TBM

ORIGINAL RESEARCH

Fitness facility staff demonstrate high fidelity when implementing an evidence-based diabetes prevention program

Tineke E. Dineen[1], Tekarra Banser[1], Corliss Bean[2], Mary E. Jung[1]

[1]School of Health and Exercise Sciences, University of British Columbia, Okanagan Campus, 3333 University Way, Kelowna, British Columbia V1V 1V7, Canada

[2]Department of Recreational and Leisure Studies, Brock University 1812 Sir Isaac Brock Way, St Catharines, Ontario L2S 3A1, Canada

Correspondence to: Mary E. Jung, mary.jung@ubc.ca

Cite this as: *TBM* 2021;11:1814–1822
doi: 10.1093/tbm/ibab039

© Society of Behavioral Medicine 2021. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

Abstract

Translating evidence-based diabetes prevention programs into the community is needed to make promising interventions accessible to individuals at-risk of type 2 diabetes. To increase the likelihood of successful translation, implementation evaluations should be conducted to understand program outcomes and provide feedback for future scale-up sites. The purpose of this research was to examine the delivery of, and engagement with, an evidence-based diet and exercise diabetes prevention program when delivered by fitness facility staff within a community organization. Ten staff from a community organization were trained to deliver the diabetes prevention program. Between August 2019–March 2020, 26 clients enrolled in the program and were assigned to one of the ten staff. Three fidelity components were accessed. First, staff completed session-specific fidelity checklists (n = 156). Second, two audio-recorded counseling sessions from all clients underwent an independent coder fidelity check (n = 49). Third, staff recorded client goals on session-specific fidelity checklists and all goals were independently assessed for (a) staff goal-setting fidelity, (b) client intervention receipt, and (c) client goal enactment by two coders (n = 285). Average self-reported fidelity was 90% for all six sessions. Independent coder scores for both counseling sessions were 83% and 81%. Overall staff helped clients create goals in line with program content and had a goal achievement of 78%. The program was implemented with high fidelity by staff at a community organization and clients engaged with the program. Findings increase confidence that program effects are due to the intervention itself and provide feedback to refine implementation strategies to support future scale-up efforts.

Keywords

Program evaluation, Implementation science, Health behavior, Prediabetic state, Diet, Exercise

Four hundred, sixty-three million adults are living with diabetes worldwide and 374 million people are at increased risk of developing type 2 diabetes (T2D) [1]. With T2D on the rise, prevention is a priority. Diet and exercise modifications remain the foundation of T2D prevention [2] with demonstrated efficacy [3]. Despite this strong evidence,

Implications

Researchers: Future researchers should conduct comprehensive fidelity assessments to ensure community-based programs are implemented with fidelity.

Practitioners: Community organizations with trained staff can deliver diabetes prevention programs with high fidelity.

Policymakers: Fitness centers may be viable community-based venues to deliver evidence-based diabetes prevention programming.

Lay summary

Interventions that focus on diet and exercise modifications have demonstrated diabetes risk-reducing results. However, transitioning these programs from research to community contexts is challenging. Implementation evaluation can offer feedback for program improvement and understanding program outcomes. Ten fitness facility staff were trained, and 26 clients engaged in the diabetes prevention program. Staff completed session-specific checklists documenting what they covered in each session. An independent research assistant listened to two audio-recorded sessions per client and completed identical checklists to check the accuracy of staff completed checklists. Finally, staff documented client goals, and goals were assessed to see whether (a) staff helped clients make goals in line with their training, (b) clients made goals in line with the program content, and (c) clients completed their goals. Staff reported providing 90% of the intervention to their clients; when compared to the audio recordings, 82%. Overall, staff helped clients create goals in line with program content and clients often achieved their goals. In sum, fitness facility staff at a community organization can be trained to deliver a diabetes prevention program as planned and clients engage with the program as it was intended to be engaged with.

Downloaded from academic.oup.com/tbm/article/11/10/1814/6272494 by guest on 15 October 2026

costly interventions designed for clinical trials are not feasible to implement in real-world settings [4]. To better reach individuals living at risk, programs must be implemented into accessible and sustainable community settings.

There have been multiple studies examining the translation of diabetes prevention programs in real-world settings (e.g., [5, 6]) and community settings (e.g., [7]), the largest to date being the dissemination of the United States National Diabetes Prevention Program [8, 9], a translation program from the original United States Diabetes Prevention Program [3]. However, many of these studies focus predominantly on physiological outcomes (e.g., weight, HbA1c). Few studies to date engaged in implementation evaluations to understand conditions that led to effective implementation and understanding the quality in which diabetes prevention programs are implemented. From the limited work that exists, the implementation of the National Diabetes Prevention Program was examined, identifying promising features of program delivery, including recruitment, program adaptations, and techniques to engage participants [8]. However, this study did not provide in-depth information on how or why such features were implemented and lacked any assessments of the fidelity of implementation. Such information is essential for program improvement and understanding mechanisms leading to program outcomes.

