Published in final edited form as:

J Subst Abuse Treat. 2018 September ; 92: 51–63. doi:10.1016/j.jsat.2018.06.012.

Assessing long-term sustainment of clinic participation in NIATx200: Results and a new methodological approach

James H. Ford II^{a,*}, Scott P. Stumbo^b, and James M. Robinson^c

^aSocial and Administrative Sciences Division, University of Wisconsin - School of Pharmacy, United States of America

^bCenter for Health Research, Kaiser Permanente Northwest, United States of America

^cCenter for Health Systems Research and Analysis, University of Wisconsin – Madison, United States of America

Abstract

Background: Sustainability frameworks differentiate between sustainability capacity and sustainment of organizational change. Multiple studies have examined sustainability capacity. Methodological approaches to assess long-term sustainment have not been explored. This study addresses this gap by describing the development of a long-term sustainment methodology and evaluating its application utilizing data from substance abuse clinics participating in a quality improvement collaborative.

Methods: The study involved clinics (n = 121) in three states (MI, NY and WA) participating in the 2007–2009 NIATx200 quality improvement (QI) intervention. It extended the primary analysis to focus on clinics' long-term sustainment of wait time, retention and admission improvements. Long-term sustainment was defined as two years post end of the active implementation period (Calendar Years 2010 and 2011). The analysis defined case exclusion criteria and spline “knot” time intervals; allowed for C_p statistic use to address clinic data volatility; established the structure of sustainment plots and explored differences between NIATx implementation strategies.

Results: Example spline and sustain plots highlight the application of the long-term sustainment methodology for NIATx200 clinics. In clinics with available longitudinal outcome data, 40.8% (n = 31 of 76 clinics) sustained improvements in wait time, 26.7% (n = 20 of 75 clinics) in retention, and 28.1% (n = 32 of 114 clinics) for admissions. Clinic assignment to a NIATx200 implementation strategy did not significantly influence a clinics' long-term sustainment except for lower wait time changes in the interest circle interventions. Thirty clinics (24.8%) sustained improvements for two outcomes and six clinics (5.0%) did so for all three outcomes. The clinics that sustained multiple outcome improvements were assigned to the interest circle (n = 12),

*Corresponding author at: University of Wisconsin – Madison, Social and Administrative Sciences Division – School of Pharmacy, 777 Highland Avenue, Madison, WI 53705, United States of America. jhfordii@wisc.edu (J.H. Ford).

[ClinicalTrials.gov](https://clinicaltrials.gov) Identifier: NCT00934141 Registered July 6, 2009. Retrospectively registered.

Availability of data and material

The data that support the findings of this study brought together existing data obtained upon request from the Michigan Department of Community Health, the Washington Department of Social and Health Services and the New York Office of Alcoholism and Substance Abuse Services but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available.

learning session (n = 10), combination (n = 8), and coaching (n = 6) implementation strategies. Guidance for applying the long-term sustainment methodology in other quality improvement settings is described.

Conclusions: Research about sustainability capacity and sustainment of change has become increasingly important in dissemination and implementation research. Assessment of long-term sustainment in a multi-organizational quality improvement collaborative (QIC) is needed to identify when program drift and intervention decay occurs. If “cut-points” indicate when effects diminish, specific sustainability modules could be developed and introduced within the structure of a QIC to improve organizational long-term sustainment. Coaches and change teams could be trained to focus on organizational change sustainment and strengthen the likelihood of institutionalization.

Keywords

Quality improvement collaborative; Long-term sustainment; NIATx200; Sustain plots; Wait time; Retention; Admissions

1. Introduction

Not every organizational change should be sustained. If a change does not improve working conditions for staff, business processes or client clinical care, the organization should determine if it is worth sustaining the change. Even if the answer is yes, evidence suggests that 40 to 70% of organizational changes are not sustained (Beer & Nohria, 2000; Scheirer & Dearing, 2011). A failure to sustain organizational change is a direct waste of invested resources, has costs associated with missed opportunities, and may affect an organization's ability to implement change in the future. Experts recognize the importance of understanding evidence-based practice sustainability in healthcare settings (Scheirer, 2005, 2013; Scheirer & Dearing, 2011; Scoville, Little, Rakover, Luther, & Mate, 2016; Stirman et al., 2012). As such, dissemination and implementation researchers have started to systematically focus their efforts to understand and learn about sustainability of organizational change.

Sustainability frameworks help to differentiate between two important concepts in sustainability research: sustainability capacity and sustainment of change within an organization (Aarons, Hurlburt, & Horwitz, 2011; Chambers, Glasgow, & Stange, 2013; Scheirer & Dearing, 2011; Schell et al., 2013). Capacity refers to “*the existence of structures and processes that allow a program to leverage resources to effectively implement and maintain evidence-based policies and activities*” (Schell et al., 2013). The frameworks suggest that sustainability capacity is attributable to (a) internal organizational attributes, (b) external environmental contextual influencers, and (c) attributes of the intervention. Organizational attributes include leadership support, champion roles, revised policies and procedures or expert support (e.g., coaching). Environmental influencers examine how financial, regulatory and policy changes impact organizational efforts to sustain change. Innovation attributes focus on understanding how benefits of the change increase the likelihood of sustainment within an organization. Examples include how the innovation addresses client needs or the ease of making changes. Multiple studies identified different

internal, external and innovation attributes informing sustainability capacity of evidence-based practices (Aarons et al., 2016; Bond, Drake, Becker, & Noel, 2016; Fleiszer, Semenic, Ritchie, Richer, & Denis, 2015a, 2015b, 2016a, 2016b; Ford et al., 2015; Hunter, Han, Slaughter, Godley, & Garner, 2017; Peterson et al., 2014).

Sustainment of change, in contrast, refers to “*the extent an evidencebased intervention can deliver its intended benefits over an extended period of time*” (Chambers et al., 2013). In other words, how long over an extended period of time is an organization able to maintain or even improve upon the benefits (e.g., reduce wait time) of the change after the intervention implementation period ends. The definition of “*an extended period of time*” is a critical part of efforts to assess sustainment of organizational change. However, time frames for assessing sustainment vary. Internal efforts to assess sustainment may focus on short time frames (e.g., four months) to determine when the impact of a change may be decaying and action is required (Pronovost et al., 2016). Other studies have assessed long-term sustainment over one to two years using interrupted time-series (Harris et al., 2016) or piece-wise linear regression (Lopatto, Keith, Del Canale, Templin, & Maio, 2014). Methodological approaches assessing sustainment of change are not well represented in the literature. One of these studies examined staff perspectives associated with quality implementation of an adolescent evidence-based practice (Huang et al., 2017). Another study utilized 19 quarters of administrative data to assess how organizational and therapist characteristics influenced sustainment of six different prevention-focused evidence-based practices (Brookman-Frazee et al., 2016).

