Integrating an early childhood development programme into Bangladeshi primary health-care services: an open-label, cluster-randomised controlled trial

Jena D Hamadani, Syeda F Mehrin, Fahmida Tofail, Mohammad I Hasan, Syed N Huda, Helen Baker-Henningham, Deborah Ridout, Sally Grantham-McGregor

Summary

Background Poor development in young children in developing countries is a major problem. Child development experts are calling for interventions that aim to improve child development to be integrated into health services, but there are few robust evaluations of such programmes. Previous small Bangladeshi trials that used individual play sessions with mothers and their children (at home or in clinics), which were predominantly run by employed women, found moderate improvements on child development. We aimed to integrate an early childhood development programme into government clinics that provide primary health care and to evaluate the effects of this intervention on child cognition, language, and motor development, growth, and behaviour in a subsample of the children.

Methods In this open-label cluster-randomised controlled trial, we recruited individuals from community clinics in Narsingdi district, Bangladesh. These clinics were randomly selected from a larger sample of eligible clinics, and they were assigned (1:1) to either deliver an intervention of 25 sessions, in which mothers of eligible children were shown how to support their child's development through play and interactions, or to deliver no intervention (control group). Participants were underweight children, defined as a weight-for-age Z score of –2 SDs of the WHO standard, who were aged 5–24 months and who lived near the clinic (defined as a walk of less than 30 min). Government health workers ran these sessions at the clinics as part of their routine work, and mothers and children attended fortnightly in pairs (instead of individual weekly home visits that were specified in the original programme). A subsample of children from each clinic was randomly selected for impact evaluation, and these children were assessed on the Bayley Scales of Infant and Toddler Development for their cognitive, language, and motor performance and for their behaviour with Wolke's ratings, before and after implementation of the intervention. The primary outcomes were the performance of this evaluation subsample on the Bayley and Wolke scales and their anthropometric measurements (weight, length or height, and head circumference) after 1 year of the intervention. This study is registered with ClinicalTrials.gov, number NCT02208531.

Findings Between Nov 29, 2014, and April 30, 2015, 12 054 children in 90 clinics were screened, and between six and 25 underweight children were enrolled from each clinic. From the 2423 (20%) underweight children, we excluded 656 (27%) children who lived more than 30-min walking distance from the community clinics, and 30 (1%) children whose mothers did not consent to participate. We therefore enrolled 1737 (72%) children from these 90 clinics. After randomisation, the control group clinics included 878 (51%) children (who all received no intervention) and the intervention group clinics included 859 (49%) children (who all received the child development programme sessions). Eight children from each clinic (360 [41%] children from the control group clinics and 358 [42%] children from the intervention group clinics) were randomly selected for inclusion in the evaluation subsample. Between Feb 24, 2016, and Sept 7, 2016, 344 (96%) children in control group clinics and 343 (96%) children in intervention group clinics were assessed for the primary outcome. 16 (5%) children in the control group clinics and 15 (4%) children in the intervention group clinics did not provide all data and were not included in final analyses. An intention-to-treat analysis showed that the intervention significantly improved children's cognition (effect size 1·3 SDs, 95% CI 1·1 to 1·5; p=0·006), language (1·1 SDs, 0·9 to 1·2; p=0·01), and motor composite scores (1·2 SDs, 1·0 to 1·3; p=0·006) and behaviour ratings (ranging from 0·7 SDs, 0·5 to 0·9; p=0·02; to 1·1 SDs, 1·0 to 1·2; p=0·007), but the intervention had no significant effect on growth (p values ranged from 0·05 to 0·74). Three (1%) children in the intervention group died, but their deaths were not related to the intervention.

Interpretation The extent and range of benefits of our intervention are encouraging. Health workers ran most of the sessions effectively and attendance was good, which is promising for scale-up of the intervention model. However, researchers trained and supervised the health workers, and the next step will be to determine whether the Bangladeshi ministry of health can perform these tasks. In future programmes, more attention needs to be paid to the nutrition of the children.

Lancet Glob Health 2019; 7: e366–75

See **Comment** page e294

Maternal and Child Health Division (J D Hamadani PhD, S F Mehrin MSc, M I Hasan MPH) **and Nutrition and Clinical Services Division** (F Tofail PhD), **International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh; Toronto, ON, Canada** (S N Huda PhD)**; School of Psychology, Bangor University, Bangor, UK** (Prof H Baker-Henningham PhD)**; and UCL Great Ormond Street Institute of Child Health, London, UK** (D Ridout MSc, Prof S Grantham-McGregor FRCP)

Correspondence to:
Dr Jena D Hamadani, Maternal and Child Health Division, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka 1212, Bangladesh
jena@icddrb.org

Funding Grand Challenges Canada (Saving Brains).

Introduction

Brain development is particularly sensitive to the environment during the first 5 years of life, and risks such as poverty can affect brain structure and function, the stress response system, and gene expression,[1] which can have long-term consequences on adult functioning. It is well established that children living in poverty have poorer cognitive and language development than their more affluent peers.[2] In Bangladesh, children in the lowest wealth quintile have cognitive deficits by age 7 months, and these deficits increase by age 5 years, when the children are substantially behind in language and cognitive development compared with children in the highest wealth quintile.[3] Children with these deficits are likely to do poorly in school and to have concomitant disadvantages over their lifetime. Many small, well run randomised controlled trials to assess the efficacy of psychosocial interventions in early childhood have shown benefits of these interventions to children's development[4–6] and a few have shown long-term benefits.[7] The 2016 Sustainable Development Goals aim for all children to have access to "quality early childhood development, care and pre-primary education" and the 2017 Series on Advancing Early Childhood Development in *The Lancet* called for large-scale early child development programmes to be integrated into government health and nutrition services;[6] however, there are inadequate data on the effectiveness of large-scale integrated programmes, and there is a need for robust evaluations of such programmes.[8]

We have previously adapted the Jamaican home visiting intervention, Reach Up,[9] which aims to help mothers promote their child's development through demonstration and coaching in home visits. We have previously done three small randomised controlled trials,[10–12] in which fewer than 225 children received an intervention; these trials used play sessions that were run by specially hired facilitators, either at home or in local clinics. All interventions showed moderate effects (effect sizes of 0·33–0·38 of standard scores; p values <0·05) on the children's Mental Development Index of the Bayley Scales of Infant Development version 2. Another study[13] in Bangladesh used a mixture of government family welfare assistants and hired facilitators to run mothers' groups, with some home visits, and this intervention was also effective.

