NIHR | Policy Research Unit Children and Families

Measuring child development at the 2-2½ year health and development review

A review of available tools, stakeholder priorities, and learning to support successful implementation of a tool for routine health care use.

UCL

Authors

Dr Jo Lysons
Rocío Méndez Pineda
German Alarcon
Dr Maria Raisa Jessica Aquino
Hannah Cann
Professor Pasco Fearon
Professor Sally Kendall
Jennifer Kirman
Dr Jenny Woodman

Funding

This project was funded by the National Institute for Health and Care Research (NIHR) Policy Research Programme through the Child and Family Policy Research Unit (PR-PRU-1217-21301). The views expressed are those of the authors and not necessarily those of the NIHR or the UK Department of Health and Social Care.

Hannah Cann's time was supported by a NIHR Local Authority Short Placement Award and NIHR Pre-doctoral Fellowship Award (award numbers: NIHR302838 PLAF; NIHR302389 LA SPARC).

Ethics

Ethical approval for this study was granted by the UCL Institute of Education and Society Research Ethics Committee (REC7124). The study was reviewed by members of the committee, who must provide impartial advice and avoid significant conflicts of interest.

Acknowledgements

We would like to acknowledge the contributions of both parents and stakeholders who took part in this study and contributed valuable insights into their views and experiences, without which this study would not be possible. We would also like to thank the team at RREAL for their contributions to the searches and screening of the rapid systematic review element of this study (https://www.rapidresearchandevaluation.com/). We also thank Dr Sarah Smith, Associate Professor in Psychology & Psychometrics at the London School of Hygiene and Tropical Medicine who gave advice on our review of tools to measure child development aged 2-2½ years.

Contents

At a glance

Measuring child development at the 2-2½ year health and developmental review

Why we did this study

Every child in England should be offered a health and development review aged 2-2½ years by their health visiting team, which includes the Ages & Stages Questionnaire (ASQ®; 3rd Edition, licensed from Brookes Publishing, the developers). The Department of Health and Social Care in England (DHSC) has selected ASQ®-3 as the population measure of early child development in England, but ASQ®-3 is also used to inform practitioners' decisions about which children should receive extra developmental support. The DHSC commissioned us to investigate the feasibility and accuracy of tools to measure child development at the 2-2½ year health and development review, including ASQ®-3, and to gather stakeholder priorities for measuring child development at this review.

What we did

We carried out a mixed methods study with two key components

1. a rapid review of the literature to identify and assess the evidence on measures of early child development for use at 2-2½ year health and developmental review. We considered tools with additional versions for use before ages 2-2½ which could be used flexibly and repeatedly to measure child development over time.
2. 15 focus groups with 63 stakeholders to identify priorities for measuring child development aged 2-2½ years.

What we found

Two tools (ASQ®-3 and CREDI Long Form (LF)) appeared feasible for use at the 2-2½ year health and developmental review, although we found no research evidence on the accuracy of these tools to measure or detect early developmental delay within a UK population. CREDI LF is open source (free to use). Both tools seem suitable for use to collect population-level data, but research is needed on both tools for a UK context. As an individual-level assessment tool, ASQ®-3 detects severe developmental delay with good to high accuracy but is only moderately able to detect mild developmental delay amongst general populations of children aged 2-2½ years (non-UK studies). No equivalent information is yet available about CREDI. Parents and practitioners wanted a tool that facilitated a holistic conversation about development, health, and wellbeing with direct observation of the child by the professional. There was a widespread lack of clarity about the purpose of using a tool to measure child development at the 2-2½ year health and developmental review. Policy colleagues at DHSC saw the benefit in a tool that serves both purposes but cautioned against separating population-level data collection from the broader developmental review. Parents, practitioners, and policy colleagues at DHSC identified some aspects of the current tool as needing improvement and questioned its appropriateness for use with children from ethnic minority groups and children with disabilities. Our qualitative study generated findings relevant to workforce and skill mix, but we have not made recommendations in these areas.

Why this is important?

There are two robustly developed tools which could feasibly be used at the 2-2½ year health and development review to measure child development, one of which is currently used (ASQ®-3). The other measure (CREDI) has the advantage of being free to use, unlike ASQ®-3. However, the health visiting infrastructure is under strain and under the wrong circumstances, implementing a new tool may exacerbate strain on the service, undermine morale and lead to patchy up-take. The delivery of the measure and the service pathways it triggers were more important to the parents and practitioners we spoke to than the specific tool used to measure child development.

What are the implications?

A tool to measure child development at 2-2½ years in England might best be embedded within an in-person holistic review of child and wider family health which includes the parent, as is currently the case. Stakeholders told us that using the tool to facilitate a holistic conversation takes professional experience and skill. The most suitable tool will have clinical utility for individual assessment of a child **and** collect population-level data which means we need new evidence to establish population distributions of child development in England and cut-offs for whichever tool is used (including ASQ®-3). The NHS England training on ASQ®-3 at the 2-2½ year health and developmental review might not be filtering through to practitioners and parents.

