



Centre for Disability Research

The Identification of Children with or at
Significant Risk of Intellectual Disabilities in
Low and Middle Income Countries:
A Review

Janet Robertson, Chris Hatton & Eric Emerson

CeDR Research Report 2009:3
July 2009

Contents

Introduction	1
Screening Test Standards	1
Method	3
Results	5
Identifying ID in LAMI Countries	5
Abbreviated Developmental Scale 1 (EAD-1)	6
ACCESS Portfolio	6
Baroda Development Screening Test for Infants	7
Developmental Assessment Tool for Anganwadis (DATA)	7
Developmental Milestone Chart (DMC)	8
Developmental Observation Card	8
Developmental Screening Inventory (DSI)	8
Disability Screening Schedule (DSS)	9
Guide for Monitoring Child Development (GMCD)	9
Indonesian Adaptation of the Vineland Adaptive Behavior Scales (IVABS)	10
Infant Neurological International Battery Test (INFANIB)	10
Malawian Developmental Screening Tool	11
Monitoring Child Development at Family & Community Level	11
National Institute for the Mentally Handicapped Developmental Screening Schedule (NIMH-DSS)	12
Parents Evaluation of Developmental Status (PEDS; Indian Validation)	12
Psychosocial Developmental Screening Test	13
The Ten Questions Screen	13
Trivandrum Developmental Screening Chart (TDSC)	18
Other Studies on the Use of Screens	18
Alternatives to Screening Tests	19
Managing Identified Disabilities	20
Discussion	22
Issues in Developmental Testing in LAMI Countries	22
Conclusion	25
References	26
Appendix One: Search Strategy	29
Appendix Two: Table of Studies Included in Review	31
Appendix Three: Summary Characteristics of Screening Tests	45

Acknowledgements: The research described in this report was funded by the World Health Organisation. However, the views expressed are the sole responsibility of the authors.

Introduction

In high income countries an important strategy for the early detection and management of intellectual disabilities (ID) has been the integration into health, education and social care systems of developmental monitoring of children (i.e. standardized screening and surveillance) (Ertem et al., 2008). Surveillance of the development of infants and pre-school children can enable the identification of children who have or at risk of developmental disabilities so that early intervention can be targeted to these children (Sonnander, 2000).

The aim of early intervention is to intervene early in children's lives to promote early child development and consequently avoid or reduce future risks to the child's health, well-being and social inclusion. Intervention programmes can either focus on a high-risk target groups of children or be universally available (Offord, 2000). Universal programmes (e.g., iodine supplementation) are appropriate if defining high risk groups is problematic and the cost of delivery is low. However, if the cost of intervention is high (e.g., sustained intensive child/family training) the costs of universal programmes may be prohibitive. They may also be inequitable, having their largest impact on those at least risk. In contrast, targeted interventions direct the available resources towards the children who are in greatest need or at greatest risk and may, therefore, provide a more efficient way of allocating available resources. Targeted interventions are designed to identify all high-risk children rather than just those for whom care is sought by parents or carers. A targeted intervention requires, however, that it is possible to accurately identify high risk groups of children (Bennett, 1998).

Methods for developmental monitoring of young children by health care providers in low and middle income (LAMI) countries are lacking (Ertem et al., 2008). This review summarises the literature on identifying children with, or at significant risk of, ID in LAMI countries.

Specifically, the review addresses the question: *What approaches are available to enable primary or secondary health care functionaries to identify children with or at significant risk of intellectual disabilities, including intellectual disabilities due to specific causes and manage or refer them when required?*

The review does not cover the issue of neonatal screening as this topic is being dealt with by WHO's existing work on prevention. A brief introduction to the issue of developmental screening tests is first given to outline some of the required characteristics of screening tests.

Screening Test Standards

Screening tests are used for many conditions to identify children who might benefit from early intervention. Ideally, screening involves a relatively brief procedure whose results can be used to select for further investigation those children who are at serious risk of developing the relevant condition. Such screening tests are typically

administered by paraprofessionals. The standards for screening test construction (from Glascoe 2007) are noted below.

Standardization: this should include a large nationally representative sample (Glascoe 2007). The WHO recommends that in populations with a high prevalence of conditions that are hazardous to child health and development (such as malnutrition or iron deficiency anemia), the reference group should be based on a “prescriptive” sample of healthy children without these risks rather than geographic, whole population references (de Onis, 2006, Ertem et al., 2008).

Reliability: information should be included on internal consistency, interrater reliability, and test-retest reliability.

Validity: this should include concurrent validity (comparison of screening measures to diagnostic tests). Criterion-related validity is the “acid test” of screening instrumentation and comprises of the following.

- *Sensitivity:* the percentage of children with disabilities who are correctly identified by the screening test (i.e. screen positive). Ideally, 70-80% of those with difficulties should be identified (Glascoe, 2007).
- *Specificity:* the percentage of children without disabilities who are correctly identified by the screen (i.e. screen negative). Specificity close to 80% or higher is desirable (Glascoe, 2007).
- *Positive predictive value (PPV):* the percentage of children who are identified as having a disability by the screening test who do indeed have a disability. Values ranging from 30 to 50% are common in developmental screening (Glascoe, 2007). However, often children overreferred on screening tests (i.e. who do not have a disability) do nonetheless have other problems which would benefit from intervention (Glascoe, 2007).
- *Negative predictive value (NPV):* the percentage of children who are identified as normally developing by the screening test who are indeed developing normally. This is less commonly presented in studies of tool validation.

Method

Searches of electronic literature databases were conducted to identify peer reviewed articles published from 1990 onwards in the English language, which included information from relevant studies in low and middle income countries. Eligibility for inclusion of countries was determined by reference to the World Bank List of Economies^a. In this, economies are divided among income groups according to 2007 gross national income (GNI) per capita, calculated using the World Bank Atlas method. The groups are: low income, \$935 or less; lower middle income, \$936–3,705; upper middle income, \$3,706–11,455; and high income, \$11,456 or more.

The databases searched were:

- Medline
- Cinahl
- Web of Science
- PsycINFO

In each database, terms for intellectual disabilities and associated synonyms were identified. The most appropriate method for searching for research from LAMI countries was also identified. These two searches were then combined to get a pool of literature dealing with ID and LAMI countries. Within this pool specific search terms relating to the review in hand were then introduced (e.g. screening; identification). Full details of the search strategies and terms employed can be found in Appendix One. All articles identified by searches were assessed for their relevance to the review objectives firstly by reading abstracts. If abstracts were unavailable, or did not provide enough detail to assess the relevance of the article, the full text of the article was obtained and relevance assessed from this. Data was extracted from the full text of articles identified as being relevant to the review.

In addition, a request for information on research relevant to all three rapid reviews carried out in this series was sent by email in April 2009 to the membership of the International Association for the Scientific Study of Intellectual Disabilities (IASSID). This enabled the identification of research literature not identified in the electronic searches, for example relevant articles which were “in press”.

A request for information was also sent to relevant organisations in LAMI countries targeted in the WHO Mental Health Gap Action Programme (mhGAP). These countries are those highlighted in bold in Appendix Two of the mhGAP document (World Health Organization, 2008). Informants were asked for information on current policy and practice in their country; descriptions of relevant services in their country; and reports that have evaluated the impact of relevant services in their country. Relevant contacts in each targeted country were identified from the Atlas-ID Compendium of Sources Used (Gabrielle Major et al., 2008). Initially, emails were

^a siteresources.worldbank.org/DATASTATISTICS/Resources/CLASS.XLS

sent to contacts but in instances where messages were returned as email addresses no longer existed, letters were sent by surface mail to the address listed in the Atlas-ID Compendium. Contacts were invited to reply either by email or by surface mail.

Results

A summary of the studies included in the results of this review are included in Appendix Two. A total of 38 articles form the basis of this review. The majority of the articles (21) were concerned with the validation of a particular screening test for use in one or more LAMI countries to identify children with disabilities. A further 8 articles reported primarily prevalence studies which employed a screening test to identify children with disabilities. Three studies looked at alternatives to the use of screening tests to identify children with disabilities (Gona, Hartley & Newton, 2006; Kuruvilla & Joseph, 1999; Thorburn, Desai & Durkin, 1991). One article described a screening test (Phatak & Khurana 1991); one talked about screening in China (Ericsson, Gebre-Medhin & Sonnander 2008); one was based on clinical trials involving use of a milestone chart (Scherzer, in press); one was based on field testing a portfolio of assessment and advice materials (Wirz, Edwards, Flower et al., 2005); one evaluated inservice training of AWWs to detect disabilities (Mathur et al., 1995); and one was a review article (Nair & Radhakrishnan, 2004).

In the following sections we describe the various screening tests that have been identified in these articles and give, where possible, data on their reliability and validity. We also present information from the three studies which have looked at alternatives to using screening tests in the identification of children with disabilities in LAMI countries. Finally, we discuss the issue of what happens when children are identified as having a disability, including ID, in terms of referral and management.

Identifying ID in LAMI Countries

The issue of identification and assessment of disabilities has been the subject of a comparatively large body of research in LAMI countries. In a systematic review of childhood disability research on screening, prevention, services, legislation and epidemiology in LAMI countries (including research published prior to 1990), of 80 articles identified approximately two thirds discussed screening tools or assessment methods (Maulik and Darmstadt, 2007). It is further noted that many of these were related to assessing cognitive dysfunction and ID.

The predominant study design is a two phase study where a simple screening tool is used in the community by field workers on large samples of children and then children who screen positive (and usually a proportion of those who screen negative) are followed up for specialised professional assessments of disabilities. These studies have been done for two main reasons: to validate screening tools by comparing positive screens (and a sample of negative screens) to the results of comprehensive assessments of disabilities; and to determine the prevalence of disabilities within a particular population. In this review we consider both studies concerned with tool validation and studies concerned with prevalence, in the latter case to identify the screening tools used to identify ID in prevalence studies. Each of the screening tools identified is described in turn and summary characteristics of the screening tools in terms of sensitivity, specificity, PPV, NPV and reliability are presented in Table One.

Abbreviated Developmental Scale 1 (EAD-1)

A two stage cross-sectional study was conducted in Bogota, Colombia to look at the prevalence of delayed neurodevelopment in pre-school children (van Meerbeke et al., 2007). In phase one, parents were interviewed and teachers wrote lists of children with possible developmental delay. In phase two the Abbreviated Developmental Scale 1 (EAD-1) battery was used (Ortiz, 1991). This was designed and validated in Colombia by Nelson Ortiz in 1991 (cited in van Meerbeke et al 2007) and has frequently been used as the primary tool for the evaluation of development in Colombian children (van Meerbeke et al., 2007). However, the validation report for this scale is in Spanish so it has not been possible to extract data on the reliability and validity of the scale. Using EAD-1 van Meerbeke et al (2007) identified neurodevelopmental delay among apparently healthy children from nurseries and kindergartens who had previously been undiagnosed and untreated. They note that the lack of evaluation of developmental milestones in children in Colombia is a substantial public health problem that will require effective intervention.

ACCESS Portfolio

The ACCESS Portfolio has two sections: identification and advice. The portfolio was field tested in Uganda and Sri Lanka (Wirz et al., 2005). The aims were to determine whether: community health workers (CHWs) could learn to use the materials in the portfolio easily; materials identified young children with disabilities; parents were satisfied with the process and advice given; healthcare workers could use the portfolio in addition to other health care duties.

In Sri Lanka 12 public health midwives (PHMs) were selected to use the portfolio for a trial period of 6 months. In Uganda, 11 primary health care nurses from 7 health centres and 2 hospitals assisted with field testing. 10% of children seen by the health workers were seen by a medical physician (Sri Lanka) and physiotherapist and occupational therapist (Uganda) and their assessments compared with screen results from the CHWs. CHWs collected data over 6 months on children age 3 or under 'who were causing their mothers concern'. Children were screened, advice given where appropriate, and referrals made to local doctor or paediatrician or local hospital where necessary. Health workers also completed questionnaires about ease of use of the portfolio and took part in focus group discussions about use of the materials and training. Parents also took part in focus groups.

Overall, 769 children were screened in Uganda and 580 in Sri Lanka. In Uganda, 44% of children seen failed the screen and in Sri Lanka 11% failed and were deemed to have a disability by the health workers. Developmental delay and difficulties with movement and self-care were the commonest forms of disability identified. Compared to expert assessments, accuracy was 82% in Uganda and 76% in Sri Lanka. However, this is based on the total number of screens found to be correct compared to expert assessment and figures are not presented for sensitivity and specificity. There were many positive comments made about how the project helped both children with disabilities and their parents. CHWs were generally very positive about

the training and use of the portfolio. There was consensus that the project had raised awareness of disability within the community and workers were pleased to have been involved in this. Parents found the advice and materials helpful.