There are more than 13 000 Bangladeshi Government-run community clinics that are used for primary health care throughout the country. To develop an early childhood intervention model that targeted undernourished children and could be integrated into the clinics at a larger scale (ie, nationwide) we modified the Reach Up intervention to increase coverage and reduce costs. We trained all health workers in the intervention clinics and, therefore, a cluster-randomised controlled trial was required to avoid contamination, and

For the *Lancet* Series on **Advancing Early Childhood Development** see https://www.thelancet.com/series/ECD2016

Research in context

Evidence before this study

We searched for reviews of trials of early childhood psychosocial interventions that were published between Jan 1, 2007, and May 30, 2018. We searched the MEDLINE and PsycINFO databases with the search terms "early childhood development", "cognitive function", "language development", "psychosocial stimulation", and "integrated child development interventions" to identify randomised controlled trials of early childhood interventions that had been integrated into the health and nutrition services and that assessed child development outcomes and that were published in English. Studies reported in other languages , integrated into services other than those in health and nutrition, and non-randomised trials were excluded. We found five published studies reporting on four trials. The best integrated intervention that used child development outcomes was in Pakistan, but the intervention group in this study was confounded by geographical region.

Added value of this study

We found that it is feasible and highly effective for the Bangladeshi clinic workers to run an early childhood development programme for children at high risk of poor development. Further research is required to determine whether the health services can also train and supervise the health workers needed to deliver the intervention.

Implications of all the available evidence

Well evaluated trials of implementation strategies for early childhood development programmes are needed, to determine the most effective ways to deliver this programme. The ability to integrate early child development interventions into health services will likely depend on the workload, educational level, supervision, and motivation of the health workers and will vary by country. In Bangladesh, it was feasible for the clinic workers to run this intervention for children at high risk of poor development; however, alternative implementation strategies will be needed for children who live far from the clinics. In our study, implementation required careful pilot studies with several adaptations.

mothers with their underweight children attended in pairs every fortnight, instead of the individual, weekly home visits used in the original model. We aimed to evaluate the effects of this early childhood development programme on child cognition, language, and motor development, growth, and behaviour in a subsample of the children given this intervention.

Methods

Study design and participants

In this open-label, cluster-randomised controlled trial, we recruited children from community clinics that are used for primary health care in Narsingdi district, Bangladesh. The main services of these clinics are maternal and neonatal health care, nutrition and health education, and treatment of minor ailments. Community health-care providers manage community clinics, supported by a family welfare assistant and a health assistant. Community health-care providers and health assistants generally have a bachelor's or master's degree, and most family welfare assistants have completed higher secondary education. At the time of our study, the male to female ratio among community health-care providers was approximately 1:2, the ratio was 1:1 among the health assistants, and all the family welfare assistants were female. We got permission from the Bangladeshi ministry of health and family welfare to involve the clinic staff for our study by use of a signed memorandum of understanding. Narsingdi was selected for this study because it had a sufficient number of community clinics required for cluster randomisation and was within 80 km of Dhaka, therefore regular visits by the research staff based in Dhaka could be made easily. Three rural sub-districts in Narsingdi that had between 18 and 102 community clinics were selected, and 90 community clinics were randomly selected from these clinics by use of a computerised random selection by a colleague who was not involved with the study. A house-to-house survey of the area surrounding each community clinic was done, and all children living in homes that the surveyor and mother estimated to be within 30 min (walking time) of the community clinic were screened for inclusion. This distance was determined in an initial 5-month pilot study, in which we examined cooperation of the health staff and mothers' attendance at community clinics. On the basis of our findings, we only included children living within this walking distance because it was difficult for mothers to bring their children from further away.

All children aged 5–24 months who lived in the catchment area were weighed with standard methods. Underweight children, defined as those with weight-for-age Z scores that were less than –2 SDs of the WHO standards,[14] were eligible for inclusion. Because of our findings in the pilot study, we restricted the sample in each community clinic to a maximum of 25 children, since it was difficult for the health staff to manage more than 25 children in the clinic. In four intervention group clinics, there were more than 25 underweight children, so 25 children were randomly selected with computerised random selection, operated by MIH, who was not involved with the intervention. Children were excluded from the evaluation sample if they had severe acute malnutrition that resulted in complications requiring close monitoring or admission to hospital, severe clinical pallor, known chronic diseases (eg, epilepsy), or were twins (or a higher order of multiple births). However, excluded children and their mothers were able to join the intervention sessions at the clinics. Children with a developmental delay or other associated illnesses were referred to the nearest health centre for treatment.

At enrolment, parents were asked to sign the written informed consent form. The proposal was approved by the Institutional Review Board of the International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b).

Randomisation and masking

After we screened and enrolled the 90 community clinics in the study, they were stratified by subdistrict. Clinics were then randomly assigned (1:1) to the intervention or control clusters by a researcher who was not part of the study, by use of a computer-generated code. Participants in the intervention group cluster were given an intervention of 25 fortnightly sessions, in which mothers were shown how to support their child's development through play and interactions. Participants in the control group clinics had no extra attention, but they used the clinic for treatment as usual if they were ill; some children were also immunised at the control group clinics. However, we did not record the attendance of children to the control group clinics.

All participants were enrolled before randomisation, but measurements were done after random allocation of clinics to groups, both at baseline and at the end of intervention, by interviewers and testers who were masked to study group. This masking was done by separating testers from the researchers involved in the intervention and concealing the intervention groups of each clinic. The measurements were taken by the testers before the start and after the end of the intervention, and so the intervention session materials were not available in the clinics at the time of the tests. However, as in all interventions of this type, a few mothers might have mentioned the intervention during testing.

