The feasibility of using smartphones to assess and remediate depression in Hispanic/Latino individuals nationally

Abhishek Pratap*
University of Washington / Sage Bionetworks
1100 Fairview Ave N
Seattle, WA 98115
apratap@uw.edu

Joaquin A. Anguera*
University of California, San Francisco
675 Nelson Rising Lane
San Francisco, CA 94158
joaquin.anguera@ucsf.edu

Brenna N. Renn
University of Washington
1959 NE Pacific St. Box 356560
Seattle, WA 98195
bnrenn@uw.edu

**Both authors contributed equally to this work*

Elias Chaibub Neto
Sage Bionetworks
1100 Fairview Ave N
Seattle, WA 98115
elias.chaibub.neto@sagebase.org

Joshua Volponi
University of California, San Francisco
675 Nelson Rising Lane
San Francisco, CA 94158
josh.volponi@ucsf.edu

Sean D. Mooney
University of Washington
850 Republican St. Box 358047
Seattle, WA 98195
sdmooney@uw.edu

Pat A. Areán
University of Washington
1959 NE Pacific St. Box 356560
Seattle, WA 98195
parean@uw.edu

Abstract

Mental Health conditions are now amongst the top five burdensome diseases in the US. Disparities in access to services and health outcomes vary due to several factors including socioeconomic status, shortage of mental health professionals, stigma and the linguistic gap between providers and non-English speaking minority population. This study explores the utility of providing Spanish (mono and bilingual) speaking Latino/Hispanics an option to access low-cost yet modern depression assessment and interventions available through smartphones. This is a fully remotely run study in order to investigate access, engagement, costs, and impact in an ecologically valid manner in a "real life" setting. The main purpose is to directly address issues of culture and language often overlooked in the digital health space, beginning with a thorough characterization of how these modern tools are perceived, utilized, and digested in the largest minority population in the US.

Author Keywords

Depression; minority population; mHealth; passive sensing; smart phone; intervention

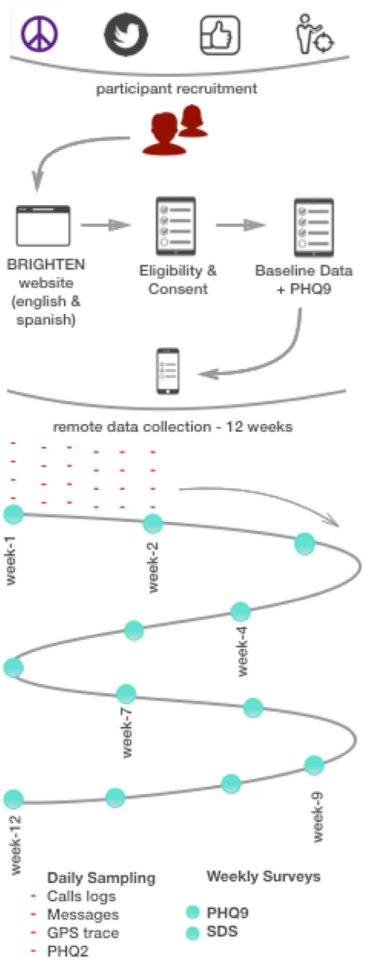

Figure 1 Overall schematic workflow for the BRIGHTEN v2 study.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

Introduction

About 20% of people in the US experience some sort of a mental health issue during any given year. The cost of providing care for this affected population tops healthcare spending and is on the rise. In 2013, around $200 billion was spent towards providing care to patients with behavioral health issues[1]. Despite this tremendous need, there exist disparities tremendous need, there exist disparities in service utilization across groups[2]. Access to care is affected by treatment barriers such as language and socioeconomic factors.[3, 4] This is highlighted in a 2005 review by the American Psychological Association[5, 6] which demonstrated that only 36% of Hispanic adults with depression received care, versus 60% of their White non-Hispanic counterparts. Also, bilingual patients were also reported to be evaluated thoroughly when the care is provided in English versus Spanish[5].