Baroda Development Screening Test for Infants

In 1983 a UNICEF aided programme for the prevention, early detection and intervention of childhood disability in urban slums was launched in Baroda, India (Phatak and Khurana, 1991). For early detection by house to house survey by CHWs, a simple, quick, cheap and precise test was required. They developed a screening test from the Bayley Scales of Infant Development (BSID) as it has been standardized on Baroda infants (Baroda norms). Only those items which were simple and easy to administer and assess and not requiring any special training, experience or equipment were selected. A total of 54 items were selected. A child who fails items in his/her chronological age group is screened out for detailed study.

The screening test was put to use in a field survey as well as in clinical practice (especially well baby clinics). It is noted that it had been used for more than 3 years by CHWs of Baroda. They have found that 5 or 6 one hour sessions are sufficient for training on screening. Information on sensitivity and specificity is reported to be 65%-95% in this paper but this is based on a personal communication. They note that although the BSID (Baroda Norms) is regularly used at 6-7 research centres in India, the DDST appears to be the better known amongst paediatricians. Both tests have been developed for non-professionals. They conclude that the Baroda Development Screening Test could have a wide application in field surveys and clinical practice in picking up children for more specific evaluation by skilled professionals.

Developmental Assessment Tool for Anganwadis (DATA)

A logical yet neglected population for early identification of developmental delay is children attending anganwadis, the largest Integrated Childhood Development Services (ICDS) in the world delivering health care to 98 million out of the 160 million children aged 2-6 in India (Nair et al., 2009). Nair et al. (2009) developed the 12-item Developmental Assessment Tool for Anganwadis (DATA) which includes milestones such as 'kick large ball', 'place bead in/under/on a cup'. Items were selected from a list of milestones from the Denver Developmental Screening Test (DDST), the Developmental Assessment Scale for Indian Infants (DASII), the Receptive-Expressive Emergent Language Test (REEL) and the Vineland Adaptive Behavior Scale (VABS). Face validity was considered to be high as the items were taken from internationally used measures. In a study of 100 toddlers in anganwadis, none of the items was assigned a score of 0 by more than 90% of parents suggesting that the items were appropriate for measuring the development of a toddler, endorsing the content validity of the scale. The internal consistency of the scale was high with a Chronbach's alpha of 0.86. Construct validity was analysed using exploratory factor analysis which yielded a 2-factor model which explained 56% of the variance. A total of 429 toddlers mean age 30.9 months from 36 randomly selected anganwadis

were recruited for its standardization. However, whilst DATA was developed for use by anganwadi workers (AWWs), in this study DATA was administered by developmental therapists and field trials with AWWs are needed. There is also no data on sensitivity, specificity or PPV for DATA.

Developmental Milestone Chart (DMC)

In a children's hospital in Cambodia the need for regular use of an appropriate milestone assessment led to the design and study of the Developmental Milestone Chart (DMC) (Scherzer, in press). Milestones were selected and modified from the existing literature (eg Denver DST, Denver II). The DMC enables brief evaluation through simple check-off of the appropriate age group (suitable for ages 1 month to 8 years). DMC is a simple culturally relevant one page chart which is designed for use at regular clinic visits. Its' use could enable recognition of delays as the basis of referral by health staff for further identification, parental education and support and enable early intervention.

Clinical trials took place in an outpatient setting for one week in 2007 and two consecutive weeks in 2008. Completing the chart based on observation and parental report took less than 5 minutes. The possibility of "delay" was considered if one or more of the age appropriate milestones were not met. In the 2007 trial, 25% of children did not achieve one or more milestone, and 32% in 2008. Multiple delays were concentrated in smaller numbers of children and these are most likely to have significant delays or deficits. One failed item may not be sufficiently sensitive to warrant the designation of delay. Further work is needed to refine the DMC in terms of cultural relevance and to evaluate the number of milestone failures that should be used to reflect delay in order to avoid excessive false positives. There is also no information yet on DMC specificity, sensitivity, test-retest and interrater reliability, or predictability.

Developmental Observation Card

In a review article, reference is made to the Developmental Observation Card (DOC) for mothers (Nair and Radhakrishnan, 2004) but articles regarding this have not been found. It is suggested that the large majority of developmental delays could be identified by using cut off points for four simple developmental milestones and the DOC presents mothers with a simple card showing the age at which these milestones should be attained. The Developmental Observation Card is available on the Child Development Center website.^b

Developmental Screening Inventory (DSI)

This existing screening inventory was validated on 128 Nigerian children aged 2-30 months against the Bayley Scales of Infant Development (BSID). The two instruments were administered sequentially to estimate their concurrent validity. Correlations between standardized scores on the DSI and BSID were significant at

^b <http://www.pediatricskerala.com/html/childdvlpcentre.htm#o2>

$p < .01$. Reliability of the DSI is reported using Chronbach's alpha which was 0.64. The authors suggest that the DSI is suitable for use in pediatric practice for the early detection of various developmental disorders (Aina and Morakinyo, 2001).

Disability Screening Schedule (DSS)

This tool was developed as a one-time screen for all major disabilities in children under 6 years of age (Chopra et al., 1999). It was validated using a two phase design in urban slums in South Delhi. Nineteen AWWs screened 3560 children aged 0-6 years from 9 urban slums following 6 days of training. Of these children, 245 were classed as impaired. To validate the screening work, 219 of the impaired children were assessed by the investigator, as well as 536 (16%) of those who screened normal. However, the exact details of the nature of this assessment are not presented. Sensitivity was found to be 0.89 and specificity was 0.98 which they note to be higher figures than other major screening tests under use. Administration time was about 5 minutes.

Guide for Monitoring Child Development (GMCD).

This tool was developed in Turkey for use by health care providers in LAMI countries to monitor the development of children 0.0-3.5 years of age (Ertem et al., 2008). The GMCD has 3 components:

1. The monitoring development component which is reported in Ertem et al 2008.
2. The supporting development component (expanded version of WHO/UNICEF Care for Development Intervention)
3. The managing developmental difficulties component, which has been adopted by the Turkish Ministry of Health and UNICEF-Turkey to be used in a nationwide training program of child development for primary health care providers.

GMCD is a brief, open-ended, pre-coded interview with the primary caregiver. The first question asks about parental concerns. There are then 6 open-ended questions (eg "tell me examples of what (child) can understand when you talk to her?") relating to developmental domains like receptive language, gross and fine motor skills, etc. For each of the 6 questions there are specific pre-coded milestones. The GMCD form is composed of two tables on each side of a single sheet. The questions are placed in rows and 8 age ranges in the columns with milestones in the cells. Milestones were chosen by pooling items from standardized instruments, selecting those that could be easily observed and reported by caregivers, and those which were universal (e.g., peek a boo) rather than culture specific (e.g., playing with a mirror).

Three studies were done in Turkey on the construction and psychometric properties of the GMCD for children age 0-24 months. Study one aimed to determine the ages of attainment of the GMCD milestones in a sample of 510 children aged 0-24

months. Item-total scale correlations ranged from 0.28 to 0.91. An age-dependent attainment pattern was seen in all of the milestones. Study two looked at interrater reliability of the GMCD based on administration of the GMCD to 92 children by pairs of medical students and administration to the same children by a child development specialist. Interrater reliability between medical-student pairs and between the child development specialist and students was high (kappa scores were 0.83–0.88). In study three, the concurrent validity of the GMCD was assessed for 79 children against a comprehensive assessment by a paediatrician involving developmental history, play observations, neurologic examination, and use of the Bayley II. The sensitivity, specificity, and positive and negative predictive values were 0.88, 0.93, 0.84, and 0.94, respectively.

The authors suggest that the GMCD advances the field conceptually by introducing a new and practical method of using an open-ended interviewing technique to get information on child development. Mean administration time by students was 7 minutes. It is proposed as an alternative to tools such as DSS, TQS and ACCESS Portfolio which they note are designed to question caregivers about severe disability.

Indonesian Adaptation of the Vineland Adaptive Behavior Scales (IVABS)

This study looked at the psychometric properties of a 245 item adaptation of a Western measure of adaptive behaviour (the Vineland Adaptive Behavior Scales Survey Form) developed for use in Indonesia (Tombokan-Runtutahu and Nitko, 1992). The Indonesian sample consisted of 43 pairs of children aged 6-18 years with and without ID matched on the basis of age, gender and socio-economic status (SES). The study is limited due to the small sample size and the authors note that it does not support the implementation of IVABS on a national basis due to the need for further validation and standardization.

Infant Neurological International Battery Test (INFANIB)

Health care specialists in Iran do not have a measurement tool for detecting developmental delay especially in mild to moderate motor delay in early years. The Denver Developmental Screening Test (DDSTII) is used, but it is not valid in the first year of life because of low sensitivity of motor delay diagnosis (Soleimani and Dadkhah, 2006). This study looks at the validity and reliability of the existing Infant Neurological International Battery Test (INFANIB) in Iran. A total of 6150 infants (aged 4-18 months) were screened. Those classed as “abnormal” (for example, with hypotonia, spastic diplegia or other types of gross motor developmental disorder) were referred to a specialist pediatric neurology unit for diagnosis. Validity was assessed against a neuro-developmental assessment by a pediatric neurologist for 153 random selected children. Sensitivity was 90%, specificity 83%, positive predictive value 79% and negative predictive value 93%. A further 54 infants were randomly selected to assess interrater reliability between occupational therapists and paediatricians. The intraclass correlation coefficient was 0.90. The INFANIB is specifically for motor delay. It can be used by non-physicians such as

physiotherapists, occupational therapists, and nurses. Mean time for scoring the INFANIB test was 8-10 mins.

Malawian Developmental Screening Tool

The aim of this work is to create a simple culturally appropriate developmental assessment tool adapted and modified from Western tools and standardised for use in rural Malawi (Gladstone et al., 2008). Items from Western tools (eg DDSST or Denver II) were examined to see which items were irrelevant to rural Malawian children. These were replaced by items appropriate to the context creating a 138 item tool. Age standardised norms were estimated based on 1130 normally developing children aged 0-6 years from a rural Malawi setting. Both inter-observer and intra-observer reliability were tested for each item based on two subsamples of 60 (inter-observer) and 28 (intra-observer) children. For inter-observer reliability, 82% of the questions had moderate to very good reliability ($\kappa > 0.4$). For intra-observer reliability the figure was 75% of items. Following a consensus meeting, 110 of the 138 items were retained in a revised instrument, with some needing further modification. They are now refining the tool further with a larger standardisation sample and creating a scoring system plus carrying out further validation. Once this new version has been created there will be a tool that could be used by community health workers in other rural settings in Africa after local validation.

Monitoring Child Development at Family & Community Level

This was a multicentre study in collaboration with WHO on the development and standardization of culturally appropriate scales of development for children age 0-6 years (Lansdown, 1996). It also looks at identifying a small number of key milestones to add to a child's home-based record. A total of 28,139 children aged 0-6 years took part in China, India, and Thailand. Milestones for the child's home based record were selected based on reliability, validity and simplicity. For illiterate families simple line drawings were used to depict the items. All three countries selected milestones and incorporated them on the weight for age home-based record. In China, 19 items were selected for the home based record and 35 test items were selected to form the South-East China Developmental Screening Test. In India, 13 milestones were selected for the home-based record and a similar selection made for the Thai record card.

On the Chinese weight-age card the milestones are presented in pictorial form and red and yellow to indicate high and moderate risk. The records are used by mothers, health workers and rural doctors who have been trained to monitor the child's development and record it on the chart. It was stated that the next objective of the multicentre study was to see if developmental screening could be used in the home, the community and primary health care to detect developmental delays early enough to do simple interventions that could improve performance and prognosis. However, it has not been possible to locate any further articles from this study.

National Institute for the Mentally Handicapped Developmental Screening Schedule (NIMH-DSS)

The aim of this study was to develop a reliable screening device for the early identification of children with ID aged 0-6 years in rural areas of India for use by Anganwadi workers (Arya, 1991). Two rural areas where the Integrated Child Development Scheme (ICDS) was being implemented were selected which had a total population of 75,000 in 63 villages of which 25 were randomly sampled for the study. A pilot study was conducted on 180 children aged 0-6 years and 600 children were included in the main study.

The screening tool was based on developmental milestones which were: simple, low cost and relevant to rural culture; measure abilities valued by parents and village workers; measure discrete and observable behaviour with a high degree of reliability rather than parental report with a clear pass or fail mark; and measure relevant aspects of development. Initially, 98 items were selected from which 10 items were selected to form the National Institute for the Mentally Handicapped Developmental Screening Schedule (NIMH-DSS) based on pilot testing on 180 children aged 0-6 years.

To validate the NIMH-DSS, 20 Anganwadi workers screened a total of 600 children aged 0-6 years, with alternate children (n=300) being assessed by a psychologist. Of the 600 children 3.2% screened positive, 100% of whom were confirmed as having developmental delay by professional assessment. One per cent of screens were false positives, and 0.8% were false negatives. 95% were correctly identified as normally developing. Sensitivity was 0.79 and specificity was 0.99. Overall screening accuracy was 0.98. The authors conclude that the NIMH-DSS can be considered an effective tool for screening pre-school children with ID in rural areas of India.

