FINAL PROGRESS REPORT: AHRQ R01 Project

Project Title: i-Matter: Investigating an mHealth texting tool for embedding patient-reported data into diabetes management

Organization: New York University Grossman School of Medicine

Project Period: 09/01/2018 – 06/30/2023

Principal Investigator: Antoinette M Schoenthaler EdD, Devin Mann MD

Team Members: Hayley Belli (Co-I).

Federal Project Officer: Christine Dymek, EdD

Grant Award Number: R01HS026522

This project was supported by grant number R01HS026522 from the Agency for Healthcare Research and Quality (AHRQ). The content is solely the responsibility of the authors and does not necessarily represent the official views of the AHRQ.

STRUCTURED ABSTRACT (250 words)

PURPOSE: To compare the efficacy of a text-messaging patient-reported outcome (PRO) tool [herein iMatter] versus usual care (UC) on HbA1c reduction and adherence to self-care behaviors.

SCOPE: PROs are increasingly being used in management of T2D to integrate data on patients' perspective into clinical care. To date, most PRO tools have lacked patient and provider involvement in their development thus, they fail to meet the unique needs of end-users, and lack the technical infrastructure to be integrated into the clinic workflow, limiting their impact.

METHODS: Using a mixed-methods design, we conducted a two-phase study: 1) a formative phase, using a user-centered design approach; and 2) a clinical-efficacy phase. For the formative phase, we used qualitative methods to: a) adapt iMatter to the needs of providers and T2D patients; b) integrate iMatter into the electronic health record and the lives of patients with T2D; and c) evaluate the usability of iMatter. For the clinical efficacy phase, we evaluated in a randomized control trial, the efficacy of iMatter versus UC on reduction HbA1c at 12-months, among 282 patients with T2D followed in primary care practices.

RESULTS: Using an iterative user-centered process resulted in the identification of the six PRO messages that were relevant to patients and providers: medication adherence; dietary behaviors; physical activity; sleep quality; quality of life; and healthy living goals. Results of the completer's analysis showed that there was no statistically significant difference in HbA1c reduction between the intervention and control groups at 12 months. Smoking status statistically improved in the intervention group.

Key Words: Type 2 diabetes, patient-reported outcome, mHealth, primary care

PURPOSE

This study used a two-phase approach, informed by the Technology Acceptance Model (TAM) and Capability-Opportunity-Motivation Model of Behavior (COM-B) to address the following aims: (1) conduct a randomized control trial to evaluate, among a sample of 282 patients with uncontrolled T2D, the efficacy of iMatter versus UC on reduction in the levels of HbA1c (primary outcome) and adherence to self-care behaviors (secondary outcome) and (2) examine the potential mechanisms of the effect of iMatter on HbA1c and adherence to self-care behaviors (secondary outcome).

SCOPE

Prevalence

Uncontrolled T2D is a significant public health problem in the US, particularly among vulnerable populations.(1, 2) Annually, patients with T2D incur about $250 billion in healthcare costs and lost productivity, thus representing a significant social and economic burden.(3) Despite recent improvements in the proportion of

adults with T2D achieving desired HbA1c targets of <7%, only 53% achieve this level of control.(4, 5) The number of patients who fail to meet these goals is even higher in resource-limited primary care practices - a place where most vulnerable populations receive their care.(6, 7) About 40% of Blacks and 50% of Latinos have poor glycemic control.(3, 8) Disparities in socioeconomic status (SES) place patients at an even greater risk, with recent data suggesting a widening of the SES gap in T2D prevalence.(9)

Background

To date, the care of patients with T2D has focused largely on patient self-management (adoption of healthy lifestyle, medication adherence) and the use of clinical parameters such as HbA1c levels to determine treatment effectiveness. Although several meta-analyses support the link between improvements in patient self-management and reduction in HbA1c, their effect sizes are modest (-0.43% reduction in HbA1c), particularly among vulnerable populations (only -0.31% reduction in HbA1c).(10) A major limitation of these studies is the lack of focus on patients' perspectives of the physical and psychosocial impact of T2D on their daily lives as well as their ability to manage the disease and adhere to the recommended treatment regimen. Measures of PROs are a *standardized and quantifiable* approach that allows the collection and integration of data on patients' perspective into the clinical management of T2D. Despite the central role PROs play on T2D patients' ability to manage their disease, providers overestimate how frequently they assess PROs in the clinic visit, as compared to patients (76% vs. 55%, respectively).(11) However, if we are to make an impact on the growing burden of T2D on patients and the broader society, treatment strategies must seek to balance primary care providers' (PCP) pursuit of glycemic control against patients' emotional, physical, and social experiences with their disease.

Context

To date, a majority of T2D studies have used PROs to assess patient tolerance of new treatments in clinical drug trials.(12) Of the practice-based studies, many use long batteries of PROs and only have patients report PROs on a single occasion, most often immediately before their clinic visits.(13, 14) Such reporting introduces a recall bias because patients are asked to approximate changes in their symptoms and functioning over several months. While a growing number of studies are utilizing mobile platforms that enable real-time collection of PROs in order to facilitate patient self-care outside the clinic environment, they are not theoretically driven; rarely adhere to current guidelines or link to scientifically-proven concepts, and lack clinical effectiveness, thus limiting their validity.(15, 16) Moreover, evidence of their effectiveness is inconclusive, in part due to small sample sizes (24-180 patients), short study duration (mean: 24 weeks), low patient compliance, limited integration with clinical practice, and exclusion of vulnerable populations that would benefit most from mHealth interventions.(17) The PROs collected in existing tools also lacked patient and provider involvement in their conceptualization, and focused primarily on assessing patient self-care behaviors (e.g., dietary behavior) with little consideration of the psychosocial factors (e.g., diabetes quality of life) that drive poor outcomes in patients with T2D. As a result, existing tools are not customized to address the unique needs and preferences of patients, and they lack the technical infrastructure to support their integration into the clinic workflow. Thus, there is a paramount need for the development of mHealth solutions that integrate systematic collection, and use of PROs [that extend beyond measures of self-care] into the management of T2D in primary care practices. Such solutions must: 1) provide data that is actionable by patients and providers; 2) support patient engagement in their health in real time; 3) be integrated within clinic workflow (as part of the EHR); and 4) improve clinical outcomes.

Setting

This study was conducted within the NYU Langone Health (NYULH) network of primary care practices.

Participants

Providers: Providers were enrolled if they: a) were a primary care provider (MD/DO, NP) practicing at the participating primary care practices and b) provided care to at least five patients with T2D.

Patients: Eligible patients must have: (a) had a provider diagnosis of T2D for ≥6 months; (b) uncontrolled T2D defined as HbA1c >7% documented in the electronic health record (EHR) on at least two consecutive visits in the past year; (c) been fluent in English or Spanish; (d) been willing to send/receive text messages; and (e) been ≥ 18 years of age. Patients were excluded if they (a) refused or were unable to provide informed consent; (b) had acute renal failure, end stage renal disease (ESRD) or evidence of dialysis, renal transplantation, or

other ESRD-related services documented in the EHR; (c) were participating in another T2D study; (d) had significant psychiatric comorbidity or reports of substance abuse (as documented in the EHR); (e) were pregnant or planning to become pregnant within 12 months; or (f) planned to discontinue care at the primary care practice within the next 12 months.