All children who were enrolled in intervention were used to assess whether providing such an intervention at community clinics was feasible. However, we also took a randomly selected subsample of eight mother–child pairs from each of the intervention and control community clinics, to evaluate the effects of this intervention on the children (evaluation subsample). Selection was done using computerised random selection by MIH who was not involved with intervention.

Procedures

The intervention was phased in by district. Two mothers at a time brought their children to a 40-min to 60-min play session at the clinic every fortnight for 1 year. The community health-care provider, who managed the community clinics 6 days per week, ran an average of three play sessions per week, whereas the family welfare assistant and health assistant, who support the clinic 2 days per week, ran an average of one weekly session each. We subsequently refer to these three cadres as health workers, to facilitate the description of the intervention.

The research team trained 130 health workers in how to conduct the intervention sessions in groups of 12–15 people for 10 days each. Supervisors with a master's degree in psychology or another related subject were recruited and trained to mentor the health workers for 20 days. Each supervisor monitored four community clinics, and there were between one and three health workers per community clinic. The supervisor met with each health worker twice a month and observed a play session with a checklist of desired activities for guidance. If a health worker's performance was poor, they were visited more frequently. The supervisors provided feedback to the health workers at the end of the session. We also ran a half-day refresher workshop for the health workers every 3 months.

Before we started the intervention, meetings were organised in each community clinic with parents, family members, and community leaders. Information on early child development was presented and details of the project discussed to motivate the community. In an initial survey, information about demographic, socioeconomic, and quality of home stimulation using Family Care Indicators was collected by the field workers from all enrolled families. We used the Family Care Indicators to measure the quality of home stimulation. The Family Care Indicators measure was developed by UNICEF,[15] and it has been validated in Bangladesh.[16]

See Online for appendix

We based the intervention sessions on those of the Jamaican Reach Up Programme. There was a manual with the curriculum, in which sessions were arranged in developmental order for children from ages 6 months to 42 months. Children were initially placed at their chronological age level on the curriculum and progressed from one session to the next, more difficult one. If necessary, the child's position on the curriculum was changed to match the child's rate of progress. Play sessions with mother and child were participatory and followed a set format that comprised a review of the previous home activities with discussion, a local song, activities with a picture book, a developmentally appropriate toy, language activities, nutritional messages, and a review of activities to be continued at home. Activities and materials were specified for every session, and they were planned to match the child's appropriate developmental level such that they were challenging but that the child could complete them successfully. Some activities were done with both mothers and children, whereas the toy activities were done with each mother–child dyad alone to ensure that developmentally appropriate activities were used.

For the **Reach Up Programme** see http://www.reachupandlearn.com

The health workers demonstrated activities and interactions with the children to the mothers, then they encouraged the mothers to repeat and expand the activities and to suggest new activities. Mothers were asked to continue the activities at home and to use everyday activities and materials for further learning opportunities. Mothers were lent the book and toy to take home, and these materials were exchanged for new ones at each session. We placed an emphasis on the health workers developing good relations with the mothers. Mothers were encouraged to respond to their child's interests and vocalisations and to give them praise and positive feedback. We aimed to make the sessions enjoyable. Further details on the unadjusted Jamaican Reach Up Programme have been published online and in previous work.[9,17,18]

We adjusted the Reach Up Programme for Bangladesh. These adjustments included redrawing all books and pictures to reflect the local context (eg, mothers were drawn in sarees—the traditional Bangladeshi clothing for women—and houses were drawn as mud houses seen in villages), printing them locally, and replacing songs and games with local, traditional versions. Some toys were made from waste material and, where the available waste was different, we occasionally modified the toys. The methods and underlying concepts of the Jamaican curriculum were unchanged; further details of this curriculum are shown in the appendix.

In the randomly selected evaluation subsamples, we assessed several measures at baseline and after 1 year of the intervention, in the intervention and control groups. First, we evaluated the children's composite scores on the Bayley Scales of Infant and Toddler Development, third edition.[19] These scores are standardised on a US population to have a mean of 100 and SD of 15, and this scale includes language (expressive and receptive language scales, combined), motor (fine and gross motor scales, combined), and cognitive scores. The Bayley Scale assessment was translated and adapted for Bangladesh by redrawing culturally inappropriate pictures (for example, pictures of a washing machine and a vacuum cleaner were changed to washing clothes by hand and sweeping the room with a broom, but we did not change underlying concepts or the order of the items).

Second, the children's behaviour was rated during the test on five Wolke scales: approach to the tester in the first 10 min, emotional tone, cooperation with the tester, vocalisation, and activity level throughout the test. These ratings have been used in several Bangladeshi studies.[10,20,21] Children were tested at baseline and after 1 year at the community clinics in the presence of their mothers by one of eight testers. All testers had a master's

degree in psychology or social sciences. Before the study began, each tester attained satisfactory interobserver reliabilities with the trainer (r=1·0 on composite scores, range 0·60–1·0 on scaled scores; on eight to 16 tests per tester). Finally, the children's weight, length or height, and head circumference were measured by the testers after completion of the Bayley test and behaviour ratings by use of standard methods.[22] The anthropometric measures were converted to Z scores by use of WHO anthroplus.[14]

After the test, we also assessed maternal knowledge of child rearing with a specially designed instrument that was used in previous studies.[9] Six questions on maternal depression were also asked as part of the Family Care Indicators assessment, which were taken from Center for Epidemiological Studies-Depression scale.[23] The shortened versions of this scale have previously been found to discriminate between depressed and non-depressed adults.[24]

Outcomes

The primary outcomes were assessed on the evaluation subsample of children in the control and intervention group clinics, at baseline and after 1 year of the intervention. These primary outcomes were the composite language, motor, and cognitive scores on the Bayley Scales of Infant and Toddler Development, behaviour (rated on five Wolke scales), and the anthropometric measurements of the children (weight, length or height, and head circumference). The secondary outcomes, which were also assessed in this evaluation subsample, were maternal knowledge of child rearing, quality of home stimulation (by use of the FCI), and symptoms of maternal depression (by use of the shortened version of the Center for Epidemiological Studies-Depression scale).

