Smartphone-based support for early childhood development in a rural low-resource setting: An individually-randomized pilot feasibility trial from Guatemala.

Scott Tschida[1], Eva Tuiz[1], Duglas López[1], Meylin Canú[1], Vilma Boron[1], Javier Zarazua[2], Maria del Pilar Grazioso[2], Beth A. Smith[3,4,5], Peter Rohloff[1,6,7]

[1] Center for Indigenous Health Research, Maya Health Alliance, Tecpán, Guatemala
[2] Proyecto Aigle Guatemala, Guatemala City, Guatemala
[3] Developmental Neuroscience and Neurogenetics Program, The Saban Research Institute, Children's Hospital Los Angeles, Los Angeles, USA
[4] Division of Developmental-Behavioral Pediatrics, Children's Hospital Los Angeles, Los Angeles, USA
[5] Department of Pediatrics, Keck School of Medicine, University of Southern California, Los Angeles, USA
[6] Division of Global Health Equity, Brigham and Women's Hospital, Boston, USA
[7] Department of Global Health and Social Medicine, Harvard Medical School, USA

Correspondence to: Scott Tschida, Wuqu' Kawoq, 2da Avenida 3-48 Zona 3, Barrio Patacabaj, Tecpán, Chimaltenango. Email: scott.tschida@wuqukawoq.org. Tel: 011 502 4958 1253

ABSTRACT

Objective: To test the feasibility of a smartphone app, BebeApp, that provides evidence-based early childhood development (ECD) guidance in rural Guatemala and pilot procedures for an adequately powered randomized controlled trial.
Design: We conducted an individually-randomized pilot feasibility trial to compare BebeApp to printed ECD guidance.
Setting: This trial was conducted in Tecpán, Chimaltenango, a semi-rural community that is 95% Kaqchikel Indigenous Maya.
Patients: First-time primary caregiver-infant (0-28 days) dyads.
Interventions: BebeApp provides age-dependent, evidence-based caregiver guidance for breastfeeding and complementary feeding, sleep, and developmental support.
Main outcome measures: A mixed-methods evaluation of BebeApp's implementation and acceptability using the RE-AIM framework. App usability, usefulness, and satisfaction were assessed via questionnaires, semi-structured interviews, and app interaction data.
Results: 41 infant-caregiver dyads were enrolled with 40 completing the study. Engagement was high with caregivers opening BebeApp a median (IQR) 11 (6-21) times per month. Usability, was found to be acceptable but no difference was found between pre- and post- measurements. In interviews, caregivers expressed some initial difficulty using BebeApp but were able to gain confidence with a training session. App usefulness and satisfaction responses were positive. Caregivers often noted that BebeApp was their only source of ECD information outside of their family.
Conclusions: We found that caregivers in rural Guatemala responded positively to a smartphone ECD app and engaged with it throughout the trial. Given the urgent need of ECD programs in low- and middle- income settings, smartphone apps may be a method to deliver services directly to caregivers.

- **What is already known on this topic** – There is an urgent need for early childhood development programs (ECD) in low- and middle- income countries (LMICs) with 250 million children at risk of not reaching their developmental potential. Mobile health interventions, such as smartphone apps, are a way to potentially expand ECD program coverage by bringing evidence-based guidance directly to caregivers. However, their feasibility in rural LMICs is not established.
- **What this study adds** – We have extensively evaluated the feasibility of a smartphone-based app to provide evidence-based guidance directly to caregivers in rural Guatemala using mixed-methods and the RE-AIM framework. Although our findings demonstrate that it is feasible, usability is a potential barrier that should be carefully considered.
- **How this study might affect research, practice or policy** – Our study provides evidence that mobile health interventions for caregivers are feasible and are a promising way to increase access to ECD support in LMICs. However, adequately powered trials are needed to determine their effectiveness.

INTRODUCTION

Providing nurturing care to support optimal early childhood development (ECD) is a pressing need globally.[1] Over 40% of children less than five years of age in low- and middle-income countries (LMICs) are at risk of not reaching their developmental potential due to factors such as malnutrition, poverty, and lack of early stimulation.[2,3] Poor ECD has extensive global consequences such as reduced human capital and higher public costs in health and education.[4]

ECD interventions that are scalable in low-resource settings are urgently needed to reach the estimated 250 million children at risk. Mobile health technology designed to provide support directly to caregivers may reduce strain on health workers where ECD services already exist or provide much needed guidance in areas where no ECD services are available. In high-income settings, smartphone apps are ubiquitous and widely used by caregivers for advice on ECD. However, they have been criticized for inclusion of non-evidence based advice as well as being influenced by commercial interests for selling private data to third parties or inclusion of commercially influenced content.[5,6] Smartphones are now common even in rural areas of LMICs where other basic services, including health services, and especially ECD services, are lacking.[7] Therefore, smartphone apps for caregivers have especially high potential for uptake and impact, and for many communities may be their first and only source of ECD support. In Guatemala, 70% of children are at risk of suboptimal development and there is little to no ECD support available in the either the public or private sectors.[3] A few studies in other LMICs have now examined the design and usability of caregiver smartphone apps to support ECD in LMICS, although child-centered outcomes and rigorous effectiveness assessments are still lacking.[8,9]

In this study we conducted an individually-randomized pilot feasibility trial of a smartphone-based ECD app, BebeApp, in a low-resource rural area of Guatemala. BebeApp provides age-based recommendations for children under 2 years of age in the areas of breastfeeding and complementary feeding, sleep, and developmental support. It also provides an interface for caregivers to enter and visualize their infant's data. A pilot feasibility trial was necessary before an adequately powered randomized controlled trial (RCT) due to the lack of detailed implementation data on smartphone caregiver apps in LMICs and the need to validate trial and assessment procedures, including of recruiting and retention of subjects into a non-smartphone control arm.[10] Therefore, our primary outcomes were focused on whether a smartphone ECD app for caregivers could be implemented in rural Guatemala. While our team has extensive experience adapting and administering the Bayley Scales of Infant and Toddler Development version 3 in Guatemala, this was the first time we administered version 4. Therefore, we included effectiveness as a secondary outcome to pilot version 4 of the Bayley Scales of Infant and Toddler Development (BSID4). We have specified milestones to transition to a future, adequately powered effectiveness RCT *a priori*.

