Implementation Science

RESEARCH

Open Access

Outcomes of the ROSE Sustainment (ROSES) Study, a sequential multiple assignment randomized implementation trial to determine the minimum necessary intervention to sustain a postpartum depression prevention program in agencies serving low-income pregnant people

Jennifer E. Johnson[1*], Shannon Wiltsey-Stirman[2], Alla Sikorskii[3], Ted Miller[4,5], Ellen Poleshuck[6], Tiffany A. Moore Simas[7], Laura Carravallah[8], Raven Miller[9] and Caron Zlotnick[10,11]

Abstract

Background This Sequential Multiple Assignment Randomized Trial (SMART) was conducted to determine minimum implementation support needed for agencies serving pregnant people on public assistance to adopt and sustain the ROSE (Reach Out, Stay Strong, Essentials for mothers of newborns) postpartum depression (PPD) prevention program.

Methods Enrolled prenatal agencies (*N*=98) received thorough initial implementation support (initial training + written sustainment planning). Agencies were identified as at risk for non-sustainment within the first 15 months (*N*=56) were randomized to: (1) no additional implementation support (*N*=12), or (2) quarterly implementation support (coaching and feedback; *N*=44). If agencies receiving quarterly implementation supports were still at risk and within the first 15 months (*N*=29), they were randomized to: (1) continued quarterly support (*N*=14), or (2) monthly implementation support (*N*=15). No implementation support occurred after 18 months. Follow-ups occurred quarterly and then at 18, 24, and 30 months. Outcomes included sustainment of core program elements, agency PPD rates, reach, and costs/cost-effectiveness of each sustainment step.

Results Twice as many agencies as expected (41 of 98; 42%) delivered ROSE with fidelity for 15+ months after receiving thorough initial implementation support only. For agencies at risk for non-sustainment, no effects of adding quarterly implementation supports were observed. However, adding monthly supports (versus quarterly) for agencies still at risk resulted in higher monthly percent of core ROSE elements sustained and more months ROSE was sustained with fidelity with large (Cohen's d = 0.73 and 0.80) effect sizes, and improved reach over 30 months. Many agencies

*Correspondence:
Jennifer E. Johnson
JJohns@msu.edu
Full list of author information is available at the end of the article

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

did not consistently collect PPD rates, making results difficult to interpret. Mean implementation costs (including implementation support and agency staff time) per agency were $1,849 (SD $1,429) for agencies receiving initial implementation support only, $2,699 (SD $1,837) for those receiving initial and quarterly implementation support, and $4,059 (SD $1,763) for those receiving initial, quarterly, and ultimately monthly implementation support.

Conclusions The cost of agency-wide ROSE implementation is far less than the cost of a single untreated case of PPD ($33,484). We suggest implementing ROSE through thorough training and written sustainment planning. For agencies not sustaining, adding monthly support can promote sustainment and improve reach.

Trial registration Registered June 14, 2018 at clinicaltrials.gov, NCT03267563 (https://clinicaltrials.gov/study/NCT03267563).

Keywords Implementation, Sustainment, Cost-effectiveness, Postpartum depression, Prevention, Public assistance, Prenatal care

Contributions to the literature

- More prenatal agencies serving low-income pregnant people than expected (42% vs. 20%) were able to implement and sustain a 5-session evidence-based postpartum depression prevention program (ROSE) with just initial logistical problem-solving, training, and written sustainment planning.
- Monthly (but not quarterly) subsequent implementation support improved sustainment and reach for agencies at risk for non-sustainment.
- Agency-wide ROSE implementation cost a fraction ($1,859-$4,059, depending on level of support) of the cost of a single untreated case of postpartum depression ($33,484).
- Results help build an evidence base for choosing implementation intervention intensity to achieve sustainment while optimizing resource use.

Background

A program is sustained where its core elements continue at sufficient fidelity, intended health benefits continue, and capacity for continuation of core elements is maintained [1–3]. Despite more than a decade of calls for more research on sustainment, an expert consensus report concluded, "Little is known about how well or under what conditions health innovations are sustained and their gains maintained once they are put into practice" [4]. The report prioritized conducting return on investment studies to quantify gains when effective programs are sustained, and cost-benefit trade-offs for effort required to sustain [4]. Limited information on methods and benefits of sustainment can result in: (1) discontinuation despite significant investment in initial implementation or (2) policymakers' uncertainty about the value of devoting resources to implementation and scale-up [4]. The ROSE (Reach Out, Stay Strong, Essentials for mothers of newborns) Sustainment (ROSES) Study [5] was designed to determine the minimum necessary implementation support needed for agencies serving pregnant people on public assistance to adopt and sustain a postpartum depression (PPD) prevention program.

Clinical context

PPD is common and can have lasting consequences, especially among low-income pregnant people and their infants [6–13]. Timely and effective intervention during pregnancy can *prevent* PPD (rather than treat it after birth) [14, 15]. However, professional societies and caregivers have primarily focused on identifying and treating perinatal depression *after* its onset, [16, 17] rather than preventing it.

ROSE is an evidence-based program, recommended by the U. S. Preventive Services Task Force, [18] that prevents half of PPD cases among low-income and racially/ethnically diverse people [19–23]. ROSE is education (rather than treatment), minimizing stigma and improving reach. ROSE teaches interpersonal psychotherapy-based skills for improving communication and social support, identified protective factors for PPD [24–26] ROSE often consists of 4 group sessions during pregnancy and a post-delivery individual booster session. However, ROSE can be offered individually; sessions can be grouped together or split into shorter sessions and offered in a different order. Paraprofessionals and non-mental health professionals (e.g. nurses, health educators) can deliver ROSE with fidelity, increasing flexibility of ROSE delivery [21].

Study rationale and aims

When an implementation intervention does not produce a desired outcome, it can be given more time or intensified. The ROSES Study used a sequential multiple assignment randomized trial (SMART) design in which all agencies received thorough initial implementation support and then were randomized to receive additional implementation supports (or not) if they were found at risk for non-sustainment. The goal was to determine

the minimum support needed to implement and sustain ROSE in agencies providing prenatal services to pregnant people on public assistance, and to determine the costs and benefits of this implementation support. Implementation support was offered using a stepwise approach (i.e., thorough initial implementation support, then quarterly implementation supports if at risk, then monthly implementation support if still at risk) to determine the minimum dose necessary for sustainment. Specific aims were to:

1. Compare effectiveness of each sustainment step (initial, quarterly, and monthly implementation support) for the following outcomes:
    a Percent sustainment of core program elements at each time point (primary), and total length of time that any ROSE services: (i) were provided, and (ii) were provided with adequate fidelity to core elements.
    b Health impact (agency-level PPD rates over time) and reach (percent of patients beginning and percent completing ROSE).
    c Return on investment (costs and cost-effectiveness of each sustainment step).
    d Hypothesized mechanisms including sustainment of: (i) clinical and organizational capacity to deliver core elements, and (ii) engagement/ownership by agency staff.
2. Explore which agencies need which level of support:
    a Explore agency characteristics associated with response to initial and quarterly implementation supports to inform future intervention intensity tailoring.

Innovation

Few randomized implementation studies have identified sustainment as their primary outcome. Identifying the "minimum necessary to sustain" is a novel study methodology [27]. Agency and contextual factors lead to heterogeneity in response to implementation interventions. In this SMART, we addressed heterogeneity by adjusting intervention intensity based on risk of non-sustainment. Results build a needed evidence base for choosing implementation intervention intensity to achieve sustainment while optimizing resource utilization.

This is also the first implementation study of PPD prevention in outpatient prenatal care agencies. Few outpatient preventive mental health interventions not requiring mental health clinicians have been the subject of implementation research [28].

Methods

Participating agencies and agency staff

We enrolled 98 outpatient agencies providing prenatal care or services (such as education, nutrition, support, etc.) in the United States or Puerto Rico between 10/5/2018 and 2/25/2021. The last agency follow-up assessment occurred on 8/20/2023. We chose to include a range of agencies for which ROSE would be appropriate to inform scale-up. Agencies were required to be: (1) outpatient, (2) provide prenatal services, (3) estimate at least 30% of their pregnant patients receive some kind of public assistance (such as cash assistance, food stamps, subsidized housing, and/or Medicaid), (4) average at least 3 new pregnant people per month (i.e., enough patient flow to run ROSE), and (5) agree to study procedures. To support recruitment, initial agency inclusion criteria [5] were expanded early in the trial from 6 states to national, from 50% on public assistance to 30%, from 10+ new people per month to 3+, and from medical clinics exclusively to any agency providing health-related services to pregnant people (e.g., prenatal education, doula organizations, nurse home visiting programs, nutrition programs). Inclusion criteria were for the agency itself; once the agency was included, agencies chose who received ROSE. Most agencies (~70%) provided or offered ROSE to everyone [29].

