Validation of the International Guide for Monitoring Child Development demonstrates good sensitivity and specificity in four diverse countries

Ilgi Ozturk Ertem[1,*], Vibha Krishnamurthy[2], Mphelekedzeni C. Mulaudzi[3], Yanina Sguassero[4], Burcu Bilik[1], Roopa Srinivasan[2], Hakan Balta[1], Ozlem Gulumser[1], Geliang Gan[5], Lisa Calvocoressi[5], Benjamin Johnson[5], Veronika Shabanova[6], Brian W.C. Forsyth[6]

[1]Developmental–Behavioural Paediatrics Division, Department of Paediatrics, Ankara University School of Medicine, Ankara, Turkey

[2]Ummeed Child Development Centre, Mumbai, India

[3]Kalafong Hospital, University of Pretoria, Pretoria, South Africa

[4]Centro Rosarino De Estudios Perinatales (Perinatal Study Centre of Rosario), Rosario, Argentina

[5]Yale Centre for Analytic Science, Yale University School of Public Health, New Haven, CT, USA

[6]Department of Paediatrics, Yale University School of Medicine, New Haven, CT, USA

*Correspondence

Ozturk Ertem, Ankara University School of Medicine Department of Pediatrics Developmental Pediatrics Division Cebeci, Ankara 06100, Turkey.

Tel: +90–533–311–0496 |

Fax: +90-312-3191440 |

Email: ertemilgi@yahoo.com

Abbreviations

CDA, Comprehensive developmental assessment;

CI, Confidence interval;

CrI, Credible interval;

GMCD, Guide for Monitoring Child Development;

WHO, World Health Organization.

Abstract

Aim: It is of critical importance to have internationally constructed tools to address early childhood development. The aim of this second phase of a two-phase study was to examine the sensitivity and specificity of the Guide for Monitoring Child Development (GMCD) in identifying developmental delay in four diverse countries.

Methods: The first phase of this 2011–2015 back-to-back study included 4949 children up to 42 months of age from primary healthcare centres in Argentina, India, South Africa and Turkey. Distribution curves were generated to show the ages when the children attained GMCD milestones and those that could be used across sexes and countries were placed in age ranges corresponding to the 85th and 97th percentile point estimates. Phase two examined a separately recruited sample of children in those countries to determine sensitivity and specificity of the GMCD.

Results: The validation phase of the 85 milestones in the GMCD identified delayed development in 30% of the 1731 children in the four countries. The sensitivity and specificity ranged from 0.71–0.94 and 0.69–0.82, respectively, for the total sample and the different age groups.

Conclusion: The GMCD standardised in four diverse countries has appropriate accuracy for identification of children with developmental delay.

Key notes

- International tools are needed to monitor early childhood development, and this paper covers the validity testing of the Guide for Monitoring Child Development (GMCD), developed by the authors.
- This phase of the study was carried out conducting assessments on 1731 children who were recruited from four very different countries: Argentina, India, South Africa and Turkey.
- The GMCD showed good sensitivity and specificity and was successfully used across countries to identify early developmental delays.

Introduction

Optimising early childhood development is now a United Nations Sustainable Development Goal, but major disparities exist in how this is addressed in health systems around the world [1, 2]. In high-income countries, early identification of children with developmental difficulties[3] is an integral component of child health care and children who are identified receive individualised, family-centred and comprehensive early interventions [4, 5]. In contrast, child development is rarely addressed during health care in low-income and middle-income countries, [3, 5, 6] and nonindividualised interventions involve promoting nurturing care [2]. While these universal strategies are crucial, complementary individualised approaches are also urgently needed to help close the unethical equity gap for children around the world [7.]

An important barrier to individualised early childhood development interventions in low-income and middle-income countries is the lack of universally applicable methods to assess children's development [8.] Most screening tools are developed for a single country and this has mostly happened in high-income countries [8]. To be applicable in other countries, these tools require the time, funds and research capacity for restandardisation and revalidation. Furthermore, a screen and refer approach is not considered appropriate for child development and is not feasible in many low-income and middle-income countries, where there may be

few resources to refer children to [3, 6-9]. Any tools used to identify children with developmental difficulties must also give service providers the skills they need to work with caregivers to promote the child's development, by planning interventions and monitoring progress [7-9].

The Guide for Monitoring Child Development (GMCD), produced in Turkey by Ertem et al. [10], is a comprehensive package that comprises three components: monitoring, supporting early childhood development and early intervention [10-14]. The theoretical conceptualisation of the GMCD, research on its standardisation, reliability, validity and applicability in Turkey and its international use has previously been reported [10-14]. The GMCD is based on a family-centred, strengths-based philosophy and it differs from screening tools in a number of ways. First, it uses an open-ended interview technique, rather than testing the child or asking caregivers questions that require yes or no answers. This is important, as it enables the healthcare providers using the GMCD to build a rapport with the child's caregiver and address challenges such as low literacy, limited knowledge of child development and any fear of stigma related to developmental disabilities. Furthermore, the GMCD does not just provide a score or categorisation. It has a unique format that allows the user to view the developmentally progressive, functional milestones in each of its seven domains, which are as follows: expressive and receptive language, gross and fine motor skills, relating to others, playing and self-help. This practical format focuses on a wide range of activities that are related to promoting the optimal development of the child. These include understanding and interpreting how well the child is functioning, together with strengths and difficulties, and identifying any delays at an early stage. The monitoring component of the GMCD is built on bioecological theory and the World Health Organization (WHO) International Classification of Functioning framework. The monitoring component links seamlessly to the supporting early childhood development component of the GMCD, enabling the healthcare provider to assess biopsychosocial risk factors and strengths, share findings with caregivers, anticipate what guidance will be needed and plan individual recommendations for each child. The GMCD can also be used to individualise interventions, such as the WHO/UNICEF Care for Child Development Intervention. The early intervention component is designed to be used for children with identified risk factors or developmental delay. The training for the GMCD includes identifying and supporting the development of children who are at risk of adverse outcomes, but not yet displaying developmental delay, as well as those already demonstrating developmental delay.

The GMCD is free of charge and has been reported to be one of few instruments that has adequate psychometric and feasibility criteria to be used in low-income and middle-income countries [7]. It has generated worldwide demand and clinicians from more than 30 countries have so far been trained in its use [13]. The WHO/UNICEF Nurturing Care Framework highlights developmental monitoring and contains a reference to the GMCD with regard to partnering with caregivers. It can be used to enhance the strengths of children, families and communities, address biopsychosocial risk factors and provide additional individualised support and services when needed [15].

This study to develop the international GMCD was conducted in four countries that are culturally and linguistically different and its aim was to standardise and validate the GMCD so that it can be used universally, without the need for restandardisation and revalidation. We have previously reported on the sociodemographic characteristics of the standardisation sample enrolled in phase one of this study and the GMCD milestones that were reliably applied across the countries. The results of this phase showed that the median ages of

attainment were the same across genders and countries for the majority of the milestones in the GMCD [14]. The aim of this second phase of the study was to determine the sensitivity and specificity of the GMCD by comparing it to a comprehensive developmental assessment.

