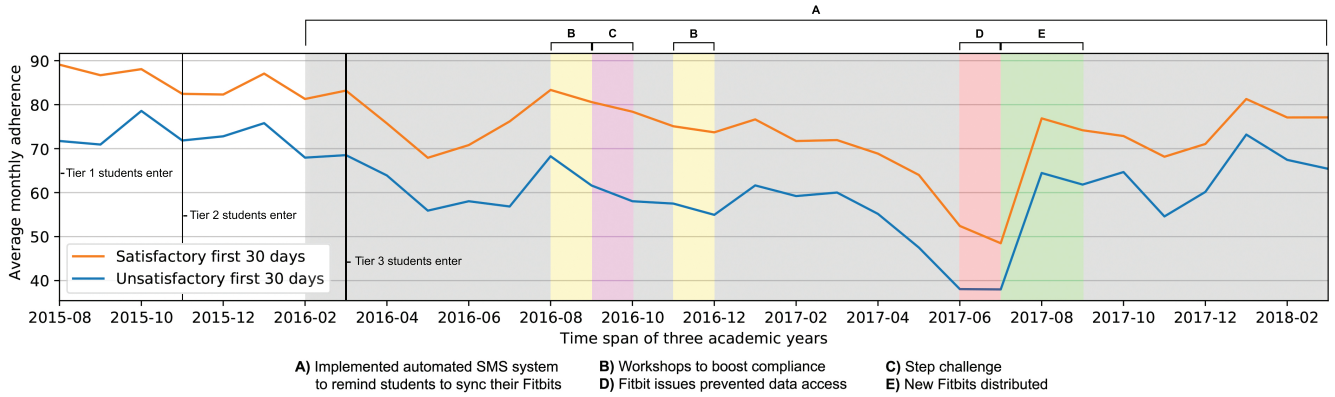


**Figure 4: Average adherence across three years for participants who did not drop the study. Interventions conducted to boost adherence and issues encountered are highlighted to provide context for major shifts in adherence**

for early adherence would still provide a strong indication of future adherence and likelihood of dropout. Therefore, we repeated the analysis for our two research questions using only the first 10 days participants were in the study. Separating the 617 participants for the survival analysis by the adherence threshold resulted in 316 participants with satisfactory early adherence and 301 with unsatisfactory early adherence. Among the 301 participants who did not drop the study, 186 participants had satisfactory early adherence and 115 had unsatisfactory early adherence. Effects of early adherence were not as strong compared to those from the 30 day analysis, however, significant associations were still present for likelihood of dropout (aOR: 1.82; 95% CI: 1.45, 2.28) and long-term adherence (aCoef -7.19; 95% CI: -11.26, -3.11), see Table 4.

## 6 DISCUSSION

In this section, we reflect on the findings from our two research questions, addressing their implications and how they may be exploited towards more effective long-term study designs involving wearable devices. We then extend our discussion to insights based on the NetHealth study and how these may further supplement future study design.

### Insights from RQ1 & RQ2

The following highlights how our findings from the research questions posed in this paper might be leveraged to increase the effectiveness of long-term study designs in retaining participants.

***Early adherence behaviors.*** The findings from RQ1 and RQ2 provide evidence for the stability of adherence behaviors: an individual who begins a study with high adherence is more likely to carry on with high adherence than an individual with poor adherence and vice versa. While this association seems obvious enough, it no doubt carries important implications. Specifically, the stability of adherence behaviors can be incorporated into long-term study designs through recruiting participants on a rolling-basis.

By introducing an early identification process into study designs, researchers can identify participants most likely to comply long-term. Through this probationary period, researchers can decide, based on the type of study, whether they wish to proceed only with those who show the most promise for future adherence, or if bias is of concern, concentrate retention efforts on those they identify as *at risk* of non-adherence or abandonment. Further, by gauging adherence early, investigators can more appropriately determine how many additional participants should be recruited based on the percentage of their current study population considered *at risk*.

While the notion of recruiting participants only to remove them after a month's observation may seem unappealing, for studies spanning multiple years, this is a small investment for reducing abandonment and non-adherence long-term. Overall, a rolling-recruitment approach would prove most effective when tailored toward the nature, budget, and goals of the given study.

***Big-five personality traits.*** By including personality traits in our model to adjust for previously known factors associated with adherence, we provide evidence supporting prior work and offer additional insights towards effective future study design [1, 6, 12, 13].

Outlining each of the big-five personality traits, we find *Extraversion* was negatively correlated with adherence, suggesting highly extroverted participants would be less compliant than those more introverted. Given an extrovert is more motivated by social interactions, the motivations for maintaining adherence may be less important to them as opposed

to introverts. *Openness* was also negatively correlated with adherence which may be in part due to those with higher levels of openness being less likely to stick with a routine, making Fitbit maintenance a lower priority. *Conscientiousness* was positively correlated with adherence, an intuitive finding as conscientious participants are more likely to take their obligations to a study seriously. And finally, *agreeableness* was also positively correlated with adherence, another intuitive finding as adherence is considered a facet of agreeableness [16].