Conducting implementation evaluations can help close this gap. Such an evaluation examines the extent an intervention is delivered as intended per protocol and increases confidence that intervention results are due to the intervention [10, 11]. Knowing exactly what was implemented, and how, is critical to understanding the internal and external validity of translating evidence-based interventions into practice [10]. The National Institute for Health Behavior Change Consortium identified five intervention fidelity components: (a) study design, (b) training providers, (c) delivery of treatment, (d) receipt of treatment, and (e) enactment of treatment [10]. Together, these five components provide a thorough understanding of how and why an intervention produces (or does not produce) intervention effects. First, the study design must adequately test the active ingredients. Second, staff must be satisfactorily trained to deliver the program to clients. Third, the program must be delivered as intended per protocol (delivery of treatment). Fourth, clients must understand the content provided in the program and/or perform intervention skills (receipt of treatment). Finally, clients must put the skills and information learned through the program into action in their daily life (enactment of skills; [10]). Assessing these five components is the best practice for improving the reliability and validity of complex health behavior intervention research to be translated into real-world settings [10]. Despite the value of implementation evaluations, they are rarely conducted [12]. Researchers typically only investigate treatment delivery [13–15]; intervention receipt and enactment are rarely assessed [12, 16]. Specifically, fidelity assessments are limited in the field of diabetes prevention. A systematic review on critical factors for implementing real-world diabetes prevention programs identified only one study from 38 that vaguely mentioned program components were delivered more frequently than omitted [5]. To our knowledge, there is only one similar study that assessed fidelity in a community-based, volunteer-led diabetes education program [14]. However, they did not specifically target those at risk for developing T2D and did not examine staff self-reported fidelity, client receipt or engagement. Comprehensive fidelity assessments that target multiple fidelity components are needed. Therefore, the current study addressed these gaps in the literature. As in prior research, intervention receipt and enactment are collectively referred to as intervention engagement [12].

The purpose of this research was to examine the delivery of, and engagement with, an evidence-based diet and exercise diabetes prevention program when delivered by a community organization. Prior research has studied the efficacy of the program within a laboratory setting on health-related outcomes [17] and evaluated the program training [18]. The current study evaluates the three remaining fidelity components outlined by Bellg [10]. Two guiding questions were asked: (a) to what extent do trained staff implement the program as planned; (b) to what extent do clients engage in the program? Based on the nature of the research questions, no hypotheses were proposed. Understanding how interventions are delivered and engaged with can provide feedback on program implementation, give confidence to intervention results and support future scale-up.

METHODS

Program description

After Small Steps for Big Changes demonstrated efficacy and effectiveness on health-related outcomes (e.g., improved cardiorespiratory fitness, moderate to vigorous physical activity, body fat percentage and waist circumference measured at 6- and 12-month post-program) in a laboratory setting [17], the program was translated into a community setting. To support uptake and sustainability, a 1-year collaborative planning process with the community partner, a local not-for-profit fitness facility with over 17,000 members, was undertaken. This 1-year knowledge translation process resulted in an implementation plan to translate the program into two local community sites, as a feasibility test for future scale-up [19]. Small Steps for Big Changes

Downloaded from academic.oup.com/tbm/article/11/10/1814/6272494 by guest on 15 October 2026

is a brief diabetes prevention program for individuals aged 18+ who are at risk for developing T2D based on HbA1c of 5.7%–6.4% or completion of the American Diabetes Association risk questionnaire with a score of ≥5 [20]. The goal of the program is to help clients make lasting changes to their diet and exercise behavior to lower the risk of developing T2D. Clients were recruited through a local physician referral system, social media posts, promotional materials (e.g., pamphlets, posters), program website, word-of-mouth, and community events. The program consists of six, one-on-one in-person sessions with trained staff over a 3-week period (three sessions in week 1, two sessions in week 2, one session in week 3). This structure allowed clients time to practice their self-management skills (e.g., self-regulation, routine physical activity) between sessions. Clients were encouraged to track their diet and exercise on a mobile tracking application (app) over the 3-week program period. Sessions ran between 70–90 min in duration (30-min exercise portion, 35–50-min counseling portion, 5–10-min logistics/paperwork portion for staff).

Staff training

Staff learned of the program during the one-year planning process and those interested voluntarily enrolled in the training. Any staff not available for the first training was offered to attend the second training, in addition to any new staff wanting to enroll at that later date. To sufficiently prepare for implementation, fitness facility volunteers (N =1) and staff (N = 12) attended a three-day training workshop. Attendees were given an implementation manual, including standard operating procedures, scripts, and checklists for each of the six sessions, and watched videos of research team members deliver session content. Finally, a research team member shadowed new staff with their first client and provided feedback after each session. As a pragmatic trial, all staff were retained regardless of quality. Monthly meetings were held with all trained staff and the research team. Meetings were used as an in-person community of practice to provide support, the opportunity for staff to learn from one another and to discuss program delivery, ongoing challenges, and lessons learned.