Study participants included three groups of agency staff: (1) someone chosen by the agency to respond to quarterly survey questions about clinical delivery of ROSE (the "clinical respondent"); (2) someone chosen by the agency to respond to quarterly surveys about operational (billing, scheduling) aspects of ROSE delivery (the "operational respondent," which could be the same or a different person than the "clinical respondent"); and (3) all ROSE facilitators for the agency (who completed self-rated ROSE adherence forms). Agencies often chose individuals leading or closely involved with implementation to fill clinical and operational respondent roles.

Implementation interventions

Study investigators led three additive implementation interventions that were similar in approach but differed in dose (see Table 1) [2]. Meetings and trainings in all conditions took place by videoconference or telephone. Study investigators were also available to answer questions by email.

Thorough initial implementation support

Consisted of 3 virtual meetings taking place before implementation began and covering logistics, clinical training, problem-solving, written sustainment planning, and reimbursement (see protocol paper [5] and

Table 1 Primary implementation strategies in each implementation condition

Implementation strategy	Initial implementation support	Quarterly Implementation Support	Monthly Implementation Support
Obtain formal commitments	One time		
Distribute educational materials (written materials, videos of each agency's 3 training meetings)	One time		
Conduct educational meetings	One time		
Promote adaptability	One time	Quarterly	Monthly
Assess readiness, barriers/facilitators	One time	Quarterly	Monthly
Develop/update formal implementation blueprint	One time	As needed	As needed
Identifying, preparing, and supporting champions	One time	Quarterly	Monthly
Facilitation	One time	Quarterly	Monthly
Capturing and sharing local knowledge	One time	Quarterly	Monthly
Provide technical assistance	One time	Quarterly	Monthly
Make billing easier (e.g., advice/help)	One time	Quarterly	Monthly
Audit and feedback		Quarterly	Monthly
Learning collaboratives		Quarterly	Monthly
Provide clinical supervision		Quarterly	Monthly
Provide ongoing consultation		Quarterly	Monthly
Reexamine the implementation		Quarterly	Monthly

The developer of ROSE (CZ) led clinical trainings and clinically focused implementation meetings. The lead implementation scientist (JJ) led operationally focused implementation meetings. They were joined by SWS, EP, TMS, and LC (implementation scientists and/or maternal health clinicians with decades of experience) for Collaborative Board meetings

implementation manuals [30]). Study investigators first met with the agency leadership team and those likely to oversee ROSE at the agency. This approximately 90-minute overview meeting included: (1) a brief clinical and operational overview of ROSE; (2) an explanation of ROSE core and adaptable elements (Table 2), and (3) logistics planning and an agency-specific customized, written plan. This collaborative process included discussion of how best to fit ROSE core and adaptable elements (e.g., group vs. individual, office vs. home vs. virtual visit, timing of sessions) within agency context, needs, and resources. This process resulted in a written, tailored implementation and sustainment plan that identified who within the agency was responsible for what aspects of sustainment. The second meeting included a 4-hour training on how to conduct ROSE for all agency ROSE facilitators. Facilitators were given the ROSE Program manual, a ROSE participant workbook that included session handouts, a summary of key components, scripts for presenting ROSE to participants, and a customized description of the agency's logistics for ROSE (the agency's sustainment plan). The final meeting (30–60 minutes) included operational staff to discuss issues such as reimbursement, scheduling, and identification and

Table 2 ROSE core and adaptable elements (adapted from Johnson et al., [5])

ROSE core elements	ROSE adaptable elements
Psychoeducation on: • PPD • Managing stress in transition to motherhood • Social support as a buffer against PPD • Relevant postpartum resources Teaching: • Communication skills via role plays • Stress management skills • Building and enhancing social skills Review/reinforcement of skills and resources at postpartum session	Group vs. individual Office vs. home visit In person vs. video vs. telephone Time during pregnancy Order of sessions Open enrollment of group Missed sessions can be made up Sessions can be split into shorter pieces or lumped together Facilitator position and education Language[a] Describing ROSE to pregnant people as "postpartum depression prevention," "skills training," "stress management and support," or "part of our agency program"

[a] ROSE facilitator and participant materials were provided to agencies in English and Spanish (see https://www.womenandinfants.org/rose-program-postpartum-depression). Some agencies also used the English facilitator materials to guide them in offering ROSE in other languages (e.g., Haitian Creole, Vietnamese)

referral procedures. The median time between the first and last training meetings was 21 days (range 2 – 110, with an outlier at 211 days). We provided agencies with recordings of their trainings and their written, agency-specific sustainment plan to replenish staff turnover.

Quarterly implementation support

Added low-intensity coaching and feedback [5] via one clinical and one operational telephone "booster" meeting with the agency and one virtual Collaborative Board meeting with other agencies and study investigators each quarter. Clinical and operational support meetings identified challenges to conducting ROSE with fidelity, collaboratively solved recruitment and attendance problems, discussed re-customization of delivery if needed, and developed an action plan to address barriers. Subsequent meetings reviewed implementation progress and updated the action plan based on new data, experiences, and discussion. During meetings, study staff also provided information about agency fidelity to core ROSE elements based on the ROSE Session-by-Session Adherence Scale (see primary outcome section) and any changes in the agency's rates of PPD for the previous quarter. In the quarterly Collaborative Board virtual meetings, agencies described their successes and challenges sustaining ROSE, asked questions of each other, encouraged each other, and shared information.

Monthly implementation support

Agencies in this condition received the same supports as the quarterly implementation support condition (i.e., clinical support meetings, operational support meetings, and Collaborative Board meetings) on a monthly, rather that quarterly, basis [5].

Characterizing implementation interventions

Every training and implementation meeting was documented in an electronic implementation case note and audio or video recorded. The note included encounter length, a checklist of implementation strategies used, [31] a checklist of discussion topics (e.g., billing options), and free response sections to describe agency responses [5].

Research design

Randomization (see Fig. 1)

After the baseline assessment, all agencies received thorough initial implementation support. At the first time[1] agencies were at risk for operational (defined as no ROSE sessions in 3 months and none scheduled) and/or clinical (defined as less than 75% fidelity to ROSE core elements; see Assessments) non-sustainment (i.e., at 3, 6, 9, 12, 15 months, N=56), they were randomized to receive either no additional (N=12) or quarterly (N=44) implementation support. If agencies receiving quarterly implementation support were still at risk at subsequent assessments up to 15 months (N=29), they were randomized to either continue with quarterly (N=14) or receive monthly (N=15) implementation support. Ratios of 3.7:1 and 1:1 for the first and second randomizations were pre-specified, with more agencies allocated to add quarterly support than to receive no additional support in the first randomization to power the second randomization. The study statistician generated the random allocation sequence. Study research assistants and project coordinator enrolled participants. The study project coordinator emailed the study statistician when an agency needed to be randomized. Randomization used via computerized minimization that assured allocation concealment and balanced trial arms by time of entry into randomization (3, 6, 9, 12, or 15 months) and whether the agency was a Federally Qualified Health Center (FQHC). Additional study follow-ups occurred at 18, 24, and 30 months, but the study team did not provide implementation support after 18 months. The study did not use masking to condition assignment.

Randomization was paused (i.e., all agencies stayed in their assigned conditions, but the assessment timeclock continued) from 3/15/2020 until 7/31/2020 because many agencies closed or limited services when the COVID-19 pandemic started. Because the pause affected all study conditions equally, it did not affect effectiveness conclusions.

Statistical power

Statistical power was based on the primary outcome: monthly average percent sustainment of core ROSE elements. We started with powering the comparison created by the second randomization (monthly vs. quarterly implementation support) using the effect size of Cohen's d=0.48 based on the literature [32–34]. We adjusted for repeated measures to arrive at the required N=19 per group for power of 0.80 at α=0.05 in two-tailed tests. Assuming 2/3 of agencies would still be at risk after receiving quarterly support, the remaining 1/3 or N=9 agencies would be deemed low risk and continue quarterly support. At the first randomization, the quarterly support group would then have size N=57. The size of the initial support only in the first randomization of N=15 would allow us to detect the target effect size with power of 0.89. Assuming 80% of agencies would be at risk after

[1] The assessment and randomization time clock (i.e., for counting 3, 6, 9 months) began when the baseline assessment was completed. Some early agencies completed training before completing the baseline assessment (including the agencies with >90 day training periods). However, early in the trial, we switched to requiring the baseline assessment to be completed before trainings occurred. Therefore, trainings for most agencies occurred during the first assessment quarter.