With access for mental health issues lacking for this significant minority population, there is an immediate need for further research to develop and evaluate new solutions for mental health care that are economically viable, scalable and focused on engaging users ultimately informing timely clinical interventions.

Smartphones offer a means to cut across language barriers by delivering medical information, assessment and intervention tailored in local dialects[7]. Also, their ability to continuously sample data from onboard sensors have provided an unusual opportunity for the field of medicine. The ability to learn from real-time data collected from people to inform individualized and early clinical interventions to scale has attracted clinicians and technologists at large[8-10].

Study Overview

The purpose of this study was to explore the feasibility of conducting a fully remote randomized controlled trial that screens, consents, assesses depression and deploys interventions using smartphones in depressed adults belonging to an underserved Hispanic/Latino population. We wanted to assess the feasibility of recruiting participants in this ethnic group with depression given our previous success in doing so in English-speaking individuals [11] as a first step to reduce access-based disparities in mental health treatments. In particular, we sought to understand how such technology-based tools are perceived and used by a minority population. To address language barriers, all the tools developed as part of the first BRIGHTEN v1[11] study were translated made available in both English and Spanish for this population.

In this short paper, we summarize some of our initial key findings about our study cohort demographics, participant recruitment and engagement, and discussion of specific challenges in this fully remote trial.

Methods

Ethical approval for the trial was granted by the UCSF Committee for Human Research.

Recruitment: Four different recruitment approaches were used: a) traditional online classified advertisement (using craigslist.org), b) social networking (ads on Facebook and Twitter), c) contextual targeting (via www.trialspark.com), and printed fliers distributed

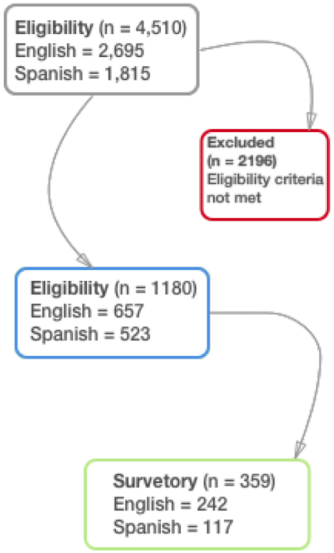

Figure 2 Summary study consort diagram

through Spanish ministries located at Catholic diocese in the largest cities in every state. Contextual targeting pushed ads to potential participants based on the recruitment methods directed potentially interested participants to the study websites (*www.brightenstudy.com for English-speakers or www.brightenstudy.com/spa for Spanish-speakers*).

Participant Eligibility: Adults (18 years and older) who owned a smartphone and endorsed mild to severe depressive symptoms as determined by a Patient Health Questionnaire [PHQ-9] score ≥ 5 were eligible to join the study.

Screening & Consent: Potential participants were directed to the study websites to explain the study purpose and procedures. Interested participants completed a brief on-line screening consisting of questions about mobile device ownership, demographics, and potential depression severity.

Procedure: Participants (*N* = 1180) were randomized to one of three mobile apps for depression: (1) *Project: Evo™*, a cognitive training app designed to address deficits in cognitive control associated with depression; (2) *iPST*, an app based on an evidence-based psychotherapy for treating depression called problem-solving therapy (PST); and (3) *Health Tips*, a psychoeducational app that was intended as an active control condition. Participants were asked to engage in remote treatment by using their assigned app for one month. Remote assessment of participant characteristics were collected at baseline; remote assessment of self-reported mood was conducted via daily push updates to complete a brief two-item depressive symptom public health questionnaire (PHQ-2)[12] and a longer, nine-item inventory on a weekly basis (PHQ-9)[13]. The PHQ-9 asks participants to rate the presence and frequency of ten depressive symptoms during the last week on a four-point rating scale (0 = *not at all*; 3 = *nearly every day*). Possible scores range from 0-30, with higher scores suggesting higher levels of symptom severity. The PHQ-2 assesses depressed mood and loss of pleasure or interest in daily activities ("anhedonia") using a response scale modified for daily tracking (1 = *not at all*; 5 = *most of the day*; total possible scores 2-10). See Figure 1 for a brief overall study flow schematic.