The associations between personality and adherence provide further evidence that limiting a sample to only those who are most compliant may introduce a bias. As such, it may be necessary when retaining only highly compliant participants, to administer personality assessments in order to test for such a bias.

In studies where personality attributes may be less important to the findings, such as device testing, investigators would likely benefit from assessing both personality and early adherence to ensure minimal participant attrition.

### Insights from the NetHealth study

Although outside the scope of the research questions posed in this manuscript, we find the results of the NetHealth study offer additional valuable insights to supplement future long-term studies.

***Strategies for maintaining adherence***. Throughout the three years of the NetHealth study, several different strategies were implemented to maintain adherence. Despite these attempts, adherence scores, on average, continued to diminish through each intervening period, suggesting these events did little to motivate participant engagement. Given the majority of these interventions were based on additional monetary incentives, participants in this cohort may not have found these incentives worthwhile.

The exception to these interventions, however, was the distribution of new Fitbit Charge HR 2's through August and September 2017 which increased adherence scores by 20%, on average, and maintained higher adherence scores throughout the remainder of the study. This suggests that for device-dependent studies, introducing updated models may be an effective approach for maintaining long-term adherence as it reignites the novelty of the device. However, such an approach is likely to be highly expensive. Therefore, finding cost-effective ways in which to exploit novelty such as software updates featuring new interfaces or additional functions may prove beneficial to maintaining adherence.

We aim to address how individuals may have been motivated by these different interventions in future work.

***Documented reasons for dropout and non-adherence***. The NetHealth study documented several reasons for dropout and non-adherence from participants. However, given the non-compliant nature of participants who dropped the study, non-response was a common barrier, resulting in few responses. Among those collected, commonly cited issues included frustration with technical issues such as the Fitbits breaking. While such issues could be fixed by attending weekly-held office hours, this observation suggests that when attempting to maintain adherence, minimizing participants responsibilities may be of greater importance than incentives. As such, investigators would benefit from thoroughly testing devices and data collection platforms, possibly even with test participants, prior to starting the study to ensure technical issues are kept to a minimum.

Regarding dropout, some participants left as they lost interest in the study or no longer wanted to wear the Fitbit. Given this lack of interest, interventions may have been more effective if they had instead emphasized the importance of participation and value participants provided to the study. For example, sharing completed research with participants throughout the three years may have been a stronger motivator for engagement than monetary-incentives as participants could directly observe the greater benefits of their participation.

### Limitations

While interventions may have been necessary to maintain participants' engagement, they also prevented the observation of unbiased adherence behaviors, potentially limiting the generalizability of our results. However, the seemingly ineffective nature of these interventions and consistency of a 10% gap in adherence between *satisfactory* and *unsatisfactory* groups may be representative of only minimal bias introduced.

As variation in socioeconomic background for NetHealth participants was minimal and all participants were observed in the same environment, additional studies across different age groups and backgrounds are necessary before determining the extent to which these results can be generalized. However, we should emphasize the value in first studying a college cohort. As universities commonly draw upon their student bodies for research, the current findings are more readily applicable to a population already utilized in a large percentage of observational studies.

## 7 CONCLUSION

The innovations in wearable technology stand to provide the field of personal health and well-being with objective, highly granular, long-term data. However, prior attempts at these large-scale and long-term studies have suffered from two primary limitations: user abandonment and non-adherence. In this manuscript, we sought to address these limitations by investigating the stability of adherence behaviors. Our

findings suggested participants' adherence behaviors from as early as 10 days were indicative of their likelihood of abandonment and future adherence behaviors. By exploiting this stability through modifications to study designs, such as the adoption of a rolling-recruitment, investigators may be able to mitigate data loss through the early identification of participants at risk for non-adherence or abandonment. In turn, boosting adherence to wearable devices among at-risk populations and ultimately, providing opportunities for prolonged engagement in personal health.

## ACKNOWLEDGEMENTS

The research reported in this paper was supported by the National Heart, Lung, and Blood Institute (NHLBI) of the National Institutes of Health under award number R01HL117757. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## REFERENCES

[1] Malin Axelsson, Eva Brink, Jesper Lundgren, and Jan Lötvall. 2011. The influence of personality traits on reported adherence to medication in individuals with chronic disease: an epidemiological study in West Sweden. *PloS one* 6, 3 (2011), e18241.

[2] James Clawson, Jessica A. Pater, Andrew D. Miller, Elizabeth D. Mynatt, and Lena Mamykina. 2015. No Longer Wearing: Investigating the Abandonment of Personal Health-tracking Technologies on Craigslist. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 647–658. https://doi.org/10.1145/2750858.2807554

[3] Lynn Coorevits and Tanguy Coenen. 2016. The Rise and Fall of Wearable Fitness Trackers. (08 2016).