Outcome improvement at specific time points, such as the end of a quality improvement initiative active intervention phase, may not predict long-term sustainment, particularly when the process improvement requires a change to organizational culture (Alharbi, Ekman, Olsson, Dudas, & Carlström, 2012; McAlearney et al., 2013; Novins, Green, Legha, & Aarons, 2013; Smith, Wright, Queram, & Lamb, 2012). While interviews provide insights about organizational sustainability capacity, the utilization of longitudinal data in sustainment research addresses issues associated with right censoring of data and latency effects associated with organizational change uptake. Right censoring occurs when the measurement definition of the outcome of interest (e.g., retention in treatment) is right censored at the date of last follow-up or designated experimental time (Lash, Sorensen, Bradbury, & Tavis, 2014). Failure to address right censoring can impact slope estimations in an analysis of long-term sustainment (Jaffa, Lipsitz, & Woolson, 2015; Lagakos, 1979; Liu & Ying, 2007; Wu & Carroll, 1988). Latency effect occurs when a change requires more time to take hold in the organizations or when incomplete access to data or insufficient data limits a planned analysis or masks improvements (Salmon et al., 2012; Singh, Rana, & Singhal, 2013). The inclusion of sufficient data in a longterm sustainment analysis addresses latency and right censoring issues. It also provides an opportunity to examine how outcome changes over time; to explore “program drift” or differences in the length of time that changes are sustained across different organizations (Chambers et al., 2013). Identification of these “cut-points” in time after which the effects of a quality improvement collaborative (QIC) diminish could inform when to introduce specific sustainability modules into the structure of a QIC. Finally, the use of longitudinal data enhances the ability to examine how different attributes (e.g., organization vs. innovation) affect sustainment across different

treatment settings (Scheirer & Dearing, 2011; Stetler, McQueen, Demakis, & Mittman, 2008), thus, relating the findings back to organizational sustainability capacity. However, no studies appear to describe a methodological approach utilizing longitudinal state administrative data from multiple states or to assess long-term sustainment of multiple outcomes within a single organization that participated in a quality improvement collaborative.

1.1. NIATx200

The NIATx200 cluster-randomized controlled trial built upon the success of prior NIATx research (Ford et al., 2007; Hoffman, Ford, Choi, Gustafson, & McCarty, 2008; Hoffman, Ford, Tillotson, Choi, & McCarty, 2011; McCarty et al., 2007; Roosa, Scripa, Zastowny, & Ford, 2011). A total of 201 addiction treatment clinics in five states (Massachusetts, Michigan, New York, Oregon and Washington) were randomized to one of four interventions: a) interest circle calls (n = 49), b) learning sessions (n = 54), c) coaching (n = 50) or d) a combination of all three (n = 48). The active implementation period lasted 18 months. Interest circles involved multi-clinic teleconferences (n = 18) where change teams could receive advice from peers and learn new skills. Learning sessions consisted of three face-to-face multi-day sessions held approximately every 6 months. The sessions were led by a core faculty team who utilized a common curriculum to offer didactic and experiential learning opportunities. Coaching consisted of a one-day site visit and monthly (n = 18) calls between the coach and change leader, executive sponsor and change team to review change projects; discuss successes; and identify ideas for future change projects. The structure of the NIATx200 intervention is described in more detail elsewhere (Grazier et al., 2015; Gustafson et al., 2013; Quanbeck et al., 2011).

Primary outcomes were: 1) mean days between first contact and first treatment (*wait time*), 2) percent of patients retained from first to fourth treatment session (*retention*), and 3) annual number of patients (*admissions*). Organizational attributes related to attributes such as nonprofit status or the number of FTEs was collected at randomization (Quanbeck et al., 2011). The clinic's overall management score and their score related to quality improvement were also assessed (McConnell, Hoffman, Quanbeck, & McCarty, 2009).

For six-month periods during the active intervention period, clinics focused either on implementing organizational changes to improve wait time (Months 1 to 6), retention (Months 7 to 12) or admissions (Month 13 to 18). The 18-month active implementation period was followed by a 9-month short-term sustainment period. The primary analysis used mixed-effect regression models to determine which implementation strategy was most effective in improving outcomes and to determine which strategy was most cost effective (Quanbeck et al., 2011). Clinics assigned to the coaching and combination strategies showed significant reductions in average wait-time and a significant increase in admissions at the end of the six-month intervention period targeting improvements for the outcome (Gustafson et al., 2013). The changes differed significantly from clinics assigned to the interest circle strategy. The coaching strategy was more cost-effective as compared to interest circles for wait time and admission improvement (Gustafson et al., 2013). The primary analysis indicated that no NIATx implementation strategy improved retention in treatment.

Differences in attributes of the clinics (e.g., size) as co-variables were not associated with improvements in the outcomes.

Although no significant improvements in retention were found, an exploratory analysis measured retention from first request for service to fourth treatment to account for early drop-off in treatment (i.e., a client not making it to the first treatment session). The results showed improvements in retention similar to wait time for providers assigned to the coaching, combination and learning session strategies suggesting that these strategies may improve retention depending on how the outcome is defined (Gustafson et al., 2013). A second exploratory analysis suggested that retention improved for 83 of the 169 clinics (49%) with available data in the nine-month period after the intervention ended. However, 54 of these 83 clinics (65%) did not experience an improvement during the active intervention period. The findings suggest a potential latency effect, due to lags in clinic data reporting to the state, may exist in adoption of retention related changes. The right censoring of data in the original analysis may have masked actual retention improvements for some clinics.

The NIATx200 initiative represents one of the largest randomized control trials involving a quality improvement (QI) intervention. Its structure and scope provide a unique opportunity to methodologically explore and examine long-term sustainment across multiple outcomes within a single organization. A separate paper described themes associated related to long-term sustainment of the NIATx200 QI practices (Stumbo, Ford, & Green, 2017). The objectives of the current paper are to (1) describe a new methodological approach utilizing longitudinal data to assess long-term sustainment of organizational changes; and (2) characterize its application to assess sustainment of wait time, retention and admission outcomes for substance abuse clinics in Michigan, New York and Washington participating in the NIATx200 quality improvement initiative. Through these objectives, we will discuss the challenges of assessing sustainment of multiple outcomes within a single organization and identify potential strategies to enhance efforts to sustain change within a QI collaborative.

2. Methods

2.1. Study setting

The study was conducted in collaboration with the Department of Community Health (MI), Department of Social and Health Services (WA) and Office of Alcoholism and Substance Abuse Services (NY) for this study. Since this study extended the primary NIATx200 analysis, each state used existing administrative data systems to provide an updated NIATx200 outcomes longitudinal dataset for all client admissions from participating NIATx200 clinics. The new dataset included data covering the original study period (July 2007 to Dec 2009) and an additional forty-two month period after the end of the NIATx200 study (January 2010 to June 2013) to explore long-term sustainment. The final sample size included 121 clinics from the three states; clinics were the unit of analysis. The data reporting systems utilized for the NIATx200 initiative in Oregon and Massachusetts were not sustained and NIATx200 clinics in these states were excluded from the analysis.

2.2. Outcomes

The analysis used the same outcome definitions as the primary NIATx200 analysis (Gustafson et al., 2013). The wait time (elapsed time from 1st contact to 1st treatment) and retention (percent of clients retained from 1st to 4th treatment sessions) outcomes were standardized by being expressed on a per client basis. Thus, these outcomes could reasonably be compared for large or small clinics. Raw admission counts or changes in admission counts were not comparable between different size clinics and were normalized, as done in the primary analysis, to account for differences in organizational size. The slope of the log-transformed admission counts was roughly equivalent to the percentage change in admissions, which was comparable from site to site.

2.3. Analytical Approach

Efforts to model and assess long-term sustainment required a multistep data review process to determine appropriate study time intervals (measured in months) and establish case exclusion criteria. The longitudinal data review utilized time series graphs by clinic and was conducted to review outcome counts and data trends and to determine the sustainment time intervals. The process allowed for modification of the interval start and stop dates. An iterative process was utilized to review different time interval combinations. For example, we considered limiting active implementation to the six-month period when a clinic was working to improve an outcome (e.g., wait time). However, coach notes and presentations at learning sessions and interest circles suggested that clinics experienced delays in implementing changes suggesting that sustainment trends could represent continued outcome improvement. The NIATx200 intervention period was utilized to determine the baseline period and active intervention period for each outcome and to establish the pivot points (“knots”) for the time intervals (Fig. 1). The collection of an additional six months of data (January to June 2013) helped reduce the effects of right censoring in the earlier periods.