Data collection

Staff and clients

All clients and staff provided written informed consent. One staff and one volunteer dropped out of the training for unrelated reasons. One staff completed the training but did not facilitate any clients through the program due to reduced hours. Thus, 10 staff completed the training and facilitated clients through the program. Staff were on average 32.60 ± 9.06 years of age (range 24–51), were predominantly women (90%), and self-identified as Caucasian (90%). Staff were employees of the community organization for <1 year (10%), 1–5 years (40%), and 10+ years (50%). Most staff had a University certificate, diploma or degree (60%), other non-university certificate or diploma (30%) or prefer not to answer (10%). Twenty-six clients enrolled between August 2019 and March 2020 and were assigned to one of the 10 trained staff. Clients were on average 59.46 ± 7.69 years of age (range 43–70), were predominantly women (62%), and self-identified as Caucasian (85%).

Staff self-reported session fidelity

Staff completed six session-specific checklists that covered intervention components, divided into four categories: (a) logistics, (b) counseling, (c) exercise, (d) optional tools. The fidelity checklists were developed to reflect key elements from each session in the program and staff were instructed to complete them post-session. One (1) point was awarded if the item was completed and zero (0) points were awarded if the item was not completed. Items left blank were labelled "missing" and no points were awarded. This conservative approach for treating missing data has been used in one prior fidelity study [21]. Points were manually awarded if staff documented a valid reason for not completing an item (e.g., if a client was not able to exercise due to safety protocol, all exercise fidelity points were awarded). The session checklists contained between 19 and 28 compulsory items and 2–3 optional tools. Program tools (e.g., mobile tracking app, client notebook, sample nutrition labels) were designed to complement the compulsory intervention material and offered to each client. To support client autonomy, the use of program tools were optional. Feedback on checklist readability and understanding was provided by staff during their first monthly meeting and minor editorial revisions were made to ensure accurate recording of session fidelity. Completed staff checklists were entered into Qualtrics and exported to SPSS.

Independent coder counseling session fidelity check

All sessions were audio-recorded. In line with similar fidelity assessments of audio recordings [22, 23], and to ensure feasibility of this research, two sessions (33.3% of client sessions) were selected to undergo independent coder fidelity checks, staff were unaware of session selection. Sessions three (physical activity focused) and four (diet focused) were chosen as these sessions are in the middle of the intervention (rapport has been built) and both focus on a specific intervention behavior. By selecting two sessions, as opposed to a random assortment, staff comparison is facilitated as selected sessions had specific information to exchange and accounts for staff delivery of two different behavioral topics. Selecting specific sessions to complete a fidelity check has been done in previous research [23]. Only the counseling portion of the appointment was audio-recorded; thus, only the counseling and optional

Downloaded from academic.oup.com/tbm/article/11/10/1814/6272494 by guest on 15 October 2026

tools section of the fidelity checklist were completed for the independent coder fidelity check. The independent coder also answered two questions after listening to the session: (a) Please rate the client's engagement in the session and (b) Please rate the overall quality of session delivery. Both questions were rated on a scale from 0 (client did not engage in the session) to 5 (client engaged in all of the session with maximal interest; took active interest in discussion) and 0 (staff did not cover any of the session content as defined by the checklist) to 5 (staff delivered all of the session content as defined by the checklists, went over and above delivering the content in a clear and concise manner and supported client exploration of the topics). Completed independent coder fidelity checklists were entered into Qualtrics and exported to SPSS. To ensure reliability and internal consistency, 20% of the audio-recorded sessions were tested for inter-rater agreement by a trained research assistant (first author).

Independent coder goal fidelity check

Staff were trained to engage in brief action planning [24] with their client at the end of every session and document client goals on the session-specific checklists. Ideas for goals came from the client, with staff providing support to build an action plan. Clients could choose how many goals they wanted to set or decide not to set a goal. All goals were extracted from the session checklists and entered into a spreadsheet (n = 285). Each goal was assessed for (a) staff goal-setting fidelity, (b) client intervention receipt, and (c) client goal enactment independently by two coders (outlined below). All disagreements were discussed by two coders (first and second authors), and independent insight was consulted when needed (third and fourth authors).

Successful brief action planning should result in clients creating SMART (specific, measurable, attainable, relevant, time-based) goals [24]. To assess staff goal-setting fidelity, all goals were assessed according to SMART criteria. A goal qualified as *specific* if it answered at least two of the following questions: what, why, where, with who, how. A goal was *measurable* if the success or lack of success could be quantitatively measured. A goal was *attainable* if the client indicated a confidence level of ≥7 (staff were trained to ask and document client confidence levels for goals). If a client had not reported a confidence level, they received a "no" as we did not want to assume if a goal was attainable for the client. A goal was *relevant* if it was applicable to the program topics (diet, tracking, behavioral regulation and/or exercise). A goal was *time-based* if the client noted a specific date or time to achieve their goal. Each letter in the SMART acronym was dichotomized into a yes and no coding system, with a score of yes receiving one (1) point and a score of no receiving zero (0) points. A fully SMART goal resulted in a score of 5, such as, "Walk in the local park Thursday afternoon for 30 minutes."