Parents Evaluation of Developmental Status (PEDS; Indian Validation)

Parental concern has been found to be a useful clinical tool in the West but nothing was known about the importance of parental concerns in the Indian context. Parental concerns were elicited using the PEDS (Malhi and Singhi, 2002). Developmental status was assessed by the Developmental Profile II which gives an IQ equivalent score. They also administered the Indian Adaptation of the Vineland Social Maturity Scale. Sensitivity was 62% and specificity 65%, both of which were lower than values found for North American children (75% and 74% respectively). They suggest that PEDS should not be used as an alternative to standardized developmental screening but may be used as a pre-screening instrument in a busy outpatient setting to identify children who may require more in depth developmental screening. However, the sample size was small (79 parent-child dyads) so the results need confirming with a larger sample.

Psychosocial Developmental Screening Test

This screening test was developed by the Indian Council of Medical Research (Vazir et al., 1994). The test contains 66 milestone items which form a simple, culturally appropriate screen for psychosocial development for administration by CHWs. The test was standardized on a rural, tribal and urban sample of over 13,000 children under 6 years of age from 3 regions of India. Interrater reliability between supervising psychologists and CHWs ranged from 95-98%. Test-retest reliability based on 1% of the sample was from 95-99%. Seven centile levels for each of the 66 milestones are presented and age of attainment at the 50th centile used for age placement of that item. The authors propose that some, or all, of the items could be used at the community level to identify children with developmental delay. Alternatively, it could be used at a second referral level (e.g. health centres) as an intermediary between preliminary detection and multidisciplinary assessment for diagnostic purposes.

This test has also been used in a cross sectional study of psychosocial development in an urban slum in central Delhi (Malik et al., 2007). 202 infants were examined and their mothers interviewed at their home. Interviews collected information on socio-demographics and also used the Psychosocial Developmental Screening Test (Vazir et al., 1994). The authors note that the research is limited and the results of the study are not sufficient to plan interventions to improve the development of children in such settings. Malik et al conclude that further research is needed with larger sample sizes.

The Ten Questions Screen

In their systematic review, Maulik and Darmstadt (2007) noted that the most commonly used tool to assess disability in large populations in LAMI countries was the Ten Questions Screen (TQS). Since this review was published, the TQS has also been included as a disability module in the UNICEF Multiple Indicator Cluster Survey (MICS). During the 2005-2008 round of the MICS, the disability module was administered to over 200,000 children across 20 participating countries (UNICEF, 2008). Between 14% and 35% of children screened positive in most of the participating countries (UNICEF, 2008).

The TQS is a brief questionnaire designed to be administered to parents by non-professionals as a personal interview. Five of the questions are designed to detect cognitive disability, two relate to movement disability, and there is one question each on seizures, vision and hearing respectively. The target age group is 2-9 years. The TQS is intended as a rapid low cost method of case-finding in communities such as those in LAMI countries where many or most seriously disabled children have never received professional services. Three features of the questionnaire design are intended to enhance its appropriateness and measurement qualities under diverse cultural and socio-economic conditions:

the questions are simple with a yes-no response format; they focus on universal abilities that children in all cultures normally acquire rather than culturally specific behaviours; and they ask the parent to compare the child to others of the same age and cultural setting (Durkin et al., 1995).

It has been noted that the TQS, if validated, would permit epidemiological studies in populations not previously studied and would facilitate the referral of children to programs now being developed

throughout the world for Community Based Rehabilitation (CBR) (Durkin et al., 1994). To address the validity of the TQS a large collaborative two phase study was conducted involving over 22,000 2-9 year olds in Bangladesh, Jamaica and Pakistan (Durkin et al., 1994, Thorburn et al., 1992, Zaman et al., 1990). When judged against clinical evaluation by psychologists and physicians, it was found that the specificity of TQS as a screen for serious disability was high in all populations: 0.92 Bangladesh; 0.85 Jamaica; and 0.86 Pakistan. Sensitivity for cognitive disability was: 0.82 Bangladesh; 0.84 Pakistan; but only 0.53 Jamaica. In all 3 populations, sensitivity was 1.00 for severe cognitive disability. Durkin et al (1994) note that the value of the TQS for identifying disability in underserved populations is limited to that of a screen; more thorough evaluations of children screened positive are necessary to

Ten Questions (TQ) Screen for Child Disability

1. Compared with other children, did (name) have any serious delay in sitting, standing, or walking?
2. Compared with other children does (name) have difficulty seeing, either in the daytime or at night?
3. Does (name) appear to have difficulty hearing?
4. When you tell (name) to do something, does he/she seem to understand what you are saying?
5. Does (name) have difficulty in walking or moving his/her arms or does he/she have weakness and/or stiffness in the arms or legs?
6. Does (name) sometimes have fits, become rigid, or lose consciousness?
7. Does (name) learn to do things like other children his/her age?
8. Does (name) speak at all (can he/she make himself/herself understood in words; can he/she say any recognizable words)?
9. a. Ages 3–9: Is (name)'s speech in any way different from normal?
b. Age 2: Can he/she name at least one object (animal, toy, cup, spoon)?
10. Compared with other children of his/her age, does (name) appear in any way mentally backward, dull or slow?

distinguish true- from false-positive results and to identify the nature of the disability if present.

The results of this two phase study of the TQS were also analysed to assess the reliability of the screen (Durkin et al., 1995). To assess test-retest reliability, repeat screenings were carried out for 101 children in Bangladesh and 52 children in Pakistan. Kappa coefficients for total scores were 0.58 in Bangladesh and 0.83 in Pakistan. Internal consistency was measured using Chronbach's alpha and found to be 0.60 for Bangladesh and Jamaica and 0.66 for Pakistan. Factor analysis revealed that factor loadings of the ten questions were consistent across the three populations with high levels of reliability for 7 of the 10 items in all three populations. Item characteristic curves were constructed for all 10 items which demonstrated some consistency across the three cultures although the characteristic curve for question 10 (on slowness) differed strikingly across countries. Durkin et al (1995) conclude that the TQS is a reliable questionnaire and indicators of reliability are comparable across populations that differ in culture and level of socioeconomic development, although question 10 appears to "over-identify" children as seriously disabled in Jamaica.

The TQS has also been validated in a two stage study in a rural setting in Northern India (Singhi, 2007). Judged against clinical evaluation by a pediatrician and a psychologist, they found that the TQS was a sensitive tool (100%) for the detection of significant disabilities in children age 2-9 years. The positive predictive value was 50% indicating that it would lead to a number of over referrals. However, some of these would have benefited from referral as 23% of the false positives had mild delay due to malnutrition. Singhi (207) notes that the TQS is a low-cost-quick screening tool that can be used by community workers but not as an assessment tool. The positive predictive value was higher for boys (61%) than for girls (31%) which may be because parents in India display more concern for the health of sons than daughters.

Finally, a two stage study looking at the validity and reliability of the TQS for detecting moderate to severe neurological impairment (NI) in 6-9 year olds in rural Kenya has been reported (Mung'ala-Odera et al., 2004). The TQS was administered by field interviewers in one of the poorest areas of Kenya and all children who screened positive were assessed by a clinician and psychologist, as well as 1 in 12 of those who screened negative. The TQS was translated into Kigiriama. A total of 10,218 children were screened, of whom 955 (9%) were positive on TQQ. Of these, 810 (85%) were assessed, and of those who tested negative 766 (8%) were assessed. Sensitivity ranged from 0.70 (cognitive impairment) to 1.00 (epilepsy). Specificity ranged from 0.71 (cognitive impairment) to 0.98 (for both motor and visual impairment). Positive predictive values ranged from 0.11 to 0.33 (0.24 for cognitive impairment). Negative predictive values ranged from 0.97 to 1.00 (0.95 for cognitive impairment). Test-retest reliability was assessed by readministration of the TQS to 270 children and kappa coefficients for individual items ranges from 0.2 to 1.0.

In addition to these validation studies, the TQS has also been used in the following prevalence studies:

Bashir et al (2002). A cohort of 1476 children born during 1984-87 in populations with varying socioeconomic conditions in and around Lahore, Pakistan were followed up at age 6-7 years at which point 649 children were identified (Bashir et al., 2002). First phase screening using the TQS identified 132 children who screened positive. Second stage assessment by specialists including a clinical psychologist involved the Wechsler Intelligence Scale for Children, the Griffiths Mental Developmental Scales, and the Harris Good Enough Drawing Test. The overall prevalence of mild mental retardation (MMR) among 6-10-y-old children was 6.2%. The distribution of MMR was uneven, with 1.2% among children from the upper-middle class, 4.8% in the village, 6.1% in the urban slum and 10.5% in the poor periurban slum area.

Christianson et al (2002). This two stage study documents the prevalence of ID in rural South Africa for children aged 2-9 years (Christianson et al., 2002). The TQS was used to screen 6692 2-9 year old children. Of these, 722 (10.8%) had a pediatric evaluation including the Griffiths Scale of Mental Development if clinicians considered the child at risk of developmental delay. ID was diagnosed in 238 children giving a ID prevalence of 3.6%/35.6 per 1000 (or 1 in 28). Severe ID was in 43 (0.6%) and mild ID in 195 (2.9%). The severe ID rate of 6 per 1000 approaches twice the rate of between 3 and 4 per 1000 estimated to be present in industrialised nations (Christianson et al., 2002).

Couper (2002). A two stage study was used to look at disability in 0-9 year olds in rural KwaZulu-Natal (Couper, 2002). The TQS was amended with the addition of 6 extra questions to cover children under 2 years old. In the first stage, the amended TQS was administered by 12 CHWs who visited a total of 736 homesteads, screening 2036 children. In the second stage a team of two OTs, a therapy assistant and a community rehabilitation facilitator confirmed the disabilities reported. All children confirmed with a disability were followed up and treated by the rehabilitation team at the hospital or nearest clinic. Of the 2036 children: 163 were reported with a disability; 158 of these were followed up; and 122 of these were confirmed to have disability (i.e. 6%). The most prevalent disabilities were mild perceptual or learning disability (17/1,000), followed by cerebral palsy (10/1,000), hearing loss (10/1,000), moderate to severe perceptual disability (6/1,000) and seizure disorders (4/1,000)

Durkin et al (1998). A two stage study was used to look at the prevalence of ID in 2-9 year olds in Karachi, Pakistan (Durkin et al., 1998). In phase one 6,365 children were screened using the TQS administered by social work students. Phase two involved clinical evaluation of children referred from phase one which was done by a team of local psychologists and physicians. Diagnosis of ID was made consensually by a physician and psychologist after they had independently assessed the child. Psychological assessment of ID was based on nonverbal scales of the Stanford-Binet intelligence test and an adaptive behaviour scale developed (and normative for) children in Pakistan. Of the 6,365 children screened, 936 (14.7%) screened positive on the TQS. Of these, 818 (87%) and 545 (10%) who screened negative were

clinically evaluated in phase 2. Overall prevalence estimates were 18.97/1000 for serious ID and 65.33/1000 for mild ID. Only 3.7% of children with ID had been previously evaluated for ID or had received any services for ID. The percentage of children attending school was much lower for children with ID: 8.4% serious ID; 48.2% mild ID; and 77.3% non-ID.

Kromberg et al (2008). A two stage study looked at the prevalence of ID in a rural population in South Africa (Kromberg et al., 2008). In stage one the TQS was administered by local field workers to 6,692 children. Those who screened positive on this were examined by one of three paediatricians all with some neurodevelopmental expertise who also administered Griffiths scale of mental development. Of the 6,692 children, 722 (10.8%) had a pediatric assessment. The most common disorder with ID (3.6%). 4.3% had one or more of the five selected disabilities. 0.6% had severe ID and 2.9% had mild ID.

Mung'ala-Odera et al (2006). This article reports prevalence rates for NI in children aged 6-9 years in one of the poorest areas of Kenya based on the data from Mung'ala et al. (2004) described above (Mung'ala-Odera et al., 2006). The prevalence for moderate/severe NI was 61/1000. The most common domains affected were epilepsy (41/1000), cognition (31/1000), and hearing (14/1000). Motor (5/1000) and vision (2/1000) impairments were less common. Of the 251 neurologically impaired children, 56 (22%) had more than one impairment. The authors conclude that there is a considerable burden of moderate/severe NI in this area of rural Kenya, with epilepsy, cognition, and hearing being the most common domains.

Pongprapai et al 1996. This three stage study looked at the prevalence of disability in a rural community in southern Thailand (Pongprapai et al., 1996). In stage one, 12 local field workers used a modified form of the TQS to detect disabilities in children under 15 years of age with 4366 children being screened and 185 (4.2%) screening positive. In stage two, all children who were determined as having an impairment in stage one were assessed by medical students using the revised WHO Training of the Disabled in the Community Manual which reduced the number to 68 mainly due to the removal of cases of isolated febrile convulsions. They then did an in-depth investigation by a rehabilitation physician for all those confirmed as disabled in stage two at which point 13 children were excluded as non-impaired. The final prevalence figure for disability was 1.2% of whom 7.9% had ID. The authors suggest that the TQS, modified with respect to the question on fits, can provide a relatively quick and valid estimate of impairment in rural Thailand and should be more widely used.

