Use of Patient Electronic Communication in Psychiatric Evaluation and Treatment
NCT04011540
January 6, 2022

Trial Protocol and Statistical Analysis Plan

Abstract

Due to the widespread use of social media, behavioral health providers have increasingly interacted and/or reviewed patients social media activity. This project seeks to harness social media data in a clinical mental health setting by partnering with both patients and their providers. We will work with these two groups to create social media and digital data dashboard that can be shared before or during clinical encounters. Our objective is to explore and evaluate how a patient-centered personalized social media summary dashboard could be used in a behavioral health therapy session to improve depressive symptoms and anxiety and if a social media dashboard improves patient and provider therapeutic alliance and engagement.

Background

Last year, approximately 4 million adults and 2 million adolescents received any mental health care. Mental health treatment can be highly effective in reducing symptoms and improving functioning of people in psychiatric distress. The nature of the clinical encounter is such that there can be a disconnection between it and the rest of the patients life. That is, patients may meet for 50 minutes with their therapist once a month or in rarer cases once a week. The therapist relies almost solely on patient report of thoughts, mood and behavior, and the events that preceded and followed them. Recall bias, social desirability bias, and the nature of the therapeutic relationship all may affect the accuracy, completeness and quality of the information the patient shares with his/her therapists. Poor accuracy of this communication then can affect the quality and outcomes of therapy. In some cases, changes in mood and adverse life-threatening events may be missed. It is not uncommon for patients to share written communication (e.g. a letter) with their therapist that occurred between the patient and a third party. These artifacts can be helpful to the therapist in assessing the patients mental state at the time of the communication and in providing feedback and advice regarding relationships. It is increasingly likely that this written communication will be electronic. Electronic communication, including emails, texts and social media, are now ubiquitous in our society. Approximately 2.4 million emails are sent per second, 22 billion texts per day. More than 79% of US adults actively use Facebook and use across other platforms like Twitter and Instagram is growing exponentially. These communications may offer an exciting way to support the effectiveness of mental health treatment. They can offer insight into patients mood, thoughts, lifestyle, and behavior that may not be available otherwise. This area of study raises several questions such as: How can clinicians best make use of the huge volume of raw, unstructured data that are available through social media and electronic communication (i.e. texts, emails)? How do therapists incorporate this data in the most productive way, without disrupting the therapeutic process? We strongly believe that there will be value in linking individual level information on social media to health outcomes and behavior; a better understanding actual patient willingness to share online data output could help to significantly improve our understanding of applications of digital data for health research. Our assumption is that ultimately this information will have tremendous utility for supporting more measurement-based care. In our ongoing work at Penn (Penn IRB 819442) we have enrolled over 6000 patients to share access to social media and EMR data. Of individuals who use social media and are willing to participate in research, approximately 85% of patients agree to share social media and EMR data. To date we have used this data to predict differences in medical diagnoses (e.g. hypertension, depression). We have also generated patient dashboards of general information about their health and social media use (Penn IRB: 819442)

Overall Objectives

Our objective is to explore and evaluate how a patient-centered personalized social media summary dashboard could be used in a behavioral health therapy session to improve depressive symptoms and anxiety and if a social media dashboard improves patient and provider therapeutic alliance and engagement.

Aims

Primary Aims

Determine the impact of incorporating digital data and electronic communication dashboard in behavioral health treatment on patient reported health-related quality of life (SF-36) compared to usual care

Secondary Aims

Determine the impact of incorporating digital data and electronic communication dashboard in behavioral health treatment on patient-reported depressive (PHQ-8) and anxiety (GAD-7) symptoms compared to usual care.

Determine the impact of incorporating digital data and an electronic communication dashboard in behavioral health treatment on therapeutic alliance (WAI-SR) compared to usual care.

Determine the impact of incorporating digital data and an electronic communication dashboard in behavioral health treatment on the CIS checklist compared to usual care.

Primary Outcome Variable

Examine self-report of health-related quality of life (SF-36), specifically mental health (vitality, social functioning, role-emotional and emotional well-being) under the intervention and control conditions and the change over the study period.

Secondary Outcome Variables

Examine depression (PHQ-8), anxiety (GAD-7), and therapeutic alliance (WAI-SR) under the intervention and control conditions and the change over the study period.

Examine the therapeutic alliance (WAI-SR) and frequency digital data and electronic communication is discussed in therapy (CIS checklist) under the intervention and control conditions and the change over the study period.