To assess client engagement, client goals were examined for fit with intervention content (receipt) and if the client achieved their goal (enactment). Client intervention receipt used client goals as a proxy, with the assumption that a client only created a goal if they adequately understood the session content and applied the information to their diet and/or exercise behaviors. All documented goals were coded to determine if they reflect intervention content and were coded into the following themes: (a) exercise, (b) diet, (c) tracking, (d) behavioral regulation, and (e) not related. A goal involving completing a physical activity bout of any type fell under the "exercise" category. Goals in relation to nutrition were categorized under "diet." Goals that revolved around tracking diet or exercise in any form were coded under "tracking." Goals set to encourage the client to perform a behavior or learn more about a behavior were coded as "behavioral regulation." A goal was coded as "not related" if it did not fall under one of the four other respective categories (e.g., pick-up furniture on Saturday).

Clients were encouraged to track their diet and exercise on a mobile tracking app, a program tool where staff and researchers were able to view client tracking through a research portal. Client goal enactment was retrospectively accessed through the research portal. Goals were scored as either (a) achieved, (b) partially achieved, (c) not achieved, (d) not able to assess, (e) not tracking, or (f) no goal. If the client's tracking matched their goal fully then it qualified as "achieved." A goal was "partially achieved" if a client's tracking reflected working towards their goal but not completing it. "Not achieved" was coded if the client did not show any progress towards achieving the goal. If a client had <5 items tracked for 3 days in a row, the client received the code "not tracking." A goal received the code "not able to assess" if the client's goal was unable to be evaluated from the information available on the mobile tracking app (e.g., step-count goal). Finally, if staff did not record a goal on the session checklist, then the client received the code "no goal." To ensure the goals documented on session checklists accurately reflect the verbal communication between client and staff, an independent coder assessed a random 20% of audio files and extracted clients' goals from the audio-recorded sessions. Extracted goals were then SMART scored and compared to staff-documented SMART scores.

DATA ANALYSIS

Fidelity scores (overall fidelity, four individual checklist categories), were obtained as mean scores and converted into a percentage for each session. The independent coder fidelity checklist responses were compared to staff checklist responses and agreement statistics were calculated. Descriptive statistics were completed for

Downloaded from academic.oup.com/tbm/article/11/10/1814/6272494 by guest on 15 October 2026

client engagement and staff quality questions. Client's goals were analyzed using content analysis to determine a SMART goal score, client receipt of intervention, and client goal enactment. Descriptive statistics were calculated for SMART goals and client goal enactment. Based on recommendations to present multiple agreement statistics [25, 26], Cohen's kappa, prevalence, and bias adjusted kappa (PABAK) and percent agreement statistics were calculated. Both Cohen's kappa and PABAK consider chance agreement, with PABAK considering both positive (code is present) and negative agreement (code is not present) and Cohen's kappa only considering positive agreement [25]. Percent agreement does not consider chance agreement [25].

RESULTS

Staff self-reported fidelity

All clients completed the program, resulting in 156 session checklists. Table 1 presents descriptive statistics for staff self-reported fidelity. Only 5% of checklist items were missing across all six sessions, with 39% of the missing data coming from one staff. Fidelity was high with compulsory items being delivered on average 85%–93% of sessions.

Independent coder counseling session fidelity check

An independent coder listened to session three (n = 24) and session four (n = 25) audio files and completed a fidelity checklist. Two session three and one session four audio files were missing. A second coder listened to a random 20% of audio files and had high agreement with the independent coder (session 3: 94%; session 4: 89%). Table 2 compares staff self-reported fidelity and independent coder fidelity. Overall, the staff and independent coder agreed in 90% of cases. The independent coder's assessment of client engagement and staff quality was high. Client engagement averaged 4.08 ± 0.78 (range: 3–5) and 4.00 ± 0.73 (range: 3–5) and staff quality averaged 4.00 ± 0.81 (range: 2–5) and 4.00 ± 1.06 (range: 1–5) for sessions 3 and 4, respectively.

Independent coder goal enactment fidelity check

Of the 156 session checklists entered, 275 goals were extracted whereby clients made at least one goal in 94% of program sessions (M = 1.8/session); 10 checklists contained no goals. All documented goals were coded for meeting SMART criteria and descriptive statistics are presented in Table 3. Two independent coders had high percent agreement (94%). In general, staff-documented goals had high agreement with the independently coded audio-file documented goals (S 87%; M 91%; A 87%; R 100%; T 81%).

Goals reflected program topics with 43% exercise (n = 123), 31% diet (n = 89), 14% tracking (n = 39), 8% behavioral-regulation (n = 22), and 0.4% not related (n = 1) with 90% agreement between coders (Cohen's kappa = −0.05; PABAK = 0.80). Of the 275 goals created, 55% were not able to be assessed either (a) through the mobile tracking app (64%) or (b) the client not tracking on the mobile tracking app (36%). Thus, 45% of goals were assessed. Of those goals, 78% were achieved, 15% were partially achieved, and 7% were not achieved with an 87% agreement between coders.