Executive summary

Background

Every child in England should be offered a health and development review at age 2-2½ years as part of the Department for Health and Social Care (DHSC)'s Healthy Child Programme. A measure of child development called the Ages & Stages Questionnaire (ASQ®; 3rd Edition), is used routinely to collect population-level data for monitoring trends and disparities at this review. The Ages & Stages Questionnaire: Socio-emotional (ASQ®:SE) is sometimes used in addition to ASQ®-3, to measure children's socio-emotional development. However, there are known barriers to the ASQ®-3 meeting its intended function as a population measure of child development in England. For example, there is considerable regional variation in how ASQ®-3 is used, and issues around national collation of data have raised questions about how far we can compare ASQ®-3 scores between local areas. As the DHSC license use of the tool from Brookes Publishing, the developers, there is cost attached to its use. Additionally, providers have been calling for a digital version of ASQ®-3. Currently we understand that ASQ®-3 is used to inform a health visitor's professional judgement about whether to refer a child for individual support. As the ASQ®-3 has not been validated as an assessment tool in a UK population, we do not have population norms. This means it is not possible to determine the implications of using specific cut-offs within the ASQ®-3 to identify children with delay and what service pathways and support might best follow identification of different levels of developmental delay at 2-2½ years of age.

Aims and Objectives

We investigated alternative tools to the one currently used in the 2-2½ year health and developmental review in England, and on identifying tools which measure multiple domains of child development in children aged 2 to 3 years. We also investigated parents, carers and professionals' priorities and needs for the tool used to measure child development in the 2-2½ year health and developmental review, in addition to identifying barriers and facilitators of successful ASQ®-3 implementation.

Research questions

1. What does recent evidence say about the most appropriate (reliable, valid, feasible, cost-effective) tools (including ASQ®-3 and ASQ®:SE) for measuring broad indicators (all-domains) of child development for children aged 2-2½ years?
2. What are the views and policy and delivery priorities among key stakeholders for assessing child development at this age (parents, practitioners, and policy colleagues at DHSC)?
3. Which tools are most promising (if any) as an alternative to ASQ®-3, based on evidence of reliability, validity, stakeholder priorities, and feasibility of implementation in England?

Methods

We carried out: (1) a rapid systematic review of the literature from January 2012 to September 2022 to identify any new tools for use at the 2-2½ year health and developmental review. We considered tools with additional versions for use before ages 2-2½ which could be used flexibly and repeatedly to measure child development over time (2) 15 focus groups with 63 key stakeholders across five local areas in England (parents, health visiting professionals, local authority colleagues and policy colleagues at DHSC) to identify key priorities for measuring child development at the 2-2½ year health and developmental review. This part of the study generated findings relevant to workforce and skill mix, but we have not made recommendations in these areas.

Tools identified in the rapid review were evaluated according to feasibility of implementation in the 2-2½ year health and developmental review, psychometric properties, and validity. We analysed focus group data thematically.

Results

The rapid review identified six potential tools that could be used at the 2-2½ year health and developmental review; three tools that were developed for individual-level assessment of children:

- The ASQ®-3
- The Parents' Evaluation of Developmental Status tools (PEDS-R and PEDS:DM)
- The Warner Initial Developmental Evaluation of Adaptive and Functional Skills (WIDEA-FS)

And three tools developed only as population measures of child development:

- The Caregiver Reported Early Development Instruments (CREDI) Short and Long Form (SF, LF)
- The Global Scales for Early Development (GSED) Short and Long Form (SF, LF)
- The WHO Indicators of Infant and Young Child Development (IYCD)

ASQ®-3 and CREDI LF appear most feasible for use at the 2-2½ year health and developmental review as they can be completed without extensive training, including by a parent, and are relatively brief. The ASQ®-3 and CREDI LF also have the advantage of producing domain-specific scores (unlike CREDI SF or GSED). ASQ®-3 and CREDI also have evidenced validity against other reference tests. Whilst the ASQ®-3 is associated with licencing costs, the CREDI is free to use. Our review was not able to locate any evidence evaluating the accuracy of these tools' use in a UK setting. The non-UK evidence in our review suggests that ASQ®-3 is only moderately able to detect true mild developmental delay but that it is better at detecting severe developmental delay amongst at-risk populations. As CREDI was designed as a population-level measure, it is not known how well it performs as a tool for individual-level assessment.

Summary: we identified two feasible tools for collecting population level data on child development at the 2-2½ year health and developmental review (ASQ®-3 and CREDI LF) but with limited or no data on the accuracy of these tools for individual assessment.

Stakeholder priorities

1. Just a part of the puzzle: Implementation of the tool

Having a conversation about health and wellbeing across the family system was the principal priority for both parents and health visiting professionals. The tool was perceived to be most useful as a conversation starter rather than a tick-list. Parents, health visitors, and nursery nurses felt that a group review format made it harder for parents and practitioners to engage in this holistic conversation, and practitioners reported that whether the tool is used as a conversation-starter or as a rigid tick-list depends on the experience and confidence level of the practitioner. In addition, parents and practitioners felt strongly that a parent-reported digital tool should enhance and not replace the in-person health and developmental review.

We also heard that direct observation of the child by a professional is a priority among parents and health visiting professionals, and that professionals view the parent-reported ASQ®-3 as a way of understanding how the parent sees the child and how much help the parent needs to be able to support their child's development. Focus groups revealed a widespread lack of clarity around the purpose of the 2-2½ year health and developmental review, which led to parents having unrealistic expectations and/or feeling disappointed. As parents didn't understand why the ASQ®-3 questions ask what they ask, they worried about what ASQ®-3 scores meant for their child and what level and type of intervention they should expect as a result of the health and developmental review. Professionals emphasised the need to promote the purpose and benefits of the child development tool and the Healthy Child Programme to improve attendance and manage parents' expectations.

2. Dual use: Data versus Individual assessment

All but two of the 24 health visiting professionals seemed unaware of ASQ®-3's intended function to provide population-level data and viewed the ASQ®-3's primary function as an individual assessment tool. Parents also wished for clarity on this point. Parents and professionals

demonstrated interest in having access to regional and national-level data on child development. Policy colleagues at DHSC wanted a tool that serves both purposes. They acknowledged the dual purpose may be difficult to achieve in practice but cautioned against separating population-level data collection from the broader developmental review.