METHODS

We prepared this manuscript according to the Consolidated Standards of Reporting Trials (CONSORT) extension to randomized pilot and feasibility trials reporting guidelines (Appendix, S6).[11]

Study context, ethics oversight, and funding

This study was conducted in collaboration with Maya Health Alliance, a primary care organization working in rural Indigenous Maya communities in Guatemala. This study was conducted in the municipality of Tecpán, Chimaltenango (population 95 000), a semi-rural community where more than 95% of the population is of Indigenous Kaqchikel Maya ancestry. This study was conducted according to the principles in the Declaration of Helsinki, prospectively registered (NCT05106894), and approved by the Institutional Review Boards of Children's Hospital Los Angeles (CHLA-21-00168) and Maya Health Alliance (WK 2021 002). This work was funded by a grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development to BAS and PR(5R33HD107983-04).

Smartphone app design

We employed an agile design framework to develop an Android-based app. Agile design is a software development philosophy in which a software product is iteratively developed with end users actively participating throughout the development. We have previously used this design philosophy to develop software for lay midwives in rural Guatemala.[12] The design team included app programmers, investigators, clinical staff at Maya Health Alliance, and volunteer caregivers/end users. We first conducted separate focus groups with caregivers and clinical staff at Maya Health Alliance to enquire about their experience using smartphone apps and their opinion on caregiver support smartphone apps. We presented existing ECD apps from high-income settings and explained our interest in developing a similar app. Based upon feedback from these preliminary sessions we developed a prototype and conducted focus groups again to elicit feedback. This iterative process of frequent updates based upon end user feedback was repeated throughout the first year of development. Once BebeApp's core features were finished, caregivers tested it in-home. Although target participants were primarily speakers of Kaqchikel Maya, the app was developed in Spanish as Kaqchikel Maya is not widely written or read. Participants who speak Kaqchikel Maya typically read and write in Spanish. App design included built-in multilingual support to permit future additional languages, and as many images as possible. Evidence-based caregiver recommendations in the app for breastfeeding and complementary feeding, sleep, and developmental support were selected from evidence-based sources (Table S1). An overview of app design, version release, and design changes is given in Figure S1 and Table S2.

Trial design and participants

This was a single-center, individually-randomized (1:1 allocation ratio), parallel-group pilot feasibility trial to compare use of the smartphone app to printed educational materials (Figure S2). Feasibility trials evaluate whether a future, adequately RCT, can realistically be done.[11] Pilot trials are a subgroup of feasibility trials that conduct at least part of a future effectiveness RCT. First-time primary caregivers and their infant were eligible to participate if the infant was 0-28 days old at the time of recruitment and a singleton full-term (≥ 37 weeks) birth. Infant exclusion criteria included: acute malnutrition (weight-for-length/height Z score ≤ −2 SD using WHO standards), severe medical illness, or medical contraindication to breastfeeding. Caregiver exclusion criteria were inability to read or speak Spanish. Eligible caregivers who did not possess a functioning

Android smartphone were given a phone for the duration of the study. This was necessary for 9 out of 21 participants. Iterative changes to the app were made throughout the study based on user feedback as described in Table S2.

Participants were recruited by trained research nurses within usual clinical care settings after referral by healthcare providers working at Maya Health Alliance. Verbal informed consent was obtained following Maya Health Alliance protocol for a minimal-risk study with Indigenous populations as obtaining signatures is not a culturally common practice. Study visits and data collection occurred in the home by trained research nurses, except for the secondary outcome variable (Bayley Scales of Infant and Toddler Development, Version 4; BSID4; Pearson Assessments, San Antonio, TX, USA) which was collected by consulting psychologists in clinical spaces at Maya Health Alliance.

After enrollment and randomization, caregivers received either the smartphone app, installed with assistance from a study nurse, or printed materials. Brief monthly visits (~15 minutes) were conducted monthly for 5 consecutive visits by a study nurse to resolve questions about the app or, in the comparator arm, to resolve any caregiver questions about the printed materials. BSID4 data was collected at enrollment and study exit at six months post enrollment.

Outcomes

The primary outcomes were a mixed methods evaluation of app implementation and acceptability using semi-structured interviews, questionnaires on usability, usefulness, and satisfaction, and on-phone interaction data. We followed the Reach, Effectiveness, Adoption, Implementation, Maintenance (RE-AIM) evaluation framework as outlined in Table 1.[13]

We pre-defined three milestones to transition to a subsequent, adequately powered effectiveness RCT: (1) an app failure rate/crash rate of less than 2%, (2) more than 50% of caregivers use the app at least twice a month, and (3) more than 50% of caregivers find the app useful (measured via questionnaire).