Methods

This study was conducted in two back-to-back phases between March 2011 and May 2015, in typical health clinics in four metropolitan areas of Argentina, India, South Africa and Turkey. The primary institutions involved in each country were as follows: Centro Rosarino de Estudios Perinatales Rosario, Argentina; Ummeed Child Development Centre, Mumbai, India; University of Pretoria, South Africa and Ankara University, Ankara, Turkey. Yale University, USA, was involved in the coordination of the study and training.

In phase one, from March 2011 to October 2014, we used the healthy sample approach applied in the WHO Multicentre Growth Reference [16] and the WHO Motor Development Studies [17] to recruit 10 246 children aged 0–42 months. They were seen for routine care or minor illnesses in 22 government community health centres in the greater urban and peri-urban regions of the four countries. In Mumbai, children were also recruited from private paediatricians' offices to ensure an adequate number of children that met the health criteria for the study. Data were collected on all the recruited children, but we excluded children with various health-related factors from the standardisation analyses. These were a birth weight of less than 2500 g, perinatal complications requiring hospitalisation, undernutrition at the time of the study (weight for age, height for age or weight for height that was below -2 *Z*-scores from the median on the WHO Child Growth Standards) or a history of undernutrition. We also excluded known chronic health or developmental problems, such as congenital heart disease, autism, a history of anaemia or a haemoglobin of less than 10.5 g/dL at recruitment. This resulted in a sample size of 4949 children included in the standardisation analyses in phase one.

Phase two was conducted from November 2014 to June 2015 using a different sample of children recruited for the validity phase in the four countries after phase one was completed. The number of children recruited for the second phase of the study, to analyse the sensitivity and specificity of the GMCD, was 1731, including 593 children from Turkey, 467 from India, 361 from South Africa and 310 from Argentina. These children were specially recruited for phase two of the study so that they were completely independent of the 4,949 children in phase one.

Children aged 6–42 months were recruited in a similar manner from the health sites to phase one, but data on all recruited children, including those that did not meet the health criteria, were used. Unlike phase one, children under the age of six months were not included in phase two because development is very rapid in the earlier months, and the GMCD is intended to be used as an assessment tool after age six months. The reporting of phase two complied with the 2015 Standards for Reporting Diagnostic Accuracy recommendations. The GMCD's International Advisory Committee which provided recommendations for the study comprised of renowned child development experts and experts from the WHO and UNICEF. The participating caregivers provided written informed consent and the internal review boards of each study site and of Yale University approved the study.

Instruments

The monitoring component of the GMCD is a 10-minute interview that asks structured, open-ended questions about each domain. The functional milestones that caregivers reliably described in response to the questions have previously been identified and reported [10, 14]. The GMCD has a table format (Figure 1), with the seven domains arranged in rows and the columns indicating the age intervals. Typically developing children attain all milestones on the interval that correspond to their completed age. We used here the term delay if a child did not attain one or more of the milestones on or before the interval that corresponds to his or her completed age. When a delay is identified, biopsychosocial risk factors are addressed, the delayed domain is supported, monitoring is repeated and community-based early intervention or rehabilitation is added when needed. We have previously reported on the translation, training and reliability of the international GMCD and pointed out that nine milestones were removed from the 125 examined, because they were not internationally reliable [14].

The comprehensive developmental assessment (CDA) that was used as the reference standard for validation of the GMCD was based on the conceptualisation of an ideal detailed developmental assessment [18]. There is no internationally standardised and validated gold standard for developmental assessment, but the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III) [19] is the most widely used tool. A single test score is never considered to provide a gold standard for developmental assessment [18, 19]; therefore, in addition to the Bayley-III, we also included in the CDA a detailed history and observations of the child and caregiver playing. Research in low-income and middle-income countries indicates that Bayley-III scores may be affected by a child's refusal or inability to complete unfamiliar, culturally inappropriate milestones [20]. Therefore, the professionals conducting the CDA provided their clinical judgement of whether or not the child was delayed before the Bayley-III scores were computed. The developmental professionals administering the CDA had backgrounds in developmental paediatrics, occupational therapy, early intervention or psychology and were experienced in assessing young children. They were retrained as a group by experts at the Yale Child Study Centre and achieved at least 90% agreement on 10 consecutive cases with the trainer using clinical judgement and the Bayley-III subscale scores. Ongoing reliability on the CDA was assured at each site.

Procedures

During phase two, trained research assistants from each site administered the open-ended questions in the GMCD, coded the milestones as attained or not attained based on the caregiver responses and obtained the sociodemographic information. A clinician then examined the child and completed the health checklist. Anthropometry was performed using the WHO standards, and haemoglobin levels were determined using the HemoCue (HemoCue AB, Ängelholm, Sweden). On the same day that a research assistant administered and recorded the GMCD, a developmental clinician who was blinded to the GMCD results administered the CDA.

Data analysis

We have previously reported the phase one sample size estimates, methods for generating developmental milestone curves and for depicting ages of attainment for each milestone in detail [14]. This involved using regression models to generate Bayesian point estimates (PE) and surrounding 95% credible intervals for a percentile of age of attainment of each