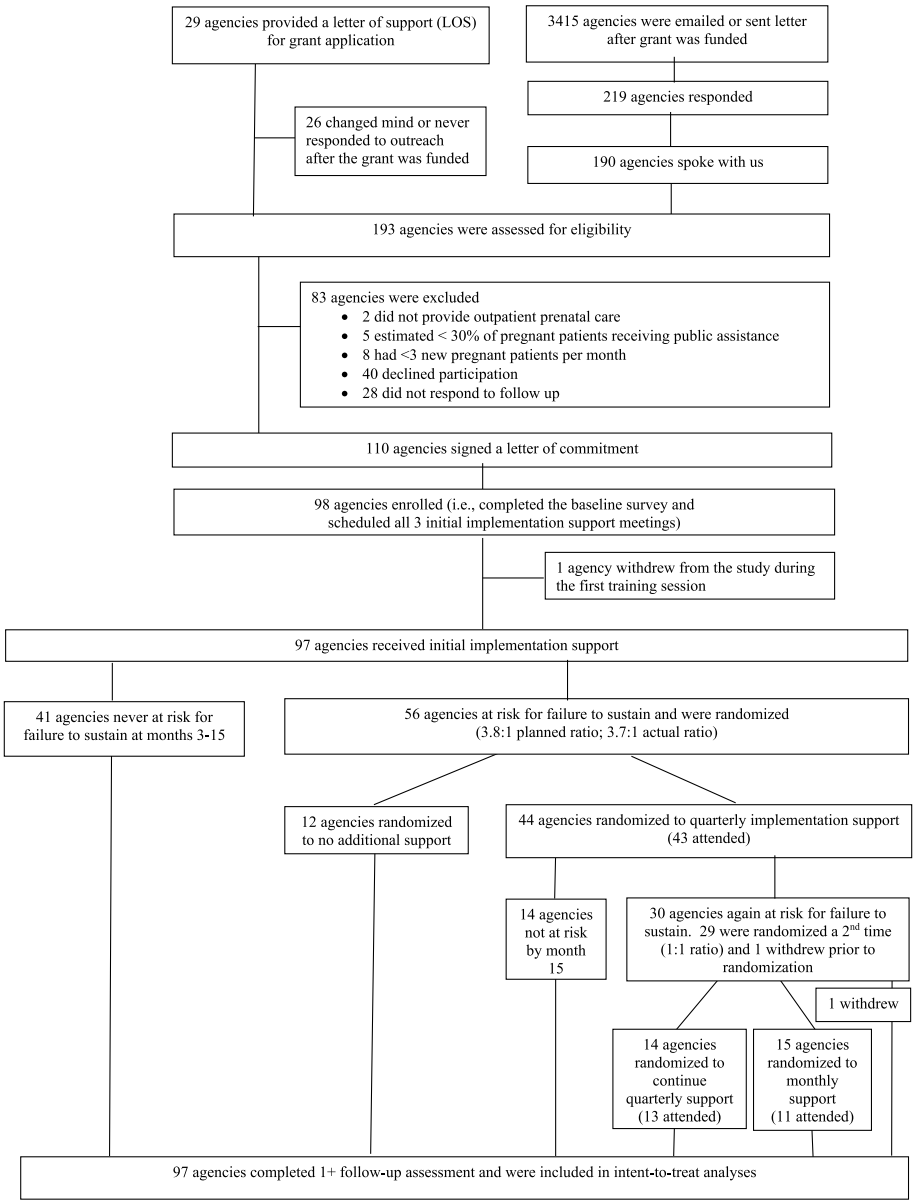

Fig. 1 Agency-level CONSORT Diagram

the thorough initial implementation support only, [34, 35] we planned to enroll n=90 [5]. Because recruitment of agencies often took 4–6 months from first conversation to training, we recruited until we were certain we would train at least 90 agencies, resulting in a sample size of 98 agencies. Johnson et al. [5] contains a diagram of estimated risk and randomization.

Assessments

Assessments were conducted at 0, 3, 6, 9, 12, 15, 18, 24, and 30 months via electronic survey; electronic submission of adherence forms was continuous. We defined *operational non-sustainment* as no ROSE sessions in 3 months and none scheduled. We defined *clinical non-sustainment* as less than adequate fidelity to ROSE core elements (i.e., an average of < 75% of core elements for each session delivered, as measured by the ROSE Session-by-Session Adherence Scale). Used in previous ROSE trials, [20–23] this set of session-specific checklists lists 4–7 items per session (rating of present/absent) assessing whether key tasks (e.g., did the ROSE facilitator explain PPD? Did the facilitator have group members practice communication skills through role plays?) were completed [5]. Self-reported checklists of

mental health intervention fidelity have shown excellent validity when compared to observer-rated scales [36–38].

Aim 1a: Sustainment (primary)

We examined: (a) percent sustainment of core program elements at each time point (primary), and (b) total number of months any ROSE services (i) were provided, and (ii) were provided with adequate fidelity to core elements. Monthly percent sustainment of ROSE's *core program elements*, [5] assessed using the ROSE Session-by-Session Adherence Scale, [21] was defined as the mean percent of core elements delivered that should have been delivered at each ROSE session each month (zero if no sessions were completed). We also used Adherence Scale forms to determine number of months agencies were *offering ROSE* and *offering it with adequate fidelity* (defined as 75%+ core elements present, averaged across all ROSE sessions and facilitators). Time periods were defined using agency enrollment and ROSE session dates. If no forms were returned within 30 days of the sixth reminder, we assumed (and often verified) no ROSE sessions occurred during the time period.

Aim 1b: Health impact and reach

Health impact was defined as PPD rates over time at each agency. Quarterly surveys asked each agency to report the following agency-wide numbers (for all patients, not just those referred to ROSE): (1) number of people who should have come for their 6-week postpartum appointment; (2) number who came; (3) number screened for PPD as part of routine clinical care; and (4) number screened positive for PPD. We also (retrospectively) collected this information for 12 months prior to baseline. We used these numbers to calculate agency-level PPD rates for each period [5]. We also computed PPD rate where its estimate was reliable (10+ people screened per time period), leaving other values missing (where <10 people were screened). If there was no screening in a study period, PPD rate was treated as missing. *Number of people enrolled in and completing ROSE (i.e., reach)* [4]. Electronic surveys assessed the number of: (1) patients attending at least 1 session, and (2) patients attending at least 3 of the 5 sessions ("completing" ROSE), each quarter. Total numbers attending ROSE over 30 months were annualized (divided by 2.5), then divided by number of pregnant people served in the 12 months prior to baseline and multiplied by 100% to estimate percent of pregnant patients at each agency who received ROSE over 30 months.

Aim 1c: Return on investment

We analyzed 4 measures of the cost-effectiveness of implementation support: (1) a primary clinical outcome, number of PPD cases averted, estimated as the change in PPD rate at the agency (post-pre)*(agency's caseload), (2) another clinical outcome, number of quality-adjusted life years (QALYs) saved per PPD case averted, computed from the previous outcome using Luca's $33,484 medical and productivity costs per PPD case [39] and Hewitt's average of 0.26 QALYs lost [40] per PPD case, (3) an implementation process outcome, cost per person served, and (4) a sustainment outcome, cost per month of additional service delivery with fidelity. Our project accounting captured our costs to provide initial, quarterly, and monthly implementation support. As detailed in Supplementary Material 1, we assessed agency costs to receive initial, quarterly, and monthly implementation support using hours that agency staff spent in support meetings, associated direct costs (e.g., printing), and staff salaries, fringes, and overheads. We also assessed the direct service delivery cost of ROSE itself. At each agency delivering ROSE in Months 10–12 after initial training, we had agency staff record their ROSE-related hours for two weeks using a time sheet.

Aim 1d: Mechanisms

Potential mechanisms were assessed at each time point using surveys completed by agencies' clinical and organizational respondents [5]. *Proposed mechanism 1: Clinical and organizational capacity to deliver ROSE*. The primary measure was the Organizational Capacity subscale of the Program Sustainability Assessment Tool (PSAT) [41]. Secondary measures included number of people trained who have time to deliver ROSE and respondents' perceptions that they were able to manage space/scheduling and to bill/get reimbursed for ROSE.

Proposed mechanism 2: Ownership and engagement by agency staff was assessed using the sum of other relevant subscales of the PSAT (primary): Communications, Partnerships, Political Support, and Strategic Planning [41]. The Staff section of the National Health Service's Sustainability Model and Guide (4 questions reflecting staff involvement, staff attitudes, senior leadership engagement, and clinical leadership engagement in sustaining the change) [42, 43] and investment in addressing PPD were secondary measures.

Aim 2a: Predictors

Organizational context was assessed using Aarons' Implementation Climate Assessment [44]. *State policy context* was assessed through two measures: (1) Enacted state legislation about PPD (0 = no enacted state legislation related to PPD, 1 = awareness-related

PPD legislation, 2 = legislation mandating PPD education and services, 3 = legislation with money attached for PPD education/services); and (2) 2018–2021 state-level maternal mortality rate [42].