TABLE 1: OVERVIEW OF PARTICIPANT-LEVEL MIXED-METHODS IMPLEMENTATION EVALUATION USING THE RE-AIM FRAMEWORK.

RE-AIM Dimension	Evaluation method/outcome definitions
Reach	Quantitative: - Enrollment and drop-out - Comparison of demographic characteristics of trial participants to overall typical population served by Maya Health Alliance Qualitative: - Semi-structured interviews with caregivers: barriers to app use
Effectiveness	Quantitative: - Exploratory comparison of BSID4 composite scores at 6 months and correlation between BSID4 scores at 0 and 6 months Qualitative:

	- Semi-structured interviews with caregivers: Usefulness of app content
Adoption	Not applicable: single-center pilot without involvement of multiple providers or health care units.
Implementation	Quantitative: - Number of visits completed, time between visits, and missing data - System usability scale[14] - Usefulness and satisfaction questionnaire[a] - User app interaction data: frequency of opening app; duration of app sessions; response to push notification; frequency of infant data entry Qualitative: - Semi-structured interviews with caregivers: acceptability, usability, interest of the app
Maintenance	Qualitative: - Semi-structured interviews with caregivers: interest in continuing app use

[a]Questionnaire was based upon multiple surveys in the literature.[15,16] Survey items were tested and tailored for contextual relevance.
Abbreviations: BSID4, Bayley Scales of Infant Development, Version 4

Additional details on the development of the mixed-methods approach and accompanying instruments are given in Supplementary Methods 1-4 and Table 1.

Secondary outcomes were the BSID4 language, motor, and cognitive composite scores, calculated as defined in the manufacturer's user manual. We chose the Bayley Scales of Infant and Toddler Development as it has been extensively used in Guatemala by our team and others. Details about our adaptation of version three of the Bayley for administration in Guatemalan Spanish and Kaqchikel Maya have been published previously.[17] As this study was the first time our team used version 4 of the Bayley, we conducted a similar adaptation to ensure contextual appropriateness. This included: translation, back translation, and cultural adaptations. Ten of the infants were below 16 days of age at the first BSID4 assessment; an age of 16 days was used to calculate their composite scores. BSID4 assessments were conducted by skilled clinical psychologists in Spanish. Psychologists were blinded to study arm assignment. Covariables included basic demographic and infant/caregiver clinical information, household poverty probability using the Guatemalan Simple Poverty Scorecard, and infant anthropometrics. All anthropometric measurements were completed in triplicate. Weight was measured to the nearest 0.1 kg with the use of a Seca 310 hanging scale (Seca, Hamburg, Germany), and length/height was measured to the nearest 0.1 cm with a locally-constructed portable length board.[18]

Sample size, randomization, and enrollment

As a pilot feasibility trial, as recommended by Leon et al.,[19] no confirmatory hypothesis testing was planned and therefore no formal power or sample size calculations were performed. We aimed to enroll 40 infant-caregiver dyads as this was reasonable from recruitment, study period, and study resource perspectives and would allow us to evaluate intervention feasibility while piloting trial procedures before an adequately powered RCT. A block randomization sequence (blocks of 4, 1:1 allocation) was generated in advance using an online program (https://www.sealedenvelope.com/) by an individual not associated with participant recruitment or study procedures. This list was then uploaded in digital data capture software (REDCap, Vanderbilt University, Nashville, TN) which

automatically disclosed subject allocation to the study nurse upon completion of informed consent procedures.

For semi-structured interviews, trial participants from the intervention arm were selected based upon app engagement. All participants were separated into quintiles based upon how many times per month they had opened the app. Initially, we decided to interview 4 participants based on our experience in prior similar studies in Guatemala where small samples sizes of 4-6 interviews typically reach thematic saturation. Two participants in the lowest quintile and two users in the highest quintile were selected at random. After a fifth interview elicited little to no new information, we determined that we had reached thematic saturation.[20]

Data analysis

Quantitative

Quantitative statistical analysis was conducted using Stata Version 17. Demographic and clinical characteristics of study participants were summarized using means and standard deviations; median and interquartile range; or percentages, as appropriate. For secondary BSID4 outcomes, raw differences between intervention and control arms were summarized and adjusted differences calculated using linear regression with adjustments chosen *a priori* for gender of child, caregiver educational level, poverty score, birth weight and change in length for age Z score from 0 to 6 months.

Qualitative

Qualitative analysis was conducted using Dedoose (Sociocultural Research Consultants, Manhattan Beach, CA). We carried out a thematic analysis using a deductive approach. The codebook was developed *a priori* based upon the dimensions of the RE-AIM framework with selected implementation themes derived from the Consolidated Framework for Implementation Research (Supplementary Methods 4).[21] ST coded all interviews and PR reviewed all codes. Selected representative quotes are presented in the Results section.

Mixed methods

We integrated our quantitative and qualitative findings using an explanatory sequential joint display technique to help elicit meta-inferences from the quantitative and qualitative sections.[22]

Patient and Public Involvement

Patients or the public were not involved with the design, conduct, reporting, or dissemination plans of this research.

RESULTS

Reach

Quantitative

Participants were recruited from January 2023 to March 2023. Final study participants completed the study in July 2023. In total, 163 subjects were referred to the study team of whom 39% had inaccurate contact information and were unreachable and 35% were ineligible, principally due to being older than the eligible infant age range or not a first-time caregiver. A total of 41 primary caregiver-infant dyads were enrolled and underwent randomization (control arm 21, intervention arm 20) (Figure 1). Baseline demographic and clinical features of participants in the two study arms are given in Table 2. Caregivers in the control arm tended to have lower educational attainment, and children tended to be younger at enrollment and have higher length-for-age Z-scores, although these differences were not statistically significant. One subject in the intervention arm withdrew consent and was not included in the analysis; resulting in a final, analyzed sample of control arm = 21 and intervention arm = 19.

