



OPEN ACCESS

# Counting outcomes, coverage and quality for early child development programmes

Kate M Milner,<sup>1,2</sup> Sunil Bhopal,<sup>3,4</sup> Maureen Black,<sup>5</sup> Tarun Dua,<sup>6</sup> Melissa Gladstone,<sup>7</sup> Jena Hamadani,<sup>8</sup> Rob Hughes,<sup>3,9</sup> Maya Kohli-Lynch,<sup>1,10</sup> Karim Manji,<sup>11</sup> Victoria Ponce Hardy,<sup>1</sup> James Radner,<sup>12,13</sup> Sonia Sharma,<sup>14</sup> Fahmida Tofail,<sup>15</sup> Cally Tann,<sup>1,16</sup> Joy E Lawn<sup>1</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/archdischild-2018-315430>).

For numbered affiliations see end of article.

## Correspondence to

Dr Kate M Milner, MARCH Centre, London School of Hygiene & Tropical Medicine, Keppel St, London WC1E 7HT, UK; [kate.milner@lshtm.ac.uk](mailto:kate.milner@lshtm.ac.uk)

Received 19 September 2018

Revised 24 January 2019

Accepted 28 January 2019

## ABSTRACT

Improved measurement in early child development (ECD) is a strategic focus of the WHO, UNICEF and World Bank Nurturing Care Framework. However, evidence-based approaches to monitoring and evaluation (M&E) of ECD projects in low-income and middle-income countries (LMIC) are lacking. The Grand Challenges Canada®-funded Saving Brains® ECD portfolio provides a unique opportunity to explore approaches to M&E of ECD programmes across diverse settings. Focused literature review and participatory mixed-method evaluation of the Saving Brains portfolio was undertaken using an adapted impact framework. Findings related to measurement of quality, coverage and outcomes for scaling ECD were considered. Thirty-nine ECD projects implemented in 23 LMIC were evaluated. Projects used a 'theory of change' based M&E approach to measure a range of inputs, outputs and outcomes. Over 29 projects measured cognitive, language, motor and socioemotional outcomes. 18 projects used developmental screening tools to measure outcomes, with a trade-off between feasibility and preferred practice. Environmental inputs such as the home environment were measured in 15 projects. Qualitative data reflected the importance of measurement of project quality and coverage, despite challenges measuring these constructs across contexts. Improved measurement of intervention quality and measurement of coverage, which requires definition of the numerator (ie, intervention) and denominator (ie, population in need/at risk), are needed for scaling ECD programmes. Innovation in outcome measurement, including intermediary outcome measures that are feasible and practical to measure in routine services, is also required, with disaggregation to better target interventions to those most in need and ensure that no child is left behind.

## BACKGROUND

Improved measurement is crucial to accelerating and tracking progress towards the Sustainable Development Goals related to early child development (ECD), and the WHO, UNICEF and World Bank Nurturing Care Framework (NCF) emphasises strengthened monitoring as a crucial component of implementation for ECD programmes at national and subnational levels.<sup>1,2</sup>

However, most literature on child development metrics is focused on challenges in assessing

## Key findings

1. **WHY?** *Tracking programme impact at scale requires effective 'real world' approaches:* While a multitude of child development assessment and screening tools are available, in practice, lack of standardised, meaningful measures of ECD outcomes which can be used at scale, impede measurement of intervention impact.
2. **WHAT'S NEW?** *Measurement of coverage in ECD programmes has had relatively little attention:* This requires agreed definition of ECD interventions and target population with available data, either for all children in a given age group or to target based on child development status and risk factors.
3. **WHAT TO DO?** *Intervention quality is crucial but inconsistently defined and measured:* Lessons from approaches used in early childhood education may inform efforts to improve quality metrics in ECD programmes
4. **KEY GAPS?** *Implementation research gaps for ECD programme measurement:* Application of more standardised measures of the caregiving environment (i.e. 'home-readiness' for ECD promotion) is feasible but challenges remain especially for simpler, routine, multi-domain outcome measurement and metrics regarding coverage and quality.

outcomes in research and programmes at relatively small scale.<sup>3–11</sup> As yet, measurement of outcomes in larger scale programmes and quality and coverage have received little attention. With increased focus on moving to scale, tracking and accelerating progress in ECD will require improved approaches to monitoring and evaluation (M&E) of programmes in diverse contexts.<sup>2,6</sup>

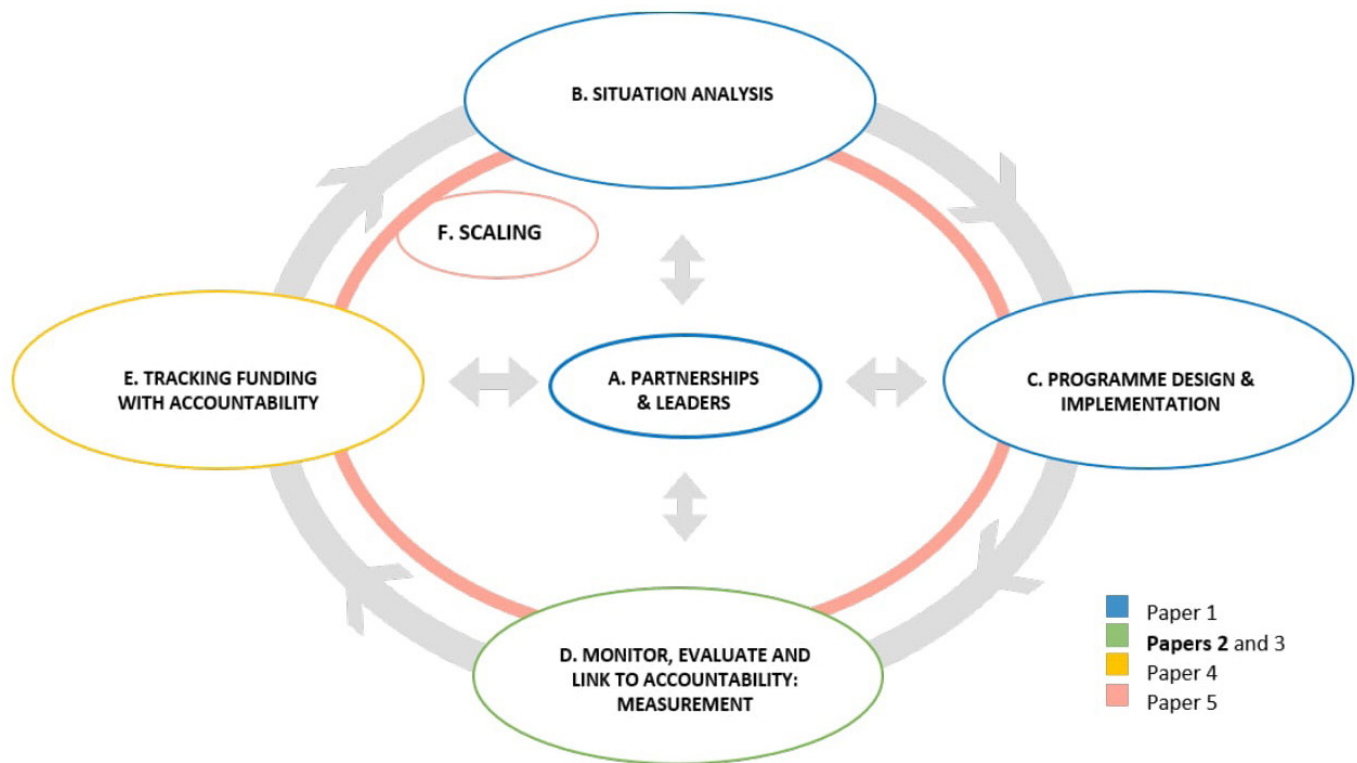
## Purpose, scope and structure of series