In the formative phase, we targeted a sample of 36 patients and 14 providers to maximize the likelihood of reaching thematic saturation for the focus groups and user testing sessions. A new sample of 282 patients with uncontrolled T2D [who did not participate in the formative phase] was targeted for the clinical efficacy phase

Recruitment: We used 2 approaches to recruit patients and providers into both study phases. First, potentially eligible patients were identified through a review of the EHR using the diagnosis-related group codes indicating the presence of T2D and receiving care from a primary care provider at one of NYULH practices. After retrieving a list of potentially eligible patients, research assistants (RAs) reviewed patients' EHR to determine if the patient met the eligibility criteria. Patients that met these criteria were contacted via telephone to confirm eligibility. During the telephone call, the RA gave patients a description of the study, including their role as participants in the study. If the patient remained interested in participating, they were given the option to either complete the informed consent procedures in-person in a private room or via a remote session using the secure Webex conferencing platform. Providers were sent emails from the study principal investigators inviting them to provide feedback on the development of an interactive mHealth tool that could help enable patients with T2D to take a more active role in their diabetes management. Providers who did not opt out of the study were sent the electronic consent form for review and signature. All patients and providers provided written informed consent before participation in either phase of the study.

METHODS

Study Design

Using a mixed-methods design, this study was conducted in two phases: 1) A formative phase, based on user-centered design and 2) a clinical-efficacy phase.

The goals of the formative phase were three-fold: First, to use focus groups to adapt iMatter to the needs of the PCP and T2D patients, including Spanish-speaking patients. Second, to conduct a design workshop to integrate iMatter into the EHR system and patient's lives. Third, to evaluate the usability of iMatter in a subset of T2D patients and their PCPs in order to optimize the tool's performance and workflow integration prior to the clinical efficacy trial. For the clinical efficacy phase, we conducted a randomized control trial to evaluate, among a sample of 282 patients with uncontrolled T2D, the efficacy of the adapted iMatter intervention versus UC on reduction in the levels of HbA1c (primary outcome) and adherence to self-care behaviors (secondary outcome). In addition, we examined the role of diabetes knowledge, self-efficacy, outcome expectations, and patient-provider communication as potential mechanisms of the effect of the iMatter on HbA1c and adherence to self-care behaviors (secondary outcome).

Data Sources/Collection

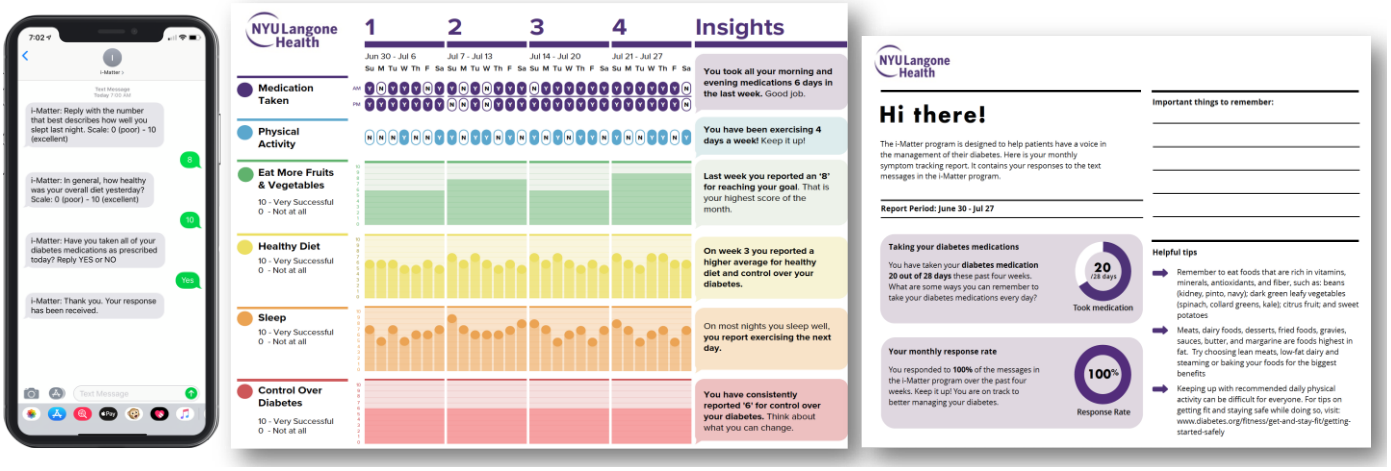
Data for this trial were divided into four categories: (1) outcomes, (2) covariates, (3) mediators, and (4) process measures. All measures were obtained by trained RAs using a standardized procedure, in the patient's preferred language. With the exception of the process measures collected in the formative phase, study measures and timing of administration were the same for both arms. Data collection forms were identified only with IDs; relating of ID code to names required information kept under lock and key, and supervised by a designated high-level staff member. None of the analyses permitted individual identification. Only ID numbers were used for communication with the RAs in the event of data anomalies. The clinical/research barrier remained intact throughout the trial, in that it was not necessary for the data-processing staff to know the identity of the participants.

Interventions

i-Matter, is a theoretically grounded technology solution informed by the COM-B model, which posits that for any behavior to occur, a person must have the capability, opportunity, and motivation to perform the behavior. i-Matter comprises three key features (**Figure 1**): (1) SMS text messages that capture patients' self-reported PROs in real time, (2) data-driven feedback and motivational SMS text messages based on responses to the PROs, and (3) dynamic visualizations of the PROs that are shared in personalized PDF reports and integrated into the clinic EHR.(18)

Over the 12-month study, patients randomized to iMatter received up to 5 daily PRO text messages that assessed their sleep quality, physical activity, diet, and medication adherence. PRO messages were sent in the morning (sleep, diet, and medication adherence) and evening (physical activity and an optional second medication adherence message). The messages required patients to respond with either a yes or no response or a response on a scale ranging from 0 to 10 within a 2-hour period. Patients also received 2 weekly messages in the afternoon that assess their diabetes, quality of life, and attainment of a personalized health goal. Personalized PDF reports were sent to patients monthly and visualize patients' responses to the PRO messages over time.

Figure 1. iMatter Intervention



Usual Care: Patients randomized to the UC group received standard diabetes treatment recommendations as determined by their PCP.

Measures

Formative Phase

Process measures

Patient use of mobile technology: before the focus groups, patients completed a survey created for this study that assessed the frequency of mobile phone use, capabilities of their mobile phones (e.g., Wi-Fi connection, Bluetooth, and mobile data plan), the most commonly used functions (e.g., text messaging, phone calls, email, and apps), comfort with using their mobile phone to manage T2D, interest in enrolling in a text messaging diabetes program, and challenges to using their mobile phone for diabetes self-management.

Use behavior: these data were extracted from the iMatter platform at the end of the user testing sessions and included the following metrics: number of mobile phone inputs, time-on-task, task success, number of missed responses to PRO questions, and number of responses by patients outside the response window.

TAM3 survey: following the 2-week user testing period, patients completed questions derived from the well-validated TAM3 survey that assessed the perceived ease of use, usefulness, and quality of iMatter; the likelihood of using iMatter in the future and recommending it to others (i.e., behavioral intention); and perceived benefits of discussing iMatter data with providers to help manage their diabetes (i.e., communication). The internal consistency of this scale ranged from 0.86 (communication) to 0.94 (perceived usefulness).

Clinical Efficacy Phase

Outcome measures

HbA1c level (primary outcome): was assessed as the difference between HbA1c at baseline and 12 months. HbA1c levels were extracted from patient's clinic EHR by a data manager blinded to the study conditions.

Patient adherence to self-care behaviors (secondary outcome) was assessed with the well-validated Summary of Diabetes Self-care Activities Measure. For this measure, patients were asked to indicate how many of the past seven days (response range 0 -7 days) they practiced the following self-care behaviors: follow a general diet, follow a diabetes specific diet, be physically active, monitor blood glucose, engage in foot care, and

smoke (scored as a yes or no response). In the analyses, we examined each behavior separately due to studies showing that engaging in one of the self-care behaviors does not correlate with practicing another behavior. The Cronbach's alpha for the scales range from 0.69 to 0.84).