Islam et al (1993). The data from the TQS validation studies (reported above) have also been used to look at the relationship between socioeconomic status (SES) and the prevalence of ID in Bangladesh (Islam et al., 1993). The prevalence of mild ID in low SES groups was nearly three times that in middle or upper SES groups but the prevalence of severe ID was not significantly related to SES.

Training on Use of the TQS. Very little is mentioned in the foregoing studies regarding training on use of the TQS with the majority simply stating that field workers, or other non-professionals such as community health workers, were trained in the use of the TQS (e.g. Durkin et al 1994, 1995; Pongprapai et al 1996; Zaman et al 1990). In other cases it is not possible to determine how much of the training received related to the TQS per se. For example, Mung'ala-Odera et al (2004) note that field interviewers were given one week's training but this included both training in field methods and the use of the TQS. Similarly, Thorburn et al (1992) specifically mention training involving two weeks of formal classroom training followed by three weeks of field training but this study involved the use of five questionnaires in a house to house survey so it is not possible to ascertain the extent of training on the TQS. Couper (2002) notes that training sessions included piloting the TQS in the community to ensure that the questions were understood and the forms filled in correctly but the extent of training sessions is unclear.

Trivandrum Developmental Screening Chart (TDSC)

The Trivandrum Developmental Screening Chart (TDSC) was designed at the Child Development Center in India (Nair et al., 1991). Seventeen test items were chosen to include mental and motor developmental milestones over the first 2 years of age. The range for each test items was taken from the norms given in the Bayley Scales of Infant Development (Baroda norms). A vertical line is drawn, or a pencil kept vertically, at the level of the chronological age of the child being tested. If the child fails to achieve any item on the left side of the line they are considered to have developmental delay.

TDSC was validated against the DDST in a two stage study in India. The TDSC showed clinically acceptable sensitivity of 66.7% and specificity of 78.8% against DDST as gold standard. The chart is recommended as a mass screening test for detection of developmental delay in children under 2 years of age. It can be done in 5 minutes by a health worker. It was noted by the authors that the screening chart was being field tested for use by AWWs in a major community study but it has not been possible to locate further substantive articles in relation to this.

Other Studies on the Use of Screens

Other articles outline the practical application of screening in LAMI countries but do not present any research data associated with this screening. In China, from 1995 two screening surveys were undertaken in 14 counties in eight of China's 30 provinces with a total population of 6.6 million (Ericsson et al., 2008). Using proportional population sampling more than 100,000 children aged 0-6 years were screened by professionals or paraprofessionals in medical and health work with children. The Denver Developmental Screening Test (DDST) was used as there was a Chinese version standardized in 6 urban areas of China. A "train the trainer" approach was employed, with people trained for a week who then set up their own training programs at county and village level. Around 400 developmental screening teams were set up and over 1500 people involved in screening. Children who

screened positive were given a complete developmental evaluation, the China Neuropsychological Developmental Scale for Children.

Alternatives to Screening Tests

Whilst there has been wide use of cross-sectional surveys employing screening tests (mainly the TQS) for identifying children with disabilities in LAMI countries, it has been argued that these methods have been time consuming, expensive and have not resulted in better services for the people identified as having a disability (Kuruville and Joseph, 1999, Gona et al., 2006). A small number of studies have looked at alternative methods of identifying children with disabilities in LAMI countries and these are outlined below.

A study in Jamaica compared the efficacy of key informant and community survey methods for identifying children with disability (Thorburn et al., 1991). In the key informant method, 130 key informants took part in a 2-day workshop giving information on signs of disability and available services. Questionnaires were given to the informants which were used to refer children with disabilities. In the survey method, eight community workers completed a house-to-house survey using the TQS to screen 5475 children aged 2 to 9 years. Seventeen referrals were made by the key informants and of these, two were found to have disabilities. Of the 821 children who tested positive on the TQS in the community survey, 193 had disabilities. Thorburn et al (1991) concluded that the key informant method would not be a satisfactory way of identifying cases of childhood disability.

A study by Kuruville et al (1999) compared the community survey method with rapid rural appraisal (RRA) in identifying disability in all age groups in a rural population in India. RRA used social mapping, semi-structured interviews and direct observation to identify people with disabilities. No child under the age of two was identified using either method and children under the age of 5 were only identified if they had a severe disability. They suggest the RRA can facilitate community awareness of disability and participation in rehabilitation but that a combination of methods would be the most effective approach to identifying people with disabilities. They suggest that a simple screening tool such as the TDSC would be needed for all children under 2 years of age (Kuruville and Joseph, 1999).

Finally, an alternative approach to identifying children with disabilities is Participatory Rural Appraisal (PRA). This is described by Gona et al (2006) as a research approach which involves local communities as active analysts of their own situation so that they can set their own priorities on how to change their situations. Gona et al (2006) employed PRA to identify children with disabilities in rural Kenya. This involved 12 focus groups of 12 people (including village leaders, women's groups and teachers) whose discussions explored perceptions of disability. Social mapping was then performed where the area was mapped and symbols placed on the map (e.g. a leaf for disabilities in hearing) to identify children with disabilities. Gona et al (2006) found that much disability was attributed to evil spirits or witchcraft. A total of 237 children aged 9-15 years were identified giving a disability

prevalence rate of 69/1000. The PRA study took 2 months and may thus be a fast and inexpensive method of screening in LAMI countries. Gona et al (2006) note that for each child identified in a house to house survey the cost is from US\$7 to US\$14 whilst the cost per child using PRA was equivalent to US\$1.20. They suggest that it also raises awareness of disability and facilitates local participation which may have a positive effect on future community involvement in rehabilitation programmes.

Managing Identified Disabilities

It has been argued that serious ethical issues may arise unless screening leads to early intervention or rehabilitation (Ericsson et al., 2008). As noted by Sonnander (2000): “The advantages of early identification must, however, always be assessed in terms of the availability of resources to provide interventions as well as the effectiveness of these interventions” (Sonnander 2000, p18). Despite this, very few of the above studies discuss interventions provided to the children identified during screening.

The lack of services available to manage identified cases has been a substantive point in a number of prevalence studies. For example, Durkin, Hasan & Hasan (1998) found that only 3.7% of children with ID had previously been evaluated for ID or received any services for ID, and only 8.4% of those with serious ID attended school (compared to 77.3% for those without ID and 48.2% for those with mild ID). They note the need for improved recognition and provision of services for ID in LAMI countries. Couper (2002) notes that a disability prevalence rate of 6% has significant implications for the delivery of health, welfare and educational services to these children where resources are limited. A study in rural northern India found that only 8% of parents/guardians were aware of rehabilitation programmes for disabled children being run in their area (Singhi, 2007). Finally, in southern Thailand (Pongprapai et al., 1996) almost half of the children confirmed as having a disability had received virtually no assessment and care from ‘Western’ medical services. This was due to both the inaccessibility and cost of such services and to traditional beliefs and practices of their culture. Two thirds of the children would have been expected to receive a definitive and practical advantage from modern rehabilitation and/or surgical service.

The lack of available services has led to the suggestion that there is no harm in missing out borderline cases of developmental delay (i.e. low sensitivity) as large scale community intervention programs are still not available and thus high specificity may be preferable (Nair et al., 1991). However, a small number of studies do make some passing mention of post-screening management and these are noted below:

- In rural KwaZulu-Natal, all children with a disability were followed up and treated by the rehabilitation team at the hospital or nearest clinic (Couper, 2002).
- An Indian study looked at the role of AWWs in an Integrated Child Development Service (ICDS) project for detection and prevention of disability

in child below 6 years of age (Mathur et al., 1995). AWWs were given inservice orientation and training to detect disabilities in children below 6 years (4 hours a day for 6 days at a medical college). They were then given a pre-tested pro-forma for the detection of disabled children. Those who screened positive were examined by doctors from the department of paediatrics who provided measures to prevent handicaps (including immunization, supplementary nutrition, iron to anemic children, vitamin A or D in those with deficiency).

- In Southern Thailand (Pongprapai et al., 1996) the physician who assessed children who had been identified through screening also assessed treatability and preventability of the conditions and gave counselling at the end of the assessment.
- The ACCESS portfolio (Wirz et al., 2005) provides CHWs with materials that enable them to offer simple advice to parents on the basis of screening results. The study also notes that children were referred to a doctor, paediatrician or the local hospital if necessary.
- Screening surveys in China (Ericsson et al., 2008) were linked to a project which aimed to offer services for rehabilitation of children identified during the screening process and develop a framework for rehabilitation training. A main focus of the project was staff training and support was given to families to help them understand disability and contribute to the development of their children. Whilst rehabilitation centres were also set up, those who lived too far away from these could only receive local and informal support.

Discussion

This review has identified a number of screening tests that have been constructed for use in LAMI countries to identify disabilities in children. The major focus of work in LAMI countries has been the identification of generic disabilities through the administration of short screening tools by community based “grass roots” workers, such as community health workers (CHWs) and anganwadi workers (AWWs), or other grass roots workers depending on the cultural context. There is little research regarding the identification of disabilities by health professionals such as doctors or nurses in clinical settings, the notable exceptions being conducted in Cambodia (Scherzer, in press) and Iran (Soleimani and Dadkhah, 2006). The focus on generic disabilities means that the identification of ID is mostly done within the context of identifying a range of childhood disabilities.

Studies reporting on the field testing of screening tools for day to day use, as opposed to studies focussing on validation or prevalence, are rare, the notable exception being a field test of the ACCESS portfolio (Wirz et al., 2005). It seems likely that tools are being used routinely but that their use has not been evaluated. For example, a large WHO collaborative study involving over 28,000 children in three LAMI countries notes that the next objective of the study was to see if developmental screening can be used in the home, the community and primary health care to detect developmental delays early enough to do simple interventions (Lansdown, 1996). However, no subsequent papers or reports have been identified in relation to this proposed work. Similarly, in a review article it is noted that TDSC is routinely used by the AWWs of Kerala (India) to screen for developmental delay (Nair and Radhakrishnan, 2004).

Issues in Developmental Testing in LAMI Countries

A number of problems have been noted in relation to the appropriateness of Western developmental tests for use in LAMI countries. Assessment in developed countries often uses Western developmental tools (e.g. Bayley scales, Griffiths, McCarthy scale, and the Denver II) which have been designed and validated in Western countries (Gladstone et al., 2008). These may be tailored for use in non-Western settings and often translation into another language is all that is done. However, translation alone may not allow for local expressions and customs, leading to the misinterpretation of results (Gladstone et al., 2008). For example all domains of Western tests have some items which are culturally inappropriate for rural Africa such as “prepares cereal”, “play board games” and other uncommon activities. Gladstone et al (2008) note the vivid example of children screaming with terror when they saw the pink doll in the DDST test kit. Further, “naming questions” have pictures that may be unfamiliar such as car or horse, and children may have never seen a book or pictorial representations of many objects. There is also an issue of cleanliness of objects in tests (Ertem et al., 2008).

Other problems with Western test items have been found in urban slums in India. In a validation study of the Revised Prescreening Denver Questionnaire (R-PDQ), it was

found that some questions were inappropriate for urban slums in Lucknow, India (Awasthi and Pande, 1997). For example, “pedals tricycle” was inappropriate as there are no tricycles in the setting, “gives first and last names” was inappropriate as in the slums last names are not used by mothers, and “copies circles” was inappropriate in a setting with high levels of illiteracy. It was concluded that the R-PDQ could not be used for first stage screening in this setting. They found that the use of the Denver Development Screening Test in this setting was more appropriate.

One suggestion is that instruments used across LAMI countries need to have universal not culture-specific concepts in child development (Ertem et al., 2008).

Further, it has been noted that there can be no single universal test of psychosocial skills and individual countries should be encouraged to devise their own culturally appropriate scales with their own normative data (Lansdown, 1996). Even the attainment of skills such as sitting and walking may vary between cultures. For example, practices such as swaddling reduces the baby’s movement and delays motor milestones such as rolling over. Hence, locally derived reference values are essential (Lansdown, 1996). The importance of producing locally based norms has been emphasised in some studies. For example, the 50th percentile for “is able to use a cup” varied from 35 months in urban Indian children to 10 months in Thailand (Lansdown, 1996).

Attitudes to ID in LAMI countries may also have an influence on the accuracy of testing. Parents or guardians may be reluctant to say their child has a disability in a culture where such disabilities may be highly stigmatizing. Further, in some countries it has been suggested that there may be a tendency to over-report problems in boys and under-report problems in girls due to a cultural preference for boys which leads to parents displaying more concern for the health of sons than daughters (Singhi, 2007, Zaman et al., 1990). Finally, it cannot be assumed that parents will know exactly how old their child is and this has been reported as one problem in the use of developmental milestones (Chopra et al., 1999). To address this problem, one approach has been to train CHWs to assess age using a local events calendar. These may include information such as weather patterns, crop seasons, festivals, public holidays and special events such as elections which may be remembered by families as occurring at the time of the birth of the child (Lansdown, 1996, Vazir et al., 1994).