DISCUSSION

Implementation evaluations are rarely conducted, and client engagement and delivery fidelity are often assumed [12]. Without understanding whether interventions are delivered as planned (by staff) and engaged with (by clients), it is difficult to fully understand whether an intervention is effective, and why or why not. This comprehensive evaluation assessed the implementation of an evidence-based diabetes prevention program by trained staff within a non-profit community organization. To our knowledge, this is the first in-depth fidelity assessment in a community-based diabetes prevention

Table 1 | Descriptive statistics for fidelity checklist data per session and sub-scale (N = 26)

	Session 1 *M* (range)	Session 2 *M* (range)	Session 3 *M* (range)	Session 4 *M* (range)	Session 5 *M* (range)	Session 6 *M* (range)
Logistics (%)	90 (45–100)	93 (40–100)	93 (44–100)	91 (0–100)	94 (50–100)	88 (12–100)
Counselling (%)	88 (38–100)	88 (71–100)	90 (75–100)	87 (0–100)	82 (29–100)	76 (0–100)
Exercise (%)	85 (0–100)	94 (25–100)	100 (100-100)	96 (0–100)	88 (0–100)	88 (0–100)
Optional tools (%)	94 (50–100)	85 (33–100)	68 (0–100)	72 (0–100)	58 (0–100)	54 (0–100)
Total score (%)	88 (33–100)	92 (57–100)	93 (71–100)	90 (0–100)	89 (32–100)	85 (7–96)
Total score with optional tools (%)	89 (38–100)	91 (63–100)	90 (71–100)	98 (0–100)	86 (38–100)	83 (7–97)
Missing* (%)	6	3	2	5	3	6

Note. *If an item was left blank on the checklist, it was deemed missing and no point was awarded. As a conservative measure, missing data was not subtracted from the total count, therefore missing data is equivalent to selecting no.

Downloaded from academic.oup.com/tbm/article/11/10/1814/6272494 by guest on 15 October 2026

Table 2 | Agreement between staff and independent coder on delivery of intervention components

	Agreement		Disagreement	
Checklist items	Staff = yes coder = yes	Staff = no coder = no	Staff = yes coder = no	Staff = no coder = yes
Session 3				
I checked in or helped revise the client's goal(s)	24 (100%)	0	0	0
Information exchange: physical sensations during exercise	22 (91.7%)	0	2 (8.3%)	0
Information exchange: talk test	23 (95.8%)	1 (4.2%)	0	0
Information exchange: physical sensations after exercise	23 (95.8%)	0	1 (4.2%)	0
I helped the client create goal(s) (depending on client readiness)	22 (91.7%)	0	1 (4.2%)	1 (4.2%)
I wrapped up the session	16 (66.7%)	3 (12.5%)	5 (20.8%)	0
Total counselling checklist	130 (90.3%)	4 (2.8%)	9 (6.3%%)	1 (0.7%)
I offered to look through: mobile tracking app with the client	19 (79.2%)	4 (16.7%)	1 (4.2%)	0
I offered to look through: client workbook with the client	9 (37.5%)	8 (33.3%)	6 (25.0%)	1 (4.2%)
Total optional tools	28 (58.3%)	12 (25.0%)	7 (14.6%)	1 (2.1%)
Total session 3	158 (82.3%)	16 (8.3%)	16 (8.3%)	2 (1.0%)
% Agreement				90.63
Kappa				0.59
PABAK				0.81
Session 4				
I checked in or helped revise the client's goal(s)	25 (100%)	0	0	0
Information exchange: carbohydrate choices	24 (96.0%)	0	1 (4.0%)	0
Information exchange: above ground/below ground vegetables	23 (92.0%)	2 (8.0%)	0	0
Information exchange: portion sizes: the plate model	21 (84.0%)	0	4 (16.0%)	0
I helped the client create goal(s) (depending on client readiness)	23 (92.0%)	0	2 (8.0%)	0
I wrapped up the session	16 (64.0%)	3 (12.0%)	6 (24.0%)	0
Total counselling checklist	132 (88.0%)	5 (3.3%)	13 (8.7%)	0 (0.0%)
I offered to look through: mobile tracking app with the client	15 (60.0%)	4 (16.0%)	2 (8.0%)	4 (16.0%)
I offered to look through: client workbook with the client	11 (44.0%)	10 (40.0%)	4 (16.0%)	0
I drew and/or looked at the plate model	21 (84.0%)	1 (4.0%)	3 (12.0%)	0
Total optional tools	47 (62.7%)	15 (20.0%)	9 (12.0%)	4 (5.3%)
Total session 4	179 (79.6%)	20 (8.9%)	22 (9.8%)	4 (1.8%)
% Agreement (%)				88.44
Kappa				0.54
PABAK				0.77
Overall total	337 (80.8%)	36 (8.7%)	38 (9.1%)	6 (1.4%)

Table 3 | SMART goal coding ($n = 275$)

	Specific	Measurable	Attainable	Relevant	Timebound
Yes (M%)	86	82	50	99	54
No (M%)	15	18	51	1	46
Kappa	0.56	0.63	1.00	0.49	0.79
PABAK	0.80	0.80	1.00	0.97	0.79
% Agreement (%)	90	90	100	99	90
SMART score $M \pm SD$ (range)					3.70 ± 1.10 (1–5)

Downloaded from academic.oup.com/tbm/article/11/10/1814/6272494 by guest on 15 October 2026

program that assesses treatment fidelity in addition to client engagement and receipt. The evaluation examined three fidelity components (delivery, receipt, and enactment of treatment) as identified by Bellg [10] with the other two components (study design, training providers) assessed in prior research [17, 18]. The study includes multiple perspectives (client, researcher, staff) and multiple methods (fidelity checklists, audio recordings) to examine implementation delivery, receipt, and enactment. In addition, in-depth goal coding was completed to assess the degree clients set SMART goals, a key program component. Overall, the program was implemented with high fidelity when delivered by staff employed at a community organization and clients were highly engaged.