Analyses

Randomization and analysis took place at the agency level. Primary analyses were intent-to-treat. Statistical tests were two-sided with $\alpha = 0.05$ for sustainment of core program elements (primary) and health impact (secondary) outcomes, specified a priori [5].

Aim 1a: Sustainment

Repeated measures of the monthly percent sustainment of core elements were analyzed via a linear mixed effects (LME) model, with study group from the first randomization treated as time-varying: missing (not assigned) prior to entry into the first randomization, and quarterly or initial support afterwards to month 30, averaging out over the second randomization. A separate LME model was used to analyze agencies entered into the second randomization. Comparison of groups created by the first randomization was designed to answer the question about best first intervention for agencies found to be at risk after initial implementation support only (averaging over the second randomization). Comparison of groups created by the second randomization was designed to answer the question about best second intervention option (quarterly or monthly implementation support) for agencies found to be at risk after receiving quarterly support. The coefficients for the study group variable and 95% confidence intervals (CIs) reflected average differences between randomized trial arms over time after randomization. Effect sizes (Cohen's d) were estimated as differences between adjusted means of trial arms divided by the square root of residual variance. Outcomes that reflected numbers of months (of 30) rather than repeated measures (i.e., total number of months of ROSE delivery and delivery with adequate fidelity) were compared between groups created by the first randomization and separately second randomization using general linear models with covariance adjustment for time of entry into the respective randomization.

Aim 1b: Health impact and reach

Analysis followed an approach similar to Aim 1a, with LMEs for repeated measures of the PPD rate for each time period (3, 6, 9, 12, 15, 18, 24, and 30 months) adjusting for baseline PPD rate. Reach outcomes (percent of pregnant people receiving ROSE at each agency over 30 months) were compared between randomized groups using the non-parametric Wilcoxon rank sum test because highly skewed distributions and smaller sample sizes in the 2nd randomization made other methods (e.g., longitudinal analysis, normal approximation, Poisson) unreliable.

Aim 1c: Return on investment

We used an incremental cost-effectiveness analysis, showing cost-effectiveness ratios for initial implementation support, for adding quarterly implementation support, and for adding monthly implementation support. The cost-effectiveness ratio equals $\Delta C/\Delta E$, where ΔC is the difference in costs as quarterly and monthly support are added, and ΔE is the difference in the outcome measure. We computed cost per QALY saved as (ΔC – medical cost saved per PPD prevented)/(difference in PPD rate). If the cost savings exceed the cost, the added implementation effort offers a net cost saving.

Aim 1d: Mechanisms of effects of increasing implementation supports

(i.e., quarterly vs. initial only, and monthly vs. quarterly)**.** We used mixed models to explore effects of the randomized condition on repeated measures of potential mediators (measures of capacity and engagement/ ownership), adjusting for baseline value of each potential mediator at its first assessment.

Aim 2a: Predictors

Characteristics of agencies (size, percent on public assistance, yes/no FQHC), organizational context (Implementation Climate Assessment score), state policy context (rating of state PPD legislation, state-level maternal mortality rate), and hypothesized mechanisms (capacity and engagement/ ownership assessed using PSAT subscale scores) responding (vs. non-responding) to initial and to quarterly supports were compared using Fisher's exact, t- or Wilcoxon rank sum tests as appropriate based on distributions.

Results

We enrolled 98 agencies from 32 U.S. States. Agency settings ranged from urban to frontier. Table 3 shows agency characteristics. Table 4 shows participant (i.e., agency respondent and ROSE facilitator) characteristics. Figure 1 shows agency flow through the study.

Characterizing implementation interventions

Implementation support (481 meetings total) included 295 meetings with 98 agencies for initial implementation support, an additional 116 meetings providing

Table 3 Agency[a] characteristics at baseline

Characteristic (with N of responses if different from *N*=98 agencies)	N (%) or Mean (SD)
Agency FQHC status	
Yes	9 (9%)
No	89 (91%)
Number of clinical staff (range 0–5,835)	
<5	19 (19%)
5–14	34 (35%)
15–29	16 (16%)
30–64	19 (19%)
>64	8 (8%)
Missing	2 (2%)
Number of years agency has existed	
<5	17 (17%)
5–10	9 (9%)
11–20	19 (19%)
>20	48 (47%)
Missing	5 (5%)
Number of pregnant people seen per year (range 0–25,880)	
<=100	30 (31%)
101–500	32 (33%)
501-1,000	9 (9%)
>1,000	18 (18%)
Missing	9 (9%)
% of pregnant people seen at agency on public assistance (*N*=88)	76.1 (24.0)
Percent race of people served at the agency (*N*=84)	
White	40.3 (30.9)
Other races	57.5 (32.2)
Percent ethnicity of people served at the agency (*N*=83)	
Hispanic	28.2 (26.9)
Not Hispanic	66.7 (31.0)
Percent revenues from:	
Federal or other grant money for care of indigent patients (*N*=75)	27.8 (40.7)
Commercial insurance (*N*=80)	9.7 (16.8)
Medicare or Medicaid (*N*=83)	35.7 (38.4)
No insurance (*N*=78)	8.2 (22.0)
Other (*N*=72)	21.3 (37.4)
PPD rate (*N*=63)	0.18 (0.15)
Hypothesized mechanism 1: Clinical and organizational capacity to deliver ROSE	
PSAT Organizational Capacity subscale (*N*=90, possible range 1–7)	4.81 (1.30)
# of people trained who have time to deliver ROSE at month 3 (*N*=78)	3.66 (4.43)
Perception of agency as able to manage space/scheduling for ROSE at month 3 (*N*=86, possible range 1–7)	5.68 (1.22)
Perception of agency as able to bill/get reimbursed for ROSE at month 3 (*N*=86, possible range 1–7)	3.43 (1.61)
Hypothesized mechanism 2: Ownership and engagement by agency staff	
PSAT other 4 relevant subscales: Communications (d), Partnerships (c), Political Support (b), and Strategic Planning (e) (*N*=95, possible range 1–7)	4.70 (1.52)
Baseline staff section score of the NHS Sustainability Model and Guide	
Staff involvement and training to sustain (*N*=97, possible range 1–4)	3.29 (0.91)
Staff attitudes toward sustaining change (*N*=97, possible range 1–4)	3.57 (0.72)
Senior leadership engagement (*N*=97, possible range 1–4)	3.62 (0.78)
Clinical leadership engagement (*N*=96, possible range 1–4)	3.77 (0.51)
Attitudes toward PPD – awareness (possible range 1–4)	3.04 (0.57)

Table 3 (continued)

Characteristic (with N of responses if different from *N*=98 agencies)	N (%) or Mean (SD)
Attitudes toward PPD – concern - (possible range 1–4)	3.86 (0.18)
Implementation climate (possible range 0–4)	2.48 (0.55)
State policy context (*N*=98)	1.63 (1.24)
No laws mentioning postpartum depression or perinatal mental health	25 (25%)
Awareness-related laws only	22 (22%)
1+ law mandating screening or services	13 (13%)
1+ law mandating screening or services with $ attached	38 (39%)
State maternal mortality rate # of deaths per 100,000 in 2018–2021 (*N*=98)	23.22 (6.95)

[a] Agencies included obstetrics and gynecology offices, attending and resident practices, health systems, nurse home visiting programs, Healthy Start agencies, FQHCs, Women Infants and Children offices, health or childbirth education programs, and doula organizations

Table 4 Participant[a] characteristics

	Agency survey respondents (*N*= 157[b])	ROSE facilitators completing fidelity forms (*N*= 634[b])
Sex (N, %)		
Male	1 (0.7)	17 (2.9)
Female	146 (99.3)	566 (96.9)
Race (N, %)		
American Indian or Alaskan Native	0 (0)	5 (0.9)
Asian	4 (3.3)	16 (2.8)
Black or African American	24 (19.8)	144 (25.4)
Native Hawaiian/Pacific Islander	0 (0)	3 (0.5)
White	82 (67.8)	323 (57.0)
More than One Race	7 (5.8)	34 (6.0)
Other	4 (3.3)	42 (7.4)
Ethnicity (N, %)		
Hispanic or Latino	23 (19.2)	157 (28.5)
Not Hispanic or Latino	97 (80.8)	394 (71.5)
Age (M, SD)	42.8 (10.7)	39.9 (11.0)
Highest degree completed? (N, %)		
High School	12 (8.4)	64 (11.2)
Associates	8 (5.6)	43 (7.5)
Bachelors	33 (23.1)	204 (35.6)
Masters	69 (48.3)	191 (33.3)
Doctorate	21 (14.7)	71 (12.4)
Years of post-degree experience (M, SD)	13.3 (10.0)	9.3 (8.9)
Years in current position (M, SD)	11.2 (6.0)	10.6 (5.6)

[a] Clinical and operational respondents included agency executive and program directors, patient educators, nurses and nurse managers, office managers, care managers, patient navigators, midwives, doulas, and mental health professionals, among others. ROSE facilitators included community health workers, nutrition educators, nurses, doulas, social workers, and trainees, among others

[b] Numbers may not add up due to missing data

quarterly support to 43 agencies, and 70 meetings providing monthly support to 15 agencies. Some quarterly and monthly implementation support meetings included multiple agencies by design. Most used implementation strategies as catalogued using checkboxes in implementation case notes are shown in Table 5. Common discussion topics included recruiting and retaining people in ROSE, including alternate methods of explaining ROSE, and how best to integrate ROSE into current services.