Data availability (e.g., number of cases per month) and volatility caused by unbounded outcomes (e.g., wait times) present challenges in longitudinal data analysis. As such, the determination of case-exclusion criteria to account for these issues represented a crucial process step in the development of the long-term sustainment methodology. In the NIATx200 study, wait time and retention outcomes were subject to extreme volatility in clinics with a small number of valid cases ($n = 5$). As such, only months with at least five cases per month were included in the analysis. Substance treatment seeking clients might have long intervals between 1st contact and 1st treatment, typically due to an incarceration event, which led to large wait times for some. Similar to the primary NIATx200 analysis, wait time outliers > 273 days were excluded. However, outliers were not an issue with retention rates, which must fall between 0 and 1, or monthly admissions. When the level of reported activity fell permanently below median monthly level for the clinic, the month when the event occurred, and all subsequent months, were dropped from the analysis. This approach removed observations when reliable data was no longer being reported by the clinic. Otherwise, wait times could appear shorter and retention rates could appear to drop even if true wait time and retention rates were unchanging resulting in a false impression of long-term sustainment.

The analysis then utilized a combination of piecewise-linear time series graphs and sustainment scatter plots to assess long-term sustainment by clinic. The fitted piecewise-linear time series graphs were obtained by regressing monthly outcome values for each clinic against line segments (“linear splines”) for each study time interval. This approach allowed the fitted line to change slope at the boundary (“knot”) between study intervals. The p-values of estimated slopes ($\alpha = 0.05$) and slope changes (i.e., the regression spline coefficients) were used to assess the statistical significance of observed changes in outcomes over time for each clinic. The resulting spline graph represented specific clinic results over time by outcome and the sustainment scatter plots summarized selected spline graph characteristics. Each plotted value on the sustainment scatter plot represented two aspects of a site’s spline graph, early improvement (plotted along the horizontal axis) versus longer-term improvement (plotted along the vertical axis). Findings from this analysis were utilized for clinic quadrant placement in the sustainment scatter plots and to assess NIATx200 long-term sustainment for each outcome.

2.3.1. Linear spline time series graphs—The analysis focused separately on each clinic’s behavior. Specifically, standard unweighted least squares regression was used to estimate the spline slopes and create spline graphs for each clinic by outcome (wait time, retention or log-admission) by month. This flexible approach differs from survival modeling because our approach let us to see how the observed intervention effect for each clinic continues to grow, plateaus or eroded over time by allowing the fitted line to respond to changes in the level or slope of the observed data. In addition, the fitted piecewise linear curve did not “blow-up” on either end of the plot, as can occur with a polynomial fit to the data, while still allowing sharp turns at each existing spline “knot” point. Mallow’s C_p statistic was used to determine which knots and the associated estimated slope changes were retained, balancing model fit and parsimony (Cantoni & Ronchetti, 2001; Kou, 2003). If observations were sparse in later months, the C_p criteria dropped knots to obtain more parsimonious and stable slope estimates. The fitted spline was not allowed to continue up or down after the last observation, rather it was forced to plateau at the last observation month. The spline time series graphs included a table of slope t-ratios, at each knot and for combinations of knot segments (e.g., time period AB, CD, and AD). Knot-level t-ratios were informative, but not directly utilized, when building the fitted spline. Rather, the approach allowed the C_p criteria to decide which slope changes to set to zero.

2.3.2. Sustainment scatter plots—The sustainment scatter plots provided a graphical visualization of clinic performance across two specific intervention periods. The plots showed the average monthly rate of change during period AB (x-axis) versus the cumulative average monthly rate of change for periods AD (y-axis) for a specific clinic. Clinics with no early improvement during the active intervention period were isolated and displayed to the left of the vertical axis. Clinics with missing data for the x-axis and y-axis periods were not included on these plots (see Table 1).

The plots showed clinics in the same study arm as a cross where the cross center represented the best estimate of the X and Y values for the clinic with its location circled on the plot. The size of the crossbars was proportional to the confidence in the horizontal and vertical

placement of the center and was inversely proportional to the standard deviation of the estimate of the center's location. Larger crossbars implied smaller standard deviations and greater confidence on the estimate location of the center. Other intervention clinics (i.e., those in other study arms) were also represented on the plot but only as points. The combination period t-ratios assessed whether a cross plotted on the sustainment plot is significantly off the horizontal or vertical axis thus indicating if the change had been sustained.

The sustainment plot also divided quadrants I and III into three sections differentiating the level of sustainment within the quadrant. The quadrant descriptions (Appendix A) detailed how to interpret the sections in Quadrants I and III. This methodology allows users to change the study periods included on the x and y axis (e.g., show Period A (x-axis) versus Period B (y-axis)), and to “switch” off the display of clinics in other implementation strategies.

2.3.3. Assessing NIATx200 long-term sustainment—Clinic placement on the sustainment plot was utilized to tabulate the number of clinics in each quadrant by outcome. Long-term sustainment was determined when a clinic achieved improvement during the intervention period and sustained that improvement (in whole or in part) during the two-year post-intervention period. The average change per month for the Period AD outcome was calculated based on clinic intervention assignment and quadrant placement on the sustain plots. Means were compared within quadrant to assess if participation in a specific NIATx200 intervention resulted in greater outcome improvements and two-year sustainment. The number of clinics that sustained at least two or three outcomes was determined. The methodology also compared clinic sustainment status in Period AD (which varied by outcome) to period E (Jan to Dec 2012) to assess if clinics continued to improve and sustain for another year.

The study was approved by the University of Wisconsin Health Science Minimal Risk Institutional Review Board (2014–1048) and the Washington State Institutional Review Board (D-072414-S).

3. Results

Of the 201 clinics in the NIATx200 study, 121 clinics were included in this analysis. Clinics included in the analysis were more likely to be part of a multi-site organization and less likely to be a nonprofit clinic (results not shown). The distribution of the 121 clinics by intervention arm was interest circles (n = 29), coaching (n = 31), learning sessions (n = 31) and combination (n = 30) and by state was Michigan (n = 42), New York (n = 41) and Washington (n = 38).

3.1. Linear spline time series graphs and sustainment scatter plots

Results revealed that clinic-specific long-term sustainment results varied by outcome, and over time, within a particular clinic. Detailed information from one clinic, MI-32 [assigned to the combination implementation strategy], demonstrates the type of results available via this new methodological approach. During the active intervention period (Dec 07 to Mar 09,

period A) and the follow-up short-term sustainment period (Apr 09 to Dec 09, period B), the clinic experienced a clinically significant wait time reduction from its baseline of 9 days at the rate of 0.29 days per month, a reduction of 4.6 days over 16 months, which was followed by a significant increase in wait time of 0.59 days per month, an increase of 5.3 days over 9 months (see Fig. 2). The net result was no significant change in wait time during the active implementation period. In Period C (Jan to Dec 2010), the clinic significantly reduced their wait time by 0.36 days per month and then sustained those improvements over the next two years (Periods D and E). The clinic's retention rate fluctuated over time but did not show significant improvement (see Appendix B, panel A). The clinic also experienced a significant improvement in the number of admissions (approximately 2 new admissions per month) by the end of the active implementation period (Oct 08 to Mar 09). Although the admissions decreased slightly in the nine-month follow-up period (Apr to Dec 09) and fluctuated over time in subsequent years, these changes did not reach significance. As a result, the clinic was able to sustain its admission improvements following participation in NIATx200 (see Appendix B, panel B).