In a post-hoc analysis, we also measured the feasibility of the approach in the whole sample by the proportion of health staff who attended the training and ran sessions and the number of sessions they ran and the proportion of mothers with eligible children who agreed to enroll and their attendance at the sessions.

Statistical analysis

We calculated that we needed to recruit eight children from each of 90 clinics to detect a 0·25 SD difference in Bayley scores between the groups at 80% power and 5% significant levels with an intracluster correlation of 0·01.

All data were checked for normality, and sociodemographic characteristics were compared between the control and intervention groups with a two-sample *t* test. The baseline outcome variables were compared between the groups with a multi-level random effects model, with child nested within community clinics, which were nested within subdistricts. The tester was included in the model when appropriate. We computed a wealth factor from the sum of the assets that the family owned and a housing factor from the quality of materials used for the wall, roof, and floor of the house, presence or absence of electricity at home, and a sanitary toilet from baseline information for all participants.

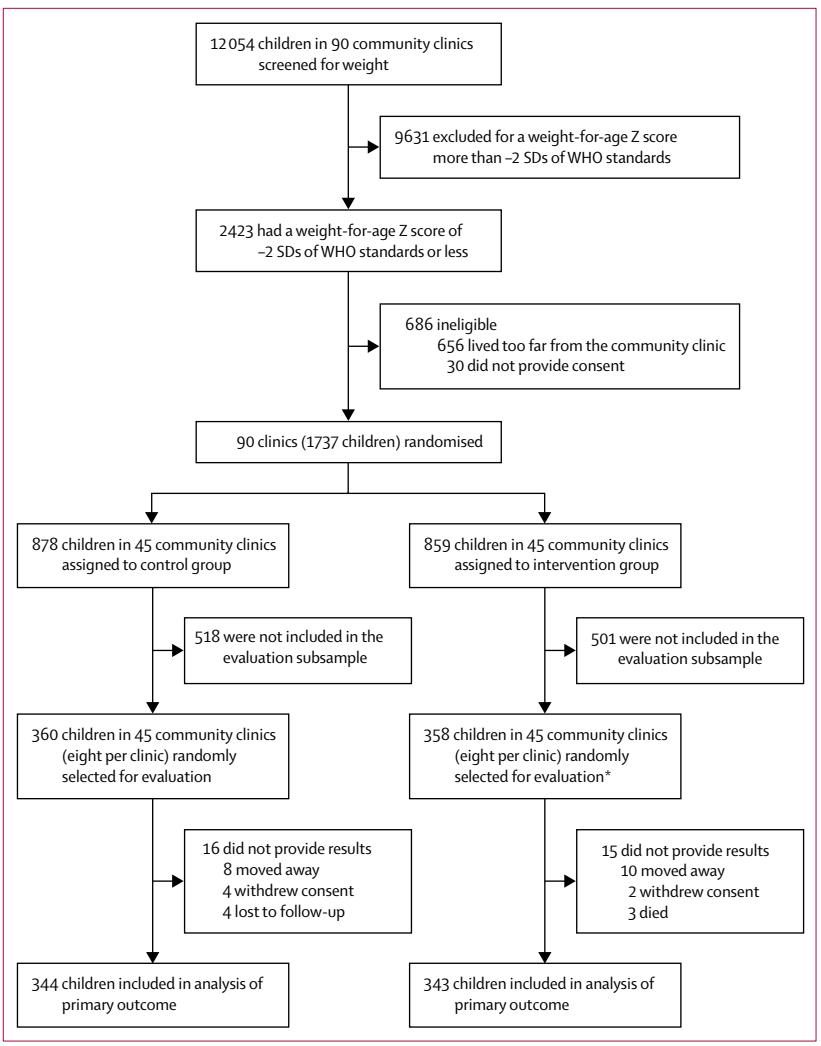

***Figure*: Trial profile**
*One community clinic only had six eligible children.

For each of the outcome measures at the final study assessment, we fitted a similar multi-level random effects model, with child nested within community clinic, nested within subdistrict. Group was included as a binary factor, and we adjusted all models for age at follow-up, sex, and relevant baseline measures. We also adjusted for the testers in analyses that involved the Bayley scores and Wolke behaviour ratings. We investigated interactions for group with parental education, sex, and socioeconomic status (eg, wealth and housing factors).

Multiple imputation (assuming data was missing at random) was used to account for missing data with regard to maternal knowledge of child rearing and maternal depression. The imputation models included the socio-demographic variables (eg, parental education, assets

	Control group (n=344)	Intervention group (n=350)	p value
Age at baseline, months	15·8 (5·2)	16·1 (4·9)	0·538
Sex of the child	..	..	0·514
Male	188 (55%)	188 (54%)	..
Female	156 (45%)	162 (46%)	..
Duration of father's education, years	..	..	0·999
0–4	133 (39%)	136 (39%)	..
5	77 (22%)	73 (21%)	..
>5	130 (38%)	133 (38%)	..
Duration of mother's education, years	..	..	0·313
0–4	102 (30%)	89 (25%)	..
5	72 (21%)	77 (22%)	..
>5	166 (48%)	176 (50%)	..
Wealth factor, quintile	..	..	0·160
First	66 (19%)	63 (18%)	..
Second	85 (25%)	76 (22%)	..
Third	53 (15%)	49 (14%)	..
Fourth	77 (22%)	76 (22%)	..
Fifth	58 (17%)	77 (22%)	..
Housing factor	−0·012 (1·1)	−0·0004 (1·0)	0·882

Data are mean (SD) or n (%). The housing factor score evaluated the quality of materials used for the wall, roof, and floor of the house and the presence or absence of electricity and a sanitary toilet.