Table 5 Most used implementation strategies (catalogued using checkboxes in implementation case notes)

Most used implementation strategies across initial implementation support meetings (% of meetings using the strategy)	Most used implementation strategies across quarterly and monthly implementation support meetings (% of meetings using the strategy)
Distributing educational materials (76%)	Facilitation (98%)
Conducting educational meetings (65%)	Audit and feedback (56%)
Promoting adaptability (64%)	Capturing and sharing local knowledge (54%)
Assessing barriers/facilitators/readiness (39%)	Assessing barriers/facilitators/readiness (54%)
Developing a formal sustainment blueprint (36%)	Promoting adaptability (47%)
Identifying and preparing champions (35%)	Providing clinical supervision (21%)
Obtaining formal commitments (32%)	Purposely reexamining the implementation (20%)
Capturing and sharing local knowledge (20%)	Using learning collaboratives (Collaborative Board meetings; 15%)
Making billing easier (15%)	Providing technical assistance (15%)

Effects of initial, quarterly, and monthly implementation support

Aim 1a. Sustainment

More than twice as many agencies as the 20% expected (41 of 98; 41.8%) delivered ROSE with adequate (75+%) fidelity for at least the first 15 months after receiving only initial implementation support and were therefore not randomized (see Fig. 1). Total months of sustainment ranged from 0 to 30. A diagram showing patterns of sustainment across agencies, time, and condition is shown in Fig. 2. Of the 85 total randomizations (56 first randomizations and 29 second randomizations) for failure to sustain, 79 (93%) were for not offering ROSE and only 6 were for inadequate fidelity.

We did not observe benefits of adding quarterly implementation supports (relative to initial implementation supports alone) at agencies' first point of risk (Table 6). However, for agencies continuing to be at risk, adding monthly (relative to quarterly) support resulted in significantly higher average monthly percent sustainment of core ROSE elements (the study primary outcome) and more months ROSE was sustained with fidelity, both with large (d = 0.73 and 0.80) effect sizes (Table 6).

Aim 1b.

Health impact Because we enrolled many kinds of agencies (Table 3), only 63 of 98 enrolled agencies, 30 of 56 agencies entering the first randomization, and 15 of 29 agencies entering the second randomization screened for PPD as part of their regular services; even fewer screened enough people (10+ per quarter) to provide reliable estimates (Table 6). Therefore, PPD rate analyses were underpowered and results unreliable. We observed no statistically significant effects of adding quarterly to initial support or of adding monthly to quarterly support on agency-level PPD rates.

Reach A total of 8282 people over the 98 agencies attended at least 1 ROSE session (agency median = 23.5, range = 0–1793) and 3670 attended at least 3 (agency median = 9, range = 0–764). In the 41 agencies sustaining well with initial support alone, 5852 total people attended at least 1 session (a median of 7.01% of each agency's patients over 30 months; Q1 = 2.40%, Q3 = 34.47%); 2769 total people attended at least 3 sessions. Among the 56 at risk agencies, non-parametric tests comparing percent of pregnant people receiving ROSE showed no significant differences when adding quarterly support. Adding monthly support significantly increased percent receiving 3+ sessions ROSE (Table 6).

Aim 1c. Return on investment (costs and cost-effectiveness of each sustainment step)

Cost effectiveness of implementation supports Over an agency's 18 months of implementation support, total implementation costs per agency for initial support-only agencies averaged $1,849 ($886 for implementation support providers and $963 in agency staff time), for initial + quarterly support agencies averaged $2,699 ($1,709 for implementation support providers and $990 in agency time), and $4,059 for initial + quarterly + monthly support agencies ($2,583 for implementation support providers and $1,477 in agency time). Because agencies were in quarterly or monthly support conditions for varying lengths of time, we also calculated cost per quarter. Quarterly implementation support cost a mean of $412 ($322 and $90 respectively) per quarter. Monthly support cost a mean of $441 ($353 and $88 respectively) per quarter. Supplemental Material 1: Tables A1, A2, and A3 show descriptive statistics, cost per session by type of technical assistance, and details about time spent on implementation by support providers vs. agency staff.

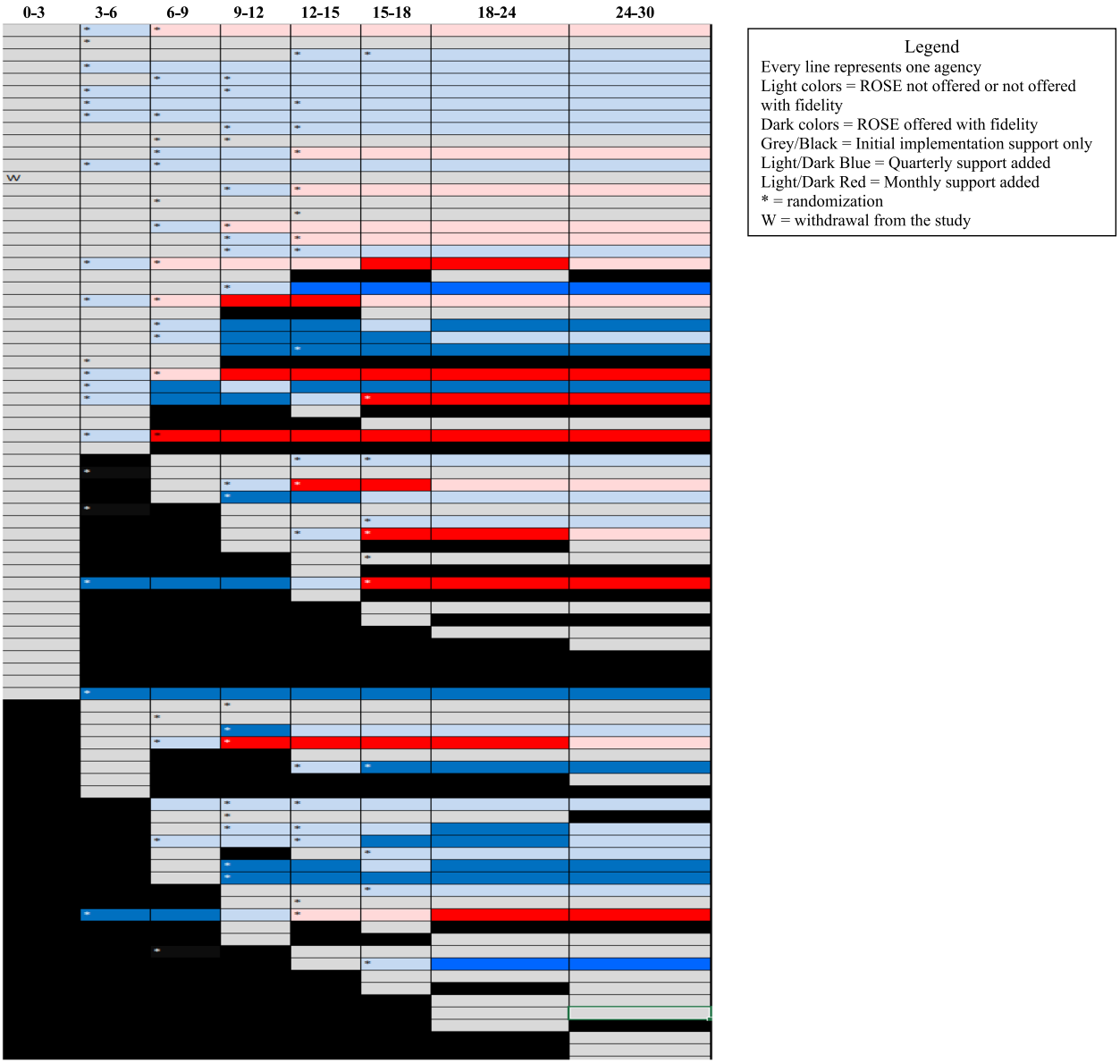

Fig. 2 Time periods in months during which ROSE was offered with adequate fidelity*. *Agencies were not randomized after a period of no ROSE if they: (1) had a ROSE session scheduled (*N*=6); (2) held a ROSE session but no one came (therefore, no fidelity; *N*=2); or (3) the period occurred during the study's randomization pause from 3/15/20 to 7/31/20 for COVID (*N*=3). In two cases, agencies were randomized who were later found to have adequate fidelity; one agency sent fidelity forms >6 months after the end of the reporting period, and another agency had a facilitator who misunderstood the fidelity checklists.