3. Barriers to universal reach

Parents and professionals suggested aspects of the tool which could be improved. Parents, professionals, and DHSC policy colleagues raised concerns about the extent to which the ASQ®-3 is appropriate for use with children from minority ethnic groups. Parents of children with disabilities valued the 2-2½ year health and developmental review but felt overlooked and neglected by the health visiting service. Whilst they did not see the ASQ®-3 as appropriate for their children, they felt that their children should be included in the population-level picture of child development and had suggestions for inclusion.

Implications of these findings

Our findings suggest that a tool to measure child development at the 2-2½ year health and developmental review might best be embedded within an in-person holistic review of child and wider family health which includes the parent, as is currently the case in England. Consideration should be given to the experience and support of the practitioner using the tool within the health and developmental review. The impact of group reviews and digital tools should be studied carefully before implementation to ascertain that these modes do not undermine the ways in which the tool and health visiting are theorised to work [1]. The most suitable tool will have both clinical utility for individual assessment and simultaneously collect population level data. ASQ®-3 and CREDI LF appear feasible to implement in the 2-2½ year health and developmental review and are likely suitable population-level measures but there is insufficient evidence on how either tool works for individual assessment of children in an English setting (despite ASQ®-3 already being used in this way in England).

Our results suggest that evidence is needed on ASQ®-3 and CREDI in a UK setting, to establish population distributions of child development in this population and to establish the most appropriate cut-offs to trigger extra support and/or referrals. Commissioning analyses of CREDI data from the Children of the 2020s birth cohort study and investigating opportunities in the other planned birth cohort study in England[1] might be an efficient way to start further research on CREDI and/or ASQ®-3. Although CREDI has the advantage of being free to use, the health visiting infrastructure is under strain; implementing a new tool has potential to exacerbate strain on the service, undermine morale and lead to patchy up-take.

The low sensitivity of ASQ®-3 for detecting mild to moderate developmental delay may reflect known difficulties in identifying meaningful developmental delay before the age of four years. It is likely that any tool which is feasible to implement in the 2-2½ year health and developmental review will encounter this problem. The Early Language Identification Measure (ELiM a tool to measure speech and language also used in the 2-2½ year health and developmental review) has been evaluated as a package with professional judgement. The same approach would be useful in further research on ASQ®-3 and CREDI, across different skill mix staff. A lack of clarity on the purpose of the tool sometimes led to anxiety and frustration in parents and professionals. Although there is NHS England training for health visiting professionals on the purpose, benefit and use of ASQ®-3 at the 2-2½ year health and developmental review, we did not see evidence in our focus groups that either parents or professionals understood the key purpose of the ASQ®-3 as a population measure or that parents understood what they might expect as a result of the ASQ®-3 being completed for their child.

Limitations

Due to the rapid nature of our systematic review, it is possible but unlikely that we missed any highly relevant and feasible tools. We were only able to review published material but know that there is in-progress work on CREDI, and this is likely the case for other tools too. We recruited our professional and parent participants from across England but had difficulty recruiting parents of children with a disability (only four recruited) and foster or kinship carers (zero recruited). Our core findings from the focus groups (63 participants) are very consistent with the previous review on this topic from 2014 [5,9] and other qualitative research [2], which increases our confidence about generalisability.

1 Early Life Cohort Feasibility Study https://cls.ucl.ac.uk/cls-studies/early-life-cohort-feasibility-study

1. Introduction

The policy in England is that every child should be offered a health and development review at age 2-2½ years, conducted by a member of the health visiting team as part of the Department for Health and Social Care (DHSC)'s Healthy Child Programme (HCP) [3]. As part of this universal health and development review at 2-2½ years, a measure of child development called the Ages & Stages Questionnaire (ASQ®; 3rd Edition), is used routinely to collect population level data for monitoring trends and disparities [4]. The ASQ®-3 covers four key domains of early development: communication, motor, problem-solving, and personal-social. The Ages & Stages Questionnaire: Socio-emotional (ASQ®:SE) is sometimes used in addition to ASQ®-3, to measure children's socio-emotional development.

However, there are known barriers to the ASQ®-3 meeting its intended function as a population measure of child development in England. There is variability in the way the ASQ®-3 is used across the country, such as whether it is administered as a parent-report or as part of a professional observation of the child [5], which raises questions about how far we can compare ASQ®-3 between local areas, whether norms or thresholds can be interpreted consistently or indeed whether the ASQ®-3 is equally valid when applied in these different ways. There are also issues with national collation of ASQ®-3 data into the Community Service Dataset (CSDS). In previous work we found that, within 33 included local authorities, only 20-30% of children aged 2-3 years old in 2018/19 had a record of ASQ®-3 in CSDS [6]. There is an alternative dataset called the Health Visitor Service Delivery Metrics, which is submitted to the Office for Health Improvement and Disparities (OHID, part of DHSC) by local authorities and is publicly available at aggregate local authority level. This data reports much higher coverage of ASQ®-3, e.g. in 2018/9, 71.4% of children aged 2½ years in the correct period had ASQ®-3 completed and ASQ®-3 was used in 90.4% of all 2-2½ year health and development reviews (figures for 2022/3 are 73.6% and 92.5% respectively [7]). The Service Delivery Metrics cannot be used to analyse disparities as there is no individual level or neighbourhood data on ethnicity or deprivation and it cannot be linked to other administrative data as CSDS can.