INTERNATIONAL GUIDE FOR MONITORING CHILD DEVELOPMENT (GMCD)			
DEVELOPMENTAL DOMAINS AND QUESTIONS	**9-11 MONTHS**	**12-14 MONTHS**	**Need for Support**
"Just as it is important to follow Ayesha's physical health and growth, it is important to follow and support her development. Children's brains develop most rapidly during the early years. It is useful to monitor their development and to see if there are any areas that need extra support. You know your child best. Let's talk for 5-10 minutes about her development. By development I mean, learning, communicating, understanding, relating to people, moving body, using hands and fingers, and also hearing and vision. **1. Caregiver's concerns. I'd like to first ask you, do you have any concerns about Ayesha's development in any of these areas?"** ☐ **No** ☐ **Yes** **What are caregiver's concerns?** *Listen to caregiver's concerns and tell caregiver you will come back to the concerns when you have learned about all domains of development.* **"Now I will ask you about how Ayesha is developing in all of these areas. Please give me examples of what she does in her daily life."**			
2. Expressive language. "How does Aisha let you know when she wants something? What kind of sounds, gestures does she use?"	☐ **Repeats syllables ("da-da")**	☐ **Has one meaningful word** ☐ **Uses arm or hand to point to people or objects**	
3. Receptive language. "How does she show you that she understands when you talk to her? For example how does she react when you tell her where is daddy? Where is ball? Come here!"	☐ **Understands names of familiar people (mummy, daddy, sister)**	☐ **Understands verbs/action words (come, take, stop)** ☐ **Understands names of objects (ball, toy)**	
4a. Gross motor (large movements). "Tell me about her movement, like holding and raising her head, sitting, walking."	☐ **Sits without support**	☐ **Pulls to stand holding on to objects** ☐ **Stands alone momentarily** ☐ **Walks holding onto objects**	
4b. Fine motor (fine movements). "How does she use her hands and fingers, like holding objects?"	☐ **Picks up small objects (like pieces of food) using pincer (thumb and index finger) aided by other fingers**	☐ **Picks up small objects using pincer (thumb and index finger) only**	
5. Relating. "How does Ayesha relate to or show interest in people she knows? What does she do to engage them? How is her eye contact? " *Wait for caregiver to respond, then ask:* **"How does she relate to strangers?" "How does she show that she knows they are strangers?"**	☐ **Shows recognition of stranger in some way (may turn away from strangers in anxiety, caution, shyness, fear, or may stare for prolonged time)**	☐ **Spontaneously seeks to share enjoyment and interest with others (cuddles caregiver, kisses, inspects toy together)**	
6. Play activities. "Tell me about Ayesha's play. How does she play with people, with objects or toys?" Ask if needed: **"What playthings/toys does she have, how does she play with them?"**	☐ **Looks for toys/objects that disappear** ☐ **Inspects toys/objects with curiosity, looks at some detail** ☐ **Imitates gestures during play (clapping hands, making face)**	☐ **Initiates game "peek-a-boo"** ☐ **Inspects how toys/objects work (how wheels turn, doll moves, bells ring, lights turn on)**	
7. Self-help activities. "What kinds of things does she do for herself, like eating?"	☼ **Children in this age range may not be expected to attain self-help milestones**	☐ **Uses fingers to feed herself (knows that it is food and feeds herself)**	
8. Nurturing care environment. "Thank you for telling me so much about Ayesha's development, you know her so well. Now please tell me about her daily life. "What do you and your family do at home, in your daily life to help her develop, learn, communicate?" *Listen to what the caregiver is telling you. Prompt by asking: "What do other family members and friends do with her?" Support caregivers by acknowledging and praising all their efforts. Provide ideas from the "GMCD Support Card" or "I Learn with You" if necessary.* **9. Developmental risks. "Sometimes caregivers may have a lot going on. For example, they may feel overwhelmed, stressed or depressed, there may be financial problems or illness in the family, and caregivers may find it hard to support their child's development. Are there such or other difficulties in your family situation?"** *Listen with empathy and identify psychosocial risk factors.* **10. Planning for interventions and follow-up. "What are some ideas or plans you have to support Ayesha's development despite these difficulties? At this early age when development is so important, what could you, your family, friends and community do to help her develop?"** *Support caregivers' efforts. If caregivers do not have ideas or plans, tell them you would like to talk further with them about these. Provide your feedback on the GMCD and plan follow-up together with the caregiver.*			
©Ertem IO, Ankara University. The GMCD is to be used only by providers who have completed the GMCD Provider Training Program given by licensed GMCD Trainers and have obtained a GMCD Provider Certificate. The GMCD cannot be used for research without completing the GMCD Researcher Training Program and obtaining a GMCD Researcher Certificate.			

Figure 1 The developmental domains, questions, examples of age ranges and milestones of the GMCD.

milestone for the total sample, both sexes and each of the countries [21]. During phase one, we identified the ages at which the 85th and 97th percentile of children attained each milestone. These percentiles used in the development of screening instruments approximate one and two standard deviations beyond the mean and were recommended by the Advisory Committee. We have previously reported that the median age of attainment had shown equivalence for 96% and 76% of milestones across sexes and the four countries, respectively [14]. We reconsidered all 116 milestones recognising that the GMCD employs age intervals instead of values. We divided the 0–42 month period into age intervals using both the data generated and expert knowledge of child development, and then fitted milestones into appropriate age intervals so that the lower and upper ages of an interval would approximate the 85th and 97th percentile PEs of the total sample, respectively.

Of the 116 milestones evaluated, 17 were omitted because their 85th percentile point estimates were older than 36 months of age and three because they were represented by other milestones attained at similar ages. We eliminated 11 milestones from the monitoring component of the GMCD and phase two analyses because there were substantial differences between the countries (Table 1): seven in the self-help domain and those relating to listening to brief stories and holding a pencil or stick using fingers. However, these milestones were retained in the supporting early childhood development component. Two gross motor milestones, climbing up and down stairs, were eliminated because the 85th percentile for South Africa was well over 36 months of age, the country differences were large as not all countries had stairs in their houses [14]. We retained five milestones (Table 2) where the 85th percentile point estimate was older than the selected age interval for individual countries, as the difference was ≤1.5 months and this was judged clinically acceptable. The final standardised GMCD had 85 milestones: 22 for expressive language, 13 for receptive language, 14 for gross motor skills, nine for fine motor skills, 11 for relating to others and 16 for play.

Figure 2 provides the example of the age of attainment curve for the milestone imitates gestures during play, including the 85th and 97th percentile point estimates for countries and the total sample. Examples of milestones for each domain and their 85th and 97th percentile point estimates for girls, boys and the countries are provided in Table 2.

Table 1 Milestones that were omitted from the GMCD and the 85th and 97th percentile ages of attainment for girls, boys, countries and total sample

	85th percentile (CrI)							
Milestones that were omitted from the GMCD	Girls	Boys	AR	IN	SA	TR	Total sample	97th percentile (CrI)
Self-help								
Uses fingers to feed herself (knows that it is food and feeds herself)	9.9 (9.5–10.3)	8.8 (8.5–9.0)	8.6 (8.4–8.9)	13.2 (12.4– 14.4)	8.8 (8.5–9.1)	9.0 (8.7–9.4)	10.3 (10.0–10.6)	12.3 (11.7–12.9)
Uses one feeding utensil	18.9 (18.0–19.9)	20.9 (20.0–21.9)	20.1 (18.9–21.4)	19.6 (18.6–20.8)	21.3 (19.3–24.2)	19.5 (18.5–20.6)	20.1 (19.4–20.8)	26.7 (25.3–28.2)
Drinks from cup	18.9 (17.8–20.2)	19.8 (18.7–21.1)	25.4 (23.5–27.9)	18.4 (17.4–19.8)	10.4 (9.8–11.1)	16.9 (16.0–17.9)	19.5 (18.7–20.3)	29.2 (27.3–31.6)
Takes a piece of clothing off	24.7 (23.1–26.8)	28.5 (26.4–31.0)	35.6 (33.1–39.0)	25.6 (24.1–27.5)	14.8 (13.4–16.7)	16.2 (15.2–17.5)	26.8 (25.4–28.4)	43.3 (39.8–47.5)
Washes hands with assistance	27.5 (26.2–29.2)	29.4 (28.0–31.0)	36.0 (33.5–39.0)	23.7 (22.3–25.5)	25.1 (22.9–27.8)	24.8 (23.7–26.1)	28.6 (27.6–29.7)	39.6 (37.4–42.1)
Toilet trained during the day	28.5 (27.6–29.5)	29.4 (28.4–30.3)	39.3 (37.6–41.3)	31.1 (29.4–33.2)	30.7 (28.7–33.2)	35.6 (34.2–37.2)	36.1 (35.1–37.2)	41.1 (39.7–42.7)
Brushes teeth with assistance	32.9 (31.2–34.9)	35.5 (33.7–37.8)	37.9 (35.3–41.5)	28.8 (27.1–31.1)	30.3 (27.8–33.7)	35.0 (32.8–37.8)	34.4 (33.1–35.9)	45.8 (43.2–48.9)
Receptive language								
Listens to brief stories or when caregivers narrate an event	20.2 (19.2–21.2)	20.7 (19.9– 21.7)	20.0 (19.0–21.1)	20.3 (19.2–21.8)	25.2 (23.3–28.0)	18.0 (17.0–19.1)	20.5 (19.9–21.2)	25.7 (24.5–26.9)
Gross motor								
Walks up stairs holding caregivers hand or rail	24.9 (23.6–26.5)	24.8 (23.6–26.2)	27.6 (25.8–30.1)	18.1 (17.2–19.1)	43.3 (37.7–52.0)	18.0 (17.2–18.9)	24.9 (23.9–25.9)	35.2 (33.1–37.8)
Walks down stairs holding caregiver's hand or rail	28.5 (26.9–30.4)	29.6 (28.0–31.5)	32.1 (29.6–35.1)	22.1 (20.9–23.6)	49.5 (42.5–61.8)	21.9 (20.9–23.1)	29.1 (28.0–30.5)	42.4 (39.7–45.8)
Fine motor								
Holds with fingers pencil or stick and scribbles	24.8 (23.4–26.3)	26.5 (25.2–28.0)	35.2 (32.4–38.9)	21.3 (20.2–22.6)	26.0 (23.6–29.3)	18.6 (17.8–19.7)	25.7 (24.8–26.8)	36.2 (34.0–38.8)