Study Design

Phase

Not applicable

Design

A randomized controlled trial for patients and their mental or behavioral health provider from Penn Medicine, Hall Mercer, and/or Horizon House or greater Philadelphia mental health practices and clinics to investigate the impact of incorporating digital data and electronic communication dashboard in behavioral health treatment session. Patient and provider participants will be randomized to receive the

treatment (having a digital health dashboard to be used at least two therapy sessions over a 2-month period) or the control (usual care). Participants who consent to share their social media and digital data, we will use a Facebook application previously designed for this purpose, Instagram API to access data, and Google take-out file. Patients will visit on an incognito browser https:// donate.centerfordigitalhealth.upenn.edu to safely and securely download social media and digital data. The data collection platform was reviewed by Penn Medicine IS team, please approval attached to this submission. Participants who agree to share their social media data will log in to their own Facebook account and download an application. When users first activate the application, they grant the right to researchers to access the information contained in their Facebook profiles. They will then be directed to a Facebook application called Penn Image Station that was designed for the purposes of this study. The Penn Image Station application will gather information from their Facebook public profile, status updates, personal description, and likes. The application will not affect their usual use of Facebook nor will it post to their page. Facebook data will be collected historically (all data that ever existed) and prospectively for 60 days. If they share their Google take-out, they will download a zip file of their google searches and YouTube activity. Google take-out data is only historical and retrospective. If the participant shares their Instagram posts, they will be directed to an Instagram application called Penn Social Mediome. They will read a brief statement, either agree or disagree with the statement, and then log on to their Instagram account. Their Instagram posts and photos will be extracted by the Penn Social Mediome application and will be collected historically and 60 days prospectively. No passwords would be collected by the study team. Patient participants will be asked to download AWARE, a free app on their smartphone. AWARE is an Android framework dedicated to instrument, infer, log and share mobile context information, for application developers, researchers and smartphone users. AWARE captures hardware-, software-, and human-based data. The data is then analyzed using AWARE plugins. AWARE is compatible on iPhone, Android, and Google phones. The use of the AWARE app in Penn research projects has been previously approved IRB #83254, The Networks of Daily Experiences (NODE) Study: Within-Person Change in Curiosity and Implications for Well-Being, (PI: Danielle S. Bassett, Ph.D). When subjects download the application, they will enter a Study ID number, which will be used to link their data with survey responses from Way to Health unique ID number. AWARE. Aware data will be included for only 60 days prospectively. This application and methodology has been previously used by investigators in this protocol for research within the World Well Being Project at the Positive Psychology Center at the University of Pennsylvania (prior Penn IRB protocols #816091, 813820, 819442, 824493). After completing the survey, donating their digital data, and downloading AWARE app, participants will be enrolled in Way To Health, a HIPAA compliant research infrastructure to receive complete day 60 and day 90 surveys. Patient participants will be contacted up to 5 times to complete each survey. If patient participants are randomized to the intervention arm, they will be prompted to complete text message prompts before each therapy session. The text message prompts entail appointment reminders and how to contribute screenshots to populate their electronic communication and digital data dashboard. The informed consent documents outline how patients will not contribute any PHI data through the Way to Health texting platform and how to only respond to the text message prompts. Additionally, if patients are randomized to the intervention arm, they will also be coached and briefed on text message confidentiality and privacy best practices. After approximately seven days from when digital data is received, the patients randomized to the intervention group will receive their first electronic communication and digital data dashboard summarizing findings from their historical social media, google take-out, and AWARE data. For example, electronic communication and digital data dashboard will include information such as average amount of time spent on smartphone during 12am-4am, physical activity callouts (i.e., most and least moved days), and most commonly used words on Google

search and YouTube. If patients want to share additional digital data, they will text in screenshots via Way to Health (safety and privacy measures are noted above). The provider participants randomized to the intervention arm will receive a brief training on the dashboard components and safety and privacy practices. Providers will watch a 3-minute video explaining how they will receive the dashboard each week and what data is being included. For example, the video breaks down each module. For the late-night phone activity module, it states late night phone activity is the amount of time your patients phone screen is on. The amount of time one spends on their phone is relative to each person and there is no right or wrong amount of time. No specific post or activity will be highlighted on the dashboard; rather trends and summaries will be displayed. The research team will work closely with the patient participant and their provider to distribute the dashboard before the patients next behavioral health appointment. The provider participants randomized to the intervention arm, will be emailed their patients dashboard via Penn +Box, a cloud-based collaboration service for securely managing and sharing files and folders within the Penn community and externally. Penn+Box also ensures that University data and intellectual property are securely protected. In addition, Penn Medicine providers can elect to have the dashboard delivered to their office, 3535 market, each morning at 9:30 am or before they see their patient. Dashboards will be delivered in a sealed envelope and delivered to the clinic front desk person or directly to the therapist. There will be no medical advice or evaluation will be included in a dashboard. If a dashboard is unable to be generated, the patient will be prompted to bring their smartphone into session and to discuss something from it, if applicable.