The purpose of this series is to inform design and implementation of ECD interventions at national and subnational level in diverse low-income and middle-income countries (LMIC). The series is structured around a programme cycle, outlined in figure 1 and built on a mixed-method evaluation of the Saving Brains, Grand Challenges, ECD portfolio as well as several additional analyses.<sup>12–15</sup> The first paper in this series detailed methods for Saving



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Milner KM, Bhopal S, Black M, et al. *Arch Dis Child* 2019;104:S13–S21.



**Figure 1** Programme cycle for design, implementation and scaling of early child development programmes.

Brains portfolio evaluation and findings related to programme design and implementation decisions in diverse LMIC.<sup>12</sup> Subsequent papers explore measurement of ECD in routine health services,<sup>13</sup> global funding for ECD<sup>14</sup> and multistakeholder perspectives on scaling.<sup>15</sup>

### AIMS AND OBJECTIVES

This paper aims to examine approaches to M&E of ECD programmes especially in LMIC.

#### Objectives

1. Review requirements for monitoring and evaluation: consider guidance for M&E in other areas of maternal, newborn and child health (MNCH) as well as implications of targeted published and grey literature review to adapt a M&E framework for ECD programmes, implemented through health programmes at national and subnational levels.
2. Describe findings of Saving Brains portfolio evaluation related to measurement: analyse approaches to M&E across the Saving Brains projects, especially measurement of intervention quality, coverage and outcomes, using the adapted M&E framework for ECD and consider implications for improved measurement.
3. Consider implications for improved M&E in ECD programmes at scale.

### METHODS

#### Objective 1: review requirements for monitoring and evaluation of ECD programmes at scale

Based on experience in other areas of MNCH plus pragmatic literature review, we adapted a generic programme impact framework for ECD programmes.<sup>16–19</sup> Our literature review used MEDLINE and Embase, with the following MeSH

terms: ‘Child development’ OR ‘Developmental Disabilities’ AND ‘Developing Countries’. Additional papers were identified through reference lists of papers retrieved. Grey literature was searched via websites of major multilateral organisations engaged in ECD programming such as the WHO, UNICEF, Save the Children Fund, the World Bank, World Vision International and other related organisations. From articles and documents retrieved, themes explored included measurement of ECD and intermediary outcomes, quality, and coverage, challenges experienced, and strategies used to address these in research and programming.

#### Objective 2: describe findings of Saving Brains portfolio evaluation related to measurement

As detailed in the first paper of this series,<sup>12</sup> a participatory mixed-method evaluation of the Saving Brains portfolio was undertaken (2016–2017) by a team at the London School of Hygiene & Tropical Medicine in collaboration with the Saving Brains platform team and innovators. A conceptual evaluation framework informed by the Medical Research Council Guidance on Evaluation of Complex Interventions<sup>20</sup> and based around a portfolio-level theory of change was developed to systematically describe and assess direct and intermediary ECD outcomes, quality and coverage measurement across the portfolio (supplementary web appendices figure A and textbox A).<sup>12</sup> Between 2012 and 2016, Saving Brains awarded 39 ‘Seed’ and ‘Transition-To-Scale’ (TTS) grants to innovators in 23 LMIC, with diverse backgrounds and variable innovation design and implementation approaches. Seed grants focused on demonstrating ‘proof of concept’ over 18–24 months, while TTS grants aimed to increase scale and sustainability in partnership with other organisations over 3 years.<sup>21</sup> Project types within the portfolio were categorised using domains in the NCF; responsive care, early learning, good health, adequate nutrition, and security and safety.<sup>12</sup>

### Box 1 Definitions of measurement levels in ECD programmes

- ▶ **Impact:** measured change in early child development (ECD) outcomes including on cognitive development and human capital.<sup>21</sup>
- ▶ **Intermediary outcomes:** factors that could be considered intermediary to ECD outcomes using an ecological model of child development (eg, child or caregiver nutrition, parental mental health and caregiving environment).<sup>21</sup>
- ▶ **Coverage:** the number of individuals receiving an intervention or service (the numerator) compared with the population in need of the intervention or service (the denominator).<sup>47</sup>
- ▶ **Quality:** variously defined including fidelity (the extent to which the delivered intervention is consistent with that intended) and/or evidence-based content and/or client satisfaction.

#### Quantitative data

Projects collected quantitative impact and process data using Grand Challenges Canada prespecified tools and reporting mechanisms that were structured around the portfolio theory of change (see supplementary web appendices figure A and table A).

#### Qualitative data

Qualitative methods were guided by the Consolidated Criteria for Reporting Qualitative Research (see supplementary web appendices textbox A and tables B,C).<sup>22</sup> Broad thematic areas of enquiry relevant to ECD programming were established based on literature review, stakeholder consultation and analysis of written portfolio documents. Themes included: impact metrics, intervention content, integration, place of delivery, human resources, coverage and quality, working in partnership, technology and sustainable financing.