AR, Argentina; CrI, Credible Intervals; IN, India; SA, South Africa; TR, Turkey.

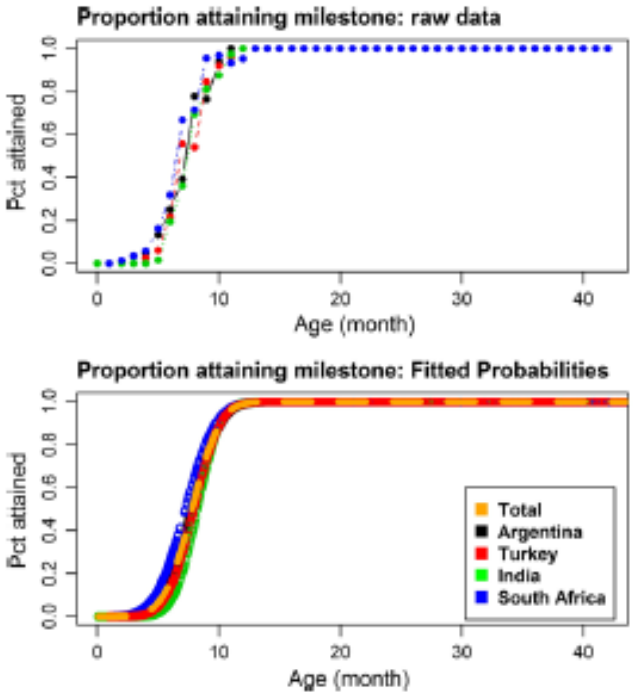
Table 2 Examples of 85th and 97th percentile ages of attainment of GMCD milestones across seven developmental domains for girls, boys, countries and total sample

Examples of GMCD milestones	85th percentile (CrI)							97th percentile (CrI)	Chosen GMCD age interval
	Girls	Boys	AR	IN	SA	TR	Total Sample		
Expressive language									
Uses gestures (shakes head in protest, lifts arms to be picked up)	6.7 (6.4–7.2)	7.4 (7.0–7.8)	7.1 (6.6–7.6)	7.9 (7.4–8.6)	6.7 (6.2–7.3)	6.3 (5.9–6.9)	7.2 (6.9–7.5)	9.6 (9.0–10.2)	7–8
Uses index finger to point	14.0 (13.5–14.7)	14.9 (14.4–15.9)	15.4 (14.5–16.3)	14.4 (13.6–15.3)	13.9 (13.0–15.2)	14.2 (13.6–14.9)	14.6 (14.2–15.0)	17.8 (17.0–18.7)	15–17
Caregivers understand some of child's communication	16.8 (16.1–17.7)	18.2 (17.5–19.1)	16.4 (15.5–17.4)	14.8 (14.0–15.8)	21.3† (19.7–23.9)	18.3 (17.4–19.3)	17.7 (17.2–18.2)	21.9 (21.0–22.9)	18–21
Uses two–word sentences like 'give water' 'mama apple'	25.3 (24.3–26.4)	28.2 (27.1–29.5)	30.5† (28.8–32.5)	23.7 (22.6–25.3)	25.6 (24.0–27.8)	26.0 (24.9–27.4)	27.0 (26.3–27.8)	33.1 (31.8–34.6)	26–29
Uses sentences with four words to communicate	33.4 (32.1–35.0)	35.9 (34.6–37.4)	35.2 (33.6–37.1)	32.5 (31.0–34.3)	31.3 (29.3–33.9)	37.2 (35.5–39.4)	34.9 (34.0–36.0)	41.6 (40.0–43.6)	36–42
Receptive language									
Responds by making sounds when caregivers talk	2.7 (2.4–3.0)	2.6 (2.4– 2.9)	2.4 (2.1–2.7)	2.5 (2.2–2.8)	3.4 (3.0–3.8)	2.4 (2.1–2.6)	2.7 (2.5–2.8)	4.5 (4.1–5.1)	3–4
Understands names of objects	12.5 (11.9–13.0)	12.8 (12.3– 13.2)	12.1 (11.4–12.9)	11.4 (10.9–12.1)	14.6† (13.6–15.7)	12.6 (12.1–13.1)	12.7 (12.3–13.0)	15.0 (14.5–15.5)	12–14
Answers simple questions ('Is mummy home?')	24.7 (23.7–25.8)	26.2 (25.2–27.4)	24.5 (23.3–26.0)	24.4 (23.1–26.1)	27.5 (25.6–30.2)	26.0 (24.8–27.6)	25.5 (24.8–26.3)	30.8 (29.5–32.3)	26–29
Understands prepositions (e.g, 'under' or 'on top')	27.5 (26.2–29.0)	28.4 (27.3–29.7)	28.1 (26.5–30.0)	25.6 (24.2–27.6)	30.9 (28.7–33.6)	27.9 (26.7–29.6)	28.0 (27.2–28.9)	36.1 (34.3–38.0)	30–35
Gross motor									
Sits with support	5.3 (5.1–5.6)	5.2 (5.0–5.5)	5.0 (4.7–5.4)	5.6 (5.3–6.0)	5.0 (4.7–5.3)	5.2 (4.9–5.5)	5.3 (5.1–5.4)	6.3 (6.0–6.5)	5–6
Walks alone	15.2 (14.6–15.9)	14.6 (14.2–15.1)	15.2 (14.5–16.0)	14.2 (13.6–15.1)	14.9 (14.1–16.0)	14.8 (14.2–15.6)	14.8 (14.5–15.2)	17.1 (16.4–17.8)	15–17
Kicks ball or another object	17.2 (16.4–18.0)	16.1 (15.6–16.7)	16.9 (16.1–17.9)	16.1 (15.3–17.0)	17.9 (16.8–19.3)	15.8 (15.1–16.6)	16.6 (16.1–17.1)	19.8 (19.0–20.6)	18–21
Fine motor									
Reaches towards objects or people with hands	4.9 (4.7–5.1)	5.3 (5.1–5.6)	4.5 (4.2–4.8)	5.8 (5.4–6.2)	4.8 (4.5–5.1)	5.1 (4.9–5.5)	5.2 (5.0–5.3)	6.4 (6.1–6.8)	5–6
Picks up small objects using pincer (thumb and index) aided by other fingers	7.9 (7.5–8.4)	8.3 (8.0–8.8)	7.5 (7.0–8.0)	8.7 (8.2–9.3)	8.0 (7.5–8.8)	8.1 (7.6–8.7)	8.2 (7.9–8.5)	10.4 (9.8–11.0)	9–11
Holds pencil or stick in any way and scribbles	15.5 (14.9–16.2)	15.8 (15.3–16.4)	15.3 (14.5–16.1)	16.8 (15.8–18.2)	16.8 (15.8–18.2)	14.9 (14.3–15.6)	15.7 (15.3–16.1)	18.2 (17.5–19.0)	15–17
Relating									
Reacts when caregiver leaves, relaxes when she reunites	7.7 (7.3–8.2)	8.0 (7.6–8.4)	8.7† (8.2–9.4)	7.5 (7.1–8.0)	7.4 (6.8–8.2)	7.3 (6.9–7.9)	7.9 (7.6–8.2)	10.1 (9.7–10.6)	7–8