Method for Assigning Subjects to Groups

Patient participants will be randomized after they donate digital data sources and complete baseline survey. We will use Way to Health's computer-generated algorithm to randomize patient participants in block sizes of 2, 4, and 6. Patient participants will be randomly assigned to control (usual care) or experimental group (use of digital health dashboard over the 2-month study period). Study assignments will be stored in Way to Health. The study will be single-blinded. The research team will know which participants are receiving the active medication or treatment and which are not. The participants will not be blinded to their assignment. Provider participants will be enrolled in the arm of their patient. There is possibility that provider participants are exposed to the intervention for 1 patient and usual care for another, despite this limitation, we will control for contamination in analyses. The study biostatistician will determine when to unblind the study.

Study Duration

This entire study will last approximately 2 years. Randomized controlled trial participants, both patient and providers, involvement is approximately 3 months, 90 days.

Facilities

Our team is diverse and multi-disciplinary. Dr. Raina Merchant is an Associate Vice President at Penn Medicine and an Associate Professor of Emergency Medicine at the University of Pennsylvania. She has secondary appointments in the Department of Internal Medicine and Department of Anesthesia and Critical Care. Dr. David Mandell is a Professor of Psychiatry at the University of Pennsylvania, where he directs the Center for Mental Health Policy and Services Research. He also is Vice-Chair for Research. He has a 30-year history of partnering with mental health providers in Philadelphia to improve care. Dr. Sharath Guntuku is a postdoctoral researcher and his research leverages large-scale social media image and text data to model social health outcomes and psychological traits. Dr. Rinad Beidas is an Assistant Professor of Psychology in Psychiatry in the Perelman School of Medicine, University of Pennsylvania.

She is the Director of Implementation Research for the Center for Mental Health Policy and Services Research and the Director of the Implementation Science Working Group at the University of Pennsylvania. Dr. Brenda Curtis is an Assistant Professor of Psychology in Psychiatry at the University of Pennsylvania. Dr. Curtis is the recipient of a research project grant (R01) awarded from the Collaborative Research on Addiction at NIH (CRAN) in which she is using social media data to predict alcohol and other drug relapse and treatment completion among patients who have recently entered community outpatient treatment programs. She is also currently serving as an investigator on several R01 grants funded by NIAAA, NCI, and NIDA including a placebo-controlled trial of bupropion for smoking cessation in pregnant women using SMS text messaging to promote medication adherence and a trial examining the impact of a smart-phone based continuing care app for alcohol dependence. Rachelle Schneider received her Bachelor of Science in Criminology from Pennsylvania State University and will start a part-time Master of Social Work course of study in the fall at the University of Pennsylvania. She has extensive experience recruiting for clinical trials across various sites. Lauren Southwick received her Master of Public Health from the Harvard T.H. Chan School of Public Health in the Department of Social and Behavioral Sciences where she concentrated in health communication and social/behavioral change theories. Prior to joining the CDH, Lauren worked at the New York University College of Global Public Health in the Division of Social Epidemiology. She graduated from Franklin & Marshall with a Bachelor of Arts in Sociology. We hold weekly team meetings to keep all team members abreast with updates and protocol. We also circulate weekly meeting minutes and actions steps to enumerate research related duties and responsibilities. We also allow adequate sufficient time for the researchers to conduct and complete the research. There are adequate facilities for the research as our team is based out of PCAM and 3535 Market street offices.

Key Inclusion Criteria

Patient Inclusion - Between 18-65 years of age - Primarily English speaking (for language analysis) Willing to share at least one digital data source this includes Facebook, Google searches, YouTube searches, or screen time by downloading a free application (app) on their smartphone. - Has a mental or behavioral health provider and currently enrolled in mental or behavioral therapy - Attends therapy at least once a month and expected to remain in therapy for the next three months - Willing to share social media dashboard with their behavioral health provider - Able to provide informed consent - Regular activity viewing and posting on social media sites, defined as once a month posting - Owns an Smartphone - Willing to download and keep an app on their phone for 3 months

Provider Inclusion - 18 years of age or older - Primarily English speaking - Provides therapy for patient participant - Willing to review a summarized social media dashboard before or in a therapy session Intends to complete post therapy weekly surveys

Key Exclusion Criteria

Patient Exclusion - Under 18 years of age - Non-English speaker - Patient is in severe distress, e.g. respiratory, physical, or emotional distress - Patient is intoxicated, unconscious, or unable to appropriately respond to questions - Not currently enrolled in therapy or does not attend at least monthly - Not expected to remain in therapy for the next three months - Not a regular social media poster, or does not use Facebook, Instagram and/or not willing to share - Unwilling to share social media summary dashboard with behavioral health provider - Patient with diagnosed psychosis -Does not own a smartphone - Unwilling to download and keep an app on their phone for 3 months.