Providers have been calling for a digital version of ASQ®-3, as the current licensed version for use in England is a paper questionnaire. The US-hosted digital version cannot be used in England (for reasons related to data protection and GDPR) and although the digital version of ASQ®-3 can be integrated into NHS or Local Authority Systems, this will come with additional costs (personal communication from DHSC).

We also know that in many cases, ASQ®-3 is also used in many local areas alongside professional judgement as a way of deciding which individual children are referred for extra support, in addition to its core purpose of collecting population level data [8]. The ASQ®-3 has not been validated as an assessment tool in a UK population so we do not have population norms. This means there is limited academic evidence to understand how well this approach works in correctly identifying children who will benefit from extra support or which ASQ®-3 scores should result in which type of support or intervention [5].

In this study, we will assess how far available tools (including the ASQ®-3) might be suitable for a population level measure of child development at the 2-2½ year health and development review in England and what we know about the tools when they are used to trigger support pathways for individual children. In addition to reviewing the feasibility of implementation, psychometric properties and validity of existing tools, a key part of this work is understanding the policy and delivery priorities from stakeholders around a universal measure of child development for children aged 2-2½ years.

This study will inform government policy on the measure of child development included in the health and development review offered to every child in England at age 2-2½ years.

2. Rapid Review

Research questions

1. What does recent evidence say about the most appropriate (reliable, valid, feasible, cost-effective) tools (including ASQ®-3 and ASQ®:SE) for measuring broad indicators (all-domains) of child development for children aged 2-2½ years?
2. What are the views and policy and delivery priorities among key stakeholders for assessing child development at this age (parents, practitioners, and policy colleagues at DHSC)?
3. Which tools are most promising (if any) as an alternative to ASQ®-3, based on evidence of reliability, validity, stakeholder priorities and feasibility of implementation in England?

The review was tasked with looking at alternative tools to the one currently used in the 2-2½ year review, and so focuses on the existing system of child development measurements under the Healthy Child Programme and identifying tools which measure multiple domains of child development in children aged 2 to 3 years. In this review, we considered tools which could be used flexibly and repeatedly to measure child development over time from birth to three years of age, including at 2-2½ years old. However, we only extracted data on the 2-2½ year version of each tool.

The review was tasked with understanding parents, carers and professionals' priorities and needs for the tool used to measure child development in the 2-2½ year review, in addition to identifying barriers and facilitators of successful ASQ implementation. Our qualitative study generated findings relevant to workforce and skill mix, but we have not made recommendations in these areas.

We undertook a rapid review with systematic searches to identify new evidence published since the last review of this evidence in 2012 [9,10] to answer the following questions:

1. Which structured tools to measure child development at the age of 2-2½ years have been developed or tested since 2012?
2. Which of these tools meet our feasibility criteria for a universal measure of child development at the age of 2-2½ years in England?
3. What do we know about validity, reliability and standardisation of tools that are feasible as a universal measure of child development at the age of 2-2½ years in England?

Methods

We conducted a rapid review to provide timely policy evidence [11–13], incorporating the recommended methods from the Cochrane Methods Group [14] (see Appendix A for more information on how we followed guidance for rapid reviews). We conducted an initial mapping stage to identify all potential tools and then an in-depth review of tools which met our feasibility criteria.

Inclusion criteria to identify tools to measure child development age 2-3 years

We included studies which: were published between 1st January 2012 and 15th September 2022, that described or tested a tool available in English language to measure child development across the major domains

- motor
- cognitive
- communicative
- emotional

for children under 5 years (see Appendix B Table 6).

Feasibility criteria to identify tools for in-depth review

Criteria for feasibility were developed via consultation with experts and with policy colleagues at DHSC (Table 1). Following consultation with policy colleagues at DHSC and health visiting experts, the decision was

made to only include tools with additional versions for use before ages 2-2½ as these stakeholders valued a tool which could be used to track child development over multiple health and developmental reviews across a child's early life. Being able to administer the tool within 30 minutes was considered ideal; between 30-60 minutes was considered of intermediate acceptability, and 60+ minutes was considered unacceptable. We introduced a 'geographical uptake' criterion in order to deprioritise tools that had been developed for a specific non-UK context (e.g., for use only in a Mongolian population), but to still include tools that have been used across multiple non-UK country contexts. Where we found research evidence or documentation about a tool's use in the UK, we categorised this tool as green in 'geographical uptake'. Where we know a tool is used (e.g., in a birth cohort study) but there is no documentation available from which to extract information, we coded this as grey (i.e., no published research evidence available Table 1). We conducted an in-depth review of tools that scored green, orange, or grey (unknown) across all feasibility criteria.

Searches

We tested our strategy against a known set of relevant studies (full details in Appendix C). We searched PubMed, PsycINFO and Web of Science for abstracts in English published between 01/01/2012 and 15/09/2022, using search terms developed from the previous review [9] which captured three concepts: child development AND measurement/tools AND childhood.

See full search strategy in Appendix B Table 7.

Study Selection

We found 13,726 publications, which were first screened for titles and abstract by three researchers from the RREAL team at University College London (GA, AK and GC, https://www.rapidresearchandevaluation.com/). Ten percent of the 13,726 studies were double screened by a member of the research team (JL and RMP) to ensure that that all study team members understood the scope of the inclusion criteria and applied them consistently. Meetings were held regularly throughout the screening process to discuss disagreements and address questions. We identified 429 publications that reported information about potentially relevant tools, of which successfully retrieved 418 full text publications (full text of 11 publications not found; see Appendix D for the PRISMA flowchart detailing flow of studies through the review). As we wanted to prioritise studies that described tools' performance, an additional criterion was applied to classify publications according to the study type (*association, tool development, implementation & acceptability*, *reliability, validation,* or *standardisation,* see Appendix F for data extraction information and study type classification).