[4] Daniel A Epstein, Monica Caraway, Chuck Johnston, An Ping, James Fogarty, and Sean A Munson. 2016. Beyond abandonment to next steps: understanding and designing for life after personal informatics tool use. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 1109–1113.

[5] Daniel A Epstein, Jennifer H Kang, Laura R Pina, James Fogarty, and Sean A Munson. 2016. Reconsidering the device in the drawer: lapses as a design opportunity in personal informatics. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 829–840.

[6] Louis Faust, Rachael Purta, David Hachen, Aaron Striegel, Christian Poellabauer, Omar Lizardo, and Nitesh V Chawla. 2017. Exploring Compliance: Observations from a Large Scale Fitbit Study. In *Proceedings of the 2nd International Workshop on Social Sensing*. ACM, 55–60.

[7] Tom Fawcett. 2015. Mining the quantified self: personal knowledge discovery as a challenge for data science. *Big Data* 3, 4 (2015), 249–266.

[8] Rúben Gouveia, Evangelos Karapanos, and Marc Hassenzahl. 2015. How do we engage with activity trackers?: a longitudinal study of Habito. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 1305–1316.

[9] Daniel Harrison, Paul Marshall, Nadia Berthouze, and Jon Bird. 2014. Tracking physical activity: problems related to running longitudinal studies with commercial devices. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM, 699–702.

[10] Daniel Harrison, Paul Marshall, Nadia Bianchi-Berthouze, and Jon Bird. 2015. Activity tracking: barriers, workarounds and customisation. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 617–621.

[11] Sander Hermsen, Jonas Moons, Peter Kerkhof, Carina Wiekens, and Martijn De Groot. [n. d.]. Determinants for Sustained Use of an Activity Tracker: Observational Study. ([n. d.]).

[12] Robert C Hilliard, Britton W Brewer, Allen E Cornelius, and Judy L Van Raalte. 2014. Big five personality characteristics and adherence to clinic-based rehabilitation activities after ACL surgery: A prospective analysis. *The open rehabilitation journal* 7 (2014), 1.

[13] Oliver P John, Eileen M Donahue, and Robert L Kentle. 1991. The big five inventory-versions 4a and 54.

[14] Youngdeok Kim, Angela Lumpkin, Marc Lochbaum, Steven Stegemeier, and Karla Kitten. 2018. Promoting physical activity using a wearable activity tracker in college students: A cluster randomized controlled trial. *Journal of Sports Sciences* 36, 16 (2018), 1889–1896. https://doi.org/10.1080/02640414.2018.1423886 PMID: 29318916.

[15] Amanda Lazar, Christian Koehler, Joshua Tanenbaum, and David H. Nguyen. 2015. Why We Use and Abandon Smart Devices. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 635–646. https://doi.org/10.1145/2750858.2804288

[16] David Matsumoto and Linda Juang. 2016. *Culture and psychology*. Nelson Education.

[17] J. Meyer, J. Schnauber, W. Heuten, H. Wienbergen, R. Hambrecht, H. Appelrath, and S. Boll. 2016. Exploring Longitudinal Use of Activity Trackers. In *2016 IEEE International Conference on Healthcare Informatics (ICHI)*. 198–206. https://doi.org/10.1109/ICHI.2016.29

[18] Jochen Meyer, Merlin Wasmann, Wilko Heuten, Abdallah El Ali, and Susanne C.J. Boll. 2017. Identification and Classification of Usage Patterns in Long-Term Activity Tracking. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 667–678. https://doi.org/10.1145/3025453.3025690

[19] Dan Peng, Fan Wu, and Guihai Chen. 2015. Pay As How Well You Do: A Quality Based Incentive Mechanism for Crowdsensing. In *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '15)*. ACM, New York, NY, USA, 177–186. https://doi.org/10.1145/2746285.2746306

[20] Lukasz Piwek, David A Ellis, Sally Andrews, and Adam Joinson. 2016. The rise of consumer health wearables: promises and barriers. *PLoS Medicine* 13, 2 (2016), e1001953.

[21] Rachael Purta, Stephen Mattingly, Lixing Song, Omar Lizardo, David Hachen, Christian Poellabauer, and Aaron Striegel. 2016. Experiences Measuring Sleep and Physical Activity Patterns Across a Large College Cohort with Fitbits. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers (ISWC '16)*. ACM, New York, NY, USA, 28–35. https://doi.org/10.1145/2971763.2971767

[22] Patrick Shih, Kyungsik Han, Erika Shehan Poole, Mary Beth Rosson, and John Carroll. 2015. Use and adoption challenges of wearable activity trackers. (03 2015).