Table 1: **Baseline sociodemographic characteristics**

owned, and number of siblings) and all baseline and end-of-study outcome measures (which we assume includes all predictors of missingness). We generated 20 datasets and ran a full multi-level random effects model, for which we used the whole dataset and implemented a bootstrap (200 samples) for each imputed dataset to correct for overfitting. The final models were derived by fitting a multi-level model with all aforementioned factors, and the estimates were combined by use of Rubin's rules.[25] Since the outcomes are recorded on different scales, we present the results for the estimated intervention effect on a standardised scale for the same models. We transformed the data by use of an internal standardisation on the whole sample and at baseline and at the study final assessment separately. This approach allowed a comparison of the relative effect of the intervention on each of the outcome measures. We used intention-to-treat analyses for all models and, to account for multiplicity comparisons at the study final analysis, p values were corrected for all ten outcomes by use of the Holm's stepdown procedure.[26] For evaluation of all the outcomes, these intention-to-treat analyses assessed all children and mothers who were randomly assigned to groups and randomly selected for the evaluation subsample.

All analyses were performed in Stata version 14 and for all significance test results presented, a p value of less than 0·05 was considered statistically significant.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. However, the funder provided an expert group that formed a platform team who helped the research group during the study with trouble-shooting, if required. For example, we consulted with this team after the initial pilot with regard to changing the number of children per clinic and limiting the eligible children to those living close to the clinics. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Between Nov 29, 2014, and April 30, 2015, children in 90 clinics who were aged 5–24 months (which included 12 054 children) were weighed (figure). We identified 2423 (20%) children as underweight, defined as a weight-for-age Z score of less than or equal to −2 SDs of the WHO standard; 9631 (80%) children were excluded because their weight-for-age Z score was more than −2 SDs of WHO standards. From the 2423 underweight children, we excluded 656 (27%) children who lived more than 30-min walking distance from the community clinics, and 30 (1%) children whose mothers did not consent to participate. We therefore enrolled 1737 (72%) children from these 90 clinics, after which the clinics were randomly assigned to deliver no intervention (control group; care as usual) or 25 sessions of our early childhood development programme (intervention group). After randomisation, the control group clinics included 878 (51%) children (who all received no intervention) and the intervention group clinics included 859 (49%) children (who all received the development programme sessions).

The evaluation subsample included 360 (41%) children in the control group and 358 (42%) children (including six children from one clinic, instead of the expected eight) in the intervention group at baseline, of which 344 (96%) children in the control group and 343 (96%) children in the intervention group remained in the study at the final analysis of the primary outcomes. There were no substantial differences in dropouts between the evaluation subsamples of the groups; 16 (5%) children in the control group clinics and 15 (4%) children in the intervention group clinics had missing data.

The children in evaluation subsamples of the control group clinics and intervention group clinics did not significantly differ at baseline with respect to wealth, housing quality, children's nutritional status, quality of home stimulation, parental education, or maternal knowledge of child rearing (table 1). However, the children in the intervention group clinics had significantly higher cognitive scores and their mothers had fewer depressive symptoms at baseline (table 2). As a

result, all analyses of child development outcomes were adjusted for the relevant baseline scores. Maternal depression was not significantly associated with any of the outcomes, and so the analyses were not adjusted for this factor.

Data completeness was good (>92%) for most outcomes, with the exception of maternal depression (82%) and knowledge of child rearing (82%). The results that used the imputed data were similar to those that used the observed data, so we present the estimated, adjusted intervention effect from the observed data for all outcomes.

The stability of the Bayley Scale scores from baseline to the final analysis in the children in the control group clinics was as expected (cognition score, r=0·24; language score, r=0·37; motor score, r=0·39; p<0·0001 for all),[27] and scores were significantly correlated with height-for-age (r=0·229; p<0·0001), wealth factor (r=0·111; p=0·004), and years of maternal education (r=0·109, p=0·004). Interobserver reliabilities for all testers with the trainer were assessed before the study and intraclass correlations were satisfactory (response to examiner, r=0·91–1·0; emotional tone, r=0·64–1·0; cooperativeness, r=0·81–0·99; vocalisation, r=0·96–1·0; and activity, r=0·67–0·97). However, one tester only attained r=0·51 for activity and was given additional training. Approximately 10% of Bayley tests were observed during the study, and interobserver reliabilities for each tester for all Bayley composite scores were more than r=0·8; for all behaviour ratings, the range of correlations was r=0·67–0·99 in 71 tests.

We found that the intervention significantly improved the Bayley cognitive (effect size 1·3 SDs, 95% CI 1·1 to 1·5; p=0·006), language (1·1 SDs, 0·9 to 1·2; p=0·01), and motor (1·2 SDs, 1·0 to 1·3; p=0·006) composite scores (table 3). We also found that the children in the intervention group clinics were more responsive to the examiner (1·1 SDs, 1·0 to 1·2; p=0·007), had a happier emotional tone (0·9 SDs, 0·8 to 1·1; p=0·01), more cooperative (1·0 SDs, 0·9 to 1·1; 0·008), and talked or vocalised more during the test (0·7 SDs, 0·5 to 0·9; p=0·02). However, we found no effects of the treatment on children's anthropometric measurements (p values ranged from 0·05 to 0·74); both groups showed improvements in their weight-for-height Z scores, with a mean increase of 0·27 (95% CI 0·19 to 0·35) but decreases in their height-for-age Z scores, with a mean decrease of –0·20 (–0·27 to –0·13) during the study (p<0·0001 for both).