Key informants were purposively selected from professional networks and included national and international programmers and policy makers, ECD researchers, Saving Brains innovators and members of the Saving Brains platform (see supplementary web appendix table B). All were invited to focus group discussions and in-depth interviews that were conducted online or face to face at Saving Brains meetings and workshops. In-depth interviews and focus group discussions were facilitated using 'topic guides' (see supplementary web appendix table C). Audio recordings of meetings were submitted to a third party for transcription.

#### Data analysis

Quantitative data were cleaned and analysed using Stata V.14 and Microsoft Excel. Descriptive statistics related to frequency and mode of outcome measurement across the portfolio were generated. Project documents, in-depth interviews and

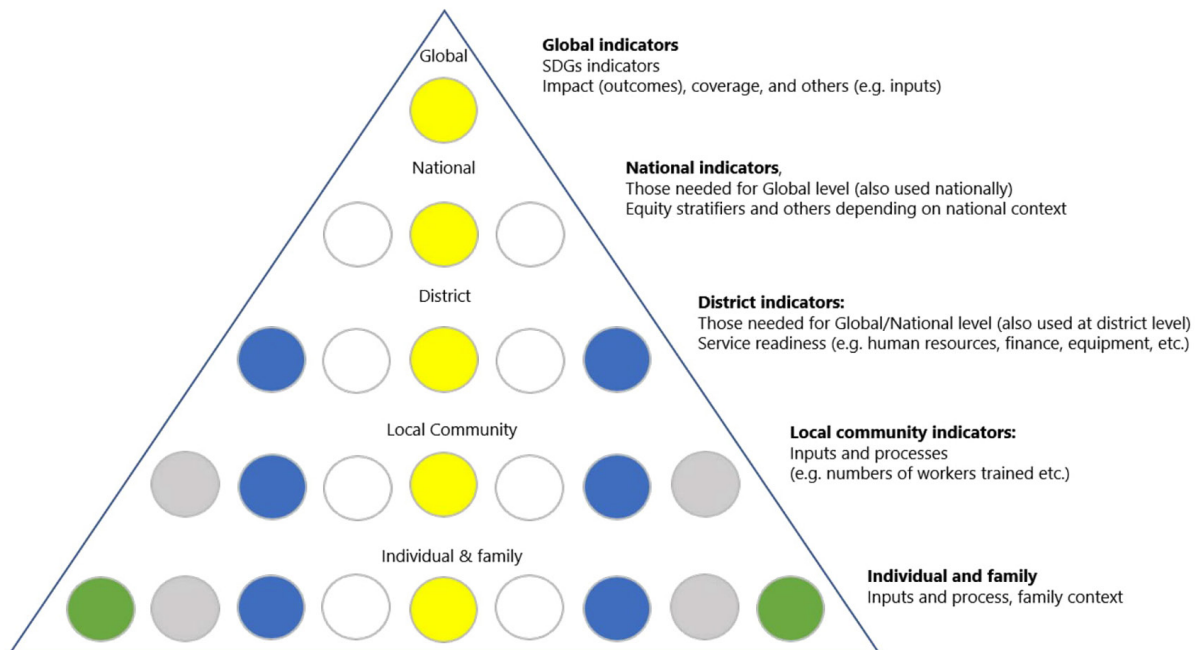
**Table 1** M&E needs in ECD programmes according to level of impact framework

Level of M&E framework	What are we trying to measure?	What is available?
Impact (outcomes)	<b>Deaths.</b> Disability. Child development – multiple domains. <b>Child health.</b> Function and participation. Education. Well-being and quality of life. Adolescent and adult outcomes.	Proxy measures of child development (eg, stunting). Limited childhood disability data. <b>Child mortality and health morbidity data.</b>
Risk factors for targeting	Biopsychosocial risk: <b>Stunting/undernutrition (2.2).</b> <b>Low birth weight (LBW), preterm birth, neonatal encephalopathy and sepsis (3.2).</b> Extreme poverty (1.2). Exposure to violence (16.2). Access to education (4.2). Maternal mental health and education.	Stunting. Neonatal estimates – LBW and preterm. Maternal mortality and birth by age 18 years. Maltreatment, harsh punishment and inadequate supervision.
Coverage*	Intervention coverage by domain of the NCF (relative to population need <b>health, nutrition, security and safety, responsive caregiving and early learning</b> )	Responsive caregiving – nil routine. <b>Healthcare and nutrition – various per RMNCH indicators.</b> Early learning – homes that have children's books, children who have support for learning, children have playthings at home and attendance in early childhood education. Safety and security – birth registration, basic drinking water and basic sanitation.
Quality	Quality of programme provision. Client experience of service provision. Home care.	Quality of intervention measures.
Inputs (service and home readiness, policy and finances)	Numbers of workers. Service readiness. 'Home readiness'. ECD policy supports. Fiscal environment and resources.	Home context. Policy – maternity leave, convention of the rights of the child, breast feeding. Education, child health and social protection grants. National government and donor finance.

Bold texts typically measured within routine MNCH programmes.

\*Defined as the number of individuals receiving an intervention or service (the numerator) compared with the population in need of the intervention or service (the denominator).<sup>47</sup>

ECD, early child development; M&E, monitoring and evaluation; MNCH, maternal, newborn and child health; NCF, Nurturing Care Framework.



**Figure 2** Data collection and use required as standard at all health system levels (priority indicators for alignment shown in yellow). Adapted from Day *et al.*<sup>48</sup>

focus group discussion transcripts were imported and coded in NVivo 11.0. An inductive approach was used to create a coding framework, and thematic content analysis was undertaken to explore themes related to outcome, quality and coverage measurement. Qualitative data were coded by two separate members of the evaluation team (KMM and MK-L) until data saturation was reached.

### Objective 3: consider implications for improved measurement in ECD programmes at scale

Results of objectives 1 and 2 were synthesised to describe and discussion challenges, opportunities and next steps for improved M&E of ECD programmes at scale [box 1](#).

## RESULTS

### Objective (1): review of measures required to monitor and evaluate ECD programmes at scale

Based on the above methods, we considered that the following measurement levels and constructs need to be considered in M&E of ECD programmes at scale ([table 1](#) and [figure 2](#)).