Mediators

Patient self-efficacy (motivation): was assessed with the 12-item Diabetes Self-efficacy Scale. For each item, patients rated their confidence in their ability to perform a recommended self-care routine using a 10-point response from "1 = not at all confident" to "4 = very confident." Responses were summed to obtain an overall self-efficacy score and divided by the total number of items to obtain a mean score, with a higher score representing greater self-efficacy. The scale has a standardized Cronbach α of 0.78, across diverse race/ethnicity populations and health literacy levels.

Diabetes distress (burden): was assessed with the well-validated and widely used 17-item Diabetes Distress Scale, which is assessed using a Likert type scale from “1= not a problem” to “6= a very serious problem.” For each item patients rated the degree of distress the potential problem a person with diabetes may experience. To score, the sum of all the patient’s responses is divided by 17. A mean score of 3 or higher is considered a level of distress that needs medical attention. In the event of an elevated score, a message will be sent to the patient’s primary care provider and be noted in the patient’s study record.

Covariates

Patient demographic data include race/ethnicity, place of birth, years in the US, primary language, age, gender, household income, education level, marital status, employment status, health insurance status, smoking and drinking behaviors, and medical comorbidity.

Characteristics of disease and medication regimen: All patients had their EHR reviewed at baseline and 12 months. Information extracted from the charts included diabetes characteristics such as clinic HbA1c readings, duration of diabetes, evidence of target organ damage, changes in diagnosis, medical comorbidity, clinic appointment attendance, and other medications prescribed and their dosages

Limitations

Singular focus on T2D: Although our intervention is designed to target patients with T2D, it is more common for patients to have 2 or more chronic diseases (i.e., multi-morbidity) than 1 disease in isolation (89.3% vs 8.5%, respectively). In fact, recent research demonstrates the negative impact of multi-morbidity on PROs such as quality of life, psychosocial health, self-efficacy, physical function, and self-management behaviors (e.g., physical activity and medication adherence)(19).

Potential for contamination: To minimize contamination, we developed quality control checks into the iMatter EHR interface to ensure that providers only accessed reports for patients randomized to the intervention arm. We also used PCP-level randomization so that all patients of a given PCP were in the same arm. Study staff also gave in-service talks to remind PCPs of the importance of complying with the protocol.

Low access to technology will inhibit delivery of iMatter: Although concerns exist about the availability and adoption of technology by vulnerable populations, recent estimates show widespread access to mobile technologies across diverse populations in the US, even among those with limited prior technology use.(20, 21) National surveys show both Black and Hispanic mobile phone users send significantly more texts per month compared to Whites.(22) US adults living in, or near poverty are more likely to live in cell phone–only households as compared to those with a fixed phone line.(23) Finally, mobile phone numbers tend to be more stable over a six-month period than home address or non-mobile devices.(23) Thus, mobile phone-based solutions may be a powerful tool for delivering behavioral interventions to vulnerable populations.

Analytic Plan

Formative Phase

Sample size estimates for the formative phase were based on best practices for maximizing the information power of qualitative research, which recommends beginning with 6 to 8 participants per qualitative method and adding to the sample, as needed(24). As with previous studies, user testing sessions were scheduled until data saturation was reached(25-27). Our previous studies suggested that we would need 2 to 3 cycles of user testing to reach saturation(25-27).

Focus groups and interviews were audiotaped, translated where necessary, and transcribed verbatim. Both data sources were analyzed using the constant comparative method, in which text was categorized into themes with the use of codes developed iteratively to reflect the data. The coding scheme was developed by the study investigators to focus on key themes identified both a priori (e.g., from the interview protocols) and those that emerged during the interviews or focus group discussions. A trained qualitative researcher coded the transcripts independently, after which the research team met to discuss the coding and resolve any discrepancies. After each round of user testing, the study team employed the best practices for instant data analysis of usability data for each PRO(28). Task success was calculated as the percentage of PRO questions that were answered correctly without errors. Time-on-task was calculated as the average amount of time in minutes and seconds that patients took to respond to each PRO question. Mobile phone inputs were calculated as raw counts of PRO questions sent by the i-Matter platform and the number of responses received by patients. Missing data were calculated as the percentage of PRO questions that had no response by patients, and late responses were calculated as the percentage of messages sent by patients that was outside a 1-hour window. In addition, we calculated frequencies for the TAM3 survey questions.

Following the analysis of use data, the research team categorized each issue with usability as either critical (abandon or remove), severe (significant delay or frustration in task completion requiring revision), or cosmetic (minor issue). Each of these issues was mapped onto the transcripts and survey responses to provide specific and detailed recommendations for refining i-Matter before proceeding to the next testing session.

Clinical Efficacy Phase

Baseline characteristics and outcomes were summarized descriptively using mean values and standard deviation or frequency descriptions. The statistical modeling of the baseline and 12-month outcome variables was based on linear mixed models (for continuous outcomes), logistic generalized linear mixed models (for binary outcomes), and random effects multinomial models (for outcomes with more than 2 levels, like adherence). In all models, clinic site, time (2 dummy variables), and intervention were included as fixed effects, and participant was the random effect. The intervention effect of interest is the treatment X time interaction in this model. Covariates were included as necessary in adjusted analyses. All tests were 2-sided 0.05 level tests.

Primary Aim: To compare the efficacy of iMatter vs. UC on HbA1c reduction at 12-months.

A generalized linear mixed effects model was used to measure HbA1c as a dependent variable. The model included a categorical indicator variable for the randomized study arm, a variable corresponding to time (baseline as reference), and an interaction term between the group assignment and time. A patient-level random intercept was also included in the model. The results represent the complete case analysis for all participants with a baseline measurement up to 180 days prior to the date of study enrollment and a follow-up measurement within 180 days (prior to or after) 12 months since the study enrollment date. Note: Further analyses will be conducted using multiple imputation as well as subgroup analyses. This analysis represents the most conservative approach to analyzing the primary outcome.

Secondary Aims:

1. To compare the efficacy of iMatter vs. UC on adherence to self-care behaviors.
2. To evaluate the potential mediators of the intervention effects on adherence to self-care and HbA1c (only if primary aim is statistically significant).

For secondary aim 1, analyses examined the intervention effect individually on each self-care behavior. Generalized linear mixed effects models were used to model each of the six domain scores as the dependent variable. Five of the six domains were modeled as continuous outcomes, whereas smoking status was modeled as a binary outcome. Each model included a categorical indicator variable for the randomized study arm, a variable corresponding to time (baseline as reference), and an interaction term between randomization group assignment and time. A patient-level random intercept was included in each model. The analysis utilized data imputation. If the patient was missing a follow-up measurement at 12 months, the measurement at 9 months was used.

For secondary aim 2, the analysis represents the complete case analysis across all 12 self-efficacy scale questions and the distress questions respectively. Generalized linear mixed effects models were used to model the responses from each of the 14 unique questions as the dependent variable. Each model included a

categorical indicator variable for the randomized study arm, a variable corresponding to time (baseline as reference, 12 months as the primary endpoint), and an interaction term between randomization group assignment and time. A patient-level random intercept was included in each model. If the primary analyses are significant, a path model will be conducted to analyze the mediating effects of self-efficacy and distress on HbA1c and adherence to self-care behaviors.