Provider Exclusion - Under 18 years of age - Non-English speaker - Not a registered behavioral health provider - Not provider for a RCT participant - Unwilling to review a social media summary dashboard before or during in a therapy session with RCT participant - Unwilling to complete weekly surveys

Subject Recruitment

Target Population

The target population is participants 18 years and older, primarily English speaking, a self-report of anxiety and/or depression, seeking mental health therapy at outpatient Penn Medicine clinics, Hall Mercer, and/or Horizon House; Center for Cognitive Therapy, Center for the Treatment and Study of Anxiety, Department of Psychiatry at the Perelman School of Medicine, or greater Philadelphia mental health practices and clinics, or public/private mental health centers/clinics in the continental U.S. Patient patients are expected to attend therapy sessions at least once a month for 3 months. Provider Participant, the target population is a provider who provides therapy sessions for a patient participant. Patient Participant The target population is participants 18 years and older, primarily English speaking, a self-report of anxiety and/or depression, seeking mental health therapy at outpatient Penn Medicine clinics, Hall Mercer, and/or Horizon House; Center for Cognitive Therapy, Center for the Treatment and Study of Anxiety, Department of Psychiatry at the Perelman School of Medicine, or greater Philadelphia mental health practices and clinics and attends therapy sessions at least once a month for 3 months. Provider Participant, the target population is a provider who provides therapy sessions for a patient participant.

Accrual

We will draw upon our working relationships with our various recruitment sites. The randomized controlled trial will be powered on SF-36 change in the patient participant. We used PASS to calculate the minimum detectable score difference required for over 80% power assuming a type I error rate of 0.05. We assumed 50 in each arm. We used the test for two means in a repeated measures ANCOVA. We assumed the correlation between repeated measures to be 0.60 (r2=0.36). The following values were used in the sample size calculations: SF36 Norms Mental Component Summary (MCS): mean = 50; sd = 10 Vitality: mean = 60.9; sd = 21.0 Social Functioning: mean = 83.3; sd = 22.7 Role-Emotional: mean = 81.3; sd = 33.0 Mental Health: mean = 74.7; sd = 18 Table: Minimum detectable score differences Sample Size 50 in each arm MCS 5 Vitality 10 Social Functioning 11 Role-Emotional 15 Mental Health 9 We aim to enroll 115 participants in a randomized controlled trial, with the anticipation that there will be 15% attrition.

Patient Subject Recruitment

We will identify participants across several settings: inpatient and outpatient Penn Medicine clinics (e.g., Emergency Department, Internal Medicine, Center for Cognitive Therapy, PIC, LGH mental health clinics), flyers at 3535 market and Penn Medicine iConnect (Use of Patient Electronic Communication in Psychiatric Evaluation and Treatment, Study Identifier: 831246), Google ads, Facebook ads, ResearchMatch.org, and provider referrals. Interested patient participants will be directly approached by a research assistant (RA) via email, telephone or BlueJeans. Patients who are receiving care in these settings will receive a digital or printed flyer about the opportunity to participate in a research study. In outpatient units, Ras or AAs will meet regularly with the nursing coordinator, office manager, charge nurse, nurse/ physician team to determine the daily census and appropriateness of approaching patients. If necessary, research assistants will also communicate directly with patients providers (physicians, residents or nurses) to determine whether it is appropriate to approach the patient.

Potentially eligible patients will be approached via email, telephone or BlueJeans to verify they meet all inclusion criteria and will then be asked to participate.

Compensation

Patient participants will receive a $50 Greenphire clincard at baseline once they complete the baseline survey, download AWARE, and donate digital data. At day 60, once they complete the day 60 survey and interview, $100 will be loaded on their Greenphire clincard. At day 90, patients will receive another $50 if they complete a survey. Provider participants in both arms will receive $50 a month on a Greenphire clincard or allocated to their clinic for completing post-therapy survey and the day 60 and 90 monthly survey. At day 60, providers will be asked to complete a debriefing interview. Provider participants will receive additional $50 loaded to their Greenphire clincard or allocated to their clinic if they complete 80% of the CIS checklist surveys, which is approximately 7 of 8 surveys

Study Procedures

Consent Process

Potentially eligible patients will be approached to verify they meet all inclusion criteria and will then be asked to participate. A RA will approach participants and explain the study in brief via email, telephone, BlueJeans or in-person. Subjects will be asked to confirm whether they will share or not share their Facebook and or Instagram data with health researchers as described in this protocol. Patients willing to share will complete a consent form. All consents will be obtained in a secure and private area on UPenns campus or University city or online. All efforts will be made to ensure that patients understand that participation is voluntary and will not affect their care or treatment in any way. The RA will obtain informed consent using the attached paper consent form or online. The RA will describe the study in detail, including the survey and the use of the Facebook/Instagram application/ plugin to extract their posts for language and photo analysis, AWARE app, and google search download instructions. The participant may refuse to answer any questions for any reason, can skip any question, and can stop participation at any time. Subjects may uninstall the Facebook/Instagram app to stop sharing their posts at any time. Or they may contact the research team as described in the consent to have their information deleted and permanently removed from the study database. The verbal consent form indicates that the participant may refuse to answer any questions for any reason, can skip any question, and can stop at any time. The research presents no more than minimal risk of harm to subjects and involves no procedures for which written consent is normally required outside of the research context. All study activities will occur in-person during enrollment or on the Way to Health, HIPAA Compliant platform, stored on PMACS. The research presents no more than minimal risk of harm to subjects and involves no procedures for which written consent is normally required outside of the research context. The study is not intervening or currently changing patient care. We will tell participants that we are not using the data for diagnostic purposes and the identifying information we will collect includes name, age, race, gender, and email address, cell-phone number, home address, and social security number.