Table 2 (Continued)

Examples of GMCD milestones	85th percentile (CrI)							97th percentile (CrI)	Chosen GMCD age interval
	Girls	Boys	AR	IN	SA	TR	Total Sample		
Shows recognition of strangers (e.g., turns away, shows caution, shyness or fear)	8.4 (7.9–9.0)	8.6 (8.2–9.1)	9.5 (8.9–10.2)	8.3 (7.8–9.0)	7.0 (6.3–7.8)	8.1 (7.6–8.8)	8.6 (8.2–8.9)	11.2 (10.7–11.7)	9–11
Spontaneously seeks to share enjoyment with others (cuddles or kisses caregivers)	11.5 (10.9–12.2)	11.3 (10.8–11.8)	10.1 (9.5–10.8)	10.4 (9.7–11.3)	14.4[†] (13.1–16.2)	11.1 (10.5–11.8)	11.4 (11.0–11.8)	15.0 (14.2–16.0)	12–14
Initiates specific interactions with people	17.6 (16.6–18.6)	17.7 (16.9–18.6)	13.3 (12.5–14.1)	22.7 (21.1–24.5)	17.8 (16.3–19.7)	16.4 (15.5–17.5)	17.7 (17.0–18.3)	24.6 (23.2–26.1)	18–25
Talks about favourite people/friends when they are not with her ('where is grandpa?')	31.6 (30.2–33.2)	33.9 (32.5–35.3)	36.7 (34.9–39.2)	28.9 (27.4–30.8)	32.5 (30.0–35.7)	31.0 (29.6–32.6)	33.0 (32.0–34.0)	41.5 (39.5–43.7)	36–42
Play									
Initiates game 'peek-a-boo'	10.8 (10.2–11.5)	11.2 (10.7–11.8)	9.9 (9.3–10.7)	11.0 (10.3–11.8)	12.9 (11.8–14.3)	10.7 (10.1–11.3)	11.0 (10.7–11.4)	14.5 (13.7–15.3)	12–14
Has simple imaginary play like feeding someone or doll, driving cars, riding animals	16.9 (16.2–17.8)	18.2 (17.4–19.1)	17.3 (16.3–18.3)	16.9 (16.0–18.0)	18.6 (17.1–20.5)	18.0 (17.0–19.3)	17.7 (17.1–18.3)	23.1 (21.9–24.4)	18–21
Has more complex pretend play involving two or more ideas such cooking a meal and feeding a doll; driving and filling gas, alone or with others	29.9 (28.4–31.7)	35.9 (34.3–37.7)	35.0 (32.3–38.9)	31.6 (29.8–34.0)	35.4 (32.3–40.0)	33.2 (31.5–35.4)	33.8 (32.6–35.2)	44.2 (41.9–47.0)	36–42

AR, Argentina; CrI, Credible Intervals; IN, India; SA, South Africa; TR, Turkey.
[†]Country 85th percentile PE is above the chosen GMCD age interval.



Point estimate (credible interval) in months for age of attainment of milestone

	Attainment Percentile	
	85 th	97 th
Argentina	9.7 (9.2–10.3)	10.8 (10.1–11.7)
India	9.8 (9.3–10.4)	10.7 (10.1–11.5)
South Africa	9.2 (8.7–9.9)	10.4 (9.7–11.2)
Turkey	9.7 (9.2–10.2)	10.8 (10.2–11.4)
Total	9.7 (9.4–10.0)	10.8 (10.5–11.2)

Figure 2. Distribution curves and percentile point estimates of ages of attainment of the milestone imitate gestures during play, such as clapping hands and making faces.

In phase two, we used the previously reported prevalence rates for developmental delay [22, 23], to conduct a power analysis for the diagnostic accuracy of the GMCD. We determined that sufficient power (80%) would be achieved with 150 children per age group–6–17, 18–29 and 30–42 months–if the prevalence of delay was 15% and 450 children if the prevalence was 5% (Figure 3, Graph A). We varied the magnitude of sensitivity (0.80, 0.75, 0.70) to examine the width of precision around sensitivity of 0.80 as measured by the 95% confidence interval (95% CI) for various sample sizes (Figure 3, Graph B). Sufficient power and precision were projected for phase two with the sample size of at least 450 children in each age group. Delayed development on the GMCD was prespecified as not attaining one or more milestones in any of the domains within the child's age interval. The children's ages were rounded to completed months and their corrected ages were used for prematurity. The Bayley-III had not been standardised in the four countries at the time of the study and controversy existed about which Bayley-III cut-off should be used to categorise developmental delay [24]. We therefore examined sensitivity and specificity for three different Bayley-III cut-off points. Subscale scores of 3 and 4 indicate less than and equal to −2 standard deviations, respectively, and subscale score of ≤5 indicates borderline delay [19]. A delay on the CDA was prespecified as a score ≤3, ≤4 or ≤5 on one or more of the Bayley-III expressive, receptive language, fine, or gross motor subscales and a delay in the clinical judgement. Sensitivity and specificity values, as well as 95% CIs, were generated for the total sample and three age ranges. We chose 6–17, 18–29 and 30–42 months because the recommended age for developmental surveillance begins at six months and includes screening at nine, 18 and 30 months [4]. Analyses were conducted with SAS version 9.4.2 (SAS Institute, North Carolina, USA) and Markov Chain Monte Carlo (MCMCpack), and the Bayesian Estimate Supersedes the t-Test (BEST) packages in R statistical software, version 3.3.1 (R Foundation, Vienna, Austria).