Cost effectiveness of implementation supports Reach. Total implementation costs per person receiving at least one session of ROSE (including both implementation provider and agency staff time) over an agency's 18 months of implementation intervention averaged $189 for initial implementation support only, $458 with only quarterly support added, and $238 with both quarterly and monthly support added. Implementation cost per ROSE attendee in the monthly support condition was lower than in the quarterly support condition because adding monthly support increased reach. Implementation cost per ROSE attendee for monthly implementation support was just $49 more than implementation cost per ROSE attendee with initial implementation support only. *Sustainment.* Relative to quarterly supports and including both implementation support and costs for agency time, monthly agency supports cost an average of $491 per additional month that the agency provided ROSE

Table 6 Outcomes of agency groups created by first and second randomizations, adjusted for month of entry into randomization

Outcome	First randomization				Second randomization			
	Initial + Quarterly Support, *N*=44 Mean (SE)	**Initial Support Only, *N*=12 Mean (SE)**	**Difference (95% CI)**	**P (ES)**	**Initial + Quarterly + Monthly Support, *N*=15 Mean (SE)**	**Initial + Quarterly Support Only, *N*=14 Mean (SE)**	**Difference (95% CI)**	**P (ES)**
Sustainment								
Monthly average percent of core ROSE elements sustained (longitudinal)	0.34 (0.05)	0.18 (0.09)	0.16 (−0.04, 0.37)	.12 (0.28)	0.43 (0.08)	0.13 (0.09)	0.30 (0.05, 0.55)	**.02 (0.73)**
Number of months ROSE was provided	7.00 (1.09)	5.07 (1.95)	1.93 (−2.49, 6.34)	.39 (0.29)	7.11[a] (1.81)	3.38 (1.87)	3.73 (−1.65, 9.10)	.17 **(0.53)**
Number of months ROSE was provided with adequate fidelity	6.02 (0.91)	4.41 (1.79)	1.61 (−2.34, 5.57)	.42 (0.27)	6.68[a] (1.45)	2.20 (1.54)	4.48 (0.16, 8.79)	**.043 (0.80)**
Health impact								
PPD rate (0 −1) (longitudinal)	0.24 (0.03) *N* = 25	0.29 (0.07) *N* = 5	−0.05 (−0.20, 0.10)	.50 (0.20)	0.26 (0.05) *N* = 12	0.36 (0.13) *N* = 3	−0.11 (−0.42, 0.20)	.47 (0.36)
PPD rate (0 – 1) where number screened is 10+ (longitudinal)	0.21 (0.03) *N* = 22	0.22 (0.07) *N* = 5	−0.02 (−0.17, 0.13)	.80 (0.10)	0.21 (0.07) *N* = 11	0.41 (0.16) *N* = 2	−0.21 (−0.60, 0.19)	.18 **(0.73)**
Reach[b]								
	Median (Q1-Q3)	**Median (Q1-Q3)**	**Test statistic**	**P**	**Median (Q1-Q3)**	**Median (Q1-Q3)**	**Test statistic**	**P**
Annualized percent of pregnant people attending at least one ROSE session over 30 months	1.07 (0–4.80) *N*=41	0.46 (0.07–11.00) *N*=11	321.50 (Z=0.67)	.51 (ES=0.09)	1.62 (0.10–7.67) *N*=14	0 (0–2.07) *N*=13	146.5 (Z=−1.74)	.08 (ES=0.33)
Annualized percent of pregnant people attending at least 3 ROSE sessions (of 5) over 30 months	0.10 (0–2.88) *N*=41	0.06 (0–10.00) *N*=11	313.00 (Z=0.48)	.63 (ES=0.07)	0.38 (0–4.27) *N*=14	0 (0–0.01) *N*=13	139.00 (Z=−2.19)	**.03 (ES=0.42)**

For health impact, also adjusted for baseline version of the outcome and repeated measures. For reach, also adjusted for # of new pregnant people per year at baseline. Between-groups *P*-values<.05 and effect sizes (ES)>=0.4 are bolded

[a] Agencies were randomized to the monthly condition between 6 and 15 months of the 30-month trial period and therefore had limited months left to implement and sustain. In addition, some never started ROSE at all (see pink lines in Fig. 2), meaning that several zeroes were included in the average of months of sustainment

[b] Total numbers attending ROSE over 30 months were annualized (divided by 2.5), then divided by number of pregnant people served in the 12 months prior to baseline and multiplied by 100% to estimate percent of pregnant patients at each agency who received ROSE over 30 months. Where numbers reported were inconsistent (e.g., agencies reporting more people attending 3+ sessions than attending 1+ sessions), we conservatively set the number attending 1+ sessions to the number attending 3+ sessions. Missing reach data were treated as zeroes

with fidelity. Agency-level PPD rates were not reliable, and therefore cost-effectiveness analyses using PPD cases averted and QALYs were not conducted. LICF alone was not cost-effective on any of the outcomes assessed.

ROSE costs Across 26 programs that served 952 clients in program months 10–12, the median agency's service delivery cost per person receiving ROSE was $124 ($63 Quartile 1, $309 Q3) and the mean was $246 (SD $291). The median agency's combined cost of implementation support and service delivery was $212 ($87 Q1, $382 Q3). However, costs were highly variable among agencies, with high implementation costs per ROSE attendee for agencies that only provided ROSE to 1 or 2 clients. Costs also varied across kinds of agencies and geographic regions.

Aim 1d: Mechanisms

We did not observe effects of adding quarterly to initial or monthly to quarterly implementation supports on hypothesized mediating variables (see Table 3 for list).

Which kinds of agencies need which level of support (Aim 2a)

Agencies that were able (vs. unable) to sustain after initial implementation support alone had higher rates of staff (but not leadership) involvement and investment (i.e., attitudes) in sustaining ROSE at baseline, and perceived the agency as more able to receive funding or reimbursement for ROSE at 3 months. Agencies able (vs. unable) to sustain ROSE after receiving quarterly implementation support were located in states with less supportive PPD laws and lower maternal mortality rates (Supplemental Table B1).

Discussion

The ROSE Sustainment Study is among the first randomized trials evaluating effectiveness and cost-effectiveness of a stepped approach to sustainment, a critical unanswered question in implementation science. Results indicated that: (1) thorough initial implementation support alone (3 meetings) worked better than expected (41/97 or 42% of cases sustained compared to expected 20% based on the literature [34, 35]); (2) no benefits accrued by adding quarterly clinical, operational, and interagency implementation support for agencies at risk for non-sustainment; but (3) among agencies still at risk after receiving quarterly support, adding monthly clinical, operational, and interagency implementation support improved sustainment of ROSE core elements over 30 months, number of months ROSE was provided with fidelity, and reach. The cost per person served with ROSE in agencies randomized to monthly support was *lower* than for those randomized to quarterly support because monthly implementation support increased reach. The cost of agency-wide ROSE implementation (means of $1,859 to $4,059 depending on level of support received) was far less than the cost of a single untreated case of PPD ($33,484) [39].

Based on these results, we suggest providing thorough training and written sustainment planning at the beginning of ROSE implementation. If the agency is not providing ROSE or not providing it with fidelity, move to monthly support, at least until ROSE delivery is stabilized. This exact sequence (adding monthly support if agency is at risk for non-sustainment and skipping quarterly) was not tested in the current SMART but may be the most efficient use of resources based on our sustainment and cost-effectiveness findings.

An encouraging finding from this trial is that strong initial implementation support, consisting of 3 meetings (logistics planning and a written sustainment plan, clinical training, and then operational/billing problem-solving, copies of recordings of all meetings with the agency) created at least 15 (and up to 30) months of sustainment in 42% of agencies. Previous research has suggested that one-time clinical training alone is usually not sufficient to promote sustained implementation [34]. However, thorough initial implementation support, which included not only clinical training but also a sustainment planning meeting culminating in a written sustainment plan, an operational support meeting, and agency-specific recordings of the 3 initial meetings/trainings, produced sustainment in almost half of agencies in this study.