Minimal Risk

This study is minimal risk and there are no known risks associated with participation in this research beyond those of everyday life.

Analysis Plan

Stata, SAS, and NVivo software will be used for analyses. All data will be summarized using descriptive statistics (mean or median, standard deviation, or interquartile range for continuous variables,

frequencies for categorical variables) and graphical techniques (histograms, scatterplot) as appropriate. Standard summary descriptive statistics to be used to analyze data from the survey. We will extract data from social media and wearable device platforms and aggregate to develop dashboards. Qualitative analysis using grounded theory will be applied for the focus group discussion, interviews, and open-ended survey questions. We will conduct an intent to treat (ITT) analysis. The ITT analysis will include all randomized patients in the groups to which they were randomly assigned, regardless of their adherence with the entry criteria, regardless of the treatment they actually received, and regardless of subsequent withdrawal from treatment or deviation from the protocol. Baseline demographic and clinical characteristics will be reported as frequency and percent for categorical variables and median and interquartile range (IQR) for continuous variables. We will compare baseline characteristics between intervention and control arms using t-tests for continuous variables and chi-squared tests for categorical variables. The goal of these comparisons will be to determine if the two arms are balanced on baseline variables after randomization. For the primary analysis, we will use linear mixed effects models to assess whether average of SF 36 scores over study follow-up differ between the two arms. We will use a random effect to account for the correlation in repeated outcome measures over time within participants. We will also account for correlation in outcome measures by therapist (as participants will be randomized within therapist), by including additional random effects to account for the nesting of patients within therapists. The primary parameter of interest will be a binary indicator for treatment arm assignment. Models will also include a variable for time of measurement. Baseline covariates that are found to be unbalanced between arms may be adjusted for in the model if they were deemed to be potential confounders a priori to adjust for potential confounding and for efficiency gain. We will use all available SF 36 scores on eligible patients from randomization through the last observation. An advantage of the linear mixed effects model is that it is robust to missing outcome measures as long as the missingness mechanism is missing at random (MAR). Patterns of missingness will be assessed carefully and multiple imputation will be used if deemed appropriate. Robust (empirical) standard errors will be used. A p0.05 will be deemed statistically significant but emphasis will be placed on point estimates and confidence intervals. In secondary analysis, we will also include an interaction term between intervention arm and time in order to assess whether the intervention affects the slope of SF 36 scores over time. We may also consider testing interactions between intervention and sex or other important covariates to test if there are differences in the effect of intervention by pre-specified sub-groups. We will use a similar approach to evaluate secondary outcomes; however, depending on the distributional form of the secondary outcome, we will use a generalized mixed effects model with the appropriate functional forms for the distribution family and link function.

This study was deemed minimal risk and there were no stopping guidelines.

Subject Confidentiality

The research team will exercise extreme caution with identifiable private information. Patients are donating historical and prospective for 60 days digital data. We are not viewing any digital or social media data in real-time. Patients randomized to the intervention arm will receive weekly messages from Way to Health. In the patients informed consent, they are instructed and agree by signing the document that they will not disclose personal health information (PHI) on the Way to Health texting platform and only respond to the text message prompts. When patient participants (both control and intervention) text into the Way to Health or complete a survey, the research team will receive email notification immediately. The research team is instructed to review the message within 1-4 hours of delivery, Monday through Friday, 8 am to 5pm. The line will be closed on Saturday and Sunday however patient