TABLE 2: BASELINE DEMOGRAPHIC AND CLINICAL CHARACTERISTICS OF STUDY PARTICIPANTS

Characteristic	Smartphone (intervention) arm (n=19)	Control arm (n=21)	P value
Maternal characteristics			
Age in years, mean ± SD	22.4±2.7	23.3±5.4	0.53[1]
Maya ethnicity, % (n)[2]	94.7 (18)	100 (20)	0.48[3]
Education primary school or less, %(n)	36.8 (7)	71.4 (15)	0.06[3]
Child characteristics			
Female sex, % (n)	52.6(10)	61.9 (13)	0.75[3]
Age at enrollment in days, median (IQR)	20 (13-24)	14 (9-16)	0.06[4]
Birth weight, kg[5]	2.89±0.08	2.99±0.07	0.35[1]
Length-for-age Z score[6]	-2.48±0.19	-2.00±0.18	0.07[1]
Weight-for-length Z score[6]	0.57±0.21	0.18±0.20	0.19[1]
Household characteristics			
Raw poverty score[7]	39.5±2.3	33.9±3.5	0.23[1]

[1]Student's t-test; [2]Self-identified ethnicity; [3]Fisher's exact test; [4]Wilcoxon rank sum test; [5]caregiver reported; [6]based on World Health Organization growth standards; [7]raw poverty score can range 0-100. A raw score of 35-39 corresponds to a 66% probability of living below the 100% national poverty line; lower scores indicate more poverty.
Abbreviations: SD, standard deviation; IQR, interquartile range

As a further evaluation of reach, several demographic characteristics of the current study population differed from those of a prior nutrition intervention study with caregivers and young children drawn from the clinical population of Maya Health Alliance in this same community.[23] In both studies participants almost universally self-identified as Indigenous Maya. However, caregiver age was lower in this study than in the prior, likely because this pilot feasibility trial only enrolled first-time caregivers (pooled mean age 22.9 ± 4.3 vs. 27.1 ± 6.8 years). Raw poverty scores were also higher, indicating less poverty (pooled mean score 36.6 ± 4.1 vs. 27.8 ±

8.0); and baseline height-for-age Z-scores were higher, indicating less stunting (pooled mean Z-score -2.23 ± 0.30 vs. -3.44 ± 0.74).

Qualitative

The main barrier to app use was that the app was challenging to use in the beginning. All five of the participants interviewed felt there was an initial learning curve for the application. However, after the training session with the study nurse, they were able to navigate the app and gain confidence over time.

> "The first time it was difficult for me to use the app because I didn't know how to do it. Although they explained everything to me it was still a bit difficult until a little later when I managed to understand everything." - Participant 2

Effectiveness

Quantitative

Of 41 randomized participants, one withdrew consent. 40 participated in BSID4 assessments at enrollment and 6 months. Correlation between enrollment (median age < 1 month, see Table 1) and 6-month post-enrollment BSID4 subscale scores assessed using a Pearson's correlation coefficient matrix were poor (Table S3). While debriefing with study psychologists, team consensus was that the BSID4 assessments at less than 1 month were unreliable due to limited number of observation items in the early infant assessment and rounding up to 16 days for those infants younger at time of assessment. Further evidence of unreliability is provided by the lack of subscale score stability across the two study time points (Figure S3). Composite BSID4 cognitive, motor, and language scores as well as unadjusted and adjusted differences between trial arms were compared at the 6-month endpoint. No significant difference for any subscale was observed (Table 3). Poverty was the only covariate with a significant interaction (Table S4).

TABLE 3: BAYLEY SCALES OF INFANT DEVELOPMENT SCORES, VERSION 4 COMPOSITE SCORES IN STUDY PARTICIPANT AT 6 MONTHS

Characteristic	Smartphone (intervention) arm[1]	Control arm[1]	Unadjusted difference (95% CI)[2]	Adjusted difference (95% CI)[3]
Cognitive composite	111 ± 9	111 ± 10	0.6 (-5.6, 6.7)	-1.0 (-7.4, 5.5)
Motor composite	105 ± 9	106 ± 11	-1.5 (-7.8, 4.9)	-2.3 (-8.3, 3.6)
Language composite	104 ± 6	104 ± 6	0.4 (-3.5, 4.2)	-1.8 (-5.6, 2.1)

[1]BSID4 evaluations occurred at 6 months after enrollment (median age in days 165.5 [IQR 155, 174.5]. BSID4 composite scales are given as means ± SD. [2]Student's T-test for unpaired samples. [3]Adjusted difference calculated using an ordinary least squares linear regression model with adjustments for gender of child, caregiver educational level, poverty score, birth weight and change in length for age Z score from 0 to 6 months.

Abbreviations: BSID4: Bayley Scales of Infant Development Scores, Version 4; SD, standard deviation; IQR, interquartile range; 95% CI: 95% Confidence Interval

Qualitative

All five interviewees found the app content useful. Three of five noted that it was especially useful for them as first-time caregivers that did not have experience raising children.

> "... [the app] helped me a lot because sometimes, as a first-time mother, you don't know what to do and these pieces of advice help, they help you say, ah yes that makes sense..." - Participant 1

Two of five interviewees stated that they would like more advice on additional topics including health, cleaning and bathing, and psychological advice for mothers.