Payment**:** Participants were paid $15 for completing the initial assessment, then $20 for attempting the assessments at the 4, 8, and 12-week marks.

Data Collection and Featurization

The study began enrolling in August of 2016 and ran for 6 months. In total 4510 participants responded to online ads (Figure 2). Of these, 1180 were enrolled, however only 359 individuals actually completed at least one assessment after a one week run-in period. We developed a mobile app (iOS and Android) called *Survetory* for collecting passive (movement-based GPS location, call and text logs) and self-reported depression assessment (active tasks) through PHQ-2/9 surveys. Participants were sent notifications on the day an active survey was due e.g. daily for PHQ-2. Raw data was processed to generate features for association with PHQ-2/9 scores. Phone usage data (call logs and texts) were aggregated at day level and stratified into incoming and outgoing groups. Phone usage data (call logs and texts) were aggregated at day level and stratified into incoming and outgoing groups. To efficiently utilize battery, GPS sensor was triggered

	Subjects (%)
Gender	
Male	287 (29)
Female	701 (70)
Age	
18-30	432 (43)
31-40	286 (30)
41-50	180 (18)
51-60	62 (6)
61-70	16 (2)
70+	2 (<1)
Bilingual	
Yes	345 (65)
No	643 (35)
Income	
< 9000	196 (20)
10 – 30k	286 (29)
30-50k	201 (20)
50-70k	126 (13)
70-90k	55 (5)
100k+	99 (10)
Device	
iPhone	894 (90)
Android	94 (10)

Table 1 Study socio-demographics variables (selected)

every one hour or at every 100 meters of displacement, whichever was first. The GPS features generated were similar to a previous study[14]. Briefly the engineered features were related to daily displacement, spurts of movement activity in a day (no moves, short moves < 1 km hourly displacement) and daily circadian displacement patterns during the morning, afternoon evening and night time.

Results

Sample Demographics

The study cohort was primarily young adults (43% of the sample between 18 to 30 years of age), with a significantly higher number (70%) of females participating. The majority (65%) of participants reported having completed a 4-year college degree or higher. Sixty-nine percent reported being employed at the time of enrollment and 49% reported making < $30,000 per year. Interestingly we saw 90% of users in the study had an iPhone; however, only 60% of them reported using an iPhone version 5 or newer. See Table 1 for more details of participant characteristics

Clinical Characteristics

The cohort was moderately depressed with a mean PHQ-9 score at the time of enrollment at 14.3(SD = 5.5). Age showed marginal but significant association with PHQ-9 ($r(980)=-0.11$, $p < .0001$), however there was no association between gender and PHQ-9 in this cohort ($F(1,980)=1.477$, $p=.22$). We also compared other baseline variables after adjusting for age and gender using F-tests. Income satisfaction was the top other demographic variables with PHQ-9 scores.

Participant Compliance

The participation compliance across different study activities varied widely (Figure 3). While passive data stream for GPS showed highest compliance overall, the serve dip in passive phone usage data is mainly due to the device differences. iOS system (90% users) does not allow for application extraction of phone usage phone logs. The compliance differential between the number of users that contributed GPS data and completed active tasks (PHQ-9, and PHQ2) is significant. Within the first week, only 40% of the participants completed the daily PHQ-2 questionnaire.