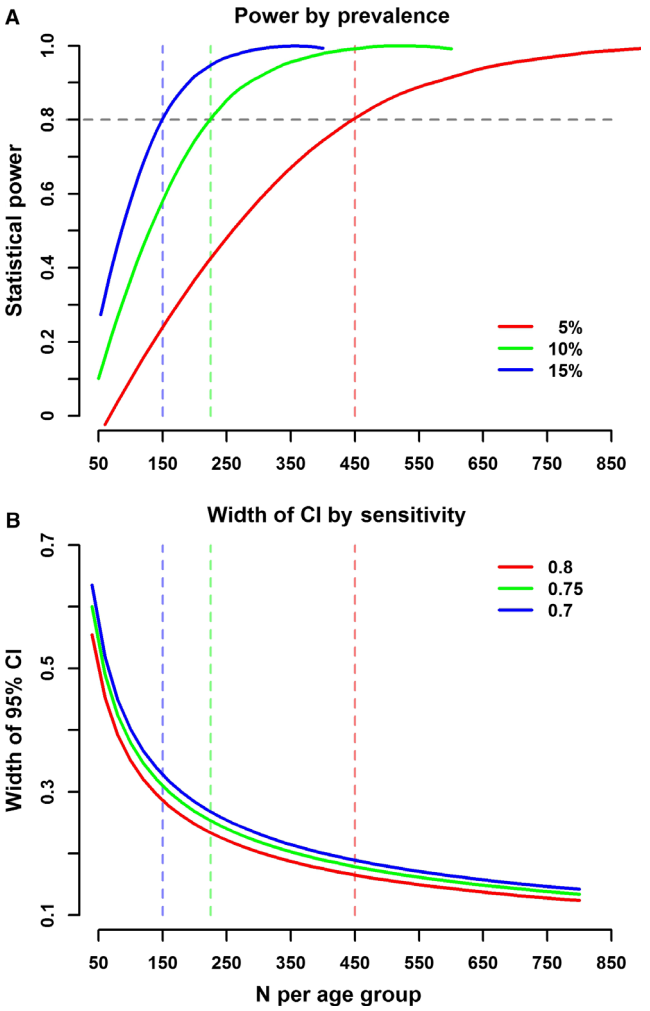


Figure 3 Graph A shows the power achieved at different prevalence levels of developmental delay (*y*-axis) by sample size (*x*-axis) to detect the minimum sensitivity of 0.80. The line graphs in red, green and blue show the power for 5%, 10% and 15% prevalence in developmental delay, respectively. The sample size of 450 per developmental age group achieved the power of 0.80, 0.98 and 1.0 for the respective prevalence. Graph B shows the width of the 95% confidence interval (95% CI) precision around sensitivity (*y*-axis) by sample size (*x*-axis). The lines in red, green and blue show the magnitude of the width in the 95% CI around sensitivity of 0.80, 0.75 and 0.70, respectively. The sample size of 450 in each developmental age group corresponds to the precision width ranging from 0.18 to 0.22 around the respective sensitivity levels.

Results

This was a two-phase back-to-back study and the second phase, which is the main subject of this paper, focused on assessing the sensitivity and specificity of the GMCD. The GMCD was constructed during phase one using data on the 4949 children (59.0% male) who met the health criteria. These represented 48.3% of the 10 246 that were originally recruited from Turkey, Argentina, India and South Africa. Further details of phase one of this study have previously been published [14].

Of the 1739 children (52% male) that were specifically enrolled for phase two of the study, eight had missing data and were excluded, leaving a final sample size of 1731, with 593 from Turkey, 467 from India, 361 from South Africa and 310 from Argentina.

The sociodemographic characteristics of the phase two sample are summarised in Table 3. The median age of the sample was 21 months (interquartile range 12-31 months). Sample sizes for the age ranges, the percentage of children identified as delayed using the GMCD and CDA, and the sensitivity and specificity results are shown in Table 4. The GMCD identified that 30.4% of children in the phase two sample were delayed. Using the Bayley-III scale scores of ≤3, ≤4 and ≤5, the CDA categorised 4.0%, 6.6% and 11.0% of children as delayed, respectively. The agreement between the clinicians' clinical judgment and the Bayley-III scaled scores of ≤3, ≤4 and ≤5 were 83.5%, 85.6% and 87.3%, respectively.

When we combined the Bayley-III cut-off of ≤4 and the clinical judgment used in the CDA, the overall sensitivity and specificity for the total sample were 0.79 and 0.73, respectively. The sensitivity increased to 0.87 when the more stringent Bayley-III cut-off of ≤3 was used for the CDA and it decreased to 0.72 when the cut-off of ≤5 was used to include more borderline cases. Specificity levels varied from 0.72 to 0.75. There were differences in sensitivity and specificity for the three age ranges. For the middle age group (18-29 months), both the sensitivity and specificity were above 0.80, except for the cut-off of ≤5 (specificity 0.74). For those over 30 months, the sensitivity for identifying delayed children was very high at all cut-off levels (1.00–0.93), while the specificity was lower, between 0.69 and 0.71. For the youngest group (six to 17 months), the sensitivity was high (0.80) when the more stringent ≤3 cut-off was used but decreased to 0.71 and 0.64 when ≤4 and ≤5 were used, respectively. In this younger age group, the specificity of the GMCD was between 0.68 and 0.70 using the three different Bayley-III cut-offs.

Discussion

The internationally constructed Guide for Monitoring Child Development featured in this study aims to monitor and support children's development and identify developmental difficulties. The previously reported first phase of this study [14] standardised the monitoring component of the GMCD, and the current study reported here examined its sensitivity and specificity in four countries with different cultural and linguistic characteristics. Our comparisons of the median ages of attainment in phase one provided indications of similarities in early child development across cultures and those findings have been previously been published [14]. However, it was not evident before the current study, whether the instrument developed could take into account similarities and differences between countries and sexes in identifying when a child was delayed.

Table 3 Sociodemographic characteristics of Phase 2

	Phase 2 sample (N = 1731)				
	Total n (%)	Argentina n (%)	India n (%)	South Africa n (%)	Turkey n (%)
Total	1,731 (100.0)	310 (18.0)	467 (27.0)	361 (21.0)	593 (34.0)
Sex[†]					
Female	825/1,726 (47.8)	144 (46.5)	210 (45.0)	182/356 (51.1)	289 (48.7)
Male	901/1,726 (52.2)	166 (53.5)	257 (55.0)	174/356 (48.9)	304 (51.3)
Child's age (months)					
6–12	487 (28.1)	124 (40.0)	83 (17.8)	97 (26.9)	183 (30.9)
13–24	543 (31.4)	111 (35.8)	144 (30.8)	107 (29.6)	181 (30.5)
25–42	701 (40.5)	75 (24.2)	240 (51.4)	157 (43.5)	229 (38.6)
Mother's age (years)[‡]					
≤19	45/1,709 (2.6)	35/307 (11.4)	0 (0.0)	10/342 (2.9)	0 (0.0)
20–34	1,287/1,709 (75.3)	215/307 (70.0)	378 (80.9)	263/342 (76.9)	431 (72.7)
≥35	377/1,709 (22.1)	57/307 (18.6)	89 (19.1)	69/342 (20.2)	162 (27.3)
Mother's education (years)					
<12	550 (31.8)	184 (59.4)	59 (12.6)	138 (38.2)	169 (28.5)
≥12	1,181 (68.2)	126 (40.6)	408 (87.4)	223 (61.8)	424 (71.5)

Data are n (%) or n/N (%) when N is different from the total N given at the top of the column.
[†]Data for 5 children missing.
[‡]Data for 22 children missing.