> "I would like [the app] to have advice about how we can regulate our emotions as mothers. It is very difficult and no one tells us how to go through raising a child. Emotions vary while raising a child. I would like psychological advice for new mothers." - Participant 4

Implementation

Quantitative

All but two (95%) of the participants participated in all scheduled check-in visits to review app usage (intervention arm) or printed stimulus materials (control arm). Two individuals in the intervention arm received two and three visits, respectively. No subjects in either arm had missing baseline or endline outcome data. The mean duration between baseline and endline assessments was 157 ± 5 days in the intervention arm and 156 ± 5 days in the control arm.

Table 4 details user engagement with the app. The median (IQR) number of times users opened the app per month was 11 (6-21). On average, 10% of push notifications about new advice were opened while 7% of push notifications requesting data entry were opened. The median (IQR) times caregivers entered data per month was 10 (3-24).

TABLE 4: USER-APP INTERACTION DATA (n=19)

Outcome	Median (IQR)
Number of times app was opened per month	11 (6-21)
Percentage of advice push notifications opened	10 (0-35)
Percentage of data entry push notifications opened	7 (0-63)
Number of times caregivers entered data per month	10 (3-24)

Abbreviations: IQR. Interquartile range

Table 5 details the results of pre- and post-surveys given to participants in the intervention arm about their experience using the app at enrollment and 6-months post-enrollment. Feedback on both app usability, usefulness, and satisfaction were high and stable over the intervention timeline. Individual item summary responses are given in Table S5.

TABLE 5: PRE- AND POST-SURVEY RESULTS ASSESSING APP USABILITY, APP USEFULNESS, AND USER SATISFACTION AMONG PARTICIPANTS IN THE SMARTPHONE INTERVENTION ARM

Survey	Pre-	Post-	P value[1]
System Usability Score[2], mean (SD)	70 (9)	66 (9)	0.25
Usefulness and satisfaction scale[3], mean (SD)	1.2 (0.4)	1.3 (0.3)	0.50

Abbreviations: SD, standard deviation. [1]Student's t-test; [2]From 0-100, higher values indicate better usability; [3]Likert scale from -2.0 to 2.0, higher levels indicate higher usefulness and satisfaction.

Qualitative

Overall, the app was acceptable for all interviewed caregivers. Interviewees expressed how the app benefited them and how it was often their only source of information regarding child development apart from their family.

> "Yes, I think it there is a need for this app not just here but everywhere in Guatemala because it goes directly to first-time mothers. We don't know if what we are doing is good or bad. The app helps us understand a bit more. We can combine the advice from the app with what our family members say. Therefore, I think it is important and very useful for first-time mothers." - Participant 1

Interviewees expressed that the design of the app was appropriate. Specifically, images used for navigation and images paired with advice were highlighted as being helpful and appreciated as being representative of their communities. However, two out of the five interviewees noted that they had some difficulties navigating the app.

Maintenance

Qualitative

All interviewees indicated they would like to continue using the app. A potential barrier to continuing use of the app, however, was the amount of currently available caregiver support messages. As noted in the effectiveness section, caregivers mentioned the desire for additional supplementary content in more subject areas.

Mixed Methods Integration

Table 6 details meta-inferences about implementation of a future adequately powered effectiveness trial that emerged for the reach and implementation domains. Overall, although two

caregivers noted that they had had some difficulties navigating the app, user engagement, satisfaction, and usability were found to be high. For many users the app was often the only source of information on child development readily available to them. Initial training to use the app was an important activation barrier, and a larger-scale implementation will require a more streamlined user training process including, most likely, an in-app tutorial.

TABLE 6: EXPLANATORY SEQUENTIAL JOINT DISPLAY FOR REACH AND IMPLEMENTATION DURING A PILOT FEASIBILITY RANDOMIZED TRIAL TO PROVIDE SMARTPHONE-BASED SUPPORT FOR EARLY CHILDHOOD DEVELOPMENT TO CAREGIVERS IN GUATEMALA

RE-AIM dimension	Quantitative findings	Qualitative findings	Meta-inferences
Reach	Baseline characteristics of intervention arm: • The mean (SD) age of caregivers was 22.4 (2.7) years • Primarily indigenous (94.7%) • High levels of poverty (66% probability of living below the 100% national poverty line) • Low level of educational attainment	The main barrier to app use identified was the initial learning curve of the app. After a one-on-one training session with the study team users were able to use the app.	Enrolled mothers in this study are comparable to first-time caregivers in the study area. Rolling out the app in the real world may require integrated training modules to replace the one-on-one training sessions that took place during the beginning of the study.
Implementation	• Caregivers engaged with the app throughout the study period • Caregivers found the app usable • Caregivers found the app satisfactory and useful	• The app was acceptable and was often the sole source of information for users on child development • The design of the app was well done and contextually appropriate • Some users had residual difficulties navigating the app	Likely due to being one of the only sources of information available on early childhood development, caregivers demonstrated high levels of engagement and satisfaction. Possible improvements in usability should be considered.

Transition milestones

All three transition milestones to support transition to an adequately powered RCT to evaluate effectiveness were reached. The app failure rate was 0%. All caregivers used the app, on average, at least twice per month and all caregivers found the app to be useful and satisfactory via questionnaires (Table S5).