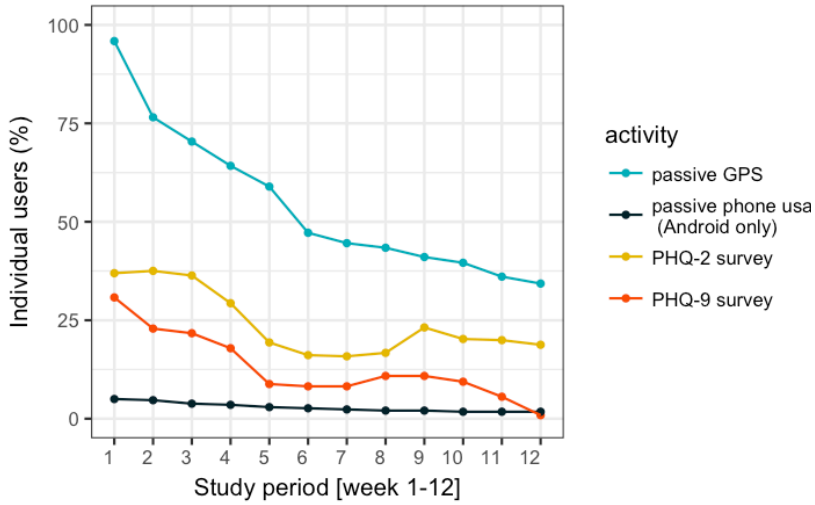

Figure 3 Study compliance faceted by active tasks (PHQ - 2/9) and passive data (GPS & Phone usage)

Cost

Total study costs included participant payments ($7,200), website/enrollment portal/database

Demographic variable	p.value (FDR corrected)
Income Satisfaction	<0.0001
Highest Education	<0.0001
Income	<0.0001
Working	0.062
Married	0.081
English fluency	0.096
Race	0.223
Hispanic/Latino	0.446
Bilingual	0.471
Spanish speaker	0.471

Table 2 Association between select study demographics variable and on-boarding baseline PHQ-9

development ($4,601.90), and total recruitment efforts ($16,590.45). $5,725 was spent on 217 Spanish Ads placed on craigslist around the country, while only $946 was spent on 33 English Ads to attain the described participant count. Thus, participant acquisition costs differed dramatically between Spanish (~$15-$22/Enrolled Participant) and English speakers (~$2).

Variations in PHQ-2 and PHQ-9 depression questionnaires at an individual level

Overall the daily PHQ-2 score reported by the participants correlated well with weekly PHQ-9 scores. (r[1669] = .58, p < .0001). However, at an individual level we observed some cases (n = 10) where some of the daily recorded PHQ-2 scores were found to the significantly different (PHQ-2 score <=3 and PHQ-9 score >=15). While this represents a small subset, it should be noted that there may be limitations in generalizing from the PHQ-2 to the broader diagnostic symptoms of the PHQ-9 in this sample.

Discussion

This study demonstrates the feasibility of conducting a large-scale mobile-based randomized trial for depression assessment and treatment among a historically underserved population of Hispanic adults with moderate depressive symptoms.

The fully-remote trial utilized both traditional and novel online recruitment techniques which yielded a large sample in a short timeframe, although clearly the cost-effectiveness between English- and Spanish-speakers is considerably different. This Hispanic sample was diverse with regards to sociodemographic variables and likely represents a population that is typically underrepresented in clinical trials. The implications of these findings are twofold: First, the use of smartphone technology has the potential to reduce access barriers, including unique targets in this population, such as language differences and issues of stigma. Second, recruitment and long term engagement approaches for this ethnic population still have considerable room for improvement given the difference in participant acquisition costs.

With regard to the remotely collected data, depressive symptoms were negatively correlated with advance age and tended to be higher in those with more financial stress. However, these daily and weekly mood ratings were poorly associated with smartphone derived features at cohort level. We believe this could be due to both large heterogeneity between individuals and data sparsity we saw in our sample.

Furthermore, for some individuals, the symptoms of depressed mood and anhedonia may not be related to an overall experience of depression as reflected in the PHQ-9. This is a priority for future trials, both remote and in-person, as the PHQ-2 is often used as a screening tool for depression but may not have the same clinical utility in Hispanic or other underrepresented samples.