### Objective 2: consider Saving Brains projects approach to M&E and implications for measurement in programmes at scale

Saving Brains portfolio evaluation included 39 projects implemented across 23 LMIC between September 2013 and November 2016. Of these, 34 were Seed and 5 were TTS projects. Most projects (n=32) included some aspect of responsive caregiving and early learning. Details of individual project details are provided in the first paper of this series.<sup>12</sup> [Table 2](#) outlines larger TTS projects and approaches taken to M&E within them.

### Outcomes

Most implementing teams reported on cognitive, language, motor and socio-emotional development ([table 3](#)). Few projects reported on broader aspects of ECD (eg, executive functioning, adaptive skills) and no projects specifically measured disability

or sensory impairments, including those following high-risk newborn populations. Given the age focus of the portfolio and short grant duration, no projects measured educational or longer-term outcomes.

Of the 37 projects that measured ECD outcomes, 51% (n=19) used a comprehensive developmental assessment. Of these, the majority used the Bayley Scales of Infant and Toddler Development (BSID II or Bayley-III), although very few formally translated and adapted this tool and none were locally validated. Regionally developed instruments including the Malawi Development Assessment Tool and the Kilifi Developmental Inventory were used in several projects in sub-Saharan countries.<sup>23–27</sup>

Among those using screening tools (49%, n=18), the Ages and Stages Questionnaires (ASQ) was the most common; ASQ was used for 22% (n=8) of projects across the whole portfolio with similar challenges in adaptation and validation. Alternative screening tools included those developed in high-income settings (eg, Parents' Evaluation of Developmental Status Survey of Well-Being of Young Children) as well as some specifically for LMIC settings (eg, Ten Questions Questionnaire, Saving Brains Early Child Development Scale [a precursor to the Caregiver Reported Early Development Index]).<sup>28–33</sup>

Forty-three percent (n=16) of project teams used a combination of tool types. Biological methods for assessment were used by one project examining event-related electroencephalograms in infants. Details of instrument translation and adaptation were reported for a minority 19% (n=7) of projects.

### Intermediary outcomes

Approximately half of projects (51%, n=20) measured child growth and/or nutritional outcomes, mostly anthropometry with a few reporting more detailed nutritional outcomes.

Caregiver capabilities, caregiver–child interactions and/or the home environment were measured in 39% (n=15) of projects. Details of specific tools were frequently unavailable; however,



**Table 2** Measurement of outcomes, coverage and quality in key Transition-to-Scale projects

Project name	Transition to scale of an integrated programme of nutritional care and psychosocial stimulation to improved malnourished children's development.	An integrated intervention targeted at deprived preschool children in rural areas.	Home visiting programmes to improve early child development and maternal mental health.	Saving Brains, Changing Mindsets.
Lead institution	International Centre for Diarrhoeal Disease Research.	Universidad de los Andes.	Faculdade de Medicina da Universidade de Sao Paulo.	Mobile Crèches for Working Mothers' Children.
Country	Bangladesh	Colombia	Brazil	India
Site	Dhaka area: Narsingdi and Kishoreganj Rural.	Central rural regions: Boyaca, Cundinamarca, Santander, Tolima.	Sao Paulo, urban slums in western area.	Delhi area, Bangalore, Ahmedabad and Chandigarh.
Vision/goal/objectives	Integrate RCEL intervention for poor, underweight children into routine government health services.	Improve quality of a pre-existing public parenting programme in a scalable fashion.	Evaluate the efficacy and cost-effectiveness of two alternative platforms for home visiting programmes.	Demonstrate scalability of workplace-based childcare for children of migrant construction workers.
No. of participating children	1597	2134	206	4845
Situational analysis and adaptation to context	Bangladesh government recognises educational and financial benefits of ECD and has incorporated ECD into national plan of action. Existing national infrastructure of government primary healthcare community clinics with supervision structure for clinic staff. Initial concerns regarding workload and motivation.	Building on existing long-standing government <i>Family, Women and Infancy Program</i> for women during pregnancy until their children are 2 years of age, combined clinic and home visits. Focused on several regions in Central Colombia, similar in terms of cultural and socioeconomic context.	Governmental support.	Pre-existing Mobile Crèches NGO programme, demand from construction companies.
Measurement approach				
Outcome domains measured	Child development outcomes: Cognition. Language. Motor. Socioemotional. Intermediary outcomes: Nutrition. Parental practices. Home environment. Parental mental health.	Direct: Cognition. Language. Motor. Socioemotional. Indirect: Nutrition. Parental practices. Home environment.	Direct: Cognition. Language. Motor. Socioemotional. Indirect: Nutrition. Parental practices. Home environment.	Direct: Cognition. Language. Motor. Socioemotional.
Quality	Detailed process evaluation.	Detailed process evaluation (see <a href="#">box 2</a> for key features measured).	Detailed process evaluation.	Detailed process evaluation (see <a href="#">box 3</a> for key features measured). Early Childhood Education Quality Assessment Scale.
Coverage (risk factor for targeting/method of measurement/population level need)	Underweight children attending community primary healthcare clinics (covered 90 of 13 000 nationally).	Socioeconomic and geographic (rural location). Population-level data on socioeconomic vulnerability, already collected by national government. Population-level survey (ELCA 2013) including Ages and Stages Questionnaires data.	Socioeconomic and geographic (urban slum).	National estimates of children of unskilled labourers (n=20 million).
References	49-51	51-52	51-53	51-54
	Nurturing Care for Early Child Development Country Profiles – <a href="http://nurturing-care.org/resources/country-profiles/">http://nurturing-care.org/resources/country-profiles/</a> .			