RESULTS

Principal Findings and Outcomes

Formative Phase

1. Adapt to patients' and providers' needs: We invited 55 patients with T2D (22 men, 33 women) to participate in the focus groups, of which 35 (63.6%) declined participation leaving 20 potential participants. Reasons for declining participation included being too busy, limitations due to other comorbid conditions, personal/family constraints, and not being interested in participating in research. Of the 20 who agreed to participate, 12 (60%) attended one of the focus groups; 1 did not attend due to scheduling conflict with work, and 7 stopped responding to the RA's outreach calls. We held four focus groups: two for English-speaking patients (n=6) and two for Spanish-speaking patients (n=6). **Table 1** describes the socio-demographic characteristics of focus group participants and their comfort with technology

Table 1. Sociodemographic characteristics and comfort with technology survey responses among focus group participants (N=12)

Sociodemographic Characteristic	
Mean age (SD)	62.5 (5.6)
Female, n (%)	8 (66.7)
Hispanic, n (%)	7 (58.3)
Race, n (%)	
White	5 (41.7)
Black	3 (25.0)
Asian	1 (8.3)
Other	4 (25.0)
Education, n (%)	
Less than high school	1 (8.3)
High school degree	4 (33.3)
Some college	2 (16.7)
College degree	5 (41.7)
Employed, n (%)	4 (33.3)
Retired, n (%)	4 (33.3)
Annual income <$25,000, n (%)	7 (58.3)
Mean HbA1c (SD)	7.95 (0.8)
Technology Survey	
Currently uses text messaging	7 (58.3)
Has an unlimited text-messaging plan	12 (100.0)
Always has mobile phone with them	9 (75.0)
Comfortable downloading apps on their mobile phone	7 (58.3)
Comfortable receiving and responding to text messages about T2D	8 (66.7)
Interested in using mobile phone to help keep track of T2D	7 (58.3)
Challenges to tracking T2D on mobile phone	
Cost of receiving messages	2 (16.7)
Unreliable internet access	1 (8.3)
Don't use mobile phone regularly	3 (25.0)
Unsure of benefit	4 (33.3)
Concerns about privacy and security	2 (16.7)

Analysis of the patient focus groups identified 4 core themes: (1) patients felt as though their lives were controlled by their blood sugar values; (2) patients' greatest fear of having T2D were vision loss, kidney failure, or risk of amputation, and avoiding these consequences served as motivators for behavior change; (3) important goals for patients were being in control of their T2D, feeling well, living a long healthy life, and

eventually not needing medications for T2D (owing to concerns about the negative long-term effects); and (4) forgetfulness, poor dietary adherence, physical inactivity, tiredness or fatigue, and poor emotional health were viewed as major barriers to keeping blood sugar in control. Patients in the Spanish-speaking focus groups also spoke about God being an important source of strength and motivation to improve their health.

We conducted six provider interviews (50% female, four primary care providers, one endocrinologist, and one general surgeon and weight management specialist). Analysis of the interviews identified the central theme that providers want PRO data that is specific and actionable and can help them focus the clinic visit on what is most important for their T2D patients' care.

2. Integrate PROs into patient lives and provider workflows: Using an iterative user-centered process resulted in the identification of the six PRO messages that were relevant to patients and providers (n=12): medication adherence; dietary behaviors; physical activity; sleep quality; quality of life; and healthy living goals. In user testing, patients recommended improvements to the wording and timing of the PRO text messages to increase clarity and response rates. Patients also recommended including motivational text messages to help sustain engagement with the program. The personalized report was regarded as a key tool for diabetes self-management by patients and providers because it aided in the identification of longitudinal patterns in the PRO data, which increased patient awareness of their need to adopt more healthful behaviors. Patients recommended adding individualized tips to the journal on how they can improve their behaviors. Providers preferred having a separate tab built into the EHR that included the personalized report and highlighted key trends in patients' PRO data across the past three months. Based on this feedback, the study team partnered with the NYULH Medical Center IT to integrate the personalized report into Epic. This included the development of security protocols that link patients' encrypted research ID to their medical record number and integration of the report image into an Epic web integration record. Web integration records are used to visually integrate external applications within Epic.

3. Evaluate the usability: We completed two rounds of user testing with patients: 7 patients completed the first round of testing (1 Spanish-speaking) and 3 patients completed the second round. **Table 2** presents results of the use behavior data for both rounds of user-testing.

Table 2. Patient text messaging use behavior during user-testing

Use behavior	User-testing Round 1 (N=7)	User-testing Round 2 (N=3)
Task Success	90.6%	93.1%
Time-on-task	44 minutes (range: 0-661.6 minutes)	20 minutes (range: 0.08 minutes-30 minutes)
Missed Responses	39.2%	15.0%
Late Responses	19.3%	7.5%
Invalid responses	9.4%	6.9%

In qualitative interviews, patients in both rounds of user-testing described the program as easy to use, not intrusive to their daily life, and helpful for managing their T2D. Similar findings were seen in the TAM3 survey responses (**Table 3**). Patients also liked the consistency in message timing; because it helped them build a habit to respond *("it becomes second nature").* Several patients commented that they felt as though a person was sending the messages to check up on them. Patients also felt the number of messages sent was adequate with two people commenting, "*No number is too many because they want to get better*." There were no differences in qualitative feedback or TAM3 responses between the English and Spanish-speaking patients.

Table 3. Response to TAM3 survey questions

PRO Questions (n=7)	n (%) agreeing with statement
I would use the i-Matter program in future	5 (71.4%)
The PRO questions are very helpful for managing T2D	6 (85.7%)
Receiving and responding to PRO questions was easy	7 (100%)
I responded to the PRO questions all of the time	5 (71.4%)
I would recommend i-Matter to friends and family	7 (100%)
My provider would be more effective managing T2D with my PRO data	5 (71.4%)
Overall, the i-Matter program is great or excellent	6 (85.7%)

Personalized Report (n=9)	n (%) agreeing with statement
I would definitely use the personalized report in the future	8 (88.9%)
The personalized report is very helpful for managing T2D	7 (77.8%)
The personalized report is easy to use	5 (55.6%)
I would share the personalized report with friends or family	5 (55.6%)
Showing my provider the personalized journal would help make clinic visits more effective	7 (77.8%)
The charts and images are great	6 (66.7%)
Overall, the personalized report is great	6 (66.7%)

Clinical Efficacy Phase

Patient Sample

Using the recruitment methods outlined above, we used the EHR to screen a total of 758 patients who met the basic eligibility criteria (e.g., having a diagnosis of T2D, HbA1c >7%, age ≥18 years, fluent in English or Spanish) and were eligible for a second level screen. Of these patients, 479 were excluded for the following reasons: 403 declined participation and 76 were unable to reach after multiple attempts. The remaining 279 potentially eligible patients were scheduled to complete the informed consent visit. Of these patients, 25 were lost to follow up and 8 patients were subsequently deemed ineligible due to a psychiatric comorbidity or a diabetes diagnosis other than T2D (e.g., Type 1). Thus, a total of 246 patients (87.2% of our total enrollment) provided informed consent and were enrolled into the trial. Of the enrolled patients, 9 withdrew from the study prior to randomization. A total of 237 patients were randomized to one of the two study arms: 117 to the iMatter intervention and 120 to the UC group.

Ninety-one percent of patients enrolled in the study completed the 3-month study visit; 90% completed the 6-month visit; 81% completed the 9-month visit; and 84% completed the final study visit at 12-months.

Baseline Characteristics

About half (49.4%) of patients were female with a mean age of 56.6 (standard deviation [SD]: 9.63) years. Approximately three-quarters (71.3%) of patients were born in the US, half were married (47.7%), and more than three-quarters had at least some college or above (79.4%). About one-third (39.4%) of patients reported being unemployed, of which about half (55.6%) were retired. A majority of patients had private insurance (62.6%) and approximately three-quarters (73%) reported an annual household income ≥$40,000. Most (84.5%) patients reported their primary language as English.