Limitations

Despite this evidence for the feasibility and potential of using remote trials to assess and treat depression in historically underserved populations, our findings also highlight some limitations that are crucial to consider as the field moves toward integrated mental healthcare and technology. A priority concern in this trial was the

poor adherence in both remote assessment and utilization of the app. Only 30% of participants completed one assessment; compliance with other aspects of the programs varied widely. The complete stratified analysis comparing adherence between treatment arms is currently in progress and will shed light on predictors and consideration of engagement. Further research and development is warranted to make their use clinically relevant. Compliance was the primary factor affecting data analysis, and limiting our ability to draw inferences or employ predictive modeling. Additional large-scale studies keeping the long-term user engagement in mind are needed to develop generalizable predictive models based on passive data feeds collected in the wild. Furthermore, there is a need to infer contextual information through passive data fusion. For example, lack of phone usage when a person is flying can inferred from the GPS signal and should not be confused with lack of social interaction. Also despite focused efforts on recruiting monolingual Spanish speaking population, we found it challenging to recruit them as most of our Spanish speakers were bilingual. Ultimately, a goal of remote assessment and treatment is to engage the end-user in treatment that he/she may not otherwise receive; further attention to these stark issues of poor engagement and user experience are warranted to best understand barriers and facilitators, particularly among those with mental health issues.

Future Directions

Depression is known to disproportionately affect people from lower socio-economic status; thus, low-cost technology-based solutions that offer the potential to both assess and intervene on mental health conditions. Such technologies also offer novel ways for researchers to capture large datasets that have the potential to link passive sensor-based data with clinical outcomes to inform diagnosis and just-in-time intervention. Finally, newer and hybrid approaches using video-over-internet protocols and showing each participant their personal journey in the study needs to be evaluated for keeping the participants engaged. The possibility to remotely engage and monitor people (both actively and passively) is a particularly fitting opportunity for both assessment and targeted and timely treatment of mental health issues.

Acknowledgements

Support for this research was provided by the National Institute of Mental Health (PAA R34-MH100466, T32MH0182607, K24MH074717), the National Institute on Aging (JAA P30-AG15272).

References

1. *Focus On Behavioral Health.* Health Aff (Millwood), 2016. **35**(6): p. 956-7.
2. McGuire, T.G. and J. Miranda, *New evidence regarding racial and ethnic disparities in mental health: policy implications.* Health Aff (Millwood), 2008. **27**(2): p. 393-403.
3. Shim, R.S., et al., *Race-ethnicity as a predictor of attitudes toward mental health treatment seeking.* Psychiatr Serv, 2009. **60**(10): p. 1336-41.
4. Altarriba, J., *A theoretical approach to conceptual nonequivalence between languages*. 2003, Sage Journals: International Journal of Bilingualism. p. 305-322

5. American Psychological Association (APA). *Closing the gap for Latino patients*. 2005 July 17, 2017]; Available from: http://www.apa.org/monitor/jan05/closingthegap.aspx.
6. Biever, J.L., Castaño, M. T., de las Fuentes, C., González, C., Servín-López, S., Sprowls, C., & Tripp, C. G, *The role of language in training psychologists to work with Hispanic clients*. 2002, American Psychological Association: *Professional Psychology: Research and Practice,*.
7. Chang, D.T., et al., *Using mobile technology to overcome language barriers in medicine.* Ann R Coll Surg Engl, 2014. **96**(6): p. e23-5.
8. Donker, T., et al., *Smartphones for smarter delivery of mental health programs: a systematic review.* J Med Internet Res, 2013. **15**(11): p. e247.
9. Chan, Y.Y., et al., *The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit.* Nat Biotechnol, 2017. **35**(4): p. 354-362.
10. Webster, D.E., et al., *The Mole Mapper Study, mobile phone skin imaging and melanoma risk data collected using ResearchKit.* Sci Data, 2017. **4**: p. 170005.
11. Anguera, J.A., et al., *Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense.* BMJ Innov, 2016. **2**(1): p. 14-21.
12. Arroll, B., N. Khin, and N. Kerse, *Screening for depression in primary care with two verbally asked questions: cross sectional study.* BMJ, 2003. **327**(7424): p. 1144-6.
13. Kroenke, K., R.L. Spitzer, and J.B. Williams, *The PHQ-9: validity of a brief depression severity measure.* J Gen Intern Med, 2001. **16**(9): p. 606-13.
14. Saeb, S., et al., *Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study.* J Med Internet Res, 2015. **17**(7): p. e175.