3.2. Summary of long-term sustainment results

Clinic sustainment over the two-years post active intervention varied by outcome (Table 1). Of the 121 clinics in the sample, 45 clinics (37.2%) did not have enough wait time data and 46 clinics (38%) had insufficient retention data and were excluded from the analysis. Twenty out of 75 clinics (26.7%) sustained retention improvements (i.e., located in Quadrant I) as compared to thirty-two out of 114 clinics (28.1%) for admission and 31 out of 76 clinics (40.8%) for wait time. Clinics not sustaining improvements are located in Quadrants II and III of the sustainment plots. Overall, a higher percentage of clinics located in Quadrants II and III either did not improve or sustain admission improvements or showed some improvement before regressing back towards their baseline value (62.5%) as compared to the wait time (46%) and retention (44%) outcomes. The second part of Table 1 provides a breakdown of clinic placement by quadrant by outcome and identifies clinics for which the rate of improvement in Period CD was greater than the improvement in Period AB. Placement in Quadrant I indicates that 1 in 4 clinics showed improvements in wait time and retention during the two-year sustainment period (CD) that exceeded their improvements in the intervention period (AB); approximately 1 in 5 clinics achieved this for admissions.

Sustainment of multiple outcomes varied. Fourteen clinics (11.6% of 121) had long-term sustainment for two of the three outcomes (Fig. 3). Nine clinics sustained improvements in wait time and retention. Three clinics were able to sustain improvements for the retention and admission outcomes and the remaining two clinics for wait time and admissions. When clinics who lagged behind in their improvements for one or more outcomes were considered (i.e., Quadrant IV did not improve and sustain until Period CD), an additional 16 clinics (13.2% of the sample) sustained improvements for two of the three outcomes. The outcomes improved and sustained for these clinics included wait time and retention ($n = 9$), wait and admissions ($n = 5$), and retention and admissions ($n = 2$). Inclusion of clinics that lagged in their improvement suggested that six additional clinics (5% of the sample) may have actually sustained improvements for all three outcomes. The interventions for multiple

sustainment clinics were learning session (n = 10), interest circle (n = 12), combination (n = 8), and coaching (n = 6).

Table 2 showed no significant differences in the average change per month for retention or admissions by intervention arms within a specific quadrant. Average wait time differed by intervention as shown in Quadrant I; clinics assigned to the interest circle intervention experienced greater improvement (reduction) in wait time per month in Period AD than the other intervention arms. Average wait time change per month was lower in the coaching as compared to the learning session intervention in Quadrant II but the difference was not significant. Otherwise no significant differences existed.

The primary long-term sustainment analysis focused on two years after the end of the active NIATx200 intervention period. Long-term sustainment was extended, using modified sustain plots comparing Period AD to Period E, to examine what occurred during the subsequent calendar year (Table 3). The table compared the average change per month per clinic for Period AD versus Period E for each outcome. For wait time, 15 clinics, 48% of the clinics with sustainment in Period AD, continued to show improvements in wait-time during period E. For the retention outcome, 50% of the clinics (n = 10) that sustained improvements through period AD continued to improve the clinic retention rate during Period E. In the 32 clinics sustaining admission improvements through period AD, 16 clinics (50%) exhibited increased clinic admissions through Period E.

4. Discussion

This research addresses a critical gap in sustainability research. It describes the development and evaluation of a methodology to assess long-term sustainment of participation in a quality improvement collaborative (Gustafson et al., 2013). Its application, despite missing clinic data, showed promise in evaluating clinic sustainment of organizational changes two years post end of the active NIATx200 intervention period in 121 substance abuse treatment clinics. Results suggest that improvements experienced by clinics during the NIATx200 intervention period do not necessarily predict long-term outcomes or sustainment. When assessing multiple outcomes, it should not be assumed that sustainment will be consistent across outcomes or be sustained for the same length of time across different organizations. These results suggest that the measure of a “successful” quality improvement intervention is more complex than simply did or did not sustain. Funders, who set accountability goals, may be unaware that the magnitude of intra-agency improvement can vary over time, or that agencies can be highly successful in one outcome domain and not others. Additional sustainment research could help define appropriate expectations and measures for long-term sustainability.

Although clinic interventions varied, sufficient evidence is lacking to infer that participation in one NIATx200 implementation strategy (e.g., coaching versus interest circle) is more likely to lead to the long-term sustainment of individual or multiple outcomes. An analysis by intervention across clinics found few significant differences in the average improvement from Period A to D by sustainment quadrant. It also supports the idea that no one implementation strategy resulted in more clinics sustaining multiple outcome improvements.

The results differed as compared to the primary analysis that showed significant differences in wait time and admission improvements by NIATx strategy. The primary analysis combined with these sustainment results indicate that, while selection of intervention type may help accelerate early improvements, longer-term improvement is not as sensitive to intervention type. These findings are partially driven by the absence of sustainment data from Oregon and Massachusetts. However, it may imply that outcome sustainment is related to the intervention activities themselves (attributes of the intervention) not just the clinic characteristics (attributes of the clinic). In other words, exposure to practices/activities, regardless of the level of implementation support or strategy utilized (e.g., coaching) seems to provide a certain level of sustainment over time. Finally, the findings suggest that sustainment is not uniform; rather, our results indicate that long-term sustainment of one outcome does not guarantee that other outcomes within the same clinic will also be sustained. The fact that sustainment occurs on a continuum rather than all or nothing suggests that strategies (e.g., coaching) that support successful implementation of change may not be the same strategies supporting successful long-term sustainment without additional interventions tailored to specific outcomes. These findings could potentially have implications for rolling out QI activities to participating clinics.

The application of our methodology versus other approaches also warrants discussion. Cubic (or higher degree polynomials) could have been fit to the data but can exhibit volatile behavior, especially at the extremes of the fitted time period, and the coefficients in such models are not easily interpretable. Linear splines are typically better behaved and adequately capture the level and trend of the data within each segment.

Secondly, the linear splines fit to some clinics were more sensitive than others in terms of model fit versus parsimony. As such, an important analytical decision to make a priori is the switch position (on or off) of the C_p statistic/criteria. Once made, it should be applied universally. Users should resist turning on the C_p statistic/criteria post hoc for some organizations and off for others because inconsistent use could bias the results.

Another consideration is that the spline plots utilized the average change per month as the monthly slope of each fitted spline segment. The total change over the segment, which is a rescaled version of the monthly slope could have been utilized in the spline plots. However, monthly slopes can be more readily compared across spline segments with different lengths unlike aggregate changes over each period. The rescaling from monthly to total change would not have affected determinations of statistical significance. Finally, survival modeling is useful in summarizing sustainment of a well-defined binary improvement (i.e., an improvement that can be observed over time as being retained or not – yes or no) for a number of subjects or clusters over time. It also accommodates right censoring (deaths and other departures from the study follow-up) and identifies factors that influence the shape of the survival curve (e.g., sex, age, race). In our study, we examined each subject (clinic) separately to assess their sustainment of the intervention effects and don't have a non-arbitrary standard for whether the clinic has or has not (yes vs. no) sustained the intervention effect. Instead, we are interested in seeing how the observed intervention effect for each clinic continues to grow, plateaus or erodes over time. The spline time series graphs are ideal

for this purpose. The sustainment scatter plots are a better option (more flexible and responsive) for summarizing observed sustainment over several clinics than a survival curve.