**Table 3** Proportion of projects reporting on child development and intermediary outcomes across Seed (n=34), Transition to Scale (n=5) and all (n=39) projects combined

Domain of child developmental	Seed grants % reporting (n=34)	Transition-to-Scale grants % reporting (n=5)	Total (n=39)
Motor/growth	82 (n=28)	80 (n=4)	78 (n=32)
Cognition	88 (n=30)	80 (n=4)	83 (n=34)
Language and communication	88 (n=30)	80 (n=4)	83 (n=34)
Socioemotional	87 (n=27)	80 (n=4)	76 (n=31)
Identity and self-efficacy: spiritual and cultural	6 (n=2)	40 (n=2)	10 (n=4)
Intermediary outcomes			
Caregiver physical health	6 (n=2)	40 (n=2)	28 (n=11)
Caregiver mental health	9 (n=3)	40 (n=2)	41 (n=16)
Caregiver capabilities	18 (n=6)	100 (n=5)	80 (n=31)
Caregiver child interaction	18 (n=6)	100 (n=5)	67 (n=26)
Home environment	15 (n=5)	100 (n=5)	56 (n=22)
Community environment or access to services	3 (n=1)	60 (n=3)	41 (n=16)
Protection from abuse or exploitation	0 (n=0)	20 (n=1)	8 (n=3)

the Home Observation for Measurement of the Environment was most commonly used as well as the Family Care Indicators.<sup>3,4</sup>

Several projects measured maternal physical (13%, n=5) or mental health (31%, n=12), particularly maternal depression using a wide range of tools.

### Coverage and equity

Within the Saving Brains portfolio most project teams measured number of children receiving an intervention, often in comparison with initial project targets. However, few projects measured coverage in terms of broader population-level need. This was likely due to the small size and emphasis on 'proof of concept' rather than scaling in the majority of Seed projects.

However, even for TTS grants that were more actively focused on scaling (table 2), population need for ECD interventions was typically defined on population-level data of risk factors (rather than ECD status per se) or implicit local knowledge. For

example, in the absence of direct population-level data on ECD, nutritional or socioeconomic status, for which population-level data were more commonly available, was used to consider both population level need and coverage. Box 2 provides an example of the use of population-level data on socioeconomic status for targeting interventions, in implementation of an enhanced Family, Women and Infancy (FAMI) Programme to deprived rural communities in Colombia.

### Quality

During focus group discussions, key informants reflected that the quality of intervention implementation was crucial to achieving and maintaining intervention impact with scaling. However, as reflected by the examples provided in boxes 2 and 3, there was no consensus about how to define quality. For example, the enhanced FAMI programme in Colombia, an integrated nutrition and parenting programme implemented through government services, developed their own definition and approach to monitoring quality. Their definition of quality included feasibility, acceptability and fidelity, and the project team developed their own approaches to measurement of these elements. By contrast, the Mobile Crèches project, implemented as an early childhood education (ECE) initiative in urban slums in India, drew on the large existing quality improvement literature within ECE plus specific, related tools to measure quality.

### DISCUSSION

In this paper, through evaluation of measurement approaches used across the largest donor-funded ECD intervention portfolio, we have highlighted the urgent need to strengthen ECD M&E frameworks for scaling.

Specifically, our evaluation, based around an adapted generic M&E framework has highlighted the importance of ensuring that that programming inputs, quality, coverage and impact are all considered, defined and measured.<sup>4</sup> Table 1 summarises these elements, building on our evaluation findings and considering indicators often available in routine government services as a basis for developing improved approaches to monitoring in existing systems.

### Box 2 Monitoring quality of an integrated nutrition and stimulation intervention for families in rural Colombia<sup>52</sup>

**Context and intervention:** in rural Colombia a pre-existing government-funded parenting programme (the Family, Women and Infancy [FAMI] programme) was enhanced through adaptation and implementation of a structured responsive caregiving curriculum, combined with nutritional education and supplementation. Participating rural families were identified by socioeconomic vulnerability according to a pre-existing government 'Proxy Means Test'.

The enhanced FAMI programme was delivered through combined group and home visits, implemented by local women who were high-school graduates but had no previous child development training.

**Approach to measurement of quality:** the intervention team considered the following elements in measuring the quality of intervention delivery: feasibility, fidelity and acceptability. Feasibility was defined as the expected ability to implement the intervention given existing resources, including personnel, context and intervention characteristics. Fidelity was defined as implementation of the intervention to compare with that intended. Acceptability was defined from front-line worker perspective. These elements were measured and tracked through all stages of intervention design and implementation.

Monitoring methods used included active supervision, in-depth interviews, focus groups and surveys of both front-line workers and participants. Process data related to quality indicators were used to make adjustments to both intervention design and implementation throughout implementation.

**Implications for measurement of quality with scaling:** based on monitoring and evaluation findings, several adaptations were considered necessary for ongoing quality with scaling. In addition, with scaling, intensive methods of monitoring, including videos and detailed qualitative methods, were considered likely to need adjustment. However, front-line worker supervision through intermittent observation and participant satisfaction surveys was considered to be both important and feasible moving forwards.

Box 3 Monitoring quality in an enhanced early learning intervention for children of construction workers in India<sup>54</sup>

**Context and intervention:** Mobile Crèches (MC) has worked for several decades in partnership with other non-government organisations (NGOs), construction industry, government, local communities to increase access to quality crèches and early childhood education services across 20 states of India. A Saving Brains, Transition to Scale grant enabled the MC team to test the feasibility, effectiveness and scalability of its workplace-based childcare programme for young children of migrant construction workers.

Within the TTS project, MC reached out to 5000 under six children with a child care programme delivered through partnerships with 11 NGOs and 24 builders at 40 crèches. The model for scaling was based on partnership to support local NGOs with rigorous supportive supervision and on-the-job training for the first 6 months while at the same time strengthening NGO capacity for ongoing monitoring and supervision. After 6 months, MC continued with less frequent monitoring according to NGO capabilities.

**Approach to measurement of quality:** quality of the daycare programmes was assessed using the Early Childhood Education Quality Assessment Scale (ECEQAS), which included domains of infrastructure, physical setting, meals, naps, learning/play aids, classroom management and organisation, personal care, hygiene and habit formation, language and reasoning experiences, fine and gross motor activities, creative activities, activities for social development and disposition of childcare workers. NGO-run centres performed well on most on the ECEQAS and observer reports were consistent with this measured findings.

**Future directions for scaling:** based on the experiences of implementing the pilot project, MC continues to invest with current partners in strengthening the childcare provisions at construction sites and gathering data against child development outcomes and has expanded into new cities. In the near future, MC plans to replicate this model across other worksites such as tea plantations, factories and brick kilns to ensure quality ECD services for vulnerable young children, through low cost crèches and day care centres.