DISCUSSION

In this study, we conducted an individually-randomized pilot feasibility trial to implement a smartphone app that provides first-time caregivers with evidence-based ECD guidance in rural Guatemala. BebeApp was well received by caregivers and scored positively on acceptability usability, usefulness, and satisfaction. We easily exceeded our prespecified endpoints on app utilization as well as perceived usefulness and acceptability. In addition, we successfully piloted procedures before an adequately powered RCT. We did not have issues with control arm recruitment or retention and we were able to pilot version 4 of the Bayley Scales of Infant and Toddler Development. Taken together, these findings suggest that mHealth interventions targeting caregivers in rural areas of LMICs are feasible and are a promising avenue for expanding coverage of ECD support.

Although the acceptability of mHealth interventions in general in LMICs is well established,[7] most are implemented with providers, and there is little research assessing their acceptability for caregivers.[24] Three studies in Peru, Brazil, and Kenya found that caregivers responded positively to interventions based on messaging platforms (Facebook, WhatsApp, SMS).[8,25,26] Our findings, which include qualitative interviews and questionnaires, support the limited evidence that mHealth interventions are well received by caregivers and provides an assessment of a dedicated smartphone app. In Guatemala, like many LMICs, ECD services are not available despite their pressing need. During qualitative interviews, caregivers recognized the lack of ECD knowledge and support in their community and therefore valued BebeApp as it was their first experience receiving ECD support outside their family. Additionally, acceptance was driven by the tailoring of app content and design to ensure contextual appropriateness. Interviewees expressed interest in additional modules such as health, cleaning, bathing, and psychological advice. As part of our agile design process, we will evaluate their viability and implement them in the future as able.

Several prior studies of caregiver mHealth interventions in LMICs have noted challenges with user engagement. [8,24,27] Conversely, we found that caregivers had high levels of engagement throughout the feasibility trial. This is likely attributed to three factors. First, our iterative co-design and testing process ensured usability and that content was adapted and relevant for the local context.[24,27] Second, caregivers had consistent reminders via push notifications and in-person monthly visits with study staff.[24] Third, as the child ages, new age-appropriate advice becomes available incentivizing caregivers to continue engaging. A concern for future long-term scale-up is how engagement would be impacted without in-person visits with study staff, which will be addressed in more detail in the future adequately powered RCT.

Despite 43% (9/21) of caregivers not owning a smartphone, usability was found to be acceptable.[28] The usability of the app is likely due to two factors. First, our iterative codesign process which included regular input with caregivers ensured that the app design was contextually appropriate. Second, study staff assisted with the initial install of the app and trained caregivers in app use.[8,29] However, there was no significant change between initial and end-point usability suggesting that the app did not get easier to use over time. This suggests that app usability can still be improved and that an integrated training module/tutorial would be beneficial. A training module/tutorial would also be a necessity for future scale-up. It is likely that usability concerns in rural LMICs will be less of an issue in the future as smartphone and app exposure continues to increase.

In preparation for an adequately powered RCT to evaluate effectiveness, we successfully piloted trial and assessment procedures. First, we did not have any issues with enrollment and retention. We were initially concerned that participants in the control arm would not be interested in participating without receiving the smartphone app and we were uncertain how study participants would respond to the BSID4 evaluation. This pilot feasibility trial was the first time our team used the BSID4, which has significant differences from the older BSID3. We found that there was a poor correlation between 0- and 6-month subscale scores. Kvestad et al. also found that BSID3 score stability was weak in early infancy but that it improved in early childhood.[30] As a result, in the future adequately powered RCT, we will focus on a single older effectiveness timepoint (18 months) and use a conservative effect size for power/sample size calculation. In addition, we will add two additional supplementary outcomes of development and nurturing care in the effectiveness trial: the Nursing Child Assessment Satellite Training Scale (NCAST) and the Home Observation Measurement of the Environment (HOME).[31,32]

This pilot feasibility trial has several limitations to consider. First, our participants come from a single area of rural Guatemala and therefore our findings may not be generalizable to other contexts. Second, as a pilot feasibility trial, the sample size and study length were not adequate to assess effectiveness and the reported BSID4 scores should not be interpreted in this way. Strengths of the study include our comprehensive mixed-methods evaluation guided by an established implementation science framework. To promote inclusion, caregivers who did not own a smartphone were given one for the study duration, a procedure we will continue in the adequately powered RCT.

In conclusion, we found that first-time caregivers in rural Guatemala valued a smartphone-based app that provided evidence-based ECD guidance. Engagement was found to be high throughout the pilot feasibility trial and in qualitative interviews caregivers shared that the app was their only source of ECD guidance outside their family. Given the absence of ECD programs in many LMICs, smartphone apps that deliver evidence-based guidance directly to caregivers are a promising method to increase accessibility to ECD services in LMICs. An adequately powered RCT to evaluate effectiveness of BebeApp is now ongoing.

Data accessibility statement

De-identified raw datasets and statistical code are available on the Harvard Dataverse (https://doi.org/10.7910/DVN/UHBCVQ).[33]

Author contributions:
Drafting of manuscript: ST PR
Design of work: ST DL MC BAS PR
Acquisition of data: ET DL MC VB JZ MDG
Analysis of data: ST PR
Interpretation of data: ST MDG BAS PR
Critical review of manuscript for intellectual content: All authors
Final approval of version submitted to publication: All authors
Agreement to be accountable for all aspects of the work: All Authors
ST is the guarantor.

Funding statement:
This work was funded by a grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development to BAS and PR (5R33HD107983-04).