A number of criteria have been outlined for appropriate screening tools for LAMI countries. They must be: quick; low cost; acceptable to the community; easy to use by grass root level workers; and have high specificity and sensitivity as false positives are costly in terms of professional time and anxiety to families, and false negatives may impact on the child’s health (Chopra et al., 1999).

Further points with regard to developmental testing in LAMI countries have been noted (Ertem et al., 2008): caregiver literacy limits the use of written questionnaires and checklists; if developmental difficulties are prevalent in the population caregivers may not know how children should develop, meaning one cannot rely on

caregivers identifying concerns as a screening method by itself; if asking about milestones, caregivers may be reluctant to say their child has not achieved the milestone as they may not believe that interventions exist or worry about the stigma related to developmental delay; and reliance on “child testing” methods is neither practical nor desirable. Ertem et al. (2008) conclude that: “Family centered methods for monitoring child development that have evolved in the West should be the methods of choice for developing countries as well” p582. Further they suggest that monitoring child development is a new concept in LAMIs and methods should be built on existing protocols such as growth monitoring and immunisations (Ertem et al., 2008).

Conclusion

Evidence relating to the validity of *specific* approaches to identifying *intellectual disability* in children in LAMI countries is of low quality. Most studies are primarily concerned with identifying child disability per se.

However, it is possible to conclude from this body of research that two *general* approaches to the development of screening measures hold promise.

1. Valid and relatively efficient screening measures based on the reported attainment of culturally-appropriate age-specific developmental milestones have been developed for use in a number of LAMI countries.
2. The Ten Questions Screen (primarily based on reported concerns about the child's *relative* development) has been shown to have acceptable levels of validity and efficiency in the identification of general child disability.

There are, however, difficulties associated with both approaches related to the stigma associated with disability or developmental delay. In addition variation in caregiver knowledge of 'normal' development and the spatial clustering of disability may reduce the effectiveness of approaches based on reported concerns about child *relative* development.

It is recommended that future research and development work to develop *specific* approaches to identifying *intellectual disability* among children in LAMI countries should focus on:

1. The development from existing developmental milestone screening tests of a small number of parallel forms that show good transportability across LAMI countries.
2. Further examination of the performance of learning/cognition related items in the TQS (e.g., through further analysis of MICS data) and their ability to identify intellectual disability

References

- Aina O. F. & Morakinyo O. (2001) The validation of Developmental Screening Inventory (DSI) on Nigerian children. *Journal Of Tropical Pediatrics*, 47, 323-328.
- Arya S. (1991) Screening of Pre-School Children for Early Identification of Developmental Disabilities in Rural Area. *Indian Journal of Clinical Psychology*, 18, 65-70.
- Awasthi S. & Pande V. K. (1997) Validation of revised prescreening Denver Questionnaire in preschool children of urban slums. *Indian Pediatrics*, 34, 919-923.
- Bashir A., Yaqoob M., Ferngren H., Gustavson K. H., Rydelius P. A., Ansari T. & Zaman S. (2002) Prevalence and associated impairments of mild mental retardation in six- to ten-year old children in Pakistan: a prospective study. *Acta Paediatrica*, 91, 833-837.
- Bennett K. J. (1998) Do measures of externalising behaviour in normal populations predict later outcome?: Implications for targeted interventions to prevent conduct disorder. *Journal of Child Psychology and Psychiatry*, 39, 1059-1070.
- Chopra G., Verma I. C. & Seetharaman P. (1999) Development and assessment of a screening test for detecting childhood disabilities. *Indian Journal Of Pediatrics*, 66, 331-335.
- Christianson A. L., Zwane M. E., Manga P., Rosen E., Venter A., Downs D. & Kromberg J. G. R. (2002) Children with intellectual disability in rural South Africa: Prevalence and associated disability. *Journal of Intellectual Disability Research*, 46, 179-186.
- Couper J. (2002) Prevalence of childhood disability in rural KwaZulu-Natal. *South African Medical Journal*, 92, 549-552.
- de Onis M. (2006) Comparison of the World Health Organization (WHO) Child Growth Standards and the National Center for Health Statistics/WHO international growth reference: implications for child health programmes. *Public Health Nutrition*, 9, 942-947.
- Durkin M. S., Davidson L. L., Desai P., Hasan Z. M., Khan N., Shrout P. E., Thorburn M. J., Wang W. & Zaman S. S. (1994) Validity of the ten questions screened for childhood disability: results from population-based studies in Bangladesh, Jamaica, and Pakistan. *Epidemiology (Cambridge, Mass.)*, 5, 283-289.
- Durkin M. S., Hasan Z. M. & Hasan K. Z. (1998) Prevalence and correlates of mental retardation among children in Karachi, Pakistan. *American Journal Of Epidemiology*, 147, 281-288.
- Durkin M. S., Wang W., Shrout P. E., Zaman S. S., Hasan Z. M., Desai P. & Davidson L. L. (1995) Evaluating a ten questions screen for childhood disability: reliability and internal structure in different cultures. *Journal Of Clinical Epidemiology*, 48, 657-666.
- Ericsson K., Gebre-Medhin M. & Sonnander K. (2008) China-Sweden partnership: Areas of ongoing development in intellectual disabilities. *Journal of Policy and Practice in Intellectual Disabilities*, 5, 112-117.
- Ertem I. O., Dogan D. G., Gok C. G., Kizilates S. U., Caliskan A., Atay G., Vatandas N., Karaaslan T., Baskan S. G. & Cicchetti D. V. (2008) A guide for monitoring child development in low- and middle-income countries. *Pediatrics*, 121, E581-E589.
- Gabrielle Major C., Lisette Dupras C. & West Montreal Readaptation Centre (2008) Atlas-ID: Compendium of Sources Used. Montreal WHO Collaborating Centre for Research and Training in Mental Health, Montreal.
- Gladstone M., Lancaster G., Jones A., Maleta K., Mtitimila E., Ashorn P. & Smyth R. (2008) Can Western developmental screening tools be modified for use in a rural Malawian setting? *Archives Of Disease In Childhood*, 93, 23-29.
- Glascoe F. P. (2007) Developmental and behavioral screening. In: *Handbook of Intellectual and Developmental Disabilities*. (Ed. (Eds. J. Mulick, J. Jacobson & J. Rojahn), pp. 353-371. Springer Publishing Co, New York, US.

- Gona J., Hartley S. & Newton C. (2006) Using participatory rural appraisal (PRA) in the identification of children with disabilities in rural Kilifi, Kenya. *Rural and Remote Health*, 6 (online), 553. Available from <http://www.rrh.org.au>.
- Islam S., Durkin M. S. & Zaman S. S. (1993) Socioeconomic status and the prevalence of mental retardation in Bangladesh. *Mental Retardation*, 31, 412-417.
- Kromberg J., Zwane E., Manga P., Venter A., Rosen E. & Christianson A. (2008) Intellectual disability in the context of a South African population. *Journal of Policy and Practice in Intellectual Disabilities*, 5, 89-95.
- Kuruwilla S. & Joseph A. (1999) Identifying disability: comparing house-to-house survey and rapid rural appraisal. *HEALTH POLICY AND PLANNING*, 14, 182-190.
- Lansdown R. G. (1996) Culturally appropriate measures for monitoring child development at family and community level: a WHO collaborative study. *Bulletin of the World Health Organization*, 74, 283-290.
- Malhi P. & Singhi P. (2002) Role of parents evaluation of developmental status in detecting developmental delay in young children. *Indian Pediatrics*, 39, 271-275.
- Malik M., Pradhan S. K. & Prasuna J. G. (2007) Screening for psychosocial development among infants in an urban slum of Delhi. *Indian Journal Of Pediatrics*, 74, 841-845.
- Mathur G. P., Mathur S., Singh Y. D., Kushwaha K. P. & Lele S. N. (1995) Detection and prevention of childhood disability with the help of Anganwadi workers. *Indian Pediatrics*, 32, 773-777.
- Maulik P. K. & Darmstadt G. L. (2007) Childhood disability in low- and middle-income countries: Overview of screening, prevention, services, legislation, and epidemiology. *Pediatrics*, 120, S1-S55.
- Mung'ala-Odera V., Meehan R., Njuguna P., Mturi N., Alcock K., Carter J. A. & Newton C. R. (2004) Validity and reliability of the 'Ten Questions' questionnaire for detecting moderate to severe neurological impairment in children aged 6-9 years in rural Kenya. *Neuroepidemiology*, 23, 67-72.
- Mung'ala-Odera V., Meehan R., Njuguna P., Mturi N., Alcock K. J. & Newton C. (2006) Prevalence and risk factors of neurological disability and impairment in children living in rural Kenya. *International Journal of Epidemiology*, 35, 683-688.
- Nair M. K., George B., Philip E., Lekshmi S., Haran J. & Sathy N. (1991) Trivandrum Developmental Screening Chart. *Indian Pediatrics*, 28, 869-872.
- Nair M. K. & Radhakrishnan S. R. (2004) Early childhood development in deprived urban settlements. *Indian Pediatrics*, 41, 227-237.
- Nair M. K., Russell P. S. S., Rekha R., Lakshmi M., Latha S., Rajee K., Preetha V. & Thajnis M. (2009) Validation of Developmental Assessment Tool for Anganwadis (DATA). *Indian Pediatrics*, 46, S27-S36.
- Offord D. R. (2000) Selection of levels of prevention. *Addictive Behaviors*, 25, 833-842.
- Ortiz N. P. (1991) Escala abreviada de desarrollo EAD-1. Colombian Ministry of Health, Bogota.
- Phatak A. T. & Khurana B. (1991) Baroda development screening test for infants. *Indian Pediatrics*, 28, 31-37.
- Pongprapai S., Tayakkanonta K., Chongsuvivatwong V. & Underwood P. (1996) A study on disabled children in a rural community in southern Thailand. *Disability and Rehabilitation: An International, Multidisciplinary Journal*, 18, 42-46.
- Scherzer A. L. (in press) Utilization of a culturally relevant developmental milestone chart for child in low and middle income countries of the world: experience in Cambodia. *Journal of Policy and Practice in Intellectual Disabilities*, in press.
- Singhi P. (2007) Utility of the WHO Ten Questions Screen for disability detection in a rural community the North Indian experience. *Journal Of Tropical Pediatrics*, 53, 383-387.

- Soleimani F. & Dadkhah A. (2006) Validity and reliability of Infant Neurological International Battery for detection of gross motor developmental delay in Iran. *Child: Care, Health And Development*, 33, 262-265.
- Sonnander K. (2000) Early identification of children with developmental disabilities. *Acta Paediatrica*, Supplement 434, 17-23.
- Thorburn M., Desai P., Paul T. J., Malcolm L., Durkin M. & Davidson L. (1992) Identification of childhood disability in Jamaica: the ten question screen. *International Journal Of Rehabilitation Research*, 15, 115-127.
- Thorburn M. J., Desai P. & Durkin M. (1991) A comparison of efficacy of the key informant and community survey methods in the identification of childhood disability in Jamaica. *Annals Of Epidemiology*, 1, 255-261.
- Tombokan-Runtutahu J. & Nitko A. J. (1992) Translation, cultural adjustment, and validation of a measure of adaptive behavior. *Research In Developmental Disabilities*, 13, 481-501.
- UNICEF (2008) Monitoring Child Disability in Developing Countries: Results from the Multiple Indicator Cluster Surveys. United Nations Children's Fund Division of Policy and Practice, New York.
- van Meerbeke A. V., Talero-Gutierrez C. & Gonzalez-Reyes R. (2007) Prevalence of delayed neurodevelopment in children from Bogota, Colombia, South America. *Neuroepidemiology*, 29, 74-77.
- Vazir S., Naidu A. N., Vidyasagar P., Lansdown R. G. & Reddy V. (1994) Screening test battery for assessment of psychosocial development. *Indian Pediatrics*, 31, 1465-1475.
- Wirz S., Edwards K., Flower J. & Yousafzai A. (2005) Field testing of the ACCESS materials: a portfolio of materials to assist health workers to identify children with disabilities and offer simple advice to mothers. *International Journal Of Rehabilitation Research*, 28, 293-302.
- World Health Organization (2007) Atlas: Global Resources for Persons with Intellectual Disabilities 2007. World Health Organization, Geneva.
- World Health Organization (2008) mhGAP Mental Health Gap Action Programme: Scaling up care for mental, neurological, and substance use disorders. World Health Organization, Geneva.
- Zaman S. S., Khan N. Z., Islam S., Banu S., Dixit S., Shrout P. & Durkin M. (1990) Validity of the 'Ten Questions' for Screening Serious Childhood Disability: Results from Urban Bangladesh. *International Journal of Epidemiology*, 19, 613-620.

Appendix One: Search Strategy

Searches were conducted using both index terms and word searches. For word searches synonyms for ID were based on the terms for ID as used in the WHO ID Atlas (World Health Organization, 2007) truncated as indicated below (*). These were:

Developmental disabilities (development* disab*)
Intellectual disabilities (intellectual* disab*)
Learning disabilities (learning disab*)
Mental deficiency (mental* deficien*)
Mental disability (mental* disab*)
Mental handicap (mental* handicap*)
Mental retardation (mental* retard*)
Mental subnormality (mental* subnormal*)

For CINAHL and MEDLINE MeSH terms were employed in searches (+ indicates exploded term).