Patient engagement with iMatter

On average, participants responded to 71% of the text messages they received each day (range: 65%-85%). The highest response rate was for the healthy diet message (85% followed by the medication adherence message (82%). We have also begun to analyze the user data for patients who have completed the trial to understand the patterns of engagement over the 12-month study. This analysis includes examining overall engagement with the program (e.g., global response rate) and the patient sociodemographic factors that influence engagement. We also examined engagement to each individual PRO question (e.g., response rate to sleep quality message, medication adherence message, etc). Latent process mixed models were implemented to examine the trajectory of users' response rate to i-Matter PRO messages. From the 39 participants, the above model classifies 24 as latent class or group 1 and the remaining 15 as group 2. Overall, the trajectories suggest user responses to PRO messages were found to be significantly different between the two groups in all cross-sections. Group 1 took lesser time to respond, responded to more messages, and were more engaged with the program than group 2. Moreover, the trajectories of user engagement suggest group 1 maintained high level of engagement throughout participation in the program, while group 2 started with slightly lower level of engagement than group 1 and their engagement declined steadily as the time progressed. Similar patterns emerged when examining responses to the individual PROs.

Results of the Primary Aim

The following results describe differences in HbA1c levels between the control and intervention participants from baseline to 12 months follow-up. A generalized linear mixed effects model was used to measure HbA1c as a dependent variable. The model included a categorical indicator variable for the randomized study arm, a variable corresponding to time (baseline as reference), and an interaction term between the group assignment and time. A patient-level random intercept was also included in the model. The following results represent the complete case analysis for all participants with a baseline measurement up to 180 days prior to the date of study enrollment and a follow-up measurement within 180 days (prior to or after) 12 months since the study enrollment date.

Table 4. HbA1c levels ANOVA Results (n = 176)

Variable	p-value
Time	0.566
Group	0.521
Time x Group	0.536

Table 5. HbA1c levels stratified by group and time (n = 176)

Group Assignment	Baseline HbA1c	Last HbA1c
Control	8.2%	8.1%
Intervention	8.3%	8.1%

These results indicate that there was no statistically significant change in HbA1c levels from baseline to 12 months, nor was there a statistically significant difference between the HbA1c levels of the intervention and control groups.

We further refined these analyses on a reduced subpopulation that limited follow-up measurements within 90 days (prior to or after) 12 months since the study enrollment date:

Table 6. HbA1c levels ANOVA Results (n = 166)

Variable	p-value
Time	0.476
Group	0.731
Time x Group	0.439

Table 7. HbA1c levels stratified by group and time (n = 166)

Group Assignment	Baseline HbA1c	Last HbA1c
Control	8.2%	8.1%
Intervention	8.3%	8.0%

While there was a greater decrease in the intervention arm from baseline to follow-up than in the control arm, the results were not statistically significant.

Results of the Secondary Aims

Diabetes Self-Care Results:

The following describe differences in diabetes self-care activities between control and intervention participants using the questions and scoring criteria of Toobert et al. 2000.(29) Diabetes self-care activities are presented across seven domains, namely diet (general and specific), exercise, blood sugar testing, foot care, smoking status, and medication adherence. In each of the figures, purple trajectories represent the control arm, and green trajectories represent the intervention.

General Diet:

Questions:

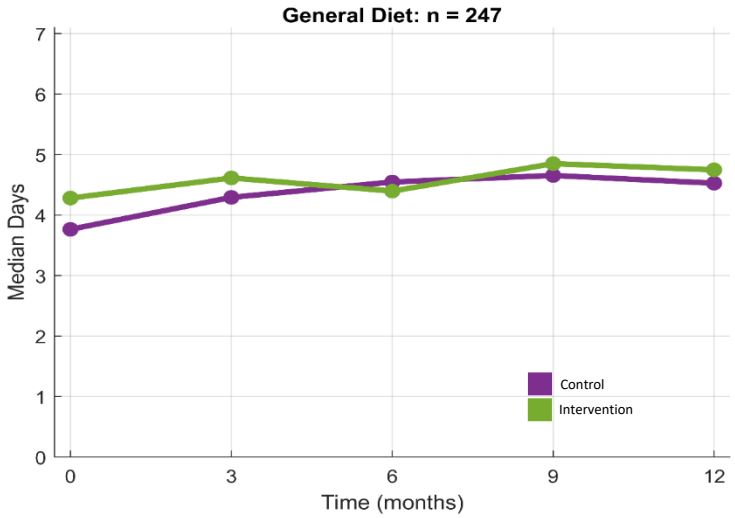
- How many of the last seven days have you followed a healthful eating plan?
- On average, over the past month, how many days per week have you followed your eating plan?

Scoring: Mean number of days across these two questions.

Table 8. General Diet ANOVA Results

Variable	p-value
Time	< 0.001
Group	0.065
Time x Group	0.332

Figure 2. Population means for general diet scores stratified by control vs. intervention groups



These results indicate that there was no statistically significant difference in specific diet between the control and intervention groups, but there was a statistically significant improvement in general diet from baseline to 12 months across both groups.

Specific Diet:

Questions:

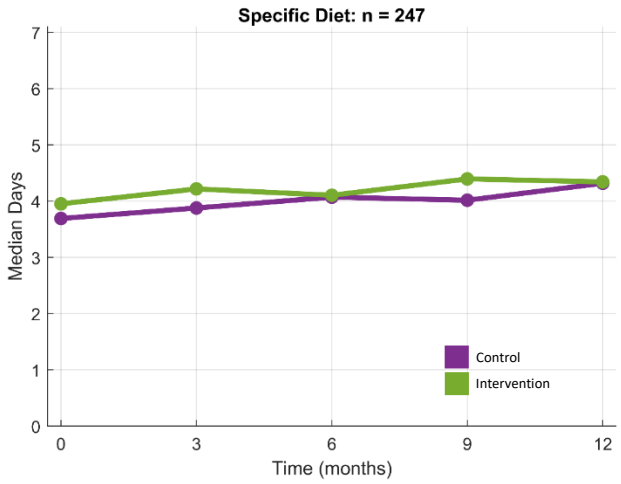
- On how many of the last seven days did you eat five or more servings of fruits and vegetables?
- On how many of the last seven days did you eat high fat foods such as red meat or full-fat dairly products?

Scoring: Mean number of days across these two questions (reversing the second question).

Table 9. Specific Diet ANOVA Results

Variable	p-value
Time	< 0.001
Group	0.214
Time x Group	0.161

Figure 3. Population means for specific diet scores stratified by control vs. intervention groups



These results indicate that there was no statistically significant difference in specific diet between the control and intervention groups, but there was a statistically significant improvement in specific diet from baseline to 12 months across both groups.