We also found that mothers in the intervention group clinics showed greater improvements in their child-rearing knowledge (effect size 1·7 SDs, 95% CI 1·5 to 1·8; p=0·005) and children in the intervention group clinics had an improved quality of home stimulation (0·8 SDs, 0·6 to 1·0; p=0·03; table 3). In the initial analysis, we noted a significant reduction in maternal depressive symptoms, but this difference was

	Control group (n=360)	Intervention group (n=358)	Difference between intervention and control means (95% CI)	p value
Age, months	15·8 (5·2)	16·1 (4·9)	0·2 (–0·5 to 1·0)	0·52
Cognitive composite scores	87·4 (11·9)	90·0 (12·2)	2·4 (0·7 to 4·2)	0·01
Language composite scores	80·3 (10.8)	82·3 (11·2)	1·4 (–0·1 to 3·0)	0·08
Motor composite scores	86·2 (12·4)	88·0 (11·8)	1·6 (–0·3 to 3·4)	0·09
Response to examiner	5·3 (0·9)	5·3 (0·9)	0·04 (–0·1 to 0·2)	0·68
Emotional tone	4·9 (0·9)	5·1 (0·9)	0·1 (–0·02 to 0·3)	0·09
Activity level	4·8 (0·9)	4·9 (0·9)	0·1 (–0·03 to 0·3)	0·12
Cooperativeness	4·9 (1·0)	5·0 (1·0)	0·1 (–0·06 to 0·3)	0·18
Vocalisation	3·4 (1·6)	3·5 (1·5)	0·1 (–0·2 to 0·4)	0·47
Knowledge of child rearing*	21·9 (5·6)	22·5 (5·4)	0·4 (–1·0 to 1·7)	0·59
Maternal depressive scores*	11·9 (8·7)	10·1 (9·0)	–1·8 (–3·3 to –0·2)	0·03
Home stimulation	6·7 (4·2)	6·9 (4·0)	0·1 (–0·7 to 1·1)	0·66
Height-for-age Z score†	–2·6 (1·0)	–2·5 (1·2)	0·03 (–0·2 to 0·2)	0·74
Weight-for-age Z score	–2·4 (0·7)	–2·4 (0·7)	0·05 (–0·06 to 0·2)	0·38
Weight-for-height Z score†	–1·5 (1·0)	–1·5 (1·1)	0·03 (–0·1 to 0·2)	0·68
Occipito-frontal circumference	43·2 (1·8)	43·4 (1·7)	0·1 (–0·01 to 0·5)	0·05

Data are mean (SD), unless otherwise indicated. Data are from participants in 45 community clinics assigned to the control group and 45 community clinics assigned to the intervention group. p values have been adjusted for clustering effects at the clinic and subdistrict levels and used *t* tests (except for maternal depressive scores, which used a Mann-Whitney U test). Cognitive, language, and motor composite scores were assessed with the Bayley Scales of Infant and Toddler Development, third edition,[19] and they have been adjusted for tester effects. Response to the examiner, emotional tone, activity level throughout the test, cooperativeness with the tester, and vocalisation were assessed with Wolke scales. Mother's knowledge of child rearing was assessed as in a previous study.[9] Maternal depression was assessed with a shortened Center for Epidemiological Studies-Depression scale.[23,24] Home stimulation was measured with the Family Care Indicators[15,16] and was assessed in 340 children in the control group and 342 children in the intervention group. *n=334 for control group and n=269 for intervention. †n=341 for intervention group at baseline.

Table 2: **Participant characteristics on outcome measures at baseline**

	Intervention effect B	Effect size, SDs	Corrected p value
Cognitive composite score	13·5 (11·8 to 15·2)	1·3 (1·1 to 1·5)	0·006
Language composite score	9·4 (7·9 to 10·9)	1·1 (0·9 to 1·2)	0·01
Motor composite score	12·6 (11·1 to 14·2)	1·2 (1·0 to 1·3)	0·006
Response to examiner	1·0 (0·8 to 1·1)	1·1 (1·0 to 1·2)	0·007
Emotional tone	0·8 (0·7 to 0·9)	0·9 (0·8 to 1·1)	0·01
Cooperativeness	1·0 (0·8 to 1·1)	1·0 (0·9 to 1·1)	0·008
Vocalisation	1·1 (0·8 to 1·4)	0·7 (0·5 to 0·9)	0·02
Knowledge of child rearing* (n=331, control group; n=270, intervention group)	15·9 (14·6 to 17·1)	1·7 (1·5 to 1·8)	0·005
Maternal depression* (n=330, control group; n=270, intervention group)	–2·9 (–4·9 to –1·0)	–0·3 (–0·6 to –0·1)	0·05
Home stimulation* (n=340, control group; n=342, intervention group)	4·1 (2·9 to 5·3)	0·8 (0·6 to 1·0)	0·03

Intervention effect B is the regression coefficient. Ranges shown are 95% CIs. Data were assessed in children in the control group versus 343 children in the intervention group with a multi-level random effects model, in which findings in children were nested within clusters of community centres, and these clusters were nested within subdistricts; these results were adjusted for age at follow-up, sex, Bayley testers, and relevant baseline measures. p values have been corrected for all ten outcomes with Holm's stepdown procedure. Cognitive, language, and motor composite scores were assessed with the Bayley Scales of Infant and Toddler Development, third edition.[15] Response to the examiner, emotional tone, cooperativeness with the tester, and vocalisation were assessed with Wolke scales. Mother's knowledge of child rearing was assessed as in a previous study.[9] Maternal depression was assessed with a shortened version of the Center for Epidemiological Studies-Depression scale.[22,23] Home stimulation was measured with the Family Care Indicators.[20,21] *Bayley testers not adjusted for in analysis.

Table 3: **Multiple regression analysis of the effects of intervention on outcome measures**

not significant after adjustment for multiple outcomes (–0·3 SDs, –0·6 to –0·1; p=0·05).

In our post-hoc analysis, we found that health workers and mothers appeared to accept the programme and compliance was good. 859 children and their mothers (the whole intervention group) attended a mean of 19 (SD 9) sessions, and 394 (46%) mothers and their children attended all sessions. The children in the evaluation subsample of the intervention group clinics attended a mean of 21 (7·4) sessions, and 192 (54%; 95% CI 49 to 60) mothers and their children in this subsample attended all 25 sessions. The evaluated subsample therefore attended significantly more sessions than those in the whole group (21 *vs* 19 sessions; 95% CI 2·2 to 4·6; p<0·0012). Compliance by staff was also good: 100% attended the training and ten (8%) of 123 health workers did not run all the sessions. Among these ten health workers, eight (80%) missed from one to nine sessions and two (20%) health workers missed 13–24 sessions. The reasons for missing the sessions were illness (five health workers), maternity leave (two health workers), and annual leave (two health workers), and only one health worker refused to run the sessions, although he had initially delivered between one and seven sessions for some of the children.