Our approach identified several key lessons to consider when assessing long-term sustainment in other healthcare settings. Table 4 offers guidance to other researchers who would like to assess long-term sustainment of their organizational change efforts. Pre-determination of the long-term sustainment analytic criteria a priori will prevent “fishing” expeditions to make the data match the desired parameters to show maximum effect on long-term sustainment. Once established, the initial analytic models can be run, results evaluated, and parameters modified to assess sensitivity of specific changes on long-term sustainment outcomes.

4.1. Limitations

The study had several limitations. Administrative data were not available for two NIATx200 states (Oregon and Massachusetts). As a result, 80 clinics were excluded from the analysis. System changes in the three remaining states (Michigan, New York and Washington) impacted consistency and availability of longitudinal data. Closed clinics or clinics without available data further reduced the ability to explore long-term sustainment and may have masked the actual sustainment of the NIATx200 outcomes across a larger clinic sample. Clinics without sufficient retention and wait time data were more prevalent in New York and Michigan. Clinics were excluded because they did not have sufficient data to be included in the analysis perhaps due to closure (not explicitly asked) or, more commonly, to not having sufficient long-term data to meet the analytical criteria. This methodology assessed clinic sustainment based on a specific set of analytical parameters. Results may vary with different parameters. The selection of sustainment periods was specific to the NIATx200 intervention. However, multiple options were available. For example, we initially considered using state or clinic fiscal years for the post NIATx200 sustainment periods (C through F). The application of the long-term sustainment methodology was specific to one healthcare setting (substance use treatment) and involved multiple implementation strategies within the context of a quality improvement collaborative. As such, an evaluation of the methodology in other healthcare settings (e.g., long-term care) or within the structure of a different quality improvement initiative (e.g., one with a single implementation strategy) is needed to further validate the applicability of the methodology to assess long-term sustainment.

4.2. Conclusions

Research about sustainability capacity and sustainment of change has become increasingly important in dissemination and implementation research (Chambers, 2017; Moore, Mascarenhas, Bain, & Straus, 2017; Shelton, Cooper, & Stirman, 2018). Efforts to systematically assess long-term sustainment represent a gap in sustainability research. The paper addresses the gap through the development and application of a methodology to successfully assess long-term sustainment for clinics that participated in the NIATx200 quality improvement collaborative. Assessment of long-term sustainment within the context of a multi-organizational quality improvement collaborative (QIC) provides a broader understanding of sustainment of organizational change. If “cut-points” indicate when the effects of QIC diminish, implementation interventionists could develop specific

sustainability modules that (a) introduce a conceptual sustainability framework, (b) teach organizations how to assess their likelihood of sustaining change, and (c) guide efforts to develop and implement a sustainability plan (Johnson, Hays, Center, & Daley, 2004; Pronovost et al., 2013; Pronovost et al., 2016; Sridharan, Go, Zinzow, Gray, & Gutierrez Barrett, 2007). It could also lead to the development of QI interventions with appropriately spaced “booster” sessions as an approach to counter program “drift”. Introduction of these modules could be utilized to evaluate their effectiveness in “extending” sustainment within the QIC. Researchers and interventionists alike would learn how to effectively train quality improvement coaches and change team to support organizational sustainment. Finally, “sustainment” cost effectiveness could be evaluated to determine if the relative costs and subsequent effects of investments such as additional coaching or sustainment focused learning sessions examined the impact sustainment on study outcomes.

Acknowledgements and disclosures

We acknowledge the staff of the Michigan Department of Community Health, the Washington Department of Social and Health Services and the New York Office of Alcoholism and Substance Abuse Services who provided the data for this study. We also would like to express appreciation to the substance abuse clinics in these states who participated in the NIATx200 study as well as members of the original NIATx200 research team. This work was supported by the National Institutes of Health (Grant number: R21 DA36700). Dr. Ford is the sole proprietor in JHFHealthcare Consulting, LLC which provides consultancy services to the California Department of Justice. All other authors do not have a conflict of interest related to this work.

Appendix A. Sample sustain plot1 with explanation on interpretation

1. The sustain plot shows the clinic location (circled) in terms of where it falls into one of four quadrants (starting in the upper right and going clockwise): improved and sustained (Quad 1), improved but regressed towards base and did not sustain (Quad 2), did not improve and also did not sustain (Quad 3), and did not improve in AB but the improved rate resulted in an overall improvement and possible sustained for these laggard clinics (Quad 4). On this graph, Period AB for wait time represents Dec 2007 to Dec 2009 and Period AD is Dec 2007 to December 2011.

Appendix B. Retention and admission spline knot interval graphs and tables for MI-32

A

B

References

Aarons GA, Green AE, Trott E, Willging CE, Torres EM, Ehrhart MG, & Roesch SC (2016). The roles of system and organizational leadership in systemwide evidence-based intervention sustainment: A mixed-method study. *Administration and Policy in Mental Health and Mental Health Services Research*, 43(6), 991–1008. [PubMed: 27439504]

- Aarons GA, Hurlburt M, & Horwitz SM (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health*, 38(1), 4–23. [PubMed: 21197565]
- Alharbi TSJ, Ekman I, Olsson L-E, Dudas K, & Carlström E (2012). Organizational culture and the implementation of person centered care: Results from a change process in Swedish hospital care. *Health Policy*, 108(2), 294–301. [PubMed: 23069131]
- Beer M, & Nohria N (2000). Cracking the code of change. *Harvard Business Review*, (May–June), 15–23.
- Bond GR, Drake RE, Becker DR, & Noel VA (2016). The IPS learning community: A longitudinal study of sustainment, quality, and outcome. *Psychiatric Services*, 67(8), 864–869. [PubMed: 27032661]
- Brookman-Frazee L, Stadnick N, Roesch S, Regan J, Barnett M, Bando L et al. Lau A (2016). Measuring sustainment of multiple practices fiscally mandated in children's mental health services. *Administration and Policy in Mental Health and Mental Health Services Research*, 43(6), 1009–1022. [PubMed: 27020796]
- Cantoni E, & Ronchetti E (2001). Resistant selection of the smoothing parameter for smoothing splines. *Statistics and Computing*, 11(2), 141–146.
- Chambers DA (2017). Commentary: Increasing the connectivity between implementation science and public health: Advancing methodology, evidence integration, and sustainability. *Annual Review of Public Health*.
- Chambers DA, Glasgow RE, & Stange KC (2013). The dynamic sustainability framework: Addressing the paradox of sustainment amid ongoing change. *Implementation Science*, 8(117), 1–11. [PubMed: 23279972]
- Fleiszer AR, Semenik SE, Ritchie JA, Richer MC, & Denis JL (2015a). An organizational perspective on the long-term sustainability of a nursing best practice guidelines program: A case study. *BMC Health Services Research*, 15, 535. [PubMed: 26634343]
- Fleiszer AR, Semenik SE, Ritchie JA, Richer MC, & Denis JL (2015b). The sustainability of healthcare innovations: A concept analysis. *Journal of Advanced Nursing*, 71(7), 1484–1498. [PubMed: 25708256]
- Fleiszer AR, Semenik SE, Ritchie JA, Richer MC, & Denis JL (2016a). Nursing unit leaders' influence on the long-term sustainability of evidence-based practice improvements. *Journal of Nursing Management*, 24(3), 309–318. [PubMed: 26081157]
- Fleiszer AR, Semenik SE, Ritchie JA, Richer MC, & Denis JL (2016b). A unit-level perspective on the long-term sustainability of a nursing best practice guidelines program: An embedded multiple case study. *International Journal of Nursing Studies*, 53, 204–218. [PubMed: 26453418]
- Ford JH, 2nd, Alagoz E, Dinauer S, Johnson KA, Pe-Romashko K, & Gustafson DH (2015). Successful organizational strategies to sustain use of A-CHESS: A mobile intervention for individuals with alcohol use disorders. *Journal of Medical Internet Research*, 17(8).
- Ford JH, II, Green CA, Hoffman KA, Wisdom JP, Riley KJ, Bergmann L, & Molfenter T (2007). Process improvement needs in substance abuse treatment: Admissions walk-through results. *Journal of Substance Abuse Treatment*, 33(4), 379–389. [PubMed: 17499961]
- Grazier KL, Quanbeck AR, Oruongo J, Robinson J, Ford JH, 2nd, McCarty D et al. Gustafson DH (2015). What influences participation in QI? A randomized trial of addiction treatment organizations. *Journal for Healthcare Quality*, 37(6), 342–353. [PubMed: 24428632]
- Gustafson DH, Quanbeck AR, Robinson JM, Ford JH, 2nd, Pulvermacher A, French MT et al. McCarty D (2013). Which elements of improvement collaboratives are most effective? A cluster-randomized trial. *Addiction*, 108(6), 1145–1157. [PubMed: 23316787]
- Harris AH, Bowe T, Hagedorn H, Nevedal A, Finlay AK, Gidwani R et al. Christopher M (2016). Multifaceted academic detailing program to increase pharmacotherapy for alcohol use disorder: Interrupted time series evaluation of effectiveness. *Addiction Science & Clinical Practice*, 11(1), 15. [PubMed: 27633982]
- Hoffman KA, Ford JH, 2nd, Choi D, Gustafson DH, & McCarty D (2008). Replication and sustainability of improved access and retention within the Network for the Improvement of Addiction Treatment. *Drug and Alcohol Dependence*, 98(1–2), 63–69. [PubMed: 18565693]