Table 4 Sensitivity and specificity of the GMCD compared to the CDA

			Bayley-III ≤3			Bayley-III ≤4			Bayley-III ≤5		
	N	GMCD +%	CDA+%	Sensitivity (95%CI)	Specificity (95%CI)	CDA+%	Sensitivity (95%CI)	Specificity (95%CI)	CDA+%	Sensitivity (95%CI)	Specificity (95%CI)
Total sample	1,731	30.4	4.0	0.87 (0.77–0.94)	0.72 (0.70–0.74)	6.6	0.79 (0.70–0.86)	0.73 (0.71–0.75)	11.0	0.72 (0.65–0.78)	0.75 (0.73–0.77)
6–17 months	719	33.9	4.2	0.80 (0.61–0.92)	0.68 (0.64–072)	7.2	0.71 (0.57–0.83)	0.69 (0.65–0.72)	12.0	0.64 (0.52–0.74)	0.70 (0.66–0.74)
18–29 months	540	23.3	5.2	0.89 (0.72–0.98)	0.80 (0.77–0.84)	8.3	0.82 (0.68–0.92)	0.82 (0.78–0.85)	14.0	0.74 (0.63–0.84)	0.85 (0.82–0.88)
30–42 months	472	33.1	2.5	1.00 (0.74–1.0)	0.69 (0.64–0.73)	3.6	0.94 (0.71–1.0)	0.69 (0.65–0.73)	5.7	0.93 (0.76–0.99)	0.71 (0.66–0.75)

GMCD+: Proportion of children that had not attained one or more milestones on the GMCD age interval.

CDA+: Proportion of children that had a Bayley-III scaled score at specified level in any subscale and a clinical decision of developmental delay.

The main strengths of phase one [14], which set the scene for this study, included the fact that the sample size was one of the largest used in instrument development studies in children [25, 26] and that stringent health criteria were applied. The cross-sectional design minimised potential for the reporting and selection biases that may be present in longitudinal designs with repeated questioning and recruitment or retention of more compliant families. The limitations of the standardisation sample have been reported in detail [14] and included our inability to involve lower income countries and rural settings, smaller numbers of older children due to exclusions for health conditions and fewer healthcare visits and our inability to exclude conditions such as iodine deficiency or iron deficiency without anaemia. We did not exclude children with psychosocial risk factors from phrase one, because factors such as poverty, low caregiver education, caregiver depression and deficiencies in the nurturing care environment are extremely prevalent in low-income and middle-income countries [1, 15]. Future studies are needed to understand the effect of psychosocial risks and protective factors on the GMCD results.

The phase two results indicate the notable accuracy of the GMCD in identifying developmental delay, when compared to a comprehensive developmental assessment, despite its brevity and simplicity. They also highlight its potential for worldwide use and the fact that it is comparable to existing screening instruments [27-29]. The accuracy was satisfactory for children with borderline delays, high for significant delays and best for children aged 18-29 months, which is a critical period for the early identification of developmental disorders. Its sensitivity was high in children over 29 months, even for those with borderline delay. The lower specificity for borderline delay in younger ages seen in our study has previously been reported and may be due to rapid development [29]. Concerns raised by the lower specificity of developmental screening have been refuted, as these children have higher rates of developmental needs, even without a diagnosis [30]. Specificity would have invariably increased if an assessment had been repeated. We recommend that when a delay is identified, biopsychosocial risk factors should be addressed, support is provided for nurturing care and the GMCD is repeated again a month later.

The phase two strengths included the sample size, which allowed us to compute total sample and age interval accuracy, and the fact that this study was one of the largest of studies to examine accuracy [25, 26, 28, 29]. Our use of a CDA and different Bayley-III cut-off scores addressed the recognised problem of only using test scores and single cut-offs [24, 27, 29]. Recruitment of children from clinics that were similar to those in which the GMCD will be used enhanced its generalisability. The prevalence of developmental delay would have been higher if we had used a sample enriched with developmentally delayed children, enabling smaller sample sizes. However, as Limbos et al. have pointed out, the accuracy of the results from such samples may not be generalisable to real-life practice [28]. The main limitation of phase two was our inability to examine accuracy separately in each country sample due to inadequate sample sizes. The difficulties involved in participating in lengthy developmental evaluations hindered recruitment and this was particularly noticeable in the smaller numbers of older children, who attend fewer healthcare appointments. The sample size estimates were based on a minimum expected prevalence of 5% with developmental disorders [22], but the prevalence was lower among children over 29 months and this resulted in the need for a larger sample size. This low prevalence of developmental delay must be interpreted with caution, as there could have been recruitment bias, due to caregivers of children with chronic illnesses or developmental delays being less likely to agree to a lengthy comprehensive developmental evaluation.

Conclusion

The international standardisation of the Guide for Monitoring Child Development, and this multicountry study, which has established its accuracy in identifying children with developmental difficulties, means that it can be used in other countries. The GMCD offers an alternative to investing energy, funds and time in restandardising and revalidating tools in every country. Its use goes beyond merely identifying developmental delays and referring children to specialist services. The uniqueness of the GMCD and its novel approach to the field is that it has been specifically constructed to encompass bioecological, strengths-based and family-centred theories. It provides open-ended questions, a conversational style that can be used to build rapport with caregivers and obtain information about a child's development and functional milestones. These elements should enable clinicians to seamlessly provide caregivers with specific information to promote their child's development. The GMCD builds capacity in frontline service providers by giving them a comprehensive, integrated approach to promoting nurturing care, monitoring, early identification, early intervention, links to other services and follow-up. We hope that this approach will lead to conceptual advances in the field and provide greater potential to address child development in low-income and middle-income countries than approaches that rely on screening with checklists and referral to services, which are frequently scarce. Research is needed to determine how the GMCD can be implemented within healthcare systems. We also need to determine whether its implementation can contribute to narrowing the gap between low- and middle-income countries and high-income countries in addressing the developmental needs of individual children.