The primary outcome measures had intraclass correlation of 0·23 (95% CI 0·16 to 0·32) for cognition, 0·17 (0·11 to 0·26) for language, and 0·11 (0·06 to 0·19) for motor composite scores at cluster level. There was no differential effect of treatment by sex, maternal education, or wealth.

Discussion

We integrated an early childhood development intervention into primary health-care clinics, and we found that this intervention significantly improved the cognitive, language, and motor development and behaviour of undernourished children who were at high risk of poor development. Home stimulation and maternal knowledge of child-rearing also improved with our intervention. The extent and range of benefits were extremely encouraging and were greater and more extensive than found in three previous Bangladeshi studies[10–12] that used a similar curriculum in individual home or clinic visits. We hypothesise that modifications to implementation of the programme that we made in this study accounted for the greater effects. The mothers appeared to be more comfortable in pairs and they interacted more with their children and with each other than in individual sessions, in which they often appeared inhibited. The health workers were generally better educated than other local women who have previously been employed as home visitors, and it is possible that the health workers in our study were more respected by the mothers than the local women in previous studies.

Our post-hoc analysis of the feasibility measures suggests that it is feasible to integrate the intervention into the clinic services. Very few mothers refused to enrol, and those who enrolled showed reasonable attendance. The children in the evaluation subsample attended significantly more sessions, which could be because the parents became more motivated due to the initial baseline measurements. The health workers and supervisors did not know who the evaluation subsample participants were, so the children in the evaluation sample were not given priority. All health workers attended training and ran most of the sessions without additional incentives. However, we only delivered the intervention to approximately 25 children per clinic, giving a maximum of six sessions a week per clinic, which seemed feasible based on the workload of the health workers and the results of our pilot study. We did not want to overburden the staff with more sessions and, unfortunately, we did not collect data to see whether our intervention interfered with their other duties. In future, it would be possible to increase the size of the groups to four mother–children pairs, which is the maximum possible due to restricted clinic space. We ran three monthly workshops for the staff, which accrued a cost that included transport and refreshments, but these workshops appeared necessary to keep staff motivated and solve problems. Very few studies have integrated child development interventions into health and nutrition services with robust evaluations, and our results compare favourably with studies in Pakistan[28] and Jamaica.[29]

The strengths of the study include the robust design that used cluster randomisation, assessment of child development with full development scales, and use of health staff to deliver the play sessions, which would facilitate expansion to larger scales.

A limitation of the study is that only 72% of targeted children served by the community clinics participated, so an alternative strategy is required to reach children who live further away. The family welfare assistants are responsible for reproductive health and family planning, and health assistants are responsible for immunisation. These health workers are supposed to make very short home visits to approximately 600–2000 households every 2 months, and it is possible that they could organise group meetings on early childhood development in the community, but this approach must first be trialled.

A notable disappointing result was the absence of improvement in the children's anthropometric measurements. We largely depended on other services to address nutrition, although we included some nutritional education in the play sessions and we provided mothers with nutrition cards with a recipe for diets for underweight children. Many clinics did not have weighing scales, and the government has begun a programme to improve nutritional care in these clinics.

Perhaps the most important limitation when considering expanding our programme on a larger scale is that the research group was highly motivated and experienced with the intervention. The research group

trained and supervised the health workers and, for future scale-up, these activities would have to be performed by the ministry of health. However, the ministry of health employs a supervisor for every three clinics and many of these supervisors have shown an interest in the programme. The next step in scaling up the intervention would be for those supervisors to take over the training and supervision of the health workers. We are presently piloting this approach and it appears to be feasible. Alongside the government, we are also exploring how to transfer the overall responsibility for the programme from researchers to the district and central health officers.

Another limitation is that the Bayley Scales were adapted but not standardised for Bangladesh. In previous studies,[3,10–12] the Bayley (second edition) scores at 18 months had good concurrent and predictive validity of intelligence quotient at age 5 years. In this study, the Bayley (third edition) scores had adequate test-retest reliability, stability over the year, and discriminant validity, and scores were significantly correlated with height-for-age, the wealth factor, and years of maternal education. We therefore think that these scales are likely to be valid.

A final limitation is that the children were assessed at the clinics, where the intervention took place, and it is possible that the children and mothers who received the intervention benefited from the familiar venue. However, in a companion randomised trial (unpublished), half the children in both groups were tested in clinics familiar to the intervention group only and half were tested in unfamiliar locations, and we have found almost identical intervention effects in both formats, which suggests that the test location is unlikely to have had substantial effects.

We were only able to work in 90 community clinics of more than 13 000 community clinics in the country. However, these 90 clinics are representative of those in other areas. Most rural areas in Bangladesh are poor and have limited resources, and therefore we might be able to generalise the data across the country.

In conclusion, we found that the intervention model was feasible and very effective, which justifies further scaling up of this programme.

Contributors

JDH and SG-M developed the original concept, drafted the project proposal for this study, developed the study design, analysed data, and drafted the initial manuscript. SFM, FT, MIH, and SNH contributed to the concept, implemented the field activities, analysed data, and reviewed the manuscript. HB-H co-wrote the project proposal, contributed to the concept and data analysis, and reviewed the manuscript. DR contributed to the study design, did the primary outcome analysis, and reviewed the manuscript. All authors have read and approved the manuscript and take full responsibility for all aspects of the work.

Declaration of interests

All authors declare no competing interests.

Data sharing

In accordance with the icddr,b policy, the principal investigator and those authorised by the principal investigator shall have an exclusive right to analyse and publish such data from the onset of the data collection and until 3 years after the protocol or study completion date. Thereafter, the data will be made accessible.