Full Text Screening

On this basis, we excluded 326 publications that investigated correlations between child development and outcomes unrelated to the present study's research questions ('association studies') and 6 publications that on full text reading did not meet our original inclusion criteria (see Appendix D for full details). Twelve additional publications were identified from citation searching, resulting in 98 publications being included for further analysis.

Table 1. Criteria for feasibility – Traffic light system

Criterion	Thresholds		
	Green	**Grey**	**Red**
Age	0-3 years*	Data not available	>3 years
English	Yes	Unclear	No
Time	0-30 min	30-60 mins	60+ min
Training related to completion and scoring	Can be administered by parents/ caregiver or practitioner	Unclear	Can only be delivered by specialist with advanced qualification in child psychology or similar such as a child psychologist
Equipment	No specialist equipment needed	Unclear	Some specialist equipment/ stimuli needed such as pictures or building blocks
	Published research evidence on use in UK	No published research evidence of use in multiple non-UK contexts	Tool developed for a specific, non-UK context

* We marked tools as green if they were a tool for use at 2-2½ year but also with additional versions for use before age 2-2½

Selecting tools for in-depth analysis and data extraction

The 98 included studies were then used to generate a long-list of potential tools. We then assessed the long-list of tools on their feasibility for use at the 2-2½ year health and development review (Table 1).

In-depth Data Extraction

For each of the tools assessed as feasible for use at the 2-2½ year health and developmental review, we extracted data on reliability, validity, standardisation, and diagnostic accuracy, where it was reported (see Appendix I for in-depth data extraction tool).

Box 1: What do we mean by reliability, validity, standardisation, and diagnostic accuracy

Reliability

Reliability tells us how consistently a measure produces similar results. Test-retest reliability measures how consistent the tool is for a child or group of children over a short period of time; inter-rater reliability measures how consistent the tool is between two different raters. Validity tells us about the degree to which a measure accurately assesses behaviours and abilities phenomena that reflect the underlying concept being tested [15].

Validity

There are various forms of validity testing, of which we have focused on three: convergent validity tells us the extent to which measurements from one tool correlates with those from another tool that measures the same construct. Conversely, discriminant validity tells us the extent to which measurements that are theoretically not supposed to be related are, in fact, unrelated. Known groups validity tells us the degree to which a tool's measurements are differentially associated with known factors that influence the underlying construct; in this case, early child development (e.g., maternal health during pregnancy, poverty, the richness of the home learning environment).

Standardisation

Standardisation of a tool to measure child development at 2-2½ years of age involves establishing norms or age-based milestones for typically developing children of this age in a specific population. This information can then be used to inform decisions about which cut-offs or threshold scores on a tool to measure child development best distinguish children who are developmentally 'on track' from those with likely developmental delay in the specific population of interest. Different cut-offs on the tool will indicate varying severities of developmental delay.

The most appropriate cut-offs or thresholds for any tool to assess an individual child's development *within a specific setting* will also be informed by the nature of the intervention to be offered to children with each threshold score on the test.

Diagnostic accuracy

A tool's diagnostic accuracy tells us how far a tool identifies true cases of developmental delay and how far it erroneously identifies developmental delay where none exists. Ideally, we want a tool that identifies true delay without erroneously classifying typically developing children as delayed (i.e., we want a tool that is accurate). To capture tool accuracy, we extracted data on sensitivity (the proportion of true positives identified by the tool) and specificity (the proportion of true negatives identified). Sensitivity and specificity of a tool will be specific to populations, influenced by prevalence of the target condition, and are determined by the cut-offs (scores) used to identify delay (see above on 'standardisation'). Threshold scores for detecting delay can be lowered to increase the proportion of all true cases of delay that are identified (increased sensitivity of the tool). However, as sensitivity increases, specificity typically decreases and vice versa. If specificity is low, there will be a high number of children who are identified with developmental delay by the tool but are in fact developing normally (high false positive rate). In this study, we report sensitivities and specificities based on cut-offs of one standard deviation away from the mean to denote mild-moderate delay and > two standard deviations away from the mean to denote severe delay (see Discussion section for limitations). Although there is not consensus, sensitivities and specificities in the range of 70-80% are generally considered adequate in the developmental screening literature [16,17].

Risk of bias assessment

We applied QUADAS-I [18] a tool for the quality assessment of diagnostic accuracy studies, to assess risk of bias in studies reporting reliability and validation of tools. For the complete results see Appendix J. QUADAS-I is not appropriate for use on the other types of studies (see row 3 of Appendix J for an application to an implementation study type). QUADAS-I methodology does not suggest using a final score for assessing quality.

Results

We included 98 unique publications in our mapping phase, describing 34 tools (Table 2). See Appendix G for the list of 34 tools. When we applied our feasibility criteria, we found that 28 of these tools were not feasible to implement at the 2-2½ year health and developmental review in England. Some of the tools in the larger set were not feasible because they did not have multiple versions for use that could be used across the 0–3-year age period (N=11, rows 7-16) and/or did not have an English language version (N=7, rows 15-21). The most common barrier to feasibility of implementation in the 2-2½ year health and developmental review was the level of training and/or equipment needed. For example, although it is widely used as a gold standard measure for the detection of early developmental delay, the Bayley Scales of Infant and Toddler Development (row 33, Table 2) is not feasible for use at the 2-2½ year health and developmental review as it must be directly administered by a highly trained practitioner, using specialised equipment and can also take up to 90 minutes to complete [19].