Acknowledgements

We would like to thank a number of people. John M Leventhal and Eugene Shapiro for their invaluable comments on the manuscript. Members of the Guide for Monitoring Child Development International Advisory Committee who provided feedback on the study concept: Joseph Avni–Singer, Elaine Borghi, Nancy Close, Bernadette Daelmans, Mercedes de Onis, James D Dziura, Tarun Dua, Deepa Grover, Frances Page Glascoe, Magdelana Janus, Betsy Lozoff, Linda Mayes, Chiara Servili, Ruth Stein, Carol C Weitzman and Barry Zuckerman. Members of the International Guide for Monitoring Child Development Research Group for data collection and management: Augusto Eezckui, Daniel Giordano, Marcia María de Luján Guerrero, Virginia Lencinas, Daniela Elvira Roldan, and Erika Sacco (Argentina); Leera J Lobo, Riddhi M Mehta, Anushree A Sane, and Sonam P Shah (India); Mamokage M L Nkadimeng, Jennifer Makin, Maria M S Sono, and Mary–Jane M M Ntima (South Africa); Sebnem Z K Coban (Turkey) and Katina Gionteris (USA).

Funding

This study was funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (award number 1R01HD057834). The study sponsors played no role in any aspect of the study or paper.

Conflicts of Interest

The authors have no conflicts of interest to declare.

References

[1.] Richter LM, Daelmans B, Lombardi J, Heymann J, Boo FL, Behrman JR, et al. Paper 3 Working Group and the Lancet Early Childhood Development Series Steering Committee. Investing in the foundation of sustainable development: pathways to scale up for early childhood development. *Lancet* 2017; 389: 103–118.

[2.] Vaivada T, Gaffey MF, Bhutta ZA. Promoting early child development with interventions in health and nutrition: A systematic review. *Pediatrics* 2017; 140: e20164308.

[3.] Ertem IO. *Developmental difficulties in early childhood: prevention, early identification, assessment and intervention in low and middle–income countries: a review*. Geneva, Switzerland: World Health Organization, 2012.

[4.] Committee on Practice and Ambulatory Medicine and Bright Futures Periodicity Schedule Workgroup. 2016 recommendations for preventive pediatric health care. *Pediatrics* 2016; 137: 1–3.

[5.] Jindal P,Mac Dermid JC, Rosenbaum P, Di Rezze B, Narayan A. Perspectives on rehabilitation of children with cerebral palsy: Exploring a cross–cultural view of parents from India and Canada using the international classification of functioning, disability and health. *Disabil Rehabil* 2017; 00: 1–11.

[6.] Aly Z, Taj F, Ibrahim S. Missed opportunities in surveillance and screening systems to detect developmental delay: a developing country perspective. *Brain Dev* 2010; 32: 90–97.

[7.] Goldfeld S, Yousafzai A. Monitoring tools for child development: an opportunity for action. *Lancet Glob Health* 2018; 6: e232–233.

[8.] Fischer VJ, Morris J, Martines J. Developmental screening tools: feasibility of use at primary healthcare level in low– and middle–income settings. *J Health Popul Nutr* 2014; 32: 314–326.

[9.] King TM, Tandon SD, Macias MM, Healy JA, Duncan PM, Swigonski NL, et al. Implementing developmental screening and referrals: lessons learned from a national project. *Pediatrics* 2010; 125: 350–360.

[10.] Ertem IO, Dogan DG, Gok CG, Kizilates S, Caliskan A, Atay G, et al. A guide for monitoring child development in low– and middle–income countries. *Pediatrics* 2008; 121: e581–589.

[11.] Ertem IO, Atay G, Bingoler BE, Dogan DG, Bayhan A, Sarica D. Promoting child development at sick–child visits: a controlled trial. *Pediatrics* 2006; 118: e124–131.

[12.] Ertem IO, Pekcici EB, Gok CG, Ozbas S, Ozcebe H, Beyazova U. Addressing early childhood development in primary health care: experience from a middle–income country. *J Dev Behav Pediatr* 2009; 30: 319–326.

[13.] Ertem IO. The international Guide for Monitoring Child Development: enabling individualized interventions. *Early Childhood Matters* 2017; 126: 83–88.

[14.] Ertem IO, Krishnamurthy V, Mulaudzi M, Sguassero Y, Balta H, Gulumser O, et al. The Development of healthy children in the first three years: similarities and differences across genders and countries: a cross–sectional observational study. *Lancet Glob Health* 2018; 6: e279–291.

[15.] World Health Organization, United Nations Children's Fund, World Bank Group. *Nurturing care for early childhood development: a framework for helping children survive and thrive to transform health and human potential.* Geneva, Switzerland: World Health Organization; 2018.

[16.] WHO Multicentre Growth Reference Study Group. Enrolment and baseline characteristics in the WHO Multicentre Growth Reference Study. *Acta Paediatr* 2006; 95 (Suppl 450): 7–15.

[17.] WHO Multicentre Growth Reference Study Group. WHO Motor Development Study: windows of achievement for six gross motor development milestones. *Acta Paediatr* 2006; 95 (Suppl 450): 86–95.

[18.] Meisels SJ, Fenichel ES (Ed). *New visions for the developmental assessment of infants and young children*. Washington, D.C.: Zero to Three/National Center for Infants, Toddlers, and Families Publications, 1996.

[19.] Bayley N. *Bayley Scales of Infant and Toddler Development*, 3rd ed. San Antonio, TX: Psychological Corp, 2006.

[20.] Hanlon C, Medhin G, Worku B, Tomlinson M, Alem A, Dewey M, et al. Adapting the Bayley Scales of infant and toddler development in Ethiopia: evaluation of reliability and validity. *Child Care Health Dev* 2016; 42: 699–708.

[21]. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian Data Analysis*, 3rd Edition. Chapman and Hall/CRC, 2013.

[22.] WHO and World Bank. *World report on disability*. Available at http://www.who.int/disabilities/world_report/2011/report.pdf)
[23]. Boyle CA, Boulet S, Schieve L, Cohen RA, Blumberg SJ, Yeargin–Allsopp M, et al. Trends in the prevalence of developmental disabilities in US children, 1997–2008. *Pediatrics* 2011; 27: 1034–1042.
[24]. Anderson PJ, Burnett A. Assessing developmental delay in early childhood–concerns with the Bayley–III scales. *Clin Neuropsychol* 2017; 31: 371–381.
[25.] Gladstone M, Lancaster GA, Umar E, Nyirenda M, Kayira E, van der Broek NR, 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.] McCoy DC, Sudfeld CR, Bellinger DC, Muhihi A, Ashery G, Weary TE, et al. Development and validation of an early childhood development scale for use in low–resourced settings. *Popul Health Metr* 2017; 15: 3.
[27.] Marks K, Glascoe FP, Aylward GP, Shevell MI, Lipkin PH, Squires JK. The thorny nature of predictive validity studies on screening tests for developmental behavioral problems. *Pediatrics* 2008; 122: 866–868.
[28.] Limbos MM, Joyce DP. Comparison of the ASQ and PEDS in screening for developmental delay in children presenting for primary care. *J Dev Behav Pediatr* 2011; 32: 499–511.
29. Steenis LJ, Verhoeven M, Hessen DJ, van Baar AL. Parental and professional assessment of early child development: the ASQ–3 and the Bayley–III–NL. *Early Hum Dev* 2015; 91: 217–225.
30. Glascoe FP. Are Overreferrals on developmental screening tests really a problem? *Arch Pediatr Adolesc Med* 2001; 155: 54–59.