- Hoffman KA, Ford JH, Tillotson CJ, Choi D, & McCarty D (2011). Days to treatment and early retention among patients in treatment for alcohol and drug disorders. *Addictive Behaviors*, 36(6), 643–647. [PubMed: 21324608]
- Huang W, Hunter SB, Ayer L, Han B, Slaughter ME, Garner BR, & Godley SH (2017). Measuring sustainment of an evidence based treatment for adolescent substance use. *Journal of Substance Abuse Treatment*, 83, 55–61. [PubMed: 29129196]
- Hunter SB, Han B, Slaughter ME, Godley SH, & Garner BR (2017). Predicting evidence-based treatment sustainment: Results from a longitudinal study of the Adolescent-Community Reinforcement Approach. *Implementation Science*, 12(1), 75. [PubMed: 28610574]
- Jaffa MA, Lipsitz S, & Woolson RF (2015). Slope estimation for informatively right censored longitudinal data modelling the number of observations using geometric and Poisson distributions: Application to renal transplant cohort. *Statistical Methods in Medical Research*, 24(6), 819–835. [PubMed: 22143404]
- Johnson K, Hays C, Center H, & Daley C (2004). Building capacity and sustainable prevention innovations: A sustainability planning model. *Evaluation and Program Planning*, 27(2), 135–149.
- Kou S (2003). On the efficiency of selection criteria in spline regression. *Probability Theory and Related Fields*, 127(2), 153–176.
- Lagakos S (1979). General right censoring and its impact on the analysis of survival data. *Biometrics*, 139–156. [PubMed: 497332]
- Lash TL, Sorensen HT, Bradbury BD, & Tavis DR (2014). Analysis of linked registry data sets In Gliklich RE, Dreyer NA, & Leavy MB (Eds.). *Registries for evaluating patient outcomes: a user's guide* [internet] (pp. 1–21). Rockville (MD): Agency for Healthcare Research and Quality (US).
- Liu M, & Ying Z (2007). Joint analysis of longitudinal data with informative right censoring. *Biometrics*, 63(2), 363–371. [PubMed: 17425632]
- Lopatto J, Keith S, Del Canale S, Templin M, & Maio V (2014). Evaluating sustained quality improvements: Long-term effectiveness of a physician-focused intervention to reduce potentially inappropriate medication prescribing in an older population. *Journal of Clinical Pharmacy and Therapeutics*, 39(3), 266–271. [PubMed: 24612175]
- McAlearney AS, Terris D, Hardacre J, Spurgeon P, Brown C, Baumgart A, & Nyström ME (2013). Organizational coherence in health care organizations: Conceptual guidance to facilitate quality improvement and organizational change. *Quality Management in Health Care*, 22(2), 86–99. [PubMed: 23542364]
- McCarty D, Gustafson DH, Wisdom JP, Ford J, Choi D, Molfenter T et al. Cotter F (2007). The Network for the Improvement of Addiction Treatment (NIATx): Enhancing access and retention. *Drug and Alcohol Dependence*, 88(2–3), 138–145. [PubMed: 17129680]
- McConnell KJ, Hoffman KA, Quanbeck A, & McCarty D (2009). Management practices in substance abuse treatment programs. *Journal of Substance Abuse Treatment*, 37(1), 79–89. [PubMed: 19195813]
- Moore JE, Mascarenhas A, Bain J, & Straus SE (2017). Developing a comprehensive definition of sustainability. *Implementation Science*, 12(1), 110. [PubMed: 28865479]
- Novins DK, Green AE, Legha RK, & Aarons GA (2013). Dissemination and implementation of evidence-based practices for child and adolescent mental health: A systematic review. *Journal of the American Academy of Child & Adolescent Psychiatry*, 52(10), 1009–1025.e1018. [PubMed: 24074468]
- Peterson AE, Bond GR, Drake RE, McHugo GJ, Jones AM, & Williams JR (2014). Predicting the long-term sustainability of evidence-based practices in mental health care: An 8-year longitudinal analysis. *The Journal of Behavioral Health Services & Research*, 41(3), 337–346. [PubMed: 23709327]
- Pronovost PJ, Demski R, Callender T, Winner L, Miller MR, Austin JM et al. Groups, N. L. C. M. W (2013). Demonstrating high reliability on accountability measures at the Johns Hopkins Hospital. *Joint Commission Journal on Quality and Patient Safety*, 39(12), 531–AP535. [PubMed: 24416944]