For ID and variants

(MH "Mental Retardation+") or
(MH "Mental Retardation, X-Linked+") or
(MH "Developmental Disabilities") or
(MH "Mentally Disabled Persons")

Low and Middle Income Countries

(MH "Africa+") OR
(MH "Asia+") OR
(MH "Europe, Eastern+") OR
(MH "Central America+") OR
(MH "South America+") OR
(MH "West Indies+") OR
(MH "Pacific Islands+") OR
(MH "Developing Countries")

In PsycINFO index terms for ID were searched as below:

(DE "Mental Retardation" OR DE "Anencephaly" OR DE "Borderline Mental Retardation" OR DE "Crying Cat Syndrome" OR DE "Downs Syndrome" OR DE "Home Reared Mentally Retarded" OR DE "Institutionalized Mentally Retarded" OR DE "Mild Mental Retardation" OR DE "Moderate Mental Retardation" OR DE "Profound Mental Retardation" OR DE "Psychosocial Mental Retardation" OR DE "Severe Mental Retardation" OR DE "Tay Sachs Disease") OR (DE "Developmental Disabilities" or DE "Fetal Alcohol Syndrome" or DE "Fragile X Syndrome" or DE

"Phenylketonuria" or DE "Prader Willi Syndrome" or DE "Rett Syndrome" or DE "Williams Syndrome")

PSYCINFO index terms for low and middle income countries were searched as below:

(ZY "afghanistan") or (ZY "africa") or (ZY "albania") or (ZY "algeria") or (ZY "american samoa") or (ZY "angola") or (ZY "armenia") or (ZY "asia") or (ZY "azerbaijan") or (ZY "bangladesh") or (ZY "belarus") or (ZY "bhutan") or (ZY "bolivia") or (ZY "bosnia-herzegovina") or (ZY "burkina faso") or (ZY "burundi") or (ZY "cambodia") or (ZY "cameroon") or (ZY "cape verde") or (ZY "cape verde islands") or (ZY "central african republic") or (ZY "central america") or (ZY "chad") or (ZY "china") or (ZY "colombia") or (ZY "commonwealth of independent states") or (ZY "comoros") or (ZY "cuba") or (ZY "democratic republic of congo") or (ZY "djibouti") or (ZY "dominican republic") or (ZY "east timor") or (ZY "eastern europe") or (ZY "ecuador") or (ZY "egypt") or (ZY "el salvador") or (ZY "equatorial guinea") or (ZY "eritrea") or (ZY "ethiopia") or (ZY "fiji") or (ZY "gambia") or (ZY "georgia") or (ZY "ghana") or (ZY "guatemala") or (ZY "guinea") or (ZY "guinea-bissau") or (ZY "guyana") or (ZY "haiti") or (ZY "honduras") or (ZY "india") or (ZY "indonesia") or (ZY "iran") or (ZY "iraq") or (ZY "ivory coast") or (ZY "jamaica") or (ZY "jordan") or (ZY "kenya") or (ZY "kiribati") or (ZY "korea") or (ZY "kyrgystan") or (ZY "laos") or (ZY "lesotho") or (ZY "liberia") or (ZY "madagascar") or (ZY "malawi") or (ZY "maldives") or (ZY "mali") or (ZY "marshall islands") or (ZY "mauritania") or (ZY "micronesia (federated states of)") or (ZY "moldova") or (ZY "mongolia") or (ZY "mozambique") or (ZY "myanmar") or (ZY "namibia") or (ZY "nepal") or (ZY "nicaragua") or (ZY "niger") or (ZY "nigeria") or (ZY "north korea") or (ZY "north vietnam") or (ZY "oceania/pacific islands") or (ZY "pakistan") or (ZY "papua new guinea") or (ZY "paraguay") or (ZY "peru") or (ZY "philippines") or (ZY "republic of congo") or (ZY "rwanda") or (ZY "samoa") or (ZY "senegal") or (ZY "sierra leone") or (ZY "solomon islands") or (ZY "somalia") or (ZY "south korea") or (ZY "sri lanka") or (ZY "sudan") or (ZY "surinam") or (ZY "suriname") or (ZY "swaziland") or (ZY "syria") or (ZY "tajikistan") or (ZY "thailand") or (ZY "tibet") or (ZY "togo") or (ZY "tonga") or (ZY "tunisia") or (ZY "turkmenistan") or (ZY "uganda") or (ZY "ukraine") or (ZY "uruguay") or (ZY "uzbekistan") or (ZY "vanuatu") or (ZY "vietnam") or (ZY "west indies") or (ZY "yemen") or (ZY "zambia") or (ZY "zimbabwe")

In all databases, the search terms employed for identifying research relevant to the identification, referral and management of children with ID were:

Diagnos*; Identif*; Assess*; Classif*; Primary; Secondary; Refer*; Test*; Manage*; Community

Appendix Two: Table of Studies Included in Review

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
Aina & Morankinyo 2001	Nigeria	LIC	Daycare centers, nursery schools, immunization clinics, religious centers and home visits in area in South West Nigeria	128 children aged 2-30 months	Validation	Developmental Screening Inventory validated against the Bayley Scales of Infant Development	Reliability based on Cronbach's correlation coefficient was found to be significantly high ($p < 0.01$) at a value of +0.64, and scores of both instruments correlated significantly indicating satisfactory validity.	Conclude that the DSI is important in the early detection of disorders such as specific developmental disorders, pervasive developmental disorders, mental retardation, autism etc. Cheap and easy to use.
Arya 1991	India	LMIC	Rural areas of India	600 children aged 0-6 years of whom 300 assessed by psychologist	Development & validation	National Institute for the Mentally Handicapped Developmental Screening Schedule (NIMH-DSS)	3.2% screened positive, 100% of whom confirmed as having developmental delay. False positive rate 1%; false negative rate 0.8%; sensitivity 0.79; specificity 0.99.	Conclude that the NIMH-DSS is an effective tool for screening pre-school children for ID in rural areas of India.
Awasthi & Pande 1997	India	LMIC	Urban slums/32 Anganwadi Centers	811 aged 2-4	Validation	Revised Prescreening Denver Questionnaire (R-PDQ); 1 in 6 also Denver Development Screening Test	R-PDQ took 19.73 mins; DDST took 22 mins. Some questions were "difficult to interpret" eg "pedals tricycle" (no tricycles to pedal); "copies circles" (mother's illiterate, would not have seen them try to copy circle); "gives first and last names" (in slums last names not used by mothers)	R-PDQ had "difficult to interpret" questions, high referral rates for further screening for developmental delay; and bad correlation with DDST. It cannot be used for first stage screening for developmental delay in urban slums of Lucknow, India. DDST may be considered for community screening in urban slums and in places with high levels of maternal illiteracy

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
Bashir, Yaqoob, Ferngren et al 2002	Pakistan	LIC	4 population groups in and around Lahore: rural village; periurban slum; urban slum; and upper middle class.	649 children aged 6-10 yrs who were part of a prospective cohort study	Prevalence study (mild mental retardation (MMR))	Ten Questions Screen (TQS) used in first phase. Second phase assessment by specialist and testing by clinical psychologist with Griffiths Mental Developmental Scales; Harris Good Enough Drawing Test; Urdu translation of the Wechsler Intelligence Scale for Children	Overall prevalence of MMR among 6-10-y-old children was 6.2%. The distribution of MMR was uneven, with 1.2% among children from the upper-middle class, 4.8% in the village, 6.1% in the urban slum and 10.5% in the poor periurban slum area.	The prevalence of MMR was found to be higher in a developing country than in developed countries. It also seemed to be related to poor socioeconomic conditions, as the prevalence in the upper-middle class was comparable to figures from developed countries, while the prevalence in children from poor population groups was much higher.
Chopra, Verma & Seetharaman 1999	India	LMIC	Urban slums of South Delhi	19 AWWs screened 3560 children aged 0-6 years	Development and validation	Disability Screening Schedule (DSS) - a one-time screen for all major disabilities administered by grass root level workers	AWWs received 6 days training. The 19 AWWs screened 3560 children from 9 urban slums. 245 classed as impaired. To validate the screening work 219 of the impaired children were reviewed as well as 536 (16%) of those who screened normal. Sensitivity was 0.89 and specificity was 0.98 which are higher figures than other major screening tests under use. Administration time was about 5 minutes.	One problem reported in using developmental milestones is that parents did not always know exactly how old their child was.
Christianson, Zwane, Manga et al 2002	South Africa	UMIC	8 villages in socio-economically deprived rural area (Bushbuckridge)	6692 children age 2-9 years from 8 villages	Prevalence study	Ten Questions Screen used in first phase. Pediatric evaluation by clinician and Griffith's Scale of Mental Development in second phase.	Of 6692 2-9 year olds screened with TQS, 722 (10.8%) had pediatric evaluation. ID prevalence was 3.56% (0.64% severe, 2.91% mild).	Factors observed which may be associated with the high rate of mild ID include poor living conditions, malnutrition, limited intellectual stimulation of infants and children, and unattended home births.

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
Couper 2002	South Africa	UMIC	Isolated rural area in KwaZulu-Natal	2,036 children age 0-9 years screened by 12 community health workers	Prevalence study	TQS adapted with addition of 6 questions to allow use with under 2 year olds. Translated into Zulu. Those positive on screen followed up with professional assessment.	163 screened positive, 158 of whom were followed up. 122 confirmed to have disability. Overall 6% had disability. The most prevalent disabilities were mild perceptual or learning disability (17/1,000), followed by cerebral palsy (10/1,000), hearing loss (10/1,000), moderate to severe perceptual disability (6/1,000) and seizure disorders (4/1,000)	They covered an extensive area in a limited time period and had to cover vast distances by foot. The survey was physically challenging for them. It proved to be a low cost method of screening for children with disabilities. The fact that 6% of rural children are disabled has serious implications for delivery of health, welfare and educational services to these children where resources are limited.
Durkin, Davidson, Desai et al 1994	Bangladesh, Jamaica, Pakistan	2 LICs; 1 UMIC	Community settings in Bangladesh, Jamaica and Pakistan	Phase one 22,125 age 2-9 years; phase two 3,983	Validation	Phase one TQS; Phase two clinical evaluation. In Bangladesh and Pakistan, used nonverbal scales of the Stanford-Binet Intelligence Scales and an adaptive behaviour scale developed for the study, or if they could not be tested with Stanford-Binet, the Denver Developmental Screening Test. Different tests used in Jamaica for 6-9 yr olds and DDST used for 2-5 year olds.	Over 22,000 children screened: 10,299 Bangladesh (8.2% screen positive), 5,461 Jamaica (15.2%), and 6,365 (14.7%) in Pakistan. Sensitivity of the screen for serious cognitive, motor, and seizure disabilities was acceptable (80-100%) in all three populations, whereas the positive predictive values range from 3 to 15%. These results confirm the usefulness of the TQS as a low-cost and rapid screen for these disabilities, although not for vision and hearing disabilities, in populations where few affected children have previously been identified and treated.	The value of the Ten Questions for identifying disability in underserved populations is limited to that of a screen; more thorough evaluations of children screened positive are necessary to distinguish true- from false-positive results and to identify the nature of the disability if present

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
Durkin, Hasan & Hasan 1998	Pakistan	LIC	Greater Karachi, Pakistan	6,365 children age 2-9 yrs	Prevalence study	Phase one TQS. Phase two for screen positives - clinical evaluation including nonverbal scales of the Stanford-Binet intelligence test and an adaptive behaviour scale developed (and normative for) children in Pakistan.	Of 6,365 children screened, 936 (14.7%) screened positive on the TQS. 818 (87%) of these and 545 (10%) who screened negative clinically evaluated in phase 2. Overall prevalence estimates were 18.97/1000 for serious MR and 65.33/1000 for mild MR. Multivariate analyses revealed lack of maternal education was strongly associated with the prevalence of both serious (odds ratio = 3.26, 95% CI 1.26-8.43) and mild (odds ratio = 3.08, 95% CI 1.85-5.14) retardation. Other factors that were independently associated with mental retardation in Karachi included histories of perinatal difficulties, neonatal infections, postnatal brain infections, and traumatic brain injury, as well as current malnourishment.	Using lack of maternal education as an indicator of socioeconomic disadvantage, prevalence rates for both serious and mild MR were associated with low socioeconomic status. Very low %s who had been evaluated or received services or been to school point to need for improved recognition and provision of services for MR in less developed countries
Durkin, Wang, ShROUT et al 1995	Bangladesh, Jamaica, Pakistan	2 LICs; 1 UMIC	Community settings in Bangladesh, Jamaica and Pakistan	22,125 age 2-9 years	Evaluation of reliability and internal structure of TQS	TQS	Using multiple methods of assessing reliability, they found that the TQS is a reliable questionnaire and indicators of reliability are comparable across populations that differ in culture and level of socioeconomic development.	One of the questions appears to "over-identify" children as seriously disabled in Jamaica.