participants will receive an out of office message if they text in; the messages notes, Thank you for reaching out, we will get back to you when we get back in the office! and a Patient Inquiry Incident is created. An exception to confidentiality is if significant suicidal ideation, is exhibited through text message. Our approach was informed by the NIMH TOOLKIT: OUTPATIENT Brief Suicide Safety Assessment and Crisis text line best practices. Using Crisis Text lines list of detecting crisis words (please visit, https://www.crisistextline.org/blog/ava2) trigger the following response, Hi PARTICIPANT_FIRSTNAME, If you are having thoughts of suicide or feel unsafe and need to talk to someone, please go to your nearest emergency room, call 9-1-1 or call the National Suicide Prevention Lifeline, 1-800-273-8255, or text HOME to 741741 for 24/7 confidential services at Crisis Text Line. The PI and enrolled therapist will be immediately alerted and prompted to review the text message. The text message will be sent via Penn Medicine email with [ENCRYPT] in the subject, high priority note, and a read recipient. Once the PI and therapist review the text message, they will determine if the participant should continue in the study. The PI or Project Manager will follow up with the provider to ensure a safety plan. If a participant reports suicidal ideation, we composed a safety concern protocol. Please see SAFETY CONCERN Email / Telephone Script and the W2H safety features attached in the submission. The safety concern protocol, it consists of two emails contacting the participants behavioral health provider alerting them of these developments. If we do not receive a response, we will report our concern to clinic director immediately. Any participant who requires a level of care exceeding customary outpatient services (i.e., psychiatric hospitalization or participation in a partial hospital program), or who otherwise report or manifest thoughts or behaviors suggesting possible safety concern due to participation in the study, will be removed from the study. In addition to the safety monitoring that takes place as part of standard care, study staff will provide clinicians with notifications about patients of potential concern through a daily email. This email will be sent through the secure Penn Medicine email system. This email will be sent each morning (Monday - Friday). Patients will be told that the notification system is active only during the business week, 8 am to 5pm and emergency situations will result in update to their therapist. The Way to Health text line is open from Monday through Friday, 8 am to 5pm. If participants text in unprompted messages outside of those hours, they will receive a message, Thank you for reaching out, we will get back to you when we get back in the office. However, if one of the messages include one of the trigger words, they will be prompted with the aforementioned message. The Subject Confidentiality section is amended to reflect this. When therapists signed the informed consent, they agreed that they will be contacted if their patient exhibits significant suicidal ideation. We will consult the PI, Dr. Raina Merchant, ED physician to manage participant safety concerns and serve as the clinical contact point person. Dr. Raina Merchant is the emergency contact person if there is a crisis during after-hours. The data security of the applications is as follows: Facebooks General Data Protection Regulation - https://www.facebook.com/business/gdpr Instagram Data privacy webpage, https://help.instagram.com/519522125107875 Google takeouts security center - https://www.google.com/safetycenter/everyone/start/security-help/ We would report to the specific digital platform anything that violates their community standards. We will also apply existing filters (e.g. not safe for work) before processing image and text data. Facebook community standards: https:// www.facebook.com/communitystandards Instagram community standards: https:// help.instagram.com/477434105621119 Google takeout standards, https://www.google.com/+/policy/ content.html. Our data scientist will manage the collection of these social media platforms. Only our research team will have access to the data through the applications. For the experimental group, the patient and provider will seem summary trends from the social media platforms. If study participants do not uninstall the app after the study ended, the apps will be closed and no additional information will be used or collected through the app.

Subject Privacy

Privacy refers to the person's desire to control access of others to themselves. Privacy concerns people, whereas confidentiality concerns data. Describe the strategies to protect privacy giving consideration to the following: The degree to which privacy can be expected in the proposed research and the safeguards that will be put into place to respect those boundaries. The methods used to identify and contact potential participants. The settings in which an individual will be interacting with an investigator. The privacy guidelines developed by relevant professions, professional associations and scholarly disciplines (e.g., psychiatry, genetic counseling, oral history, anthropology, psychology).

All encounters will take place in the outpatient clinic setting, telephone, email, BlueJeans or via Way to Health, HIPAA compliant platform. To enroll provider participants, we will discuss the trial during their team meeting and provide informed consent in-person or electronically via the Way to Health platform. To enroll patient participants, we will discuss the trial via email, phone calls, or BlueJeans or on other platforms and applications outlined on the Penn Medicine / Penn Dental / UPHS (HIPAA Covered Entities) Guidance on Remote Consent Discussion and Documentation Related to COVID-19 Concerns. During the discussion, we will give them the opportunity to sign the informed consent document online or in-person. They will then complete the baseline survey, and install AWARE on their iPhone Participants can donate their digital data in-person or virtually. Regardless of the mode, patient participants will be emailed a detailed study packet noting all procedures and schedule a time (in-person or remote) to assist them to donate their data to a secure url, http://donate.centerfordigitalhealth.upenn.edu (reviewed and approved by Penn Medicine Information Security). Patients will receive a unique Way to Health ID. Patients who choose to share their Facebook, Instagram, Google takeout or other social media data will be given a unique ID that will be linked to identifiable information. Participants who install the application to share their Facebook posts will be able to opt in and out of the application at their discretion. The Facebook application will access data that is stored in the user profiles in the Facebook platform. When participants first click on the link to use the app, they will first be greeted by the Facebook consent screen that will ask the user permission for the application to have access to their profile information and to consent to the terms and service of the application. Participants would have to agree to this screen in order to use the application. They will have to click agree to access the application. All language data will be stored on secure severe using only the unique study ID which can be used to link this data to the REDCap database. This linkage will be stored in a double locked fashion on a password protected computer will access only by the principal investigator and key study personnel. The study team will not have direct access to any participants username/password for any digital media accounts and will not have the ability to post any messages or content on behalf of the user. The study team will not communicate with participants via social media or publicly post that the participant is involved in the study. Participants will be informed that they should not have any expectation of surveillance or medical advice/medical response or that any posts, searches, physical activity data will be responded to in real time or at any time by the study team. The focus group will be instructed that there are no right or wrong answers, only differing opinions. The session will be audio- recorded and that we ask that they please speak one at a time. The discussions are being used for research and everything will remain confidential and anonymous. No responses will be connected to a participants name and if there are any questions one does not want to answer, they do not have to. We will instruct all participants that the focus groups intention to keep discussion within the group and not to share information or details once the discussion is over. Focus group audio-recordings will be transcribed verbatim and all identifiers will be removed. The recording and transcript will be kept a secure and locked area with access limited to designated researchers. After we analyze the recordings, we will destroy recordings after data analysis or completion of the study. The day 60