Staff delivered the intervention as intended with high fidelity. Average staff self-reported fidelity was 90%, which is similar to other fidelity studies (100%, [13]); 65%, [23]). The high fidelity may be related to the use of the following recommended strategies, including 3-day training workshop, intervention manual with scripts, a senior trainer observing new staff with a client, audio recording of sessions, session checklists and monthly staff meetings [10]. Future research will test a modified version of these strategies in future scale-up sites (e.g., virtual training program, practice phone calls) to understand what combination of strategies, resources and level of support is needed to support successful implementation.

The gold standard to assess treatment fidelity is to audio-record all sessions and evaluate sessions compared to a priori criteria [10, 15]. As in similar fidelity assessments, to maintain feasibility, a selection of audio-recoded sessions were coded [22, 23]. Independent coder scores for the counseling sessions (83%, 81%) were modestly lower when compared to the staff scores (91%, 89%) for sessions 3 and 4, respectively. In the case of disagreement (overall 11%), the majority were a staff reporting they completed an item (87%), but the coder did not. Although this discrepancy between staff and independent coder has been found in previous research (45%, [13]; 42%, [23]), our results show less discrepancy. Higher independent fidelity scores have been previously reported (81%, [14]; 66%, [15]). Agreement may have been enhanced by using a simple fidelity checklist [27]. Comparing these fidelity studies, lower fidelity scores seem to occur with comprehensive fidelity assessments (e.g., evaluating specific behavior change techniques (BCTs) and/or intervention behaviors [13, 15, 23]) compared to less comprehensive fidelity assessments, such as coding broad intervention techniques (e.g., goal setting, providing information; [14]). Future research is encouraged to examine what level of fidelity evaluations provide the best feedback on program implementation, whilst minimizing burden.

The current analysis reflects staff delivery of core intervention components and broad clusters of BCTs (e.g., goal setting, self-monitoring of behaviors) and does not analyze delivery of all the specific BCTs contained in Small Steps for Big Changes. Full BCT coding of the program was completed after this study and can inform a future, more comprehensive, fidelity assessment [28]. More comprehensive fidelity evaluations may increase burden on the researcher (developing and coding the intervention), training of intervention facilitators, ongoing assessment of fidelity and trained research staff to analyze data and provide feedback [13]. Conducting quality fidelity evaluations that balance researcher burden and providing useful program data is necessary.

Client receipt and enactment are two understudied factors that impact intervention success [12]. For interventions to be successful, clients must interact with the program to learn and apply the targeted skills. Overall, 96% of client-created goals were related to program content, suggesting clients received the intervention as intended (gather program information and create relevant goals to put that information into action). Only 4% of program sessions resulted in no goal formed. Of the goals that could be assessed (45% of all goals), 93% were achieved or partially achieved suggesting that clients are not solely receiving program information, but actively making behavior changes. A fidelity study by Walton et al. [21] assessed receipt and enactment through four participant self-reported responses with comparable results. Although examining program effectiveness is outside the scope of the current study, high levels of participant engagement have been shown to be associated with greater participant outcomes [23].

Brief action planning is a highly structured technique to facilitate goal setting, create a SMART goal, and foster client self-efficacy [24]. A well-defined goal can help individuals focus their intentions and evaluate their behavior to a success standard [29]. Overall, goals had a high SMART score, with *attainable* and *timebound* categories having the lowest scores. For the attainable category, only 50% of goals had a confidence rating. Asking for a confidence rating is important to gauge client self-efficacy [24] as low self-efficacy is associated with non-completion [30]. Just over 50% of goals were timebound. Within the Small Steps for Big Changes program, goals are often created to be executed between sessions. Staff may have omitted this detail when documenting goals thinking it was implied. Indeed, in the 20% audio-file check, time-bound had the lowest agreement (81%). Future staff training will reiterate the importance of SMART goal categories, specifically attainable and time-bound, to facilitate even stronger SMART goals. The operationalization and evaluation of this key program component led to useful practical feedback to improve implementation and should be considered in future implementation evaluations.

Downloaded from academic.oup.com/tbm/article/11/10/1814/6272494 by guest on 15 October 2026

High goal achievement (78%) may have been facilitated by the program through offering clients the opportunity to create their own goal(s), execute goal(s) within a short time-frame, reflect on the goal(s) at the following session, and modify goal(s) as needed [30]. High goal achievement has been found before. For example, an online health behavior change study examined self-reported weekly action plan completion over a 6-month study period and found 49% of action plans were completed, 40% partially completed and 11% not completed [30], similar to our current results. Despite the commonly reported "intention-behavior" gap [31], current results indicate clients in Small Steps for Big Changes turned their intentions into actions.