Ethical approval statement:
The trial was conducted according to the principles in the Declaration of Helsinki, prospectively registered (NCT05106894), and approved by the Institutional Review Boards of Children's Hospital Los Angeles (CHLA-21-00168) and Maya Health Alliance (WK 2021 002).

Trial registration and protocol:
https://clinicaltrials.gov/study/NCT05106894

Competing interests:
All authors declare they have no competing interests.

Figure Legend

I. Figure 1 – Enrollment flow chart

REFERENCES

1. Daelmans B, Manji SA, Raina N. Nurturing Care for Early Childhood Development: Global Perspective and Guidance. Indian Pediatr. 2021 Nov 15;58 Suppl 1:S11-S15. PMID: 34687182.
2. Lu C, Cuartas J, Fink G, McCoy D, Liu K, Li Z, Daelmans B, Richter L. Inequalities in early childhood care and development in low/middle-income countries: 2010-2018. BMJ Glob Health. 2020 Feb 4;5(2):e002314. doi: 10.1136/bmjgh-2020-002314. PMID: 32133201; PMCID: PMC7042597.
3. Lu C, Black MM, Richter LM. Risk of poor development in young children in low-income and middle-income countries: an estimation and analysis at the global, regional, and country level. Lancet Glob Health. 2016 Dec;4(12):e916-e922. doi: 10.1016/S2214-109X(16)30266-2. Epub 2016 Oct 4. PMID: 27717632; PMCID: PMC5881401.
4. Richter LM, Daelmans B, Lombardi J, Heymann J, Boo FL, Behrman JR, Lu C, Lucas JE, Perez-Escamilla R, Dua T, Bhutta ZA, Stenberg K, Gertler P, Darmstadt GL; Paper 3 Working Group and the Lancet Early Childhood Development Series Steering Committee. Investing in the foundation of sustainable development: pathways to scale up for early childhood development. Lancet. 2017 Jan 7;389(10064):103-118. doi: 10.1016/S0140-6736(16)31698-1. Epub 2016 Oct 4. PMID: 27717610; PMCID: PMC5880532.
5. Virani A, Duffett-Leger L, Letourneau N. Parenting apps review: in search of good quality apps. Mhealth. 2019 Sep 30;5:44. doi: 10.21037/mhealth.2019.08.10. PMID: 31620471; PMCID: PMC6789304.
6. Cheng H, Tutt A, Llewellyn C, Size D, Jones J, Taki S, Rossiter C, Denney-Wilson E. Content and Quality of Infant Feeding Smartphone Apps: Five-Year Update on a Systematic Search and Evaluation. JMIR Mhealth Uhealth. 2020 May 27;8(5):e17300. doi: 10.2196/17300. PMID: 32459187; PMCID: PMC7287747.
7. McCool J, Dobson R, Whittaker R, Paton C. Mobile Health (mHealth) in Low- and Middle-Income Countries. Annu Rev Public Health. 2022 Apr 5;43:525-539. doi: 10.1146/annurev-publhealth-052620-093850. Epub 2021 Oct 14. PMID: 34648368.
8. Jäggi L, Aguilar L, Alvarado Llatance M, Castellanos A, Fink G, Hinckley K, Huaylinos Bustamante ML, McCoy DC, Verastegui H, Mäusezahl D, Hartinger Pena SM. Digital tools to improve parenting behaviour in low-income settings: a mixed-methods feasibility study. Arch Dis Child. 2023 Jun;108(6):433-439. doi: 10.1136/archdischild-2022-324964. Epub 2023 Mar 28. PMID: 36977547; PMCID: PMC10313948.
9. Alam M, Hickie IB, Poulsen A, Ekambareshwar M, Loblay V, Crouse J, Hindmarsh G, Song YJC, Yoon A, Cha G, Wilson C, Sweeney-Nash M, Troy J, LaMonica HM. Parenting app to support socio-emotional and cognitive development in early childhood: iterative codesign learnings from nine low-income and middle-income countries. BMJ Open. 2023 May 16;13(5):e071232. doi: 10.1136/bmjopen-2022-071232. PMID: 37192801; PMCID: PMC10193093.
10. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, Bakken S, Kaplan CP, Squiers L, Fabrizio C, Fernandez M. How we design feasibility studies. Am J Prev Med. 2009 May;36(5):452-7. doi: 10.1016/j.amepre.2009.02.002. PMID: 19362699; PMCID: PMC2859314.
11. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, Lancaster GA; PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ. 2016 Oct 24;355:i5239. doi: 10.1136/bmj.i5239. PMID: 27777223; PMCID: PMC5076380.