Six tools appeared broadly feasible to implement at the 2-2½ year health and developmental review in England.

These tools are listed in rows one to six, Table 2, with full text description given in Appendix K, the developmental domains covered in Appendix L and a table of tool characteristics in Appendix M.

The six tools that appeared feasible to implement at the 2-2½ year health and developmental review in England covered the five domains of communication and language, motor skills, problem solving and/or behaviour and personal care (reflecting our inclusion criteria) with PEDS, CREDI and GSED including extra domains (see Table 14 in Appendix L). Of the 40 publications we retrieved, 24 (60%) described or evaluated the ASQ®-3, meaning that this is the tool about which we have most information (see Appendix M for table of tool characteristics). The ASQ®-3 and PEDS have questionnaires suitable for use with babies up to 5½ years (ASQ®) and 8 years (PEDS). The other tools are all designed for use up to 3 years of age. There are three available PEDS [20] tools to screen for early childhood developmental delay. We only focus on PEDS-R and PEDS:DM as we found evidence of them being used separately and together in the literature. PEDS-R enables professionals to view discrete types of risk for developmental delays or disorders, mental health/emotional/behavioural or a combination of both. PEDS:DM is a short developmental milestone checklist, with one item corresponding to each domain, which can be used to predict difficulties by domain. The third tool, PEDS:DM-AL, is designed for children who are at an elevated risk for problems and none of our included publications mentioned the use of PEDS:DM-AL.

Five of the six tools have a version which takes under 20 minutes to complete, including scoring. CREDI and GSED both have short and long versions (see table in Appendix M). The GSED long form is not feasible for use at the 2-2½ year health and developmental review as it requires direct administration by a professional and can take up to 75 minutes to complete. Although we did not find information about how long the sixth tool (IYCD) would take to administer, it has 100 items which is more than double ASQ®-3 (N=43 items, 20 minutes) and GSED short form (N=45[2] items, <25 minutes) and 30% more than CREDI long form (N=69 items, 20 minutes). See Appendix M for more details.

Professionals generally indicated that they would be open to the introduction of a new tool, whilst observing that *"the expectations of a child of that age are not going to change, so it's just how the questions would be worded, really."* [Health visiting professionals group 3]

2 The GSED-Short Form (SF) and CREDI-Log Form (LF) operate on a stop/start basis whereby practitioners start at a certain, age-appropriate point on the tool (i.e., start for 24-27 months on GSED-SF; start for 24-29 months on CREDI-LF) and then keep administering items until the caregiver provides five "no" or "don't know" answers in a row. The number of items given for GSED-SF and CREDI-LF therefore represent the maximum number of items a caregiver could be asked, from the 24-month start point through to the final age band.

Table 2. 34 included tools mapped to feasibility criteria for implementation within the 2-2½ year health and development review in England.

	Measure	Age	English language	Time to administer	Geographical uptake	Equipment needed	Training needed	Total no. papers
		0-3y >3y	Yes No Unclear	0-30 min 30-60 min >60 min Unclear	Evidence of use in UK Developed for a non-UK context Other	No special equipment Some special equipment Unclear	Administer by caregiver/ practitioner High level of specialism needed Unclear	
1	Ages & Stages Questionnaire (ASQ®-3)							24
2	Parents' Evaluation of Developmental Status (PEDS)							3
3	Warner Initial Developmental Evaluation of Adaptive and Functional Skills (WIDEA-FS)							2
4	Caregiver Reported Early Development Instruments (CREDI)							6
5	Global Scales for Early Development (GSED)							3
6	WHO Indicators of Infant and Young Child Development (IYCD)							2
7*	Parent Report of Children's Abilities (PARCA-R)							3
8*	Early Childhood Development Assessment Scale- Caregiver Survey (ECDAS-CS)							1
9*	Brief Early Skills & Support Index (BESSI)							1
10*	Early Childhood Development Index (ECDI)							1

	Measure	Age	English language	Time to administer	Geographical uptake	Equipment needed	Training needed	Total no. papers
11*	Early Years Toolbox (EYT)							1
12	International Development and Early Learning Assessment (IDELA)							1
13	Playful Learning Observation Tool (PLOT)							1
14	McCarthy Scales of Children's Abilities (MSCA)							1
15	The Early Human Capability Index (eHCI)							1
16	Preschool Child Development Inventory (PCDI)							1
17	Mongolian Rapid Baby Scale (MORBAS)							1
18	Taiwan Birth Cohort Study-Developmental Instrument (TBCS-DI)							2
19	The Griffiths Developmental Scales-Chinese (GDS-C)							1
20	The Toddler Language and Motor Questionnaire (TMLQ)							1
21	Cambodian Developmental Milestone Assessment Tool (cDMAT)							1
22	Malawi Developmental Assessment Tool (MDAT)							1
23	Brigance Inventory of Early Development (IED-ii)							1
24	Mullen Scales of Early Learning (MSEL)							3

	Measure	Age	English language	Time to administer	Geographical uptake	Equipment needed	Training needed	Total no. papers
25	Denver Developmental Screening Test (DDST-II)							2
26	Battelle Developmental Inventory (BDI-2)							2
27	Vineland Adaptive Behaviour Scales (VABS-II)							1
28	Rapid Neurodevelopmental Assessment (RNDA)							1
29	The Differential Ability Scales (DAS-II)							1
30	Hawaii Early Learning Profile (HELP)							1
31	The Intergrowth Neurodevelopmental Assessment (INTER-NDA)							2
32	Merrill-Palmer-Revised (M-P-R)							1
33	Bayley Scales of Infant and Toddler Development (BSID-III)							23
34	Australian Developmental Screening Test (ADST)							1
	34							98

* Rows 7-11 were marked red for age because, although they had a tool for use at 2-2½y review they didn't also have additional versions for use before age 2-2½

Three of these tools were designed primarily for individual assessment (ASQ®-3, PEDS-R and WIDEA-FS) and three for population measurement only (CREDI, GSED, and IYCD), as defined by the instrument developers.