- Pronovost PJ, Holzmüller CG, Callender T, Demski R, Winner L, Day R et al. Miller MR (2016). Sustaining reliability on accountability measures at The Johns Hopkins Hospital. *Joint Commission Journal on Quality and Patient Safety*, 42(2), 51–62. [PubMed: 26803033]
- Quanbeck AR, Gustafson DH, Ford JH, 2nd, Pulvermacher A, French MT, McConnell KJ, & McCarty D (2011). Disseminating quality improvement: Study protocol for a large cluster-randomized trial. *Implementation Science*, 6, 44. [PubMed: 21524303]
- Roosa M, Scripa JS, Zastowny TR, & Ford JH, 2nd (2011). Using a NIATx based local learning collaborative for performance improvement. *Evaluation and Program Planning*, 34(4), 390–398. [PubMed: 21371751]
- Salmon D, Yih WK, Lee G, Rosofsky R, Brown J, Vannice K et al. Gellin B (2012). Success of program linking data sources to monitor H1N1 vaccine safety points to potential for even broader safety surveillance. *Health Affairs*, 31(11), 2518–2527. [PubMed: 23129683]
- Scheirer MA (2005). Is sustainability possible? A review and commentary on empirical studies of program sustainability. *American Journal of Evaluation*, 26(3), 320–347.
- Scheirer MA (2013). Linking sustainability research to intervention types. *American Journal of Public Health*, 103(4), e73–e80.
- Scheirer MA, & Dearing JW (2011). An agenda for research on the sustainability of public health programs. *American Journal of Public Health*, 101(11), 2059. [PubMed: 21940916]
- Schell S, Luke D, Schooley M, Elliott M, Herbers S, Mueller N, & Bunger A (2013). Public health program capacity for sustainability: A new framework. *Implementation Science*, 8(1), 15. [PubMed: 23375082]
- Scoville R, Little K, Rakover J, Luther K, & Mate K (2016). Sustaining improvement, IHI white paper. Cambridge, MA: Institute for Healthcare Improvement.
- Shelton RC, Cooper BR, & Stirman SW (2018). The sustainability of evidencebased interventions and practices in public health and health care. *Annual Review of Public Health*, 39, 55–76.
- Singh V, Rana RK, & Singhal R (2013). Analysis of repeated measurement data in the clinical trials. *Journal of Ayurveda and Integrative Medicine*, 4(2), 77. [PubMed: 23930038]
- Smith MA, Wright A, Queram C, & Lamb GC (2012). Public reporting helped drive quality improvement in outpatient diabetes care among Wisconsin physician groups. *Health Affairs*, 31(3), 570–577. [PubMed: 22392668]
- Sridharan S, Go S, Zinzow H, Gray A, & Gutierrez Barrett M (2007). Analysis of strategic plans to assess planning for sustainability of comprehensive community initiatives. *Evaluation and Program Planning*, 30(1), 105–113. [PubMed: 17689317]
- Stetler C, McQueen L, Demakis J, & Mittman B (2008). An organizational framework and strategic implementation for system-level change to enhance research-based practice: QUERI series. *Implementation Science*, 3, 30. [PubMed: 18510750]
- Stirman SW, Kimberly J, Cook N, Calloway A, Castro F, & Charns M (2012). The sustainability of new programs and innovations: A review of the empirical literature and recommendations for future research. *Implementation Science*, 7(1), 17. [PubMed: 22417162]
- Stumbo SP, Ford JH, & Green CA (2017). Factors influencing the long-term sustainment of quality improvements made in addiction treatment facilities: A qualitative study. *Addiction Science & Clinical Practice*, 12(1), 26. [PubMed: 29089054]
- Wu MC, & Carroll RJ (1988). Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics*, 175–188.

Fig. 1.

Linear spline time intervals for sustainment analysis.

¹The pre-intervention baseline period is equivalent to the original NIATx200 baseline period for each outcome. The NIATx200 active implementation period starts one month after the end of the baseline period, continuing to March 2009 (end of active NIATx200 intervention period).

²The short-term sustainment period corresponds to the original study nine-month short-term sustainment period (April 2009 to December 2009).

³The partial fourth year sustainment period differs by outcome due to right censoring of wait time and retention but not admission.

Table Column Explanations

1. Start and end dates of each study period;
2. Average observed (actual) vs. LS fitted value;
3. The least squares fit which varied by outcome (log of admissions) for each time period showing the ending outcome value; change from the previous period in terms of the actual and percent change as well as the change/month; the standard deviation; and T-ratio is an indication if the change was significant.
4. The last set of data columns in the table indicated the number of calendar months included in the analysis and the number of truncated months representing the number of months from the start of the period through the lesser of the end of the period or the last unfiltered observed value for the clinic. To compute the total fitted change in the outcome for each period, we multiplied the slope estimate by the number of truncated months in the period
5. Finally, the last three rows in the table aggregated the impact of the outcome improvements across three time periods showing the data values as outlined in C and D above. The time periods represented were: study periods A and B combined, study periods C through D combined and study periods (A through D) combined

Fig. 2.

MI-32 wait time spline knot chart with explanation of associated table.

Fig. 3.

Clinics sustaining two or more outcomes by implementation strategy.

Notes: Clinics who did not initially improve or sustain a specific outcome (e.g., wait time, retention, or admissions) during the active NIATx intervention (Period AB) but showed sustained improvement during Period CD (Jan 2010 to Dec 2011) were classified as clinics who lagged behind in their overall efforts to improve and sustain a given outcome. These clinics are represented in the “Clinics with Lagged Improvement & Sustainment” group in the Figure. Clinics that improved during the active NIATx intervention and then sustained those improvements are represented in the “Clinics that Improved and Sustained” group in the Figure.

Table 1

Impact of participation in NIAT200 on two-year outcome sustainment.

Sustain plot placement	Wait time		Retention		Admissions	
	N of clinics	% of clinics with sustainment data	N of clinics	% of clinics with sustainment data	N of clinics	% of clinics with sustainment data
Total in Quadrant I	31	40.8%	20	26.7%	32	28.1%
Total in Quadrant II	6	7.9%	11	14.7%	26	22.8%
Total in Quadrant III	29	38.2%	22	29.3%	45	39.5%
Total in Quadrant IV	10	13.2%	22	29.3%	11	9.6%
Clinics with sustainment data ^a	76	62.8%	75	62.0%	114	94.2%
Clinics excluded (not enough or data) ^b	45	37.2%	46	38.0%	7	5.8%

Clinic location in Quadrant I see Fig. 2 for details)	Wait time		Retention		Admissions	
	N of clinics	% of clinics in Quadrant I	N of clinics	% of clinics in Quadrant I	N of clinics	% of clinics in Quadrant I
Total clinics in Quadrant I ^c	31		20		32	
1. Rate of change in Period CD > Period AD	8	25.8%	5	25.0%	6	18.8%
2. Improved by rate of change in Period CD not as rapid as Period AD	12	38.7%	7	35.0%	12	37.5%
3. Rate in Period CD regressed but still overall net sustainment in Period AD	11	35.5%	8	40.0%	14	43.8%

^aBased on total sample of 121 clinics.^bFor the wait time outcome, the state distribution of excluded clinics was Michigan (n = 13), New York (n = 27) and Washington (n = 5). For the retention outcome, excluded clinics distribution was Michigan (n = 13), New York (n = 29) and Washington (n = 4). For the admission outcome, the distribution of excluded clinics was Michigan (n = 3), New York (n = 2) and Washington (n = 2).^cPeriod CD represents the same time period (Jan 2010 to Dec 2011) for all outcomes. Period AD varies by outcome. Wait time: Dec 2007 to Dec 2011; retention: Apr 2008 to Dec 2011; admissions: Oct 2008 to Dec 2011.

Table 2
Comparison of average change per month during Period AD by clinic intervention and quadrant placement.