Non-sustainment in this trial was largely due to agencies not offering ROSE at all versus ROSE quality issues. The vast majority (93%) of randomizations for non-sustainment were for not offering ROSE rather than for inadequate fidelity. Furthermore, agencies needed to begin ROSE to sustain it[2]. Challenges beginning ROSE accounted for many of the randomizations for non-sustainment

[2] Implementation science scholars disagree about when implementation ends and sustainment begins, with suggestions ranging from sustainment starting two years after initial implementation begins to sustainment starting as soon as initial implementation supports end. However, most agree that it is important to focus on sustainment from the beginning of implementation because sustained implementation is the true goal. Therefore, in this study, initial implementation (which consisted of 3 meetings at the beginning of the 30-month assessment period for each agency) included explicit attention to sustainment, such as a written implementation and sustainment plan. After that, implementation supports were withdrawn until and unless: (1) ROSE was not delivered with fidelity during Months 3-18, and (2) the agency was randomized to receive additional support. This occurred early for some agencies, some agencies never started ROSE, and others sustained without additional support for at least 30 months. For simplicity and because imposing a specific cut-off period seemed arbitrary, we used the terms "sustainment" and "at risk for non-sustainment" for agencies any time after the initial training, rather than using terms like "partial implementation" or "non-implementation" to refer to part of the 30-month period and "sustainment" to refer to other parts of the 30- month period.

in this study (see Fig. 2). The pandemic does not explain these randomizations because: (1) we paused randomization for non-sustainment during the first several months of the pandemic (several agencies closed temporarily), and (2) much of our 29-month recruitment and initial training window occurred before the pandemic.

Agencies randomized to monthly support had been unable to implement consistently for 6–15 months. As suggested by Wolfenden et al., [45] these agencies experienced practical barriers and needed more frequent consultation, problem-solving, and discussion with other implementing agencies. Initial implementation included a conversation about ROSE adaptable elements (Table 2) and which delivery and referral options best fit their agency. Quarterly and monthly implementation support meetings helped agencies revisit adaptable elements when their initial choices hit obstacles, reminding them of options they had forgotten.

Predicting which agencies will need additional supports at the outset may not be feasible. We found few consistent differences between responders and non-responders to initial implementation support alone. Therefore, we suggest providing the initial training, checking in with the agency monthly (instead of waiting a quarter), and if they are not offering ROSE or do not have a project date for the first ROSE session, starting monthly assistance.

Strengths and weaknesses

Strengths of the study include the randomized design, manualized protocols for implementation supports, careful characterization of implementation processes, and transparent power and statistical analyses. The study's focus on enhancing care of low-income populations, examination of potential mechanisms and predictors of intervention response, and cost-effectiveness analyses increase its relevance.

Trial weaknesses included challenges with agencies' self-reported PPD rates. To reduce trial burden, we did not ask agencies to PPD screen for the study, but to report numbers already collected if available. Because types of agencies varied (e.g., medical, birth education, nutrition agencies), only 63 agencies screened for PPD (at all) and even fewer screened often enough to make PPD rate estimates reliable. Patterns of data at some agencies (e.g., only 2 people screened per quarter and both screens were positive) suggested that some agencies did not screen everyone. These reporting challenges, combined with more agencies than expected sustaining after initial implementation, created small sample sizes for PPD rate comparisons. Therefore, findings (or lack thereof) for agency PPD rates should be interpreted with caution.

Conclusions

In randomized trials, ROSE prevented half of PPD cases, roughly one PPD case for every six program completers [19–23]. ROSE is sustainable for most agencies at relatively minimal cost (means of $1,859 [initial implementation support only] to $4,059 [initial+quarterly+monthly] including both implementation support providers and agency staff time), far less that the cost of a single case of untreated PPD ($33,484) [39]. Almost half (42%) of prenatal agencies serving low-income pregnant people were able to implement and sustain ROSE with only initial logistical problem-solving, training, and written sustainment planning. Monthly (but not quarterly) subsequent implementation support improved sustainment among agencies at risk. Monthly implementation support had a lower cost per ROSE attendee than did quarterly support because it increased reach. Therefore, we suggest that: (1) ROSE be implemented, and that (2) implementation should provide thorough training and written sustainment planning at the beginning. If the agency does not provide ROSE at all or with fidelity, move to monthly support until delivery is stabilized.

Abbreviations

CI	Confidence interval
ES	Effect size
FQHC	Federally qualified health center
IPT	Interpersonal Psychotherapy
LME	Linear mixed effects
QALY	Quality adjusted life year
PPD	Postpartum depression
PSAT	Program Sustainability Assessment Tool
ROSE Program	Reach Out, Stay Strong, Essentials for mothers of newborns
ROSES Study	ROSE Sustainment Study
SE	Standard error
SMART	Sequential Multiple Assignment Randomized Trial

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13012-025-01420-z.

Supplementary Material 1.

Acknowledgements

We would like to thank Takeya Harris, Morgan Morgan, and Brianca Nuby, research assistants for the trial. We would also like to think Morgan Cinader for her analysis of descriptive data for Table 3. We would especially like to thank the representatives and ROSE facilitators at our 98 prenatal agencies, who completed our quarterly surveys for 30 months, who met with us, who worked to provide ROSE, and whose day-to-day efforts sustain care for low-income pregnant and postpartum people during global pandemics and everything else that life and healthcare throw their way. They are heroes.

Authors' contributions

JJ, CZ, SWS, and AS conceptualized the research project. TM designed and conducted the cost-effectiveness analyses. AS conducted statistical analyses and supervised data cleaning and management. TMS, EP, and LC helped to design agency recruitment procedures, consulted with the study team, and helped to provide implementation support to agencies. LC also conducted qualitative interviews with agencies. RM oversaw recruitment, retention, data collection, data organization, and helped to maintain relationships with the

agencies. JJ drafted the paper with help from AS and TM. All other authors reviewed the manuscript and edited as needed.

Funding
This study was funded by a grant from the National Institute of Mental Health (NIMH; R01 MH114883, PIs Johnson and Zlotnick). NIMH had no role in the design or implementation of the study, analysis of the data, or writing of the manuscript.

Availability of data and materials
ROSE training materials, videos, the written sustainment plan template, and implementation support manuals are available for free at https://www.womenandinfants.org/rose-program-postpartum-depression. Implementation support manuals are also available at cutt.ly/JohnsonFlint (click on "intervention manuals") or by contacting the first author at jjohns@msu.edu. Data are available by contacting Michigan State University's Biomedical Research Informatics Core at support@bric.msu.edu.

Declarations

Ethics approval and consent to participate
The protocol for this study was approved by the Michigan State University Biomedical and Health Institutional Review Board (protocol IP# 00056524). Consent for participation in the study will be obtained using an electronic informed consent form.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Charles Stewart Mott Department of Public Health, Michigan State University, 200 East 1st St Room 366, Flint, MI 48502, USA. [2]National Center for PTSD, Department of Psychiatry and Behavioral Sciences, Dissemination and Training Division and Stanford University, 795 Willow Road (NC-PTSD 334), Menlo Park, CA 94025, USA. [3]Department of Psychiatry, Michigan State University, 909 Wilson Rd Room 321, East Lansing, MI 48824, USA. [4]Pacific Institute for Research and Evaluation, 11720 Beltsville Drive Suite 900, Calverton, New York, MD 20705, USA. [5]Curtin University School of Public Health, Perth, Australia. [6]University of Rochester Medical Center, 601 Elmwood Ave, Box PSYCH, Rochester, NY 14642, USA. [7]University of Massachusetts Chan Medical School/UMass Memorial Health, Memorial Campus, - 119 Belmont Street - Jaquith 2.060, Worcester, MA 01605, USA. [8]Department of Pediatrics and Human Development, Michigan State University, Life Sciences, Bldg, 1355 Bogue St., Rm 240, East Lansing, MI 48824, USA. [9]Charles Stewart Mott Department of Public Health, Michigan State University, 200 East 1st St Room 372, Flint, MI 48502, USA. [10]Butler Hospital and Brown University, 345 Blackstone Blvd, Providence, RI 02906, USA. [11]University of Cape Town, Cape Town, South Africa.

Received: 9 August 2024 Accepted: 16 January 2025
Published online: 10 February 2025