ASQ®-3, PEDS-R and WIDEA-FS have been designed as tools for use with individual children to detect developmental delay using established cut-offs based on population norms (See Box 1 for explanation of cut-offs and population norms). These individual assessment tools could also act as a population-level measure of child development.

Appendix M provides characteristics for each tool which relate to feasibility of implementation, including cost, scoring, level of training needed, administration time, cultural adaptability and whether a digital version exists. Of these, the ASQ®-3 seems most feasible for implementation in our setting, as one might expect given it is currently the tool mandated and licensed for use at the 2-2½ year health and developmental review in England. PEDS also appears feasible as it is quick to administer and score and involves minimal training. We did not find enough information on WIDEA-FS to be able to assess feasibility in any detail (see Appendix M).

The other three tools (CREDI, GSED, and IYCD) have been designed and tested as tools to collect data across populations. The developers of the CREDI and GSED specifically state that their tools should not be used for an individual-level assessment of a child or to trigger action or referral pathways based on scores and cut-offs. The stated purpose of these tools is to compare child development between populations, and countries, over time and evaluate policies and interventions. These three population measures are **free to use**, without licensing requirements.

Two of these tools provide domain-specific scores: ASQ®-3 and CREDI-Long Form

ASQ®-3 (individual level assessment tool) and CREDI-Long Form (LF) (population measure) have the advantage of producing domain-specific scores. The ASQ®-3 (43 items) and the CREDI-LF (69 items) share 15 question items (see Appendix H for table showing how far questions in each of the six feasible tools overlap).

Because PEDS is a pass/fail screening test and thus cannot show where a child is on a distribution of development, and as the CREDI-SF and GSED only provide a global score (rather than a score for each developmental domain), these tools cannot provide detailed information about populations cross-sectionally or over time. We did not find enough information on WIDEA-FS or IYCD to judge suitability of scoring (see Appendix M, Table 15).

In non-UK contexts, all six tools performed reasonably well in terms of reliability and correlation with other gold-standard measures.

We did not find any studies that investigated consistency and correlation with other measures in a UK setting. From the included non-UK studies, all six tools performed well in terms of measurement reliability (see Appendix N for full details). There was good to excellent consistency between two or more 'raters' who administered the tool in the same group of children (inter-rater reliability) and for the same 'rater' using the tool at two different, closely spaced, points in time (test-retest reliability). All tools had good internal consistencies on their respective domains (i.e., Cronbach's alpha >.75, [21]), indicating reliable measurement of specific domains of child development (i.e., cognitive, gross/fine motor, communication personal-social). See Appendix N for full details of internal consistency. Where low internal consistencies were found, this tended to be in the context of validating translations of the ASQ®-3 into a different language[3].

In terms of convergent validity, all six feasible tools demonstrated significant correlations with other well-established measures of early childhood development including the Bayley Scales of Infant and Toddler Development (BSID-III), the Vineland Adaptive Behaviour Scales (VABS-3) and the INTER-NDA. Associations ranged in strength from low (i.e., <.50) to

3 Two studies found below-acceptable internal consistencies for ASQ®-3 , both of which were validations of ASQ®-3 translations (Spanish, ICCs .37-.68 by domain, [22]; Italian, ICCs .58-.72, [23] which was investigated using the Griffiths Scales of Child Development, Third Edition, the Peabody Developmental Motor Scale, Second Edition, and the Developmental Profile, Third Edition tools). In order to evaluate discriminant validity, differences between typical development children and several clinical groups have been performed. Finally, two different cut-off scores have been proposed. Results: The results showed that the questionnaires are composed of high-quality items; the original factor structure has been confirmed and strong Pearson product-moment correlation coefficients between the overall and the total for each domain (ranging from 0.73 to 0.88). One study found below-acceptable internal consistency for CREDI socio-emotional (.66) and motor (.68) domains at age 24-29 months in a sample of children from impoverished regions. [24]

acceptable (i.e., >.50) levels (see Appendix N for full details)[4]. The included studies also provide evidence that CREDI, GSED and IYCD measure child development over and above associated constructs such as children's nutritional status (indicated by height-for-age), home learning environment, household socioeconomic status and care giver education level. (i.e., acceptable discriminant and known-groups validity).

ASQ®-3 and PEDS-R miss the majority of children with both mild and moderate developmental delay but do not often misidentify delay in children who are in fact developing typically; PEDS:DM has the inverse profile.

The three tools (ASQ®-3, PEDS and WIDEA-FS) that have been designed for use with individual children to detect developmental delay have been tested for performance as a screening instrument to identify children for monitoring and/or more in-depth assessment or intervention (e.g., activities to practise at home or referral to specialist services depending on score). Seventeen of our included studies reported sensitivity and specificity of ASQ®-3, PEDS, and WIDEA-FS, although we excluded two studies which were judged to be low quality in our quality assessment ([25,26], see Appendix J for detailed quality assessment). In Appendix O, we report detailed data from non-UK studies that used general population samples and from studies that used at-risk subsamples. Where results were stratified by age, for clarity we have reported findings most relevant to the 2-2½ year health and developmental review (around 24-30 months).