Exercise:

Questions:

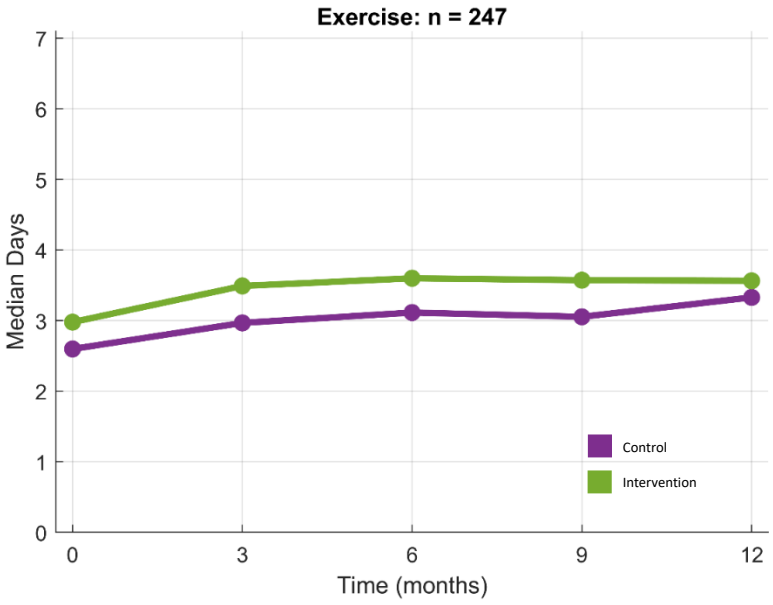
- On how many of the last seven days did you participate in at least 30 minutes of physical activity? (Total minutes of continuous acitivity, including walking)
- On how many of the last seven days did you participate in a specific exercise session (such as swimming, walking, biking) other than what you do around the house or as part of your work?

Scoring: Mean number of days across these two questions.

Table 10. Exercise ANOVA Results

Variable	p-value
Time	0.004
Group	0.213
Time x Group	0.965

Figure 4. Population means for exercise scores stratified by control vs. intervention groups



These results indicate that there was no statistically significant difference in exercise between the control and intervention groups, but there was a statistically significant improvement in exercise from baseline to 12 months across both groups.

Blood glucose:

Questions:

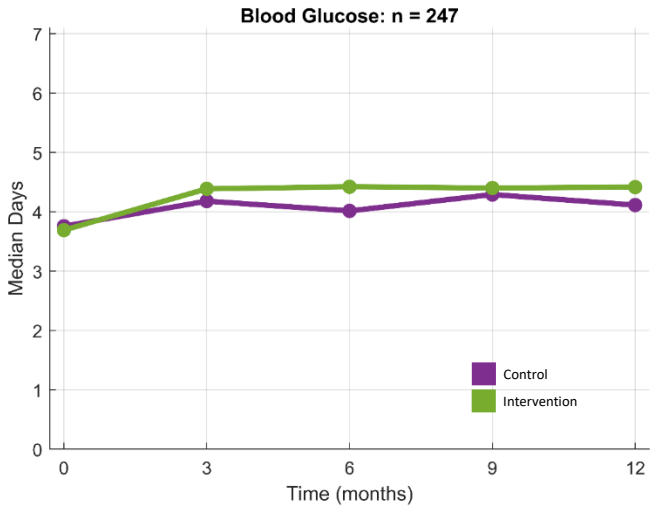
- On how many of the last seven days did you test your blood sugar?
- On how many of the last seven days did you test your blood sugar the number of times recommended by your health care provider?

Scoring: Mean number of days across these two questions.

Table 11. Blood glucose ANOVA Results

Variable	p-value
Time	0.459
Group	0.918
Time x Group	0.096

Figure 5. Population means for blood glucose scores stratified by control vs. intervention groups



These results indicate there is no statistically significant difference in blood glucose between groups and time from baseline to 12 months.

Foot care:

Questions:

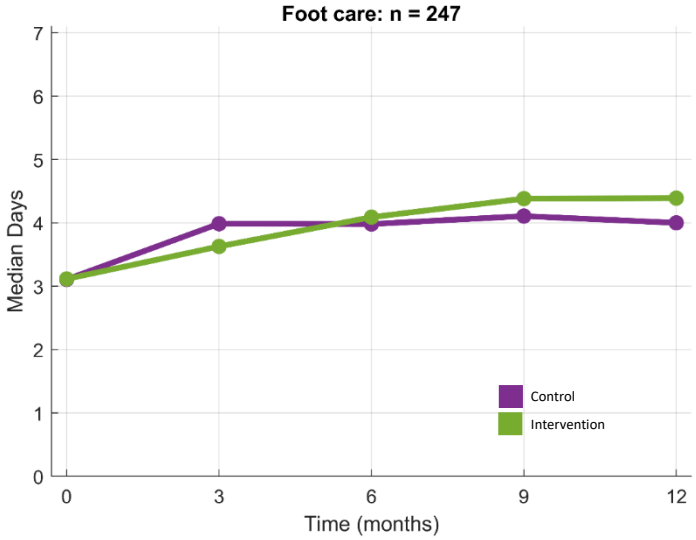
- On how many of the last seven days did you check your feet?
- On how many of the last seven days did you inspect the inside of your shoes?

Scoring: Mean number of days across these two questions.

Table 12. Foot care ANOVA Results

Variable	p-value
Time	<0.001
Group	0.706
Time x Group	0.308

Figure 6. Population means for foot care scores stratified by control vs. intervention groups



These results indicate that there was no statistically significant difference in foot care between the control and intervention groups, but there was a statistically significant difference in time from baseline to 12 months across both groups.

Smoking status:

Questions:

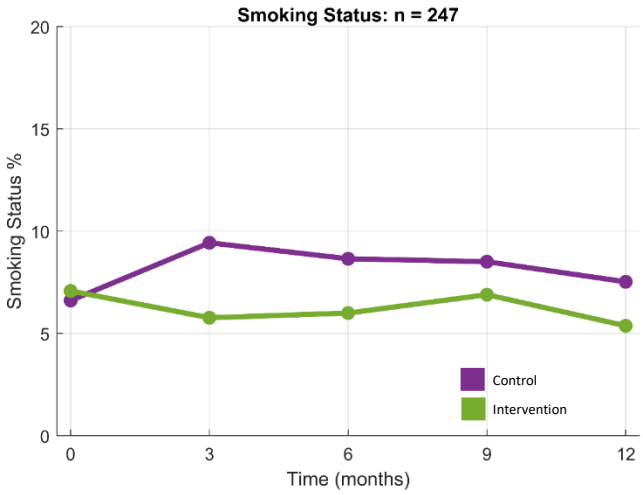
- Have you smoked a cigarette – even one puff – during the past seven days?

Scoring: Frequency among participants (0 = non-smoker, 1 = smoker)

Table 13. Smoking status ANOVA Results

Variable	**p-value**
Time	<0.001
Group	0.002
Time x Group	**<0.001**

Figure 7. Frequency of smoking status stratified by control vs. intervention groups



These results indicate that **there was a statistically significant decrease in smoking status between the control and intervention groups, with a greater reduction in the fraction of smokers in the intervention group relative to the control group**.

Medication Adherence:

Questions:

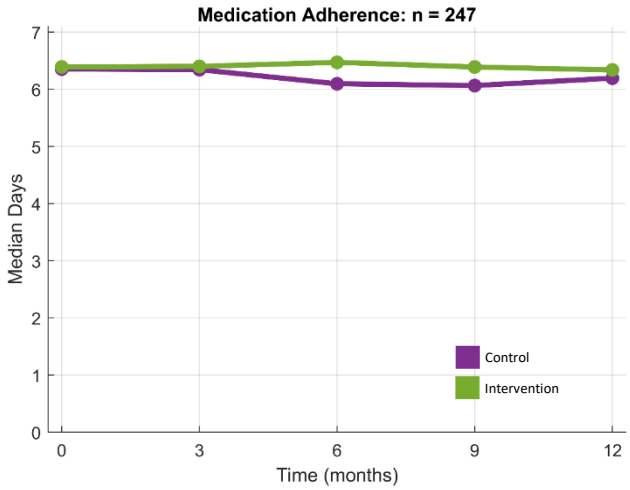
- On how many of the last seven days did you take your recommended diabetes medication?