Furthermore, as has been recognised in M&E of other areas of MNCH programming, we suggest that agreement on a select number of indicators in each of these areas will be important to meaningful monitoring and improved programme accountability for ECD moving forward.<sup>19</sup> For example, figure 2, which has been used in development of a global newborn measurement improvement roadmap, illustrates that at lower health system levels multiple indicators are typically collected by different groups for diverse purposes, but at higher health service levels, fewer indicators are collected.<sup>35 36</sup> While all health system levels are crucial, measurement needs vary. High levels of precision are needed at individual or lower system levels whereas lower sensitivity or precision for the individual is acceptable for measurement at higher system levels. Although alignment of indicators across levels is not possible, agreement on a select number, as indicated in yellow in figure 2, will be important to allow greater comparability across settings and with programme variation.<sup>19</sup>

However, our evaluation also highlights two major challenges for M&E of ECD programmes at scale.

### Challenge 1: measuring ECD outcomes in routine programmes

Consistent with previous literature and despite substantial investment, our evaluation has demonstrated challenges in measurement of ECD outcomes across LMIC, which limit comparative understanding of impact.<sup>3-11</sup> Issues highlighted include the use of screening tools to monitor the effect of interventions (used in 49% of projects), although these are not designed to measure impact. Furthermore, there was variability in translation, adaptation, piloting and standardisation of tools and emphasis on short-term cross-sectional rather than longer term measurement of outcomes. Additionally, important domains (eg, vision, hearing, functioning and disability) were often omitted, limiting holistic understanding of individual results.

As such, practical challenges to outcome measurement (eg, cost, access, staff training and cultural adaptability) need to be addressed or alternative approaches explored to improve understanding of ECD programme impact across diverse settings.

As Boggs *et al* discuss, limitations in existing outcome measurement approaches warrant further research to improve their use in routine services.<sup>13</sup> While recent developments in

population-level child development tools are encouraging, application of these tools for programme monitoring is not yet established.<sup>37</sup>

### Challenge 2: defining and monitoring quality and coverage

Our analysis has highlighted major measurement challenges related to quality and coverage of ECD programmes.

Intervention quality was considered important but variably defined and measured. Examples such as the enhanced FAMI programme in Colombia demonstrate how, in a research context, multimethod approaches can be used to assess quality and can be fed back into existing programmes to enhance their design and implementation (R Bernal, unpublished data, 2018). However, it is also important to note the observation by authors of that study that quality monitoring approaches need to continue to be adjusted to remain feasible and achievable with scaling (R Bernal, unpublished data, 2018). Furthermore, given relative underexploration of monitoring of quality within ECD programming in health, lessons can perhaps be translated from more developed quality monitoring standards and approaches in ECE.<sup>38</sup> The Mobile Crèches project provides an important example of using ECE quality standards in a low-income setting. At larger scale, initiatives such as the Measuring Early Learning Quality and Outcomes provide examples of deliberate efforts to improve feasible and meaningful measurement of quality of preprimary learning environments, which may also be informative for measurement of quality of ECD programmes for younger children.<sup>38</sup>

Within Saving Brains, emphasis on coverage measurement was mostly limited to coverage within individual projects rather than in the population more generally. For many Seed projects, emphasising proof of concept, this is not surprising. However, as projects and programmes are scaled, further evaluation of coverage related to the overall population in need will be crucial to accelerating impact. This resonates with previous literature emphasising that in other areas of MNCH, monitoring intervention coverage has been crucial to accelerating and tracking progress.<sup>39 40</sup>

However, within ECD, specific barriers to measurement of both numerator (intervention) and denominator (population

in need of intervention) pose challenges to improving coverage measurement. First, the NCF describes many effective ECD interventions, but achieving consensus on what defines a core intervention package (ie, numerator) remains challenging. The NCF has broadly defined 'responsive caregiving' but more focused definitions of interventions are needed for estimations of coverage at scale.<sup>2,41</sup>

Furthermore, the population at need (ie, denominator) needs to be more clearly defined based on population level data rather than perceptions that may be incorrect. The NCF suggests periodic population-based assessment of child development home-care practices, along with risk and protective factors for nurturing care.<sup>2</sup> As we have highlighted, many relevant indicators are available in DHS or MICS data that may support better understanding of population need. However, more research is required to see if population-based measures under development can be used for this purpose.<sup>32,37,42–44</sup>

Given current challenges in coverage measurement beyond small-scale research or projects, it may be possible to use intermediary measures or inputs as a proxy while improving more routinely useable measures of ECD. Consistent with other literature, stakeholders within our evaluation reported the value, of measuring intermediary outcomes, especially to policy makers who value short-term demonstration of change.<sup>45</sup> Furthermore, previous literature has demonstrated the predictive value of intermediary factors, such as caregiving environment, on long-term development and educational outcomes.<sup>11,46</sup>

Finally, as the NCF outlines, improving equity while scaling ECD programmes will require data disaggregated by a broad range of stratifiers including sex, age, income, wealth, ethnicity, migratory status, disability and geographic location.<sup>2</sup>

### Strengths, limitations and future research directions

To improve M&E of ECD programmes at scale, applied research that tests what is feasible and useful in routine health services, as well as approaches to intersectoral coordination is needed. Innovative ways to address well-recognised limitations of current approaches to measuring outcomes including further consideration of long-term function and well-being as well as educational outcomes are required. Greater investment in defining and measuring ECD programme quality and coverage are also crucial to accelerating progress.

### CONCLUSION

Translating calls for action into positive change for child development requires clearer recommendations on what to measure, how to track ECD programmes and course correct when indicated. Evaluation of the Saving Brains portfolio highlights the strengths of a 'theory of change' based approach to M&E, and the importance of measuring intermediary outcomes, such as changes in caregiving environment, on the pathway to impact. However, substantial challenges in measurement of child development outcomes reflect the need for further research and development to ensure that outcomes can be meaningfully and feasibly measured in routine programmes. Clearer definitions of ECD interventions, quality and the population in need are also required to understand and improve equitable coverage. Approaches to data disaggregation in M&E are needed to ensure that ECD programmes delivery on their efforts to support all children to thrive.