interviews will be audio-recorded if the participant agrees, if they prefer to not be recorded, the interviewer will take notes. No digital data or social media displayed on the dashboard will be reviewed real-time.

Data Disclosure

Data responses from participants will be linked to participant identifying information, and data will be available only for listed study personnel and the IRB if requested. However, all potential subjects will be informed that the information they provide will be held in confidence to the extent that the law allows, but that the exception to this confidentiality is any disclosure of potential for immediate harm of themselves or others, such as active suicidal or homicidal ideation or child abuse. The participants will be notified prior to participation that if any of these issues are raised, the researchers will take whatever steps are necessary to protect the subject or others, including bringing risk of harm to the attention of the proper authorities. This applies only to direct verbal disclosure by the patient and not to any data included in social media data. Specifically, the research team will not review social media data to determine disclosures of suicidal, homicidal, or child abuse information that will be acted on by the research team. Data collected for analysis is detailed in the informed consent form and specifically indicates: Data extracted from Facebook will include but not be limited to past and future content: age, gender, language, photos, News Feed content, work history, education history, hometown, interests, relationship status, current city, religious or political views, number of friends, items liked, and number and status updates. Data from Instagram will include username/ handle, number of posts, engagement metrics, text captions, and messages, Google takeout includes, My Activity and YouTube, for more details, please see, https://takeout.google.com/.

Data protection

x Name
x Street address, city, county, precinct, zip code, and equivalent geocodes
x All elements of dates (except year) for dates directly related to an individual and all ages over 89
x Telephone and fax number
x Electronic mail addresses
x Social security numbers
Medical record numbers
Health plan ID numbers
Account numbers
Certificate/license numbers
Vehicle identifiers and serial numbers, including license plate numbers
x Device identifiers/serial numbers
x Web addresses (URLs)
x Internet IP addresses
x Biometric identifiers, incl. finger and voice prints
x Full face photographic images and any comparable images
x Any other unique identifying number, characteristic, or code

Consent Process Overview

Based on the potential risks and benefits described below, we believe this study may be considered minimal risk.

Potential Study Risks

Since participation in this research project will not alter patient management in any way, we believe that there will be minimal to no risk posed to participants. The most significant risk will be breach of confidentiality. This risk is minimal given the various safeguards embedded in our data storage and management. The participants will be notified prior to participation that if any of these issues are raised, the researchers will take whatever steps are necessary to protect the subject or others, including bringing risk of harm to the attention of the proper authorities. This applies only to direct verbal disclosure by the patient and not to any data included in social media data. Specifically, the research team will not review social media data to determine disclosures of suicidal, homicidal, or child abuse information that will be acted on by the research team. Data collected for analysis is detailed in the informed consent form and specifically indicates: Data extracted from Facebook will include but not be limited to past and future content: age, gender, language, photos, News Feed content, work history, education history, hometown, interests, relationship status, current city, religious or political views, number of friends, items liked, and number and status updates. Data from Twitter will include: handle, number of followers/following, tweets, geography, time/date, profile. Data from Instagram will include username/ handle, number of posts, engagement metrics, messages, and photos. The use of REDCap will also help minimize the risk of confidentiality. Additionally, the survey items are expected to cause no to minimal emotional distress. We do not expect physical, emotional, economic, psychological or societal harms that may accrue to others not in the trial. The study is designed to minimize the risk of harm to all participants participating by providing secure surveys and direct communication access to the research team and Principal Investigator.

Potential Study Benefits

Participants in this study may benefit directly from having a social media summary dashboard in their therapy sessions; in particular it may uncover interesting trends and correlations that could improve patient-provider communication and their engagement in treatment. It could also provide additional knowledge about how they communicate and use social media platforms, these summaries may be revealing and aid healthy behavior change and/or self-management. In addition feedback on the dashboard will help us better tailor future iterations to better meet the needs of our sample, both the patient and provider.

Alternatives to Participation

Considering the context of this study, there are no alternative to participation available to the participants outside of the research context.