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
Ericsson, Gebre-Medhin & Sonnander 2008	China	LMIC	Two screening surveys in 14 counties in eight of China's 30 provinces.	Over 100,000 children aged 0-6 years were screened	Screening surveys	Used the Denver Developmental Screening Test (DDST) as there was a Chinese version standardized in 6 urban areas of China. Those who screened positive given a developmental evaluation using the China Neuropsychological Developmental Scale for Children	Not stated	Used a train the trainer approach to set up around 400 developmental screening teams in 14 counties.
Ertem, Dogan, Gok et al 2008	Turkey	UMIC	1. University affiliated community based well-child care clinics in Ankara 2. Medical students at University School of Medicine Pediatrics Dept	1. 510 aged 0-24 months. 2. 184 medical students/92 children.	Development, validation, determining ages of attainment of milestones	Guide for Monitoring Child Development (GMCD). A practical one sheet open-ended interview with developmental milestones for 8 age ranges from 0-24 months.	GMCD administered in average of 7 minutes. Item-total scale correlations ranged from 0.28 to 0.91. An age-dependent attainment pattern was seen in all of the milestones. Interrater reliability between medical-student pairs and between a child development specialist and students was high (kappa scores were 0.83–0.88). The sensitivity, specificity, and positive and negative predictive values (based on n=79) were 0.88, 0.93, 0.84, and 0.94, respectively.	The GMCD training program developed by the authors consists of written materials, slides and demonstration videos and has been adopted by the Turkish Ministry and Health and UNICEF-Turkey to be used in a nationwide training program on child development for primary health care providers.

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
Gladstone, Lancaster, Jones et al 2008	Malawi	LIC	Rural area of Southern Malawi with low literacy rates	1st stage piloting 20; 2nd stage piloting 20; standardisation 1130 children age 0-6.5 years	Development, validation, standardization	138 item developmental assessment tool created using items from Denver II, DDST and Griffith's by replacing culturally inappropriate items.	Face, content and respondent validity were demonstrated. At a consensus meeting 110 items were retained in the revised instrument. Items not attained by age 6 years came from the Denver II fine motor section (e.g. Draws a square).	In all domains of western tests there are some items which are culturally inappropriate for rural Africa such as prepares cereal, play board games. Children screamed with terror when they saw the pink doll in the DDST kit. They are now refining the tool further with a larger standardisation sample and creating a scoring system plus carrying out further validation.
Gona, Hartley & Newton 2006	Kenya	LIC	Rural area of Kenya	144 focus group participants	Evaluation of identification method	Participatory rural appraisal (PRA)	237 children with disabilities aged 9 to 15 years identified giving prevalence of 69/1000.	Estimated to cost US\$1.20 per child identified compared to between US\$7 to US\$14 for survey methods
Islam, Durkin & Zaman 1993	Bangladesh	LIC	Community settings in Bangladesh	Phase one: 10,299 age 2-9 years; phase two 1,626	Prevalence and relationship to measures of SES	TQS & scale of SES constructed for study	For mild MR the prevalence in lower SES was nearly three times that in middle or upper SES. The relationship of SES to severe MR was relatively weak and ns.	
Kromberg, Zwane, Manga et al 2008	South Africa	UMIC	8 villages randomly selected in socio-economically deprived rural area	6,692 age 2-9 years	Two phase screening to look at prevalence and types of disability. (Also interviews with traditional healers).	Phase one: TQS administered by local field-workers. Phase two: examination and testing by pediatricians with neurodevelopmental expertise	722 (10.8%) had a pediatric assessment. 4.3% had one or more of five selected disabilities. The most common disorder was ID (3.6%). 0.64% had severe ID and 2.92% had mild ID.	
Kuruville & Joseph 1999	India	LMIC	Community setting in rural India	5,968 (all ages including adults)	Comparison of identification methods	House to house survey versus rapid rural appraisal (RRA)	No child under the age of two was identified using either method and children under the age of 5 were only identified if they had a severe disability.	They suggest that a simple screening tool such as the TDSC would be needed for all children under 2 years of age

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
Lansdown, Goldstein, Shah et al 1996	China, India & Thailand	3 x LMICs	Community settings including rural, urban slums & rural tribal	28,139 children age 0-6 years	Development & standardization	Culturally appropriate measures for monitoring child psychosocial development at family and community level for each country.	In each country between 13 and 19 key milestones incorporated into the child's home-based record; between 35 and 67 test items devised in each country to test children at first-referral level	Line drawings used for illiterate families. Study illustrates importance of producing locally based norms eg "is able to use a cup" varied from 35.4 months in urban Indian children to 9.5 months in Thailand.
Malhi & Singhi 2002	India	LMIC	Well-child pediatric outpatient department of tertiary care teaching hospital	79 parent-child dyads (child age 24-60 months)	Evaluation of relationship between parental concern and developmental status	Parents Evaluation of Developmental Status (PEDS); compared with Developmental Profile II & Vineland Social Maturity Scale	Overall, the presence of significant parental concerns identified 61.5% of children with delayed development and 65.2% of children with normal development. The positive predictive value of PEDS was 25.8% and the negative predictive value was 89.6%.	Sensitivity in North American children was 75%, specificity 74% - as both are lower in Indian sample they suggest that PEDS should not be used as an alternative to standardized developmental screening in that setting. They suggest that PEDS may be used as a prescreening instrument in a busy outpatient setting to identify children who may require more in depth developmental screening. Need to confirm and extend these results with larger sample.
Malik, Pradhan & Prasuna 2007	India	LMIC	Urban slum of Delhi	202 infants aged 0-12 months	Screening for psychosocial development	Psychosocial Developmental Screening Test developed by the Indian Council of Medical Research. It looks at 5 areas: gross motor; vision and fine motor; hearing language and concept development; self help skills or personal skills; social skills.	Personal skills, hearing, language and concept development and motor milestones were attained by more than 90% of infants in time. Vision and fine motor and social skills were achieved in time by slightly less.	An objective evaluation of development of infants living in urban slums is necessary for early detection of developmental delay. The research is limited and the results of the present study are not sufficient to plan interventions to improve the development of children in such settings. Further research is needed with larger sample sizes.

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
Mather, Mather, Singh et al 1995	India	LMIC	Anganwadi Centers	1545 children below 6 years of age	Evaluation of training AWWs to detect disability	Inservice training on detecting disability (4 hours a day for 6 days). Screening proforma used but unspecified.	Amongst the 1545 children, AWWs identified disability in 126 subjects which were verified in 118 cases by pediatricians. The disability rate was 7638 per 100,000 population. Visual, mental, orthopedic, speech and hearing disabilities rates were 4790, 2654, 583, 518 and 453 per 100,000 population, respectively	Doctors from dept of paediatrics visited each center to help AWWs with the survey and provide measures to prevent handicaps (including immunization, supplementary nutrition, iron to anemic children, vitamin A or D in those with deficiency).
Mung'ala, Meehan, Njunguna et al 2004	Kenya	LIC	Community settings in poor rural area of Kenya	Phase one 10,218 age 6-9 years. Phase two 810 screen positives, 766 screen negatives	Validation study	TQS administered by field workers and evaluation by clinician and psychologist for phase two	Sensitivity ranged from 0.70-1.00; specificity 0.71-0.98; PPV 0.11-0.33 & NPV 0.97-1.00.	Low PPV suggests TQS should be used alongside other assessments.
Mung'ala-Odera, Meehan, Njuguna et al 2006	Kenya	LIC	Community settings in poor rural area of Kenya	Phase one 10,218 age 6-9 years. Phase two 810 screen positives, 766 screen negatives	Prevalence study on neurological disability and impairment (NI)	Ten Questions Screen administered by field workers and evaluation by clinician and psychologist for phase two	The prevalence for moderate/severe NI was 61/1000. The most common domains affected were epilepsy (41/1000), cognition (31/1000), and hearing (14/1000). Motor (5/1000) and vision (2/1000) impairments were less common. Of the neurologically impaired children (n 5 251), 56 (22%) had more than one impairment. Neonatal insults were found to have a significant association with moderate/severe NI in both the univariate (OR 1.70) and multivariate analyses (OR 1.30)	CBR services need to be instituted to support people with disabilities arising from NI

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
Nair & Radhakrishnan 2004	India	LMIC	India (Review Article)	n/a	Review article including information on tools for identifying developmental delay	Developmental Observation Card (for mothers); Trivandrum Developmental Screening Chart (TDSC for field staff such as health workers, creche workers); Child Development Center grading for motor milestones	A World Bank project in Kerala has involved Developmental Therapists training 9258 AWWs to detect developmental delay and the TDSC is routinely used by the AWWs of Kerala to screen infants for developmental delay.	see CDC website http://www.pediatricskerala.com/html/childdvlpcentre.htm#o2
Nair, George, Philip et al 1991	India	LMIC	Hospital clinic and community based sample in Kerala, India	Total 1945 children age 0-2 years (455 community sample; 1500 well child clinic sample); 141 cross validation with Denver Developmental Screening Test	Development and validation of a simple screening tool for health workers	Trivandrum Developmental Screening Chart (TDSC). 17 items eg social smile, says two words. A vertical line is drawn, or a pencil kept vertically, at the level of the chronological age of the child being tested. If the child fails to achieve any item on the left side of the line they are considered to have developmental delay.	The TDSC can be done in 5 minutes by a health worker. The TDSC showed clinically acceptable sensitivity of 66.7% and specificity of 78.8% against DDST as gold standard.	They note that there is no harm in missing out borderline cases of developmental delay as large scale community intervention programs are still not available. Thus they prefer high specificity. The screening chart was being field tested for use by AWWs in a major community study.

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
Nair, Russell, Rekha et al 2009	India	LMIC	Anganwadis	100 toddlers for validation study. 429 toddlers for standardization study	Development, validation, standardization	Developmental Assessment Tool for Anganwadis (DATA). Milestones selected from existing developmental measures eg Denver DST, Developmental Assessment Scale for Indian Infants (DASII).	12-item DATA developed. Internal consistency, face validity, content validity and construct validity found to be appropriate. DATA score between 33 and 28 suggested 'at risk' for developing developmental delays. A score of ≤ 27 suggested already delayed milestones. A score of 27 to 16 suggested a 'mild delay', a score of 15 to 5 suggested a 'moderate delay' and ≤ 4 suggested a 'severe delay' in development.	DATA was administered by experienced developmental therapists. Field trials with administration by AWWs are needed. Subgroup analysis (e.g. Gender) in relation to standardization not done.
Phatak & Khurana 1991	India	LMIC	Community settings	n/a	Scale description	Baroda Development Screening Test for infants up to 30 months. 54 items selected from the Bayley Scales of Infant Development which has been standardised on 4141 normal Baroda babies (Baroda norms)	The screening test was put to use in the field survey as well as in clinical practice (especially well baby clinics). It had been used for more than 3 years by CHWs of Baroda. 5 or 6 one hour sessions are sufficient for training on screening. Information on sensitivity and specificity is reported to be 65%-95% in this paper but this is based on a personal communication.	Although the BSID (Baroda Norms) is regularly used at 6-7 research centres in India, the DDST appears to be the better known amongst paediatricians. Conclude that the Baroda Development Screening Test could have a wide application in field surveys and clinical practice.
Pongprapai, Tayakkanonta, Chongsuvivatwong et al 1996	Thailand	LMIC	Rural communities in deprived district in Southern Thailand	4366 children under 15 years of age screened; 185 who screened positive assessed	Prevalence & validity of screen	Phase One TQS; second phase evaluation by medical students; third phase evaluation by rehabilitation physician	In stage one, 185 screened positive; in stage two 68 confirmed to have impairments; in stage 3 53 confirmed to have impairments. Overall prevalence of disability 1.21% of whom 7.9% had ID. Many false positives at stage one were due to isolated instances of febrile convulsions pointing to need to alter the question on fits.	Almost half of the children had received virtually no assessment and care from Western medical services. This was due to both the inaccessibility and cost of such services and to traditional beliefs and practices of their culture. Two thirds of the children would be expected to receive a definitive and practical advantage from modern rehabilitation and/or surgical service.