12. Martinez B, Hall-Clifford R, Coyote E, Stroux L, Valderrama CE, Aaron C, Francis A, Hendren C, Rohloff P, Clifford GD. Agile Development of a Smartphone App for Perinatal Monitoring in a Resource-Constrained Setting. J Health Inform Dev Ctries. 2017;11(1):http://www.jhidc.org/index.php/jhidc/article/view/158/212. PMID: 28936111; PMCID: PMC5604479.
13. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, Ory MG, Estabrooks PA. RE-AIM Planning and Evaluation Framework: Adapting to New Science and Practice With a 20-Year Review. Front Public Health. 2019 Mar 29;7:64. doi: 10.3389/fpubh.2019.00064. PMID: 30984733; PMCID: PMC6450067.
14. Lewis JR. The System Usability Scale: Past, Present, and Future. Int J Hum Comput Interact. 2018;34(7):577-590. doi:10.1080/10447318.2018.1455307.
15. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly. 1989;13(3):319. doi:10.2307/249008.
16. Lund A. Measuring usability with the USE questionnaire. Usability Interface. 2001;8(2):3-6.
17. Martinez B, Cardona S, Rodas P, Lubina M, Gonzalez A, Farley Webb M, Grazioso MDP, Rohloff P. Developmental outcomes of an individualised complementary feeding intervention for stunted children: a substudy from a larger randomised controlled trial in Guatemala. BMJ Paediatr Open. 2018 Oct 3;2(1):e000314. doi: 10.1136/bmjpo-2018-000314. PMID: 30306144; PMCID: PMC6173251.
18. Contreras M, Palomino C. Elaboración y Mantenimiento de Infantómetros y Tallímetros de Madera. Lima (Peru): Instituto Nacional de Salud; 2007.
19. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res. 2011 May;45(5):626-9. doi: 10.1016/j.jpsychires.2010.10.008. Epub 2010 Oct 28. PMID: 21035130; PMCID: PMC3081994.
20. Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. PLoS One. 2020 May 5;15(5):e0232076. doi: 10.1371/journal.pone.0232076. PMID: 32369511; PMCID: PMC7200005.
21. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci. 2022 Oct 29;17(1):75. doi: 10.1186/s13012-022-01245-0. PMID: 36309746; PMCID: PMC9617234.
22. Guetterman TC, Fetters MD, Creswell JW. Integrating Quantitative and Qualitative Results in Health Science Mixed Methods Research Through Joint Displays. Ann Fam Med. 2015 Nov;13(6):554-61. doi: 10.1370/afm.1865. PMID: 26553895; PMCID: PMC4639381.
23. Martinez B, Webb MF, Gonzalez A, Douglas K, Grazioso MDP, Rohloff P. Complementary feeding intervention on stunted Guatemalan children: a randomised controlled trial. BMJ Paediatr Open. 2018 Apr 27;2(1):e000213. doi: 10.1136/bmjpo-2017-000213. PMID: 29719876; PMCID: PMC5926563.
24. Jäggi L, Hartinger SM, Fink G, McCoy DC, Alvarado Llatance M, Hinckley K, Ramirez-Varela L, Aguilar L, Castellanos A, Mäusezahl D. Parenting in the Digital Age: A Scoping Review of Digital Early Childhood Parenting Interventions in Low- and Middle-Income Countries (LMIC). Public Health Rev. 2025 Jan 21;45:1607651. doi: 10.3389/phrs.2024.1607651. PMID: 39906193; PMCID: PMC11790347.
25. Altafim ERP, de Oliveira RC, Pluciennik GA, Marino E, Gaspardo CM. Digital Parenting Program: Enhancing Parenting and Reducing Child Behavior Problems. Children (Basel). 2024 Aug 14;11(8):980. doi: 10.3390/children11080980. PMID: 39201915; PMCID: PMC11352771.

26. Wambui E, Nampijja M, Okelo KO, Muendo R, Onyango S, Kimani-Murage EW, Kitsao-Wekulo P. Perceptions of the influence of a mobile phone-based messaging platform on caregiver ECD knowledge, attitudes and practices: a qualitative exploration in an informal settlement in Nairobi. BMC Prim Care. 2023 Sep 14;24(1):187. doi: 10.1186/s12875-023-02127-0. PMID: 37710157; PMCID: PMC10502968.
27. Loblay V, Ekambareshwar M, Naderbagi A, Song YJ, Ford M, Zahed I, Yoon A, Hickie IB, LaMonica HM. Enhancing equitable engagement for digital health promotion: Lessons from evaluating a childrearing app in Indonesia. Digit Health. 2023 Dec 25;9:20552076231222112. doi: 10.1177/20552076231222112. PMID: 38152442; PMCID: PMC10752113.
28. Sauro J, Lewis JR. Quantifying the User Experience. 1st ed. San Francisco (CA): Elsevier; 2012.
29. Greenall-Ota J, Yapa HM, Fox GJ, Negin J. Qualitative Evaluation of mHealth Implementation for Infectious Disease Care in Low- and Middle-Income Countries: Narrative Review. JMIR Mhealth Uhealth. 2024 Dec 13;12:e55189. doi: 10.2196/55189. PMID: 39670953; PMCID: PMC11660726.
30. Kvestad I, Hysing M, Ranjitkar S, Shrestha M, Ulak M, Chandyo RK, Strand TA. The stability of the Bayley scales in early childhood and its relationship with future intellectual abilities in a low to middle income country. Early Hum Dev. 2022 Jul;170:105610. doi: 10.1016/j.earlhumdev.2022.105610. Epub 2022 Jun 16. PMID: 35728398.
31. Bernard KE, Eyres SJ. Child Health Assessment. Part 2: The First Year of Life. Hyattsville (MD): Department of Health, Education, and Welfare; 1979.
32. Totsika V, Sylva K. The Home Observation for Measurement of the Environment Revisited. Child Adolesc Ment Health. 2004 Feb;9(1):25-35. doi: 10.1046/j.1475-357X.2003.00073.x. PMID: 32797621.
[dataset] 33. Tschida, Scott, Eva Leticia Tuiz, Duglas Lopez, Meylin Canú, Javier Zarazua, Maria del Pilar Grazioso, Beth A. Smith, and Peter Rohloff. 2026. “Replication Data for: Smartphone-Based Support for Early Childhood Development in a Rural Low-Resource Setting: An Individually-Randomized Pilot Feasibility Trial from Guatemala.” Harvard Dataverse. https://doi.org/10.7910/DVN/UHBCVQ.