Data and Safety Monitoring

As noted above, the research team will immediately contact the PI and enrolled therapist via Penn Medicine email with [ENCRYPT] in the subject, high priority note, and a read receipt. Once the PI and therapist review the text message, they will determine if the participant should continue in the study. Please see SAFETY CONCERN Email / Telephone Script and the W2H safety features attached in the submission. The safety concern protocol, it consists of two emails contacting the participants behavioral health provider alerting them of these developments. If we do not receive a response, we will report our concern to clinic director immediately. Additionally, all research staff will consult the National Institute of Mental Health (NIMH) toolkit for suicide safety assessment. Any participant who requires a level of care exceeding customary outpatient services (i.e., psychiatric hospitalization or participation in a partial

hospital program), or who otherwise report or manifest thoughts or behaviors suggesting possible safety concern due to participation in the study, will be removed from the study. In addition to the safety monitoring that takes place as part of standard care, study staff will provide clinicians with notifications about patients of potential concern through a daily email. This email will be sent through the secure Penn Medicine email system. This email will be sent each morning (Monday - Friday). Patients will be told that the notification system is active only during the business week and expectations for emergency situations will be clearly set.

Original Statistical Analysis Plan – Sent to IRB

Stata, SAS, and NVivo software will be used for analyses. All data will be summarized using descriptive statistics (mean or median, standard deviation, or interquartile range for continuous variables, frequencies for categorical variables) and graphical techniques (histograms, scatterplot) as appropriate. Standard summary descriptive statistics to be used to analyze data from the survey. We will extract data from social media and wearable device platforms and aggregate to develop dashboards. Qualitative analysis using grounded theory will be applied for the focus group discussion, interviews, and open-ended survey questions. We will conduct an intent to treat (ITT) analysis. The ITT analysis will include all randomized patients in the groups to which they were randomly assigned, regardless of their adherence with the entry criteria, regardless of the treatment they actually received, and regardless of subsequent withdrawal from treatment or deviation from the protocol. Baseline demographic and clinical characteristics will be reported as frequency and percent for categorical variables and median and interquartile range (IQR) for continuous variables. We will compare baseline characteristics between intervention and control arms using t-tests for continuous variables and chi-squared tests for categorical variables. The goal of these comparisons will be to determine if the two arms are balanced on baseline variables after randomization. For the primary analysis, we will use linear mixed effects models to assess whether average of SF 36 scores over study follow-up differ between the two arms. We will use a random effect to account for the correlation in repeated outcome measures over time within participants. We will also account for correlation in outcome measures by therapist (as participants will be randomized within therapist), by including additional random effects to account for the nesting of patients within therapists. The primary parameter of interest will be a binary indicator for treatment arm assignment. Models will also include a variable for time of measurement. Baseline covariates that are found to be unbalanced between arms may be adjusted for in the model if they were deemed to be potential confounders a priori to adjust for potential confounding and for efficiency gain. We will use all available SF 36 scores on eligible patients from randomization through the last observation. An advantage of the linear mixed effects model is that it is robust to missing outcome measures as long as the missingness mechanism is missing at random (MAR). Patterns of missingness will be assessed carefully and multiple imputation will be used if deemed appropriate. Robust (empirical) standard errors will be used. A $p \leq 0.05$ will be deemed statistically significant but emphasis will be placed on point estimates and confidence intervals. In secondary analysis, we will also include an interaction term between intervention arm and time in order to assess whether the intervention affects the slope of SF 36 scores over time. We may also consider testing interactions between intervention and sex or other important covariates to test if there are differences in the effect of intervention by pre-specified sub-groups. We will use a similar approach to evaluate secondary outcomes; however, depending on the distributional form of the secondary outcome, we will use a generalized mixed effects model with the appropriate functional forms for the distribution family and link function.

Amendments to the Statistical Analysis Plan

All amendments to the original SAP were made prior to unblinding of the data analyst and prior to access to outcome data.

Based on covariate balance between the intervention arms, we amended our original statistical analysis plan. In addition, we originally anticipated enrolling 20-30 therapist participants with several patients nested within each therapist. This would have necessitated the accounting for clustering within therapist using random effects. However, we enrolled 69 therapist participants with cluster sizes that were very small, with most therapists treating 1-2 patients. Hence, as our primary analysis, we conducted an unadjusted paired t-test (rather than our originally proposed linear mixed model) to compare mean changes on our primary outcomes from baseline to 60-day. As a secondary analysis, we did conduct a linear mixed effects model exactly as proposed in the original SAP. These results were nearly identical to the unadjusted paired t-test and would not change our conclusions. In addition, we conducted a complete case analysis as our primary analysis. After assessing patterns of missingness and determining that missing at random is a plausible mechanism in this study, we conducted multiple imputation. Again, our imputed results were nearly identical to the complete case analysis with identical conclusions.