Limitations and future directions

Study limitations must be acknowledged. A dichotomous scale (yes/no) was used to evaluate fidelity checklist components and does not reflect the extent to which material and skills were delivered by staff. Future fidelity checklists could consider a broader scale (e.g., 5-point). Client goals were evaluated from staff-completed checklists and therefore may not accurately reflect all the details described orally between the staff and client. However, staff documented goals had high agreement to audio-file goals from a 20% random independent audio-file review. Client goal enactment could not be assessed for clients who were not tracking or chose to create goals that were not able to be assessed via the app (e.g., step goal). The method for analyzing client enactment was chosen for pragmatic reasons as it minimized client burden; tracking was an optional tool for clients to engage in and researchers had access to app data. The current method is limited in that the difference between a client not achieving a goal and not being able to assess a goal was not able to be examined. Although only 45% of goals were able to be assessed using this method, we feel that it reflects client behaviors and is a novel method to assess goals as opposed to client self-report on a questionnaire as used in previous studies [21]. Future research should continue to explore methods to examine low-resource techniques to study treatment enactment such as having coaches confirm client goal achievement on fidelity checklists.

Fidelity checklists were developed to minimize burden for staff, whereby they were short and easy to complete and had a dual-purpose—an implementation tool (e.g., reminder of intervention content and skills) and a data collection tool [10]. Although fidelity checklists are low cost and easy to use, social desirability and recall bias may make them less reliable. The present analysis suggests the programs' fidelity checklist are a good measure of treatment fidelity when compared to the audio-recorded independent coder scores. The low rate of missing data on the fidelity checklists provide positive indications for checklist practicality and acceptability [12] and will be further investigated in qualitative work interviewing staff within the community organization. Qualitative work will help ensure feasibility and sustainability of the intervention and to investigate the high degree of missing data from one staff in this present study.

CONCLUSION

Small Steps for Big Changes was implemented with high fidelity in two community sites, which increases the confidence that program results are from the intervention. Working with a community partner to develop and execute an implementation plan, may have led to higher fidelity. While evidence of effectiveness in a community setting is high [32], preliminary outcomes related to the current study demonstrate success, yet are ongoing. Implementation evaluations are useful to understand how programs are being implemented, received, and engaged with. Such valuable feedback can be used to understand the program, refine implementation strategies to support future program translations and scale-up efforts, and better reach and impact at risk individuals.

Funding: This research was funded by both a Social Sciences and Humanities Research Council Doctoral Scholarship (#767-2020-2130) and a Partnership Engage Grant (#892-2018-3065), the Canadian Institutes of Health Research (#333266), and Michael Smith Foundation for Health Research Reach Grant (#18120).

COMPLIANCE WITH ETHICAL STANDARDS

Authors' Statement of Conflict of Interest and Adherence to Ethical Standards: All authors declare that they have no conflicts of interest.

Human Rights: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards (IRB: H16-02028).

Informed Consent: Informed consent was obtained from all individual participants included in the study.

Welfare of Animals: This article does not contain any studies with animals performed by any of the authors.

Study Registration: This study was not formally registered.

Analytic Plan Preregistration: The analysis plan was not formally pre-registered.

Data availability: De-identified data from this study are not available in a public archive. De-identified data from this study will be made available (as allowable according to institutional IRB standards) by emailing the corresponding author. Analytic code used to conduct the analyses presented in this study are not available in a public archive. They may be available by emailing the corresponding author. Materials used to conduct the study are not publicly available.

References

1. International Diabetes Federation. *International Diabetes Federations Diabetes Atlas*. 9th ed. 2019. Available from: https://www.diabetesatlas.org.

Downloaded from academic.oup.com/tbm/article/11/10/1814/6272494 by guest on 15 October 2026