Scoring: Number of days

Table 14. Medication adherence ANOVA Results

Variable	p-value
Time	0.369
Group	0.857
Time x Group	0.618

Figure 8. Population means for number of days diabetes medication was taken in past week



These results indicate that there was no statistically significant difference in the number of days that the recommended diabetes medication was taken between groups, nor was there a statistically significant difference in the number of days that the recommended diabetes medication was taken from baseline to 12 months.

Diabetes Self-Efficacy and Distress Results:

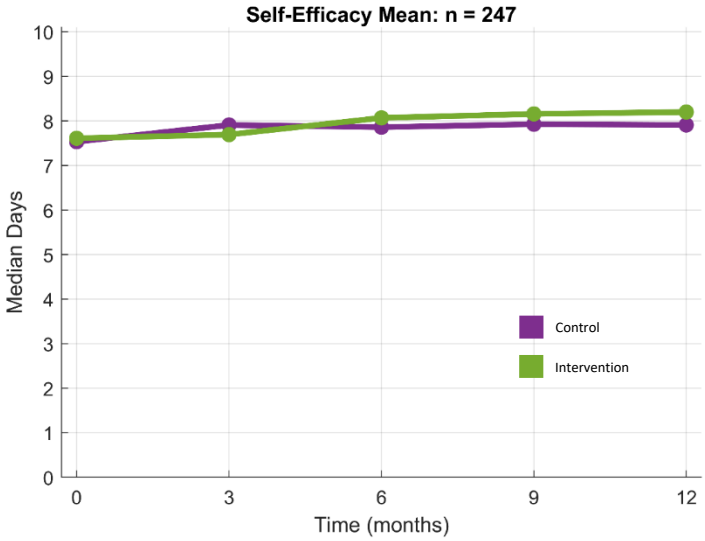
The following results describe differences in diabetes self-efficacy (motivation) responses between control and intervention participants, as well as distress with maintaining a diabetes regimen, and the burden of the demands of living with diabetes. The 12-question Diabetes Self-Efficacy Scale was assessed using an 11-point response from 0 = “Not at all confident” to 10 = “Very confident.” Diabetes distress and burden were assessed using a 6-point response scale where 1 = “Not a Problem” to 6 = “Very Serious Problem.” Purple trajectories represent the control arm, and green trajectories represent the intervention.

Single Measure Summary of Self-Efficacy Questionnaire: The following results are for the mean of all 12 self-efficacy questions.

Table 15. Summary mean of self-efficacy questionnaire

Variable	p-value
Time	0.270
Group	0.995
Time x Group	0.112

Figure 9. Population means for all 12 self-efficacy questions as a single summary statistic stratified by control vs. intervention groups



These results indicate that there was no statistically significant difference in overall self-efficacy between the control and intervention groups, nor was there a change in overall self-efficacy across time.

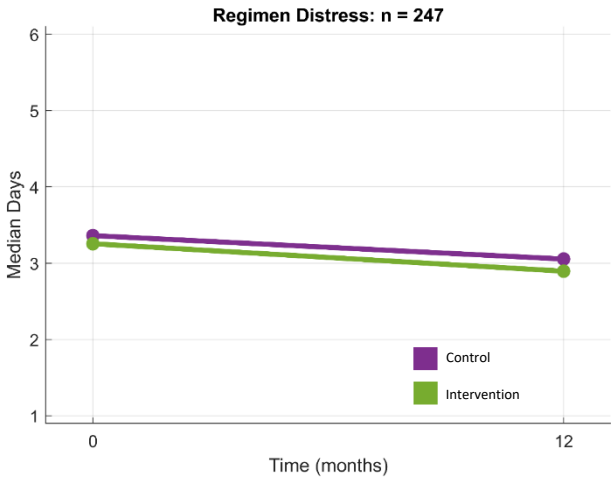
Diabetes distress:

Question: Consider the degree to which the following may have distressed or bothered you during the past month: Feeling that you are often failing with your diabetes regimen.

Table 16. Diabetes distress

Variable	p-value
Time	0.033
Group	0.619
Time x Group	0.947

Figure 10. Population means for diabetes distress stratified by control vs. intervention groups



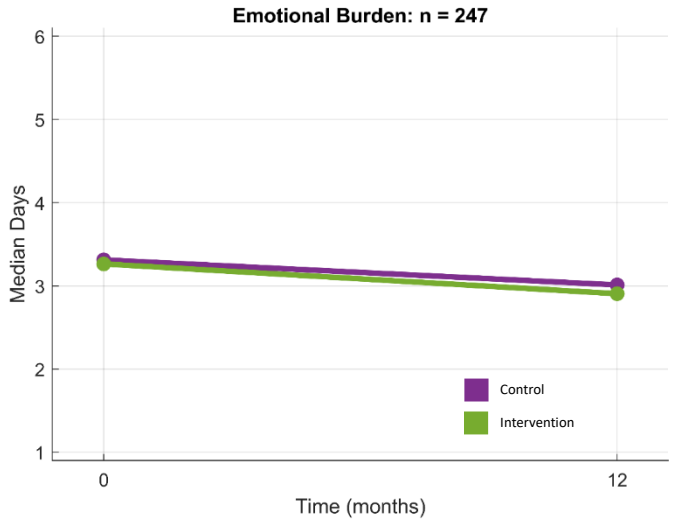
These results indicate that there was no statistically significant difference in diabetes distress between the control and intervention groups, but there was a statistically significant decrease in diabetes distress across time in both groups.

Question: Consider the degree to which the following may have distressed or bothered you during the past month: Feeling overwhelmed by the demands of living with diabetes.

Table 17. Diabetes distress

Variable	p-value
Time	0.077
Group	0.879
Time x Group	0.830

Figure 11. Population means for diabetes burden stratified by control vs. intervention groups



These results indicate that there was no statistically significant difference in emotional burden between the control and intervention groups, nor was there a statistically significant decrease in emotional burden across time.

Discussion

Although there was a small decrease in HbA1c levels from baseline to 12 months in the intervention group, this change was not statistically significant relative to the control group. These results remained non-significant statistically for a reduced subpopulation that restricted follow-up A1c measures to within 90 days of the 12-month follow-up endpoint. All analyses were on an intention to treat population and utilized complete cases only. Future analyses will explore multiple imputation for participants who are only missing either a baseline or follow-up measurement and will explore whether statistical differences in HbA1c levels are present among select patient subgroups stratified by socio-demographic and clinical characteristics.

Among the self-care results, smoking status was the only behavior to significantly decrease statistically from baseline to 12 months in the intervention arm relative to the control arm. Detection of a statistically significant mediation variable requires (1) a relationship between the study group assignment and primary outcome (HbA1c levels), (2) a relationship between the study group assignment and the mediating variable, and (3) a relationship between the mediating variable and HbA1c levels. Because there was no relationship between the study group assignment and primary HbA1c outcome, nor was there a relationship between the study group assignment and self-efficacy or diabetes distress responses, none of these measures serve as a mediating variable between the study group assignment and HbA1c levels. Future analyses will again explore stratifications by patient socio-demographic and clinical variables.

Conclusions

On a study population average level, the secondary outcome of smoking status was the only statistically significant finding (greater decrease from baseline to 12 months follow-up in the intervention group relative to the control group). Follow-up exploratory analyses will examine whether select patient phenotypes may have benefited more from the intervention.

Significance

Despite incorporation of end-users in the design, one-size-fits all models of intervention delivery may not be potent enough to make a measurable impact on clinical outcomes in patients with type 2 diabetes.