### Author affiliations

<sup>1</sup>Maternal, Adolescent, Reproductive and Child Health Centre, London School of Hygiene & Tropical Medicine, London, UK

<sup>2</sup>Murdoch Children's Research Institute, Melbourne, Australia

<sup>3</sup>Maternal & Child Health Intervention Research Group, Department of Population Health, London School of Hygiene and Tropical Medicine, London, UK

<sup>4</sup>Northern School of Paediatrics, Newcastle upon Tyne, UK

<sup>5</sup>Department of Pediatrics, University of Maryland School of Medicine, Baltimore, Maryland, USA

<sup>6</sup>Department of Mental Health and Substance Abuse, World Health Organisation, Geneva, Switzerland

<sup>7</sup>Institute of Translational Medicine, University of Liverpool, Liverpool, UK

<sup>8</sup>Maternal and Child Health Division, International Centre for Diarrhoeal Disease Research Bangladesh, Dhaka, Bangladesh

<sup>9</sup>Children's Investment Fund Foundation, London, UK

<sup>10</sup>University of Bristol, School of Social and Community Medicine, Bristol, UK

<sup>11</sup>Department of Paediatrics and Child Health, Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania

<sup>12</sup>Munk School of Global Affairs and Public Policy, University of Toronto, Toronto, Canada

<sup>13</sup>Center on the Developing Child, Harvard University, Cambridge, Massachusetts, USA

<sup>14</sup>Mobile Crèches, New Delhi, India

<sup>15</sup>Nutrition and Clinical Services Division, International Centre for Diarrhoeal Disease Research, Dhaka, Bangladesh

<sup>16</sup>Neonatal Medicine, University College London Hospitals NHS Foundation Trust, London, UK

**Acknowledgements** We would like to thank all the innovators, participants and researchers involved in projects included in the Saving Brains portfolio and evaluation. We would also like to thank Grand Challenges Canada as funder of unpublished data. We would like to thank the Early Child Development Expert Advisory Group for their guidance, and we are grateful to Claudia da Silva and Fion Hay for administrative assistance.

**Contributors** Technical oversight of the series was led by JEL and KMM. The first draft of the paper was undertaken by KMM. Other specific contributions were made by SB, MB, TD, MG, JH, RH, MK-L, KM, VPH, JR, SS, FT, CT and JEL. The Early Child Development Expert Advisory Group (alphabetically: Pia Britto, TD, Esther Goh, Sally Grantham-McGregor, MG, JH, RH, KM, JR, Muneera Rasheed, Karlee Silver and Arjun Upadhyay) contributed to the conceptual process throughout. All authors reviewed and agreed on the final manuscript.

**Funding** This supplement has been made possible by funding support from the Bernard van Leer Foundation. Saving Brains impact and process evaluation funded by Grand Challenges Canada.

**Disclaimer** The authors alone are responsible for the views expressed in this article and they do not necessarily represent the views, decisions or policies of the institution with which they are affiliated.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Commissioned; externally peer reviewed.

**Data sharing statement** All data were shared under an agreement with Saving Brains. Further information is available if required.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

### REFERENCES

- 1 United Nations. Sustainable Development Goals. 2015.
- 2 World Health Organisation. Nurturing Care Framework. Geneva: World Health Organisation. 2018. [https://www.who.int/maternal\\_child\\_adolescent/documents/nurturing-care-early-childhood-development/en/](https://www.who.int/maternal_child_adolescent/documents/nurturing-care-early-childhood-development/en/)
- 3 Fernald L, Prado E, Kariger P, et al. *A Toolkit for Measuring Early Childhood Development in Low- and Middle-Income Countries*. Washington, DC: International Bank for Reconstruction and Development/The World Bank, 2017.
- 4 About FE, Prado EL. Measuring the implementation of early childhood development programs. *Ann N Y Acad Sci* 2018;1419:249–63.
- 5 Sabanathan S, Willis B, Gladstone M. Child development assessment tools in low-income and middle-income countries: how can we use them more appropriately? *Arch Dis Child* 2015;100:482–8.
- 6 Goldfeld S, Yousafzai A. Monitoring tools for child development: an opportunity for action. *Lancet Glob Health* 2018;6:e232–e233.
- 7 Semrud-Clikeman M, Romero RAA, Prado EL, et al. [Formula: see text] Selecting measures for the neurodevelopmental assessment of children in low- and middle-income countries. *Child Neuropsychol* 2017;23:1–42.