2. Tabák AG, Herder C, Rathmann W, Brunner EJ, Kivimäki M. Prediabetes: A high-risk state for diabetes development. *Lancet.* 2012;379(9833):2279–2290.
3. Knowler WC, Barrett-Connor E, Fowler SE, et al.; Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med.* 2002;346(6):393–403.
4. Glasgow RE, Lichtenstein E, Marcus AC. Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *Am J Public Health.* 2003;93(8):1261–1267.
5. Aziz Z, Absetz P, Oldroyd J, Pronk NP, Oldenburg B. A systematic review of real-world diabetes prevention programs: learnings from the last 15 years. *Implement Sci.* 2015;10(1), 172.
6. Dunkley AJ, Bodicoat DH, Greaves CJ, Russell C, Yates T, Davies MJ, Khunti K. Diabetes prevention in the real world: Effectiveness of pragmatic lifestyle interventions for the prevention of type 2 diabetes and of the impact of adherence to guideline recommendations. *System Rev Meta-anal.* 2014;37(4), 922–933.
7. Van Name MA, Camp AW, Magenheimer EA, et al. Effective translation of an intensive lifestyle intervention for Hispanic women with prediabetes in a community health center setting. *Diabetes Care.* 2016;39(4):525–531.
8. Ely EK, Gruss SM, Luman ET, Gregg EW, Ali MK, Nhim K, Rolka DB, Albright AL. A national effort to prevent type 2 diabetes: participant-level evaluation of cdc's national diabetes prevention program. *Diabetes Care.* 2017;40(10), 1331–41.
9. Nhim K, Gruss SM, Porterfield DS, et al. Using a RE-AIM framework to identify promising practices in National Diabetes Prevention Program implementation. *Implement Sci.* 2019;14(1):81.
10. Bellg AJ, Borrelli B, Resnick B, et al.; Treatment Fidelity Workgroup of the NIH Behavior Change Consortium. Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol.* 2004;23(5):443–451.
11. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci.* 2007;2:40.
12. Walton H, Spector A, Tombor I, Michie S. Measures of fidelity of delivery of, and engagement with, complex, face-to-face health behaviour change interventions: A systematic review of measure quality. *Br J Health Psychol.* 2017;22(4):872–903.
13. Hardeman W, Michie S, Fanshawe T, Prevost AT, Mcloughlin K, Kinmonth AL. Fidelity of delivery of a physical activity intervention: Predictors and consequences. *Psychol Health.* 2008;23(1):11–24.
14. Kok Michele SY, Jones M, Solomon-Moore E, Smith Jane R. Implementation fidelity of a voluntary sector-led diabetes education programme. *Health Education.* 2018;118(1):62–81.
15. Lorencatto F, West R, Christopherson C, Michie S. Assessing fidelity of delivery of smoking cessation behavioural support in practice. *Implement Sci.* 2013;8:40.
16. Rixon L, Baron J, McGale N, Lorencatto F, Francis J, Davies A. Methods used to address fidelity of receipt in health intervention research: A citation analysis and systematic review. *BMC Health Serv Res.* 2016;16(1):663.
17. Jung ME, Locke SR, Bourne JE, et al. Cardiorespiratory fitness and accelerometer-determined physical activity following one year of free-living high-intensity interval training and moderate-intensity continuous training: a randomized behaviour change intervention trial. *Int J Behav Nutr Phys Act.* 2020;17(1):25.
18. Dineen TE, Bean C, Ivanova E, Jung M. Evaluating a motivational interviewing training for facilitators of a prediabetes prevention program. *J Exerc Mov Sport.* 2018;50(1):234.
19. Bean C, Sewell K, Jung ME. A winning combination: collaborating with stakeholders throughout the process of planning and implementing a type 2 diabetes prevention programme in the community. *Health Soc Care Community.* 2019;28(2):681–689.
20. American Diabetes Association. 2. Classification and diagnosis of diabetes: Standards of medical care in diabetes—2019. *Diabetes Care.* 2019;42 (Suppl 1):S13–S28.
21. Walton H, Spector A, Roberts A, et al. Developing strategies to improve fidelity of delivery of, and engagement with, a complex intervention to improve independence in dementia: A mixed methods study. *BMC Med Res Methodol.* 2020;20(1):153.
22. Jelsma JG, Mertens VC, Forsberg L, Forsberg L. How to measure motivational interviewing fidelity in randomized controlled trials: Practical recommendations. *Contemp Clin Trials.* 2015;43:93–99.
23. Rocchi MA, Robichaud Lapointe T, Gainforth HL, Chemtob K, Arbour-Nicitopoulos KP, Kairy D, Sweet SN. Delivering a tele-health intervention promoting motivation and leisure-time physical activity among adults with spinal cord injury: An implementation evaluation. *Sport Exerc Perform Psychol.* 2021;10(1):114–132.
24. Gutnick D, Reims K, Davis C, Gainforth H, Jay M, Cole S. Brief action planning to facilitate behavior change and support patient self-management. *J Clin Outcomes Manage.* 2014;21:17–29.
25. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol.* 1993;46(5):423–429.
26. Chen G, Faris P, Hemmelgarn B, Walker RL, Quan H. Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa. *BMC Med Res Methodol.* 2009;9:5.
27. Harting J, van Assema P, van der Molen HT, Ambergen T, de Vries NK. Quality assessment of health counseling: performance of health advisors in cardiovascular prevention. *Patient Educ Couns.* 2004;54(1):107–118.
28. MacPherson MM, Dineen TE, Cranston KD, Jung ME. Identifying behaviour change techniques and motivational interviewing techniques in Small Steps For Big Changes: a community-based program for adults at risk for type 2 diabetes. *Can J Diabetes.* 2020;44(8):719–726.
29. Bailey RR. Goal setting and action planning for health behavior change. *Am J Lifestyle Med.* 2019;13(6):615–618.
30. Lorig K, Laurent DD, Plant K, Krishnan E, Ritter PL. The components of action planning and their associations with behavior and health outcomes. *Chronic Illn.* 2014;10(1):50–59.
31. Hagger MS, Luszczynska A. Implementation intention and action planning interventions in health contexts: State of the research and proposals for the way forward. *Appl Psychol Health Well Being.* 2014;6(1):1–47.
32. Bean C, Dineen TE, Locke SR, Bouvier B, Jung ME. An evaluation of the reach and effectiveness of a diabetes prevention behaviour change program situated in a community site. *Can J Diabetes.* In press.

Downloaded from academic.oup.com/tbm/article/11/10/1814/6272494 by guest on 15 October 2026