Implications

Future research should explore whether tailored approaches that dynamically capture and share patient PRO data overtime can increase the efficacy of digital health behavioral interventions.

LIST OF PUBLICATIONS and PRODUCTS

Schoenthaler A, Cruz J, Payano L, Rosado M, Labbe K, Johnson C, Gonzalez J, Patxot M, Patel S, Leven E, Mann D. Investigation of a Mobile Health Texting Tool for Embedding Patient-Reported Data Into Diabetes Management (i-Matter): Development and Usability Study. JMIR Form Res. 2020 Aug 31;4(8):e18554. doi: 10.2196/18554.

Mandal S, Belli H, Cruz J, Mann D, Schoenthaler A. Analyzing User Engagement Within a Patient-Reported Outcomes Texting Tool for Diabetes Management: Engagement Phenotype Study. JMIR Diabetes 2022;7(4):e41140. doi: 10.2196/41140

Marini C, Cruz J, Payano L, Flores R, Arena G, Mandal S, Leven E, Mann D, Schoenthaler A. Opening the Black Box of an mHealth Patient-Reported Outcome Tool for Diabetes Self-Management: Interview Study Among Patients With Type 2 Diabetes. JMIR Form Res 2023;7:e47811. doi: 10.2196/47811

Work Cited

1. Fox CS, Golden SH, Anderson C, Bray GA, Burke LE, de Boer IH, et al. Update on Prevention of Cardiovascular Disease in Adults With Type 2 Diabetes Mellitus in Light of Recent Evidence: A Scientific Statement From the American Heart Association and the American Diabetes Association. Diabetes Care. 2015;38(9):1777-803.
2. Geiss LS, Wang J, Cheng YJ, Thompson TJ, Barker L, Li Y, et al. Prevalence and incidence trends for diagnosed diabetes among adults aged 20 to 79 years, United States, 1980-2012. JAMA. 2014;312(12):1218-26.
3. Prevention. CfDCa. National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States, 2014. . In: Promotion NCfCDPaH, editor. Atlanta, GA: U.S. Department of Health and Human Services; 2014.
4. Resnick HE, Foster GL, Bardsley J, Ratner RE. Achievement of American Diabetes Association clinical practice recommendations among U.S. adults with diabetes, 1999-2002: the National Health and Nutrition Examination Survey. Diabetes Care. 2006;29(3):531-7.
5. Ali MK, Bullard KM, Gregg EW. Achievement of goals in U.S. Diabetes Care, 1999-2010. The New England journal of medicine. 2013;369(3):287-8.
6. Boltri JM, Okosun IS, Davis-Smith M, Vogel RL. Hemoglobin A1c levels in diagnosed and undiagnosed black, Hispanic, and white persons with diabetes: results from NHANES 1999-2000. Ethnicity & disease. 2005;15(4):562-7.
7. Ali MK, McKeever Bullard K, Imperatore G, Barker L, Gregg EW. Characteristics associated with poor glycemic control among adults with self-reported diagnosed diabetes--National Health and Nutrition Examination Survey, United States, 2007-2010. Mmwr.61 Suppl:32-7.
8. McWilliams JM, Meara E, Zaslavsky AM, Ayanian JZ. Differences in control of cardiovascular disease and diabetes by race, ethnicity, and education: U.S. trends from 1999 to 2006 and effects of medicare coverage. Annals of internal medicine. 2009;150(8):505-15.
9. Beckles GL, Chou CF. Disparities in the Prevalence of Diagnosed Diabetes - United States, 1999-2002 and 2011-2014. MMWR Morb Mortal Wkly Rep. 2016;65(45):1265-9.
10. Cheng L, Sit JW, Choi KC, Chair SY, Li X, He XL. Effectiveness of Interactive Self-Management Interventions in Individuals With Poorly Controlled Type 2 Diabetes: A Meta-Analysis of Randomized Controlled Trials. Worldviews Evid Based Nurs. 2017;14(1):65-73.
11. Franch-Nadal J, Labrador Barba E, Gomez-Garcia MC, Buil-Cosiales P, Millaruelo JM, Pena ML. Patient-reported outcomes in type 2 diabetes mellitus: patients' and primary care physicians' perspectives in the Spanish health care system. Patient preference and adherence. 2015;9:1413-22.
12. Barsdorf AI, Rubinstein, E., Jaksa, A., editor Patient-Reported Outcomes (PROs) in Type 2 Diabetes Clinical Trials. Context Matters; 2014.
13. Ridgeway JL, Beebe TJ, Chute CG, Eton DT, Hart LA, Frost MH, et al. A brief Patient-Reported Outcomes Quality of Life (PROQOL) instrument to improve patient care. PLoS medicine. 2013;10(11):e1001548.
14. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. Journal of evaluation in clinical practice. 2006;12(5):559-68.

15. Hoppe CD, Cade JE, Carter M. An evaluation of diabetes targeted apps for Android smartphone in relation to behaviour change techniques. J Hum Nutr Diet. 2017;30(3):326-38.
16. Chomutare T, Fernandez-Luque L, Arsand E, Hartvigsen G. Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. J Med Internet Res. 2011;13(3):e65.
17. Baron JS, Hirani S, Newman SP. A randomised, controlled trial of the effects of a mobile telehealth intervention on clinical and patient-reported outcomes in people with poorly controlled diabetes. Journal of telemedicine and telecare. 2016.
18. Schoenthaler A, Cruz J, Payano L, Rosado M, Labbe K, Johnson C, et al. Investigation of a Mobile Health Texting Tool for Embedding Patient-Reported Data Into Diabetes Management (i-Matter): Development and Usability Study. JMIR Form Res. 2020;4(8):e18554.
19. Sasseville M, Chouinard MC, Fortin M. Patient-reported outcomes in multimorbidity intervention research: A scoping review. Int J Nurs Stud. 2018;77:145-53.
20. Kavanaugh A, Reese DD, Pérez-Quiñones M. Mobile Phone as Scaffolding Technology: How low literacy groups learn computing?
21. Center PR. Pew Internet: Mobile Fact Sheet2017.
22. Fox S DM. Mobile Health 20122012.
23. Blumberg SJ, Luke, J.V. Early Release of Estimates From the National Health Interview Survey, January–June 2015. In: Division of Health Interview Statistics NCfHS, editor.: U.S. Department of Health and Human Services; 2015.
24. Malterud K, Siersma VD, Guassora AD. Sample Size in Qualitative Interview Studies: Guided by Information Power. Qual Health Res. 2016;26(13):1753-60.
25. Chokshi SK, Belli HM, Troxel AB, Blecker S, Blaum C, Testa P, Mann D. Designing for implementation: user-centered development and pilot testing of a behavioral economic-inspired electronic health record clinical decision support module. Pilot Feasibility Stud. 2019;5:28.
26. Chokshi SK, Mann DM. Innovating From Within: A Process Model for User-Centered Digital Development in Academic Medical Centers. JMIR Hum Factors. 2018;5(4):e11048.
27. Chokshi SK, Troxel A, Belli H, Schwartz J, Blecker S, Blaum C, et al. User-Centered Development of a Behavioral Economics Inspired Electronic Health Record Clinical Decision Support Module. Stud Health Technol Inform. 2019;264:1155-8.
28. Joe J, Chaudhuri S, Le T, Thompson H, Demiris G. The use of think-aloud and instant data analysis in evaluation research: Exemplar and lessons learned. J Biomed Inform. 2015;56:284-91.
29. Toobert DJ, Hampson SE, Glasgow RE. The summary of diabetes self-care activities measure: results from 7 studies and a revised scale. Diabetes Care. 2000;23(7):943-50.