Acknowledgments

Our study was funded by Saving Brains programme of Grand Challenges Canada.

References

1 Johnson SB, Riis JL, Noble KG. State of the art review: poverty and the developing brain. *Pediatrics* 2016; **137:** e20153075.
2 Black MM, Walker SP, Fernald LCH, et al. Early childhood development coming of age: science through the life course. *Lancet* 2017; **389:** 77–90.
3 Hamadani JD, Tofail F, Huda SN, et al. Cognitive deficit and poverty in the first 5 years of childhood in Bangladesh. *Pediatrics* 2014; **134:** e1001–08.
4 Aboud FE, Yousafzai AK. Global health and development in early childhood. *Annu Rev Psychol* 2015; **66:** 433–57.
5 Rao N, Sun J, Wong JMS, et al. Early childhood development and cognitive development in developing countries: a rigorous literature review. September, 2014. http://eppi.ioe.ac.uk/cms/Portals/0/PDF%20reviews%20and%20summaries/ECD%202014%20Rao%20report.pdf?ver=2014-10-02-145634-017 (accessed Jan 28, 2019).
6 Britto PR, Lye SJ, Proulx K, et al. Nurturing care: promoting early childhood development. *Lancet* 2017; **389:** 91–102.
7 Tanner JC, Candland T, Odden WS, for the World Bank. Later impacts of early childhood interventions: a systematic review. 2015. https://ieg.worldbankgroup.org/sites/default/files/Data/Evaluation/files/ImpactsofInterventions_EarlyChildhoodonLaterOutcomes.pdf (accessed Jan 28, 2019).
8 Grantham-McGregor SM, Fernald LC, Kagawa RM, Walker S. Effects of integrated child development and nutrition interventions on child development and nutritional status. *Ann NY Acad Sci* 2014; **1308:** 11–32.
9 Grantham-McGregor S, Smith A. Extending the Jamaican early childhood development intervention. *J Appl Res Child* 2016; **7:** 4.
10 Hamadani JD, Huda SN, Khatun F, Grantham-McGregor SM. Psychosocial stimulation improves the development of undernourished children in rural Bangladesh. *J Nutr* 2006; **136:** 2645–52.
11 Nahar B, Hossain MI, Hamadani JD, et al. Effects of a community-based approach of food and psychosocial stimulation on growth and development of severely malnourished children in Bangladesh: a randomised trial. *Eur J Clin Nutr* 2012; **66:** 701–09.
12 Tofail F, Hamadani JD, Mehrin F, Ridout DA, Huda SN, Grantham-McGregor SM. Psychosocial stimulation benefits development in nonanemic children but not in anemic, iron-deficient children. *J Nutr* 2013; **143:** 885–93.
13 Aboud FE, Singla DR, Nahil MI, Borisova I. Effectiveness of a parenting program in Bangladesh to address early childhood health, growth and development. *Soc Sci Med* 2013; **97:** 250–58.
14 WHO. WHO child growth standards: methods and development. Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age. 2006. https://www.who.int/childgrowth/standards/technical_report/en/ (accessed Jan 28, 2019).
15 Kariger P, Frongillo EA, Engle P, et al. Indicators of family care for development for use in multicountry surveys. *J Health Popul Nutr* 2012; **30:** 472–86.
16 Hamadani JD, Tofail F, Hilaly A, Huda SN, Engle P, Grantham-McGregor SM. Use of family care indicators and their relationship with child development in Bangladesh. *J Health Popul Nutr* 2010; **28:** 23–33.
17 Smith JA, Baker-Henningham H, Brentani A, Mugweni R, Walker SP. Implementation of Reach Up early childhood parenting program: acceptability, appropriateness, and feasibility in Brazil and Zimbabwe. *Ann NY Acad Sci* 2018; **1419:** 120–40.
18 Walker SP, Chang SM, Smith JA, Baker-Henningham H. The Reach Up early childhood parenting program: origins, content, and implementation. *Zero Three* 2018; **38:** 37–43.
19 Bayley N. Bayley scales of infant and toddler development, third edn. Cambridge, UK: Pearson Publishing, 2005.
20 Hamadani JD, Fuchs GJ, Osendarp SJ, Khatun F, Huda SN, Grantham-McGregor SM. Randomized controlled trial of the effect of zinc supplementation on the mental development of Bangladeshi infants. *Am J Clin Nutr* 2001; **74:** 381–86.

21 Hamadani JD, Fuchs GJ, Osendarp SJ, Huda SN, Grantham-McGregor SM. Zinc supplementation during pregnancy and effects on mental development and behaviour of infants: a follow-up study. *Lancet* 2002; **360:** 290–94.

22 WHO. Measuring change in nutritional status: guidelines for assessing the nutritional impact of supplementary feeding programmes for vulnerable groups. 1983. http://apps.who.int/iris/handle/10665/38768 (accessed Jan 28, 2019).

23 Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977; **1:** 385–401.

24 Wolf AW, De Andraca I, Lozoff B. Maternal depression in three Latin American samples. *Soc Psychiatry Psychiatr Epidemiol* 2002; **37:** 169–76.

25 Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley and Sons, 2004.

26 Holm S. A simple sequentially rejective multiple test procedure. *Scand Stat Theory Appl* 1979; **6:** 65–70.

27 Pollitt E, Triana N. Stability, predictive validity, and sensitivity of mental and motor development scales and pre-school cognitive tests among low-income children in developing countries. *Food Nutr Bull* 1999; **20:** 45–52.

28 Yousafzai AK, Rasheed MA, Rizvi A, Armstrong R, Bhutta ZA. Effect of integrated responsive stimulation and nutrition interventions in the Lady Health Worker programme in Pakistan on child development, growth, and health outcomes: a cluster-randomised factorial effectiveness trial. *Lancet* 2014; **384:** 1282–93.

29 Chang SM, Grantham-McGregor SM, Powell CA, et al. Integrating a parenting intervention with routine primary health care: a cluster randomized trial. *Pediatrics* 2015; **136:** 272–80.