- 8 Rubio-Codina M, Araujo MC, Attanasio O, *et al.* Concurrent Validity and Feasibility of Short Tests Currently Used to Measure Early Childhood Development in Large Scale Studies. *PLoS One* 2016;11.
- 9 Fischer VJ, Morris J, Martinez J, *et al.* Developmental screening tools: feasibility of use at primary healthcare level in low- and middle-income settings. *J Health Popul Nutr* 2014;32:314–26.
- 10 Fernald L, Kariger P, Engle P, *et al.* *Examining Early Child Development in Low-Income Countries: A Toolkit for the Assessment of Children in the First Five Years of Life.* Washington DC: World Bank, 2009.
- 11 Frongillo EA, Tofail F, Hamadani JD, *et al.* Measures and indicators for assessing impact of interventions integrating nutrition, health, and early childhood development. *Ann N Y Acad Sci* 2014;1308:68–88.
- 12 Milner KM, Bernal R, Brentani A, *et al.* Contextual design choices and partnerships for scaling early child development programmes. *Arch Dis Child* 2019;104(Suppl 1):S3–S12.
- 13 Boggs D, Milner KM, Black M, *et al.* Rating early child development outcome measurement tools for routine health programme use. *Arch Dis Child* 2019;104(Suppl 1):S22–S33.
- 14 Arregoces L, Hughes R, Tann C, *et al.* Accountability for funds for Nurturing Care: what can we measure? *Arch Dis Child* 2019;104(Suppl 1):S34–S42.
- 15 Cavallera V, Tomlinson M, Radner J, *et al.* Scaling early child development: what are the barriers and enablers? *Arch Dis Child* 2019;104(Suppl 1):S43–S50.
- 16 World Health Organisation. *The Global Accelerated Action for the Health of Adolescents (AA-HA!) Implementation Guidance*, 2017.
- 17 World Health Organisation. *WHO recommendations on antenatal care for a positive pregnancy experience*, 2016.
- 18 World Health Organisation. Scaling-Up Nutrition. <https://scalingupnutrition.org/>
- 19 World Health Organisation. Every Newborn: an action plan to end preventable deaths, 2014. [https://www.who.int/maternal\\_child\\_adolescent/newborns/every-newborn/en/](https://www.who.int/maternal_child_adolescent/newborns/every-newborn/en/)
- 20 Moore GF, Audrey S, Barker M, *et al.* Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 2015;350:h1258.
- 21 Saving Brains. Saving Brains, Grand Challenges Canada. 2018.
- 22 Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007;19:349–57.
- 23 Abubakar A, Holding P, van Baar A, *et al.* Monitoring psychomotor development in a resource-limited setting: an evaluation of the Kilifi Developmental Inventory. *Ann Trop Paediatr* 2008;28:217–26.
- 24 Abubakar A, Holding P, Van de Vijver F, *et al.* Developmental monitoring using caregiver reports in a resource-limited setting: the case of Kilifi, Kenya. *Acta Paediatr* 2010;99:291–7.
- 25 Gladstone M, Lancaster GA, Umar E, *et al.* The Malawi Developmental Assessment Tool (MDAT): the creation, validation, and reliability of a tool to assess child development in rural African settings. *PLoS Med* 2010;7:e1000273.
- 26 Bayley N. *Manual for the Bayley Scales of Infant Development.* New York: The Psychological Corporation, 1969.
- 27 Bayley N. *Bayley Scales of Infant and Toddler Development.* 3rd edn. San Antonio, TX: ed Harcourt Assessment, 2006.
- 28 Ages & Stages Questionnaires. <https://agesandstages.com/>
- 29 PEDS Test. Available: <http://www.pedstest.com/default.aspx>
- 30 Floating Hospital for Children. The Survey of Well-being of Young Children. <https://www.floatinghospital.org/The-Survey-of-Wellbeing-of-Young-Children/Overview>
- 31 Harvard T.H. Chan School of Public Health. Caregiver-Reported Early Development Instruments (CREDI). 2018. <https://sites.sph.harvard.edu/credi/>
- 32 McCoy D, Black M, Daelmans B, *et al.* Measuring child development in children from birth to age 3 at population level. *Early Childhood Matters* 2016.
- 33 Durkin MS, Davidson LL, Desai P, *et al.* Validity of the ten questions screened for childhood disability: results from population-based studies in Bangladesh, Jamaica, and Pakistan. *Epidemiology* 1994;5:283–9.
- 34 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.
- 35 The EQUITY Project. *Using Information for Action: a manual for health workers at facility level.* Pretoria, South Africa: Department of Health, 2018. Available: [https://pdf.usaid.gov/pdf\\_docs/Pnacx654.pdf](https://pdf.usaid.gov/pdf_docs/Pnacx654.pdf)
- 36 Every Newborn. *An Action Plan to End Preventable Deaths.* Geneva: World Health Organisation, UNICEF, 2014.
- 37 Lancaster GA, McCray G, Kariger P, *et al.* Creation of the WHO Indicators of Infant and Young Child Development (IYCD): metadata synthesis across 10 countries. *BMJ Glob Health* 2018;3:e000747.
- 38 UNESCO, UNICEF, Brookings Institution, *et al.* *MELQO: Measuring Early Learning Quality and Outcomes.* Washington, DC, 2017.
- 39 De Silva MJ, Lee L, Fuhr DC, *et al.* Estimating the coverage of mental health programmes: a systematic review. *Int J Epidemiol* 2014;43:341–53.
- 40 Moxon SG, Ruysen H, Kerber KJ, *et al.* Count every newborn; a measurement improvement roadmap for coverage data. *BMC Pregnancy Childbirth* 2015;15(Suppl 2):S8.
- 41 Britto PR, Lye SJ, Proulx K, *et al.* Nurturing care: promoting early childhood development. *Lancet* 2016;389:91–102.
- 42 Jacobusse G, van Buuren S, Verkerk PH. An interval scale for development of children aged 0–2 years. *Stat Med* 2006;25:2272–83.
- 43 van Buuren S. Growth charts of human development. *Stat Methods Med Res* 2014;23:346–68.
- 44 McCoy D, Black MM, Daelmans B, *et al.* Measuring development in children from birth to age 3 at population level. In: Moreno T, ed. The Hague, The Netherlands: Bernard van Leer Foundation, 2016:34–9.
- 45 Lombardi J. What policymakers need from implementation evaluations of early childhood development programs. *Ann N Y Acad Sci* 2018;1419:17–19.
- 46 Hamadani JD, Tofail F, Hilaly A, *et al.* Use of family care indicators and their relationship with child development in Bangladesh. *J Health Popul Nutr* 2010;28:23–33.
- 47 Tanahashi T. Health service coverage and its evaluation. *Bull World Health Organ* 1978;56:295–303.
- 48 Day LT, Ruysen H, Gordeev VS, *et al.* “Every Newborn-BIRTH” protocol: observational study validating indicators for coverage and quality of maternal and newborn health care in Bangladesh, Nepal and Tanzania. *Journal of Global Health.* in press.
- 49 Hamadani JD, Mehrin SF, Tofail F, *et al.* Integrating an early childhood development programme into Bangladeshi primary health-care services: an open-label, cluster-randomised controlled trial. *Lancet Glob Health* 2019;7:e366–75.
- 50 Hamadani JD, Nahar B, Huda SN, *et al.* Integrating early child development programs into health and nutrition services in Bangladesh: benefits and challenges. *Ann N Y Acad Sci* 2014;1308:192–203.
- 51 Radner JM, Ferrer MJS, McMahan D, *et al.* Practical considerations for transitioning early childhood interventions to scale: lessons from the Saving Brains portfolio. *Ann N Y Acad Sci* 2018;1419:230–48.
- 52 Attanasio O, Baker-Henningham H, Bernal R, *et al.* Early Stimulation and Nutrition: the impacts of a scalable intervention. NBER Working Paper No. 25059, 2018.
- 53 Smith JA, Baker-Henningham H, Brentani A, *et al.* Implementation of Reach Up early childhood parenting program: acceptability, appropriateness, and feasibility in Brazil and Zimbabwe. *Ann N Y Acad Sci* 2018;1419:120–40.
- 54 Bajaj M, Sharma S. Scaling-up early childhood centres for migrant construction workers’ children in India. *Early Childhood Matters* 2016;125:74–9.