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
Scherzer in press	Cambodia	LIC	Regular outpatient clinic at children's hospital	300 (100 in 2007; 200 in 2008) age range 6 weeks to 7 years	Clinical trials of feasibility of using DMC during regular clinic visits	Developmental Milestone Chart (DMC) - single page check off chart designed for this outpatient clinic (included as appendix). Milestones selected and modified from existing literature eg Denver DST, Denver II.	In 2007 25% and 2008 32% failed to achieve one or more age-appropriate milestones. Fine motor activities such as copying a circle, square or triangle appeared to be more challenging for some otherwise age appropriate children in the clinic setting.	Further work is needed to refine the DMC in terms of cultural relevance and to evaluate the number of milestone failures that should be used to reflect delay in order to avoid excessive false positives. There is no information on DMC specificity, sensitivity, test-retest and inter-rater reliability, or predictability.
Singhi, Kumar, Malhi et al 2007	India	LMIC	House to house interview survey in three villages in rural area of North India	Pilot study with 60 children. Screening study 1763 age 2-9 years	Validation and utility assessment of screening tool	First stage TQS and second phase clinical evaluation	A total of 1763 children were screened from 3 villages with a total population of 5830. 5.1% of boys and 3.4% of girls screened positive. The sensitivity and negative predictive value were found to be 100%. The positive predictive value of the screen for significant disabilities was 50%. Positive predictive value higher for boys (61%) than for girls (31%). Only 8% of parents/guardians were aware of rehabilitation programmes for disabled children being run in their area.	Some of those classed as "false positives" would have benefited from referral as 23% of the false positives had mild delay due to malnutrition. TQS is a low-cost-quick screening tool that can be used by community workers but not as an assessment tool. Further questions about autism or ADHD could increase the scope and completeness of the screen.
Soleimani & Dadkhah 2008	Iran	LMIC	Infants referred to health network for routine care and vaccination	6150 aged 4-18 months	Validation	Infant Neurological International Battery (INFANIB). INFANIB has 20 items to assess the infant age 4-18 months for gross motor developmental delay	Mean time for scoring the INFANIB test was 8-10 mins. The INFANIB was valid for the normal and abnormal group with 90% sensitivity, 83% specificity, 79% PPV and 93% NPV. Also the reliability coefficient between the examiners (paediatrician and occupational therapists) was calculated, and the intraclass correlation coefficient was 0.90.	INFANIB only looks at motor developmental delay. Could be used in developing countries but only where there are adequate numbers of trained staff and availability of specialist neuro-developmental services.

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
Thorburn, Desai & Durkin 1991	Jamaica	UMIC	Community setting in Jamaica	130 key informants; 5475 children aged 2-9 years screened	Comparison of identification methods	Key informant versus TQS screen	Key informant method found to be unsatisfactory for identifying disabilities	
Thorburn, Desai, Paul et al 1992	Jamaica	UMIC	Community settings	Phase one 5,461 age 2-9 years; phase two 1,219	Validation	Ten Questions Screen	as reported in Durkin et al 1994	To overcome the unethical situation of identifying children without follow-up the survey was conducted in an area where a CBR programme was being established
Tombokan-Runtutahu & Nitko 1992	Indonesia	LMIC	Children with ID and non-ID children in schools	43 ID matched with 43 non-ID children aged 6-18 years	Adaptation of a Western measure of adaptive behavior to the Indonesian context and analysis of psychometric properties	Through translation and adaptation of the Vineland Adaptive Behaviour Scales (VABS) Survey Form the Indonesian VABS (IVABS) was formed consisting of 245 items.	Psychometric characteristics were similar to that of the American version of VABS.	The research does not warrant immediate implementation of IVABS on a national basis due to limitations of this study and the need for further validation and standardization

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
van Meerbeke & Talero-Gutierrez 2007	Colombia	LMIC	Convenience samples from schools and daycare centers in Bogota, Colombia. Most below poverty baseline.	2,043 preschool children (<5 years). 287 subject to neuro-development evaluation	Prevalence study	Carers completed questionnaire about possible neurological disorders in their families and teachers asked for list of children with suspected disorder. Those who were "suspect" evaluated using The Abbreviated Developmental Scale 1 (EAD-1) battery designed and validated in Colombia. Used as the primary tool for the evaluation of development in Colombian children.	Of 2,043 children, 287 evaluated using EAD-1. One or more abnormal items (alert category) were found in 67 (23.3%) children, for an estimated prevalence of 32.8 per thousand children <5 years of age, including deficits in gross motor function (9.3 per thousand), personal-social interactions (9.8 per thousand), fine motor skills (10.3 per thousand), auditory language delay (18.6 per thousand) and overall delay (10.8 per thousand).	Identified NDD among apparently healthy children from nurseries and kindergartens, who had previously been undiagnosed and untreated. Lack of evaluation of developmental milestones in children in Colombia is a substantial public health problem that will require effective intervention.
Vazir, Naidu, Vidyasagar et al 1994	India	LMIC	Community settings from 3 regions of India	13,000 children under 6 years of age	Development & standardization	The Indian Council of Medical Research (ICMR) Developmental Screening Scale for Indian Rural Children. 66 milestone items which form a simple, culturally appropriate screen for psychosocial development for administration by CHWs.	7 centile levels for each of the 66 milestones are presented and age of attainment at 50th centile used for age placement of that item. Inter-tester reliability ranged from 95-98% and retest coefficients from 95-99%.	CHWs were trained to assess age using a local event calendar for children without birth certificates. It is proposed that the scale could be used to detect children at the community level for developmental delay.

Study	Country	World Bank Classification	Setting	Sample Size	Design or purpose	Instrument	Results	Other Comments
Wirz, Edwards, Flower et al 2005	Sri Lanka and Uganda	LMIC & LIC	Community health worker run clinics in Sri Lanka and Uganda	23 community health workers; 769 children Uganda; 580 children Sri Lanka	Field testing of ACCESS Portfolio for identifying children with disabilities and offering advice. Health workers trained, children screened, 10% of screened children assessed by experts.	ACCESS Portfolio which has an identification section and an advice section. Includes 'Messages for mother'; WHO Growth Charts (local versions); TQS; Jamaican adaptation of the Denver Developmental Screening Test; WHO play materials; short screen about vision; short screen about hearing and communication	769 children were screened in Uganda and 580 in Sri Lanka. In Uganda, 44% of children seen failed the screen and in Sri Lanka 11% failed and were deemed to have a disability by the health workers. Developmental delay and difficulties with movement and self-care were the commonest cause of disability identified. Field workers identified disabilities in children over age 2 with 82% accuracy compared with professionals.	Both health workers and parents found the process clear and useful. Parents found the advice materials helpful. Some health workers thought the manual was too bulky. They felt there were some omissions eg epilepsy not covered well, and training in counselling for family members needed.
Zaman, Khan, Islam et al 1990	Bangladesh	LIC	Community settings ranging from Urban slums to affluent areas in Dhaka	Phase one: 2576 age 2-9 years; phase two 359.	Validation with focus on gender and age	TQS	Overall more boys screened positive (7.7%) than girls (5.8%) but ns. No major age or gender differences in the validity of the questionnaire were apparent. Sensitivity, specificity and negative predictive value were perfect or near perfect for severe or moderate disabilities. PPV for serious disabilities was only 22%	70% of false positives had mild disabilities or other conditions for which early detection and treatment could be beneficial. Despite the number of false positives it reduces the number of children to be evaluated by professionals from 100% to just the around 7% who screen positive on the TQS.

Appendix Three: Summary Characteristics of Screening Tests

Instrument	Study	Country	Sample Size	Design ³	Sensitivity	Specificity	PPV	NPV	Other Validity	Reliability ⁴
Abbreviated Developmental Scale (EAD-1)	Van Meerbeke, Talero-Gutierrez & Gonzalez-Reyes 2007	Colombia	2,043 children aged <60 months	CS	NK ⁵	NK	NK	NK	NK	NK
ACCESS Portfolio	Wirz, Edwards, Flower et al 2005	Sri Lanka and Uganda	769 children Uganda; 580 children Sri Lanka	CS/Field test	NK	NK	NK	NK	Accuracy 76-82%	NK
Baroda Development Screening Test	Phatak & Khurana 1991	India	n/a	Scale description	65%-95% ⁶	65%-95%	NK	NK	NK	NK
Developmental Assessment Tool for Anganwadis (DATA)	Nair, Russell, Rekha et al 2009	India	100 toddlers for validation study. 429 toddlers for standardization study	CS	NK	NK	NK	NK	Face and content validity reported to be high. Factor analysis yielded 2-factor model explaining 56% of variance.	IC 0.86
Developmental Milestone Chart (DMC)	Scherzer in press	Cambodia	300 (100 in 2007; 200 in 2008) age range 6 weeks to 7 years	Clinical trials	NK	NK	NK	NK	NK	NK
Developmental Observation Chart	Cited in Nair & Radhakrishnan 2004	India	NK	NK	NK	NK	NK	NK	NK	NK
Developmental Screening Inventory (DSI)	Aina & Morankinyo 2001	Nigeria	128 children aged 2-30 months	CS	NK	NK	NK	NK	Concurrent validity against BSID - correlations significant at p<.01	IC 0.64

³ CS = cross sectional

⁴ IC = internal consistency measured using Chronbach's alpha

⁵ NK = not known

⁶ Sensitivity & specificity is reported to be 65-95% in this paper but this is based on a personal communication

Instrument	Study	Country	Sample Size	Design ³	Sensitivity	Specificity	PPV	NPV	Other Validity	Reliability ⁴
Disability Screening Schedule (DSS)	Chopra, Verma & Seetharaman 1999	India	3560 children aged 0-6 years; 2nd phase 219 'impaired' & 536 'normal' children	CS	0.89	0.98	NK	NK	NK	NK
Guide for Monitoring Child Development (GMCD)	Ertem, Dogan, Gok et al 2008	Turkey	510 aged 0-24 months. 2 nd phase 184 medical students/92 children	CS	0.88	0.93	0.84	0.94	Concurrent validity against comprehensive pediatric assessment showed overall agreement of 91.1%	IC 0.95; interrater reliability kappa scores 0.83-0.88
Indonesian Adaptation of the Vineland Adaptive Behavior Scales (IVABS)	Tombokan-Runtutahu & Nitko 1992	Indonesia	43 ID matched with 43 non-ID children aged 6-18 years	CS	NK	NK	NK	NK	Psychometric properties of IVABS found to be similar to American VABS	Interrater reliability R .80 to .98. Test-retest reliability .86 to .99
Infant Neurological International Battery (INFANIB)	Soleimani & Dadkhah 2006	Iran	6150 aged 4-18 months	CS	0.9	0.83	0.79	0.93	NK	Interrater reliability R=0.9
Malawian Developmental Screening Tool	Gladstone, Lancaster, Jones et al 2008	Malawi	1st stage piloting 20; 2nd stage piloting 20; standardisation 1130 children age 0-6.5 years	CS	NK	NK	NK	NK	Considered to have good face and content validity based on expert assessment	Interrater reliability kappa >0.4 for 82% of items. Intrarater figure 75% of items with kappa >0.4.
Monitoring Child Development at Family & Community Level	Lansdown, Goldstein, Shah et al 1996	China, India & Thailand	28,139 children age 0-6 years	CS	NK	NK	NK	NK	NK	Interrater reliability/intrarater reliability: China 0.95/0.91; India 0.90/0.95; Thailand 0.96/0.92.
National Institute for the Mentally Handicapped Developmental Screening Schedule (NIMH-DSS)	Arya 1991	India	Piloting 180 aged 0-6 years; validation 600 age 0-6 years	CS	0.79	0.99	100%	95%	False negatives 0.8%; false positives 1%	NK
Parents Evaluation of Developmental Status (PEDS)	Malhi & Singhi 2002	India	79 parent-child dyads (child age 24-60 months)	CS	61.50%	65.20%	25.80%	89.60%	NK	NK

Instrument	Study	Country	Sample Size	Design ³	Sensitivity	Specificity	PPV	NPV	Other Validity	Reliability ⁴
Psychosocial Developmental Screening Test	Vazir, Naidu, Vidyasagar et al 1994	India	13,000 children under 6 years of age	CS	NK	NK	NK	NK	NK	Interrater reliability 95-98%. Test-retest reliability 95-99%
Ten Questions Screen	Durkin, Davidson, Desai et al 1994; Durkin, Wang, Shrout et al 1995; Zaman, Khan, Islam et al 1990	Bangladesh (B), Jamaica (J), Pakistan (P)	Phase one 22,125 age 2-9 years; phase two 3,983	CS	B 0.87 J 0.56 P 0.85 ⁷	B 0.93 J 0.85 P 0.88	B 0.09 J 0.07 P 0.18	B 1.00 J 0.99 P 0.99	Factor loadings consistent over 3 populations	IC 0.60-0.66 Test-retest reliability kappa 0.58-0.83
Ten Questions Screen	Mung'ala, Meehan, Njuguna et al 2004	Kenya	Phase one 10,218 aged 6-9 years; phase two 810 screen positives & 766 screen negatives	CS	0.70-1.00	0.71-0.98	0.11-0.33	0.97-1.00	NK	Test-retest kappa values 0.2-1.00
Ten Questions Screen	Singhi, Kumar, Malhi et al 2007	India	Pilot study with 60 children. Screening study 1763 age 2-9 years	CS	1.00	0.74	0.50	1.00	NK	NK
Trivandrum Developmental Screening Chart (TDSC)	Nair, George, Philip et al 1991	India	1945 children age 0-2 years (455 community sample; 1500 well child clinic sample); phase 2 141	CS	66.70%	78.80%	NK	NK	NK	NK

⁷ These figures are for serious cognitive, motor and/or seizure disability. Vision and hearing omitted due to low sensitivity