	Wait time ^a			Retention ^b			Admissions ^c		
	N	Average change per month	F-test	N	Average change per month	F-test	N	Average change per month	F-test
Total Quadrant I	31	-0.40	6.91 ^d	20	0.18%	0.93	32	0.95%	0.29
Coaching	7	-0.17		4	0.19%		6	1.18%	
Combination	8	-0.33		6	0.22%		8	1.05%	
Interest circle	9	-0.73		7	0.12%		10	0.76%	
Learning session	7	-0.27	5.68 ^e	3	0.21%	0.87	8	0.93%	1.03
Total Quadrant II	6	0.17		11	-0.07%		25	-0.86%	
Coaching	4	0.07		5	-0.04%		6	-0.95%	
Combination	0	N/A		4	-0.12%		5	-0.97%	
Interest circle	0	N/A	0.03	1	-0.07%	0.80	6	-0.40%	0.43
Learning session	2	0.39		1	-0.05%		8	-1.06%	
Total Quadrant III	29	0.22		21	-0.23%		46	-1.27%	
Coaching	10	0.20		8	-0.26%	0.64	15	-1.12%	1.32
Combination	6	0.22	0.15	3	-0.14%		14	-1.47%	
Interest circle	9	0.23		6	-0.30%		7	-0.87%	
Learning session	4	0.21		4	-0.14%		10	-1.51%	
Total Quadrant IV	10	-0.11	0.15	22	0.10%	0.64	11	0.52%	1.32
Coaching	3	-0.09		6	0.08%		3	0.90%	
Combination	2	-0.10		3	0.10%		1	0.52%	
Interest circle	1	-0.14		5	0.15%		5	0.24%	
Learning session	4	-0.12	0.03	8	0.08%	0.64	2	0.67%	1.32

^aWait time measured in days is the difference between the date of 1st contact and the date of the 1st treatment session.

^bRetention was measured as the percent of clients with a 1st treatment session making it to the 4th treatment session.

^cAdmissions was measured as the natural log of the monthly admissions.

^dSignificant at $p < 0.001$.

^eTrends towards significance at $p = 0.079$.

Table 3

Comparison of two year (Period AD) versus three year (Period E) long term sustainability.^a

	Number of clinics	Change per month (Period AD) per clinic	Change per month (Period E) per clinic
Wait time^b			
Improved and Sustained or Lagged Improvement in Period AD and continued to improve in Period E	15	-0.30	-0.90
Did not sustain or improved and got worse in Period AD but showed improvement during Period E	15	0.20	-0.70
Improved and Sustained or Lagged Improvement in Period AD and then did not improve in Period E	15	-0.40	1.10
Did not sustain or improved and got worse during Period AD and continued to get worse in Period E	12	0.20	0.90
Retention^c			
Improved and Sustained or Lagged Improvement in Period AD and continued to improve in Period E	10	0.10%	0.31%
Did not sustain or improved and got worse in Period AD but showed improvement during Period E	15	-0.21%	0.83%
Improved and Sustained or Lagged Improvement in Period AD and then did not improve in Period E	19	0.17%	-0.68%
Did not sustain or improved and got worse during Period AD and continued to get worse in Period E	7	-0.21%	-1.07%
Admissions^d			
Improved and Sustained or Lagged Improvement in Period AD and continued to improve in Period E	16	0.70%	1.05%
Did not sustain or improved and got worse in Period AD but showed improvement during Period E	29	-0.91%	2.18%
Improved and Sustained or Lagged Improvement in Period AD and then did not improve in Period E	27	0.92%	-2.53%
Did not sustain or improved and got worse during Period AD and continued to get worse in Period E	30	-1.02%	-1.91%

^aPeriod E (Jan 2011–Dec 2012) is the same for all outcomes. Period AD is Dec 2007–Dec 2011 (wait time), Apr 2008–Dec 2011 (retention) and Oct 2008–Dec 2011 (admissions).

^bWait time is the change per month in the time between 1st date of contact and the 1st treatment session. 68 clinics did not have enough wait time data. The change per period per clinic represents the average days per month of reduction in wait time days.

^cRetention is the percent change per month in the percent of patients retained from first to fourth treatment session. 70 clinics did not have enough retention data to be included in the analysis. Change per period per clinic represents the percent change per month in the clinic retention rate.

^dAdmissions are the change in the percent of admissions per month. 19 clinics did not have enough data to be included in the analysis. Change per period per clinic represents the percent change per month in the clinic admission rate.

Table 4

Guidance for assessing long-term sustainment of organizational change.

Methodological consideration	Description																				
1. Outcome structure	<p>The structure of the outcome will inform how it is address when assessing long-term sustainment and will inform decisions related to case exclusion criteria. An overview of how certain but not all types of outcomes might need to be addressed in terms of right censoring, exclude outliers or transform the outcome is provided below.</p> <table><tr><td>Type of organizational outcome</td><td>Subject to right censoring</td><td>Exclude outliers</td><td>Transform the outcome</td></tr><tr><td>Percentage outcomes (e.g., No-shows)</td><td>Yes</td><td>No</td><td>No</td></tr><tr><td>Time based outcomes (e.g., wait time)</td><td>Yes</td><td>Yes</td><td>No</td></tr><tr><td>Volume counts (e.g., admissions, visits)</td><td>No</td><td>Outcome dependent</td><td>Yes</td></tr><tr><td>Incidence rates (e.g., falls/1000 days)</td><td>Outcome dependent</td><td>Outcome dependent</td><td>No</td></tr></table>	Type of organizational outcome	Subject to right censoring	Exclude outliers	Transform the outcome	Percentage outcomes (e.g., No-shows)	Yes	No	No	Time based outcomes (e.g., wait time)	Yes	Yes	No	Volume counts (e.g., admissions, visits)	No	Outcome dependent	Yes	Incidence rates (e.g., falls/1000 days)	Outcome dependent	Outcome dependent	No
Type of organizational outcome	Subject to right censoring	Exclude outliers	Transform the outcome																		
Percentage outcomes (e.g., No-shows)	Yes	No	No																		
Time based outcomes (e.g., wait time)	Yes	Yes	No																		
Volume counts (e.g., admissions, visits)	No	Outcome dependent	Yes																		
Incidence rates (e.g., falls/1000 days)	Outcome dependent	Outcome dependent	No																		
2. Sustainment period	The researcher must determine the sustainment period length (e.g., 2 years) based on the organizational change efforts and/or intervention structure.If needed, team should request additional data to account for outcome right-censoring.																				
3. Review longitudinal data	An important process step which can help to identify extreme outlier values, determine the time periods for the sustainment analysis, examine case counts, and review outcome data trends																				
4. Analytical Approach	The structure of the organizational change intervention, number and type of implementation strategies and frequency of data collection, will inform different analytical parameters choices. Pre-determination a priori will prevent “fishing” expeditions to match the data and desired parameters to show maximum effect on long-term sustainment. Parameters can be modified to assess sensitivity.																				
5. Case exclusion criteria	Case exclusion criteria should be informed by the longitudinal data review. Decisions include: (a) how to handle time periods with fewer than a certain number of cases, (b) if the data drops off over time, what is the threshold for excluding all subsequent months past that time point, and (c) how will outliers be handled including the exclusion threshold.																				
6. Spline knot graphs	The structure and availability of data will inform the number and type of the different spline “knot” intervals (e.g., baseline, active intervention and sustainment); the unit of time (e.g., months, quarters) will be utilized; and initial time periods will be included in the sustainment analysis (e.g., active intervention versus two year sustainment)																				
7. C _p statistic utilization	The research team should make an a priori decision based on attributes such as organizational size or data volatility to either turn on or turn off the C _p statistic in the initial long-term sustainment analysis.																				
8. Sustain plot structure	Key decisions related to sustain plots include: (a) the x and y axis time periods, (b) the display (e.g., dots, crosses) of individual organizational points, (c) point size to represent confidence in the predictive likelihood that changes have been sustained, and (d) if applicable, the number (one or multiple) of implementation strategies to include on the sustain plot and how the different strategies be visualized on the sustain plot.																				