References

1. Wiltsey-Stirman S, Kimberly J, Cook N, Calloway A, Castro F, Charns M. The sustainability of new programs and innovations: a review of the empirical literature and recommendations for future research. Implement Sci. 2012;7:17.
2. Scheirer M. Is sustainability possible? A review and commentary on empirical studies of program sustainability. Am J Eval. 2005;23:320–47.
3. Shediac-Rizkallah M, Bone LR. Planning for the sustainability of community-based health programs: Conceptual frameworks and future directions for research, practice and policy. Health Educ Res. 1998;13(1):87–108.
4. Proctor E, Luke D, Calhoun A, McMillen C, Brownson R, McCrary S, Padek M. Sustainability of evidence-based healthcare: Research agenda, methodological advances, and infrastructure support. Implement Sci. 2015;10:88.
5. Johnson JE, Wiltsey-Stirman S, Sikorskii A, Miller T, King A, Blume JL, Pham X, Moore Simas TA, Poleshuck E, Weinberg R, Zlotnick C. Protocol for the ROSE sustainment (ROSES) study, a sequential multiple assignment randomized trial to determine the minimum necessary intervention to maintain a postpartum depression prevention program in prenatal clinics serving low-income women. Implement Sci. 2018;13(1):115.
6. Lindahl V, Pearson JL, Colpe L. Prevalence of suicidality during pregnancy and the postpartum. Arch Womens Ment Health. 2005;8:77–8.
7. Webb DA, Bloch JR, Coyne JC, Chung EK, Bennett IAM, Culhane JF. Postpartum physical symptoms in new mothers: Their relationship to functional limitation and emotional wellbeing. Birth. 2008;35:179–87.
8. Postmontier B. Functional status outcomes in mothers with and without postpartum depression. J Midwifery Womens Health. 2008;53:310–8.
9. Gavin N, Gaynes BN, Lohr KN, Meltzer-Brody S, Gartlehner G, Swinson T. Perinatal depression: A systematic review of prevalence and incidence. Obstet Gynecol. 2005;106:1071–83.
10. Sohr-Preston S, Scaramella L. Implications of timing of maternal depressive symptoms for early cognitive and language development. Clin Child Fam Psychol Rev. 2006;9:65–83.
11. Federenko IS, Wadhwa PD. Women's mental health during pregnancy influences fetal and infant developmental and health outcomes. CNS Spectrums. 2004;9:198–206.
12. Murray L. The impact of postnatal depression on infant development. J Child Psychol Psychiatr. 1992;33:543–61.
13. Murray L, Cooper PJ. Postpartum depression and child development. Psychol Med. 1997;27(2):253–60.
14. American College of Obstetricians and Gynecologists. Treatment and management of mental health conditions during pregnancy and postpartum: ACOG Clinical Practice Guideline No. 5. Obstet Gynecol. 2023;141(6):1262–88.
15. American College of Obstetricians and Gynecologists. Screening and diagnosis of mental health conditions during pregnancy and postpartum: ACOG Clinical Practice Guideline No. 4. Obstet Gynecol. 2023;141(6):1232–61.
16. Yonkers KA, Smith MV, Lin H, Howell HB, Shao L, Rosenheck RA. Depression screening of perinatal women: An evaluation of the healthy start depression initiative. Psychiatr Serv. 2009;60(3):322–8.
17. Venkatesh KK, Nadel H, Blewett D, Freeman MP, Kaimal AJ, Riley LE. Implementation of universal screening for depression during pregnancy: feasibility and impact on obstetric care. Am J Obstet Gynecol. 2016;215(4):517.e1-8.
18. U.S. Preventive Services Task Force. Intervention to prevent perinatal depression: U.S. preventive services task force recommendation statement. JAMA. 2019;321(6):580–7.
19. Morrell C, Sutcliff P, Booth A, Stevens J, Scope A, Stevenson M, Harvey R, Bessey A, Cantrell A, Dennis C, Ren S, Ragonesi M, Barkham M, Churchill D, Henshaw C, Newstead J, Slade P, Spiby H, Stewart-Brown S. A systematic review, evidence synthesis and meta-analysis of quantitative and qualitative studies evaluating the clinical effectiveness, the cost-effectiveness, safety and acceptability of interventions to prevent postnatal depression. Health Technol Assess. 2016;20(37):1.
20. Zlotnick C, Miller I, Pearlstein T, Howard M, Sweeney P. A preventive intervention for pregnant women on public assistance at the risk for postpartum depression. Am J Psychiatry. 2006;163(8):1443–5.
21. Zlotnick C, Tzilos G, Miller I, Seifer R, Stout R. Randomized controlled trial to prevent postpartum depression in mothers on public assistance. J Affect Disord. 2016;189:263–8.
22. Zlotnick C, Johnson SL, Miller IW, Pearlstein T, Howard M. Postpartum depression in women receiving public assistance: Pilot study of an interpersonal-therapy-oriented group intervention. Am J Psychiatry. 2001;158(4):638–40.
23. Crockett K, Zlotnick C, Davis M, Payne N, Washington R. A depression preventive intervention for rural low-income African-American pregnant women at risk for postpartum depression. Arch Womens Ment Health. 2008;11(5–6):319–25.
24. O'Hara MW, Swain AM. Rates and risk of postpartum depression: a meta-analysis. Int RevPsychiatry. 1996;8(1):37–54.
25. Sidebottom AC, Hellerstedt WL, Harrison PA, Hennrikus D. An examination of prenatal and postpartum depressive symptoms among women

served by urban community health centers. Arch Women Ment Hlth. 2014;17(1):27–40.
26. Beck C. Predictors of postpartum depression: An update. Nurs Res. 2001;50(5):275–85.
27. Kilbourne AM, Abraham KM, Goodrich DE, Bowersox NW, Almirall D, Lai A, Nord KM. Cluster randomized adaptive implementation trial comparing a standard versus enhanced implementation intervention to improve uptake of an effective re-engagement program for patients with serious mental illness. Implement Sci. 2013;8:136.
28. Cooper LA, Ford DE, Ghods BK, Roter DL, Primm AB, Larson SM, Gill JM, Noronha GJ, Shaya EK, Wange NY. A cluster randomized trial of standard quality improvement versus patient-centered interventions to enhance depression care for African Americans in the primary care setting: Study protocol. Implement Sci. 2010;5:18.
29. Johnson JE, Loree A, Sikorskii A, Miller TR, Carravallah L, Taylor B, Zlotnick C. Study protocol for the ROSE Scale-Up Study: Informing a decision about ROSE as universal postpartum depression prevention. Contemp Clin Trials. 2023;132: 107297.
30. Zlotnick C. The ROSE Program. Women and Infants Hospital. https://www.womenandinfants.org/rose-program-postpartum-depression.
31. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, Proctor EK, Kirchner JE. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implement Sci. 2015;10:21.
32. Moyers TB, Manuel JK, Wilson PG, Hendrickson SML, Talcott W, Durand P. A randomized trial investigating training in motivational interviewing for behavioral health providers. Behav Cogn Psychother. 2008;36:149–62.
33. Miller WR, Yahne CE, Moyers TB, Martinez J, Pirritano M. A randomized trial of methods to help clinicians learn motivational interviewing. J Consult Clin Psychol. 2004;72(6):1050–62.
34. Herschell A, Kolko DJ, Baumann BL, Davis AC. The role of therapist training in the implementation of psychosocial treatments: A review and critique with recommendations. Clin Psychol Rev. 2010;30(4):448–66.
35. Beidas RS, Kendall PC. Training therapists in evidence-based practice: A critical review of studies from a systems-contextual perspective. Clin Psychol. 2010;17(1):1–30.
36. Beidas RS, Maclean JC, Fishman J, Dorsey S, Schoenwald SK, Mandell DS, Shea JA, McLeod BD, French MT, Hogue A, Adams DR, Lieberman A, Becker-Haimes EM, Marcus SC. A randomized trial to identify accurate and cost-effective fidelity measurement methods for cognitive-behavioral therapy: project FACTS study protocol. BMC Psychiatry. 2016;16(1):323.
37. Duffy SA, Cummins SE, Fellows JL, et al. Fidelity monitoring across the seven studies in the Consortium of Hospitals Advancing Research on Tobacco (CHART). Tob Induc Dis. 2015;13(1):29.
38. McGrew JH, White LM, Stull LG, Wright-Berryman J. A comparison of self-reported and phone-based fidelity for Assertive Community Treatment (ACT): A pilot study in Indiana. Psychiatr Serv. 2013;64(3):272–6.
39. Luca DL, Margiotta C, Staatz C, Garlow E, Christensen A, Zivin K. Financial toll of untreated perinatal mood and anxiety disorders among 2017 births in the United States. Am J Public Health. 2020;110(6):888–96.
40. Hewitt CE, Gilbody S, Brealey S, Paulden M, Palmer S, Mann R, Green J, Morrell J, Barkham M, Light K, Richards D. Methods to identify postnatal depression in primary care: an integrated evidence synthesis and value of information analysis. Health Technol Assess. 2009;13:36.
41. Stamatakis KA, McQueen A, Filler C, Boland E, Dreisinger M, Brownson RC, Luke DA. Measurement properties of a novel survey to assess stages of organizational readiness of evidence-based interventions in community chronic disease prevention settings. Implement Sci. 2012;65(7):1.
42. Ford JH, Krahn D, Wise M, Oliver KA. Measuring sustainability within the veterans administration mental health. Qual Manag Health Care. 2011;20(4):263–79.
43. Maher L, Gustafson D, Evans A. Sustainability model guide NHS institute for innovation and improvement. 2007.
44. Ehrhart MG, Aarons G, Farahnak LR. Assessing the organizational context for EBP implementation: the development and validity testing of the Implementation Climate Scale (ICS). Implement Sci. 2014;9:157.
45. Wolfenden LSA, Hall A, Bauman A, Nathan N. An initial typology of approaches used by policy and practice agencies to achieve sustained implementation of interventions to improve health. Implement Sci Commun. 2024;5:21.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.