We did not find any publications reporting performance as a screening test for the three tools that are designed to only measure child development at a population level: CREDI, GSED or the WHO IYCD. This is to be expected, given that the authors of the tools specifically caution against their use as screening tests. However, we understand that work is currently being conducted by the CREDI and GSED teams to produce data on tool performance in specific populations (personal correspondence, September 2023).

ASQ®-3

The evidence on the ASQ®-3's ability to accurately detect early delay is mixed, **with findings suggesting that the ASQ®-3 may be better at detecting severe delay than mild-moderate delay.** We found two studies in general populations of English-speaking children, neither from the UK. In one study by Sheldrick et al. of 1,495 9-66-month-olds in the USA [27] the ASQ®-3 only detected 23.1% of children who were confirmed to have mild developmental delay in the younger 9–42-month subgroup, using BSID-III as the gold standard measure of delay (i.e., low sensitivity: 23.1%). In this study, the ASQ®-3 accurately ruled out mild developmental delay in 89% of the sample (i.e., good specificity: 89.4%).

In these two studies, the ASQ®-3 was slightly better at accurately detecting children with severe delay (i.e., scores ≥ two standard deviations below the mean in both ASQ®-3 and the gold standard test), with Sheldrick et al. [27] reporting correct identification in 41% of cases, and Veldhuizen et al. [28] reporting correct identification in 60% of children with severe delay. Specificity remained comparable when detecting severe delay (89.4% [27]; 82%, [28]). The low sensitivities found in these studies may be explained in part by the fact that they used broad age groups rather than stratifying by narrow age bands. Other studies have found that ASQ®-3's ability to accurately detect delay varies across age groups, with some evidence suggesting ASQ®-3 becomes more accurate as the child's age increases across the preschool period [24,29].

The two included studies of English-speaking at-risk subgroups (Noeder et al. [30]; Duggan et al. [31]) concurrent validity of the ages and stages questionnaire (ASQ®-3) reported higher sensitivity for detecting mild delay than found in the general population studies, likely because there was a higher prevalence of mild delay in these subgroups compared to other studies using the general population. In their study of 163 children with congenital heart disease, Noeder et al. [30] found the communication domain to be best at detecting true delay (90%, specificity 84%) with the other domains ranging from 65-77% (specificities 84-92%, see Appendix O). Conversely, Duggan et al. [31] concurrent validity of the ages and stages questionnaire (ASQ®-3 found the motor domain to be most sensitive (50%) amongst a sample of 278 Irish children with low birthweight. This paper also found that ASQ®-3 overall sensitivity increased with severity of delay, with ASQ®-3 identifying 45% of children who scored with mild delay by the BSID-III, which increased to 84% for children with severe delay. Specificity was relatively stable from mild (74.4%) to severe (73.2%) delay.

Findings from non-English speaking studies confirm this pattern, with the ASQ®-3 demonstrating low to moderate sensitivity for detecting mild delay (59%, Schonhaut et al. [29]; 62%, Steenis et al. [32]) and

4 Strength of associations with reference standard, across domains: ASQ®-3= .12-.67; PEDS-R= .26-.59; PEDS:DM= .22-.48; WIDEA-FS= .26-.65; CRED= .12-.93; GSED= .88-.97. ICYD: NR.

severe delay (21%, Agarwal et al., [33]; 61.5%, Steenis et al. [32]) in general population samples, and a much stronger ability to detect mild (87%, Kerstjens et al. [34]; 80-86%, Schonhaut et al. [29]) and severe (84%, Agarwal et al. [35]; 100%, Kerstjens et al. [34]; 71%, Simpson et al. [36]) delay in at-risk subsamples. Most of these studies demonstrate good (70%, Steenis et al; 76%, Kerstjens et al.) to excellent (84-86%, Schonhaut et al. [29]; 84%, Steenis et al. [32]; 92%, Simpson et al. [36]; 97%, Agarwal et al. [35]) specificities, indicating that the ASQ®-3 does not tend to incorrectly identify delay in typically developing children[5].

PEDS-R and PEDS:DM

PEDS-R demonstrated low sensitivity for detecting mild delay (28%) and good specificity (78.9%) amongst American 1-42-month-olds [27], according to the BSID-III. The same study reported PEDS-R as much more able to accurately detect severe delay (78.9%, specificity 79.6%). PEDS:DM, a shorter version of PEDS-R designed as a developmental milestones checklist, demonstrated moderate sensitivity (67.2%, 60.8%) but poor specificity (42.7%) for detecting mild and severe delay. Two studies looked at the use of PEDS-R and PEDS:DM together [27,39]. Sheldrick et al. [27] found the combined PEDS tools to have low sensitivity for detecting mild delay (22.7%) but much better sensitivity for detecting severe delay (78.9%), with good specificity (83.9%). Conversely, du Toit et al. [39] found the combined PEDS tools to have excellent sensitivity (92.6%) but low specificity (22.5%) for detecting mild delay amongst a sample of 276 South African 36–83-month-olds, according to the VABS-3, suggesting that, in this context, the combined PEDS tools identify the majority of cases of delay, but may also incorrectly identify delay where none is present.

WIDEA-FS

Only one paper reported sensitivity and specificity data for the WIDEA-FS, from a sample of North American 10-36-month-olds who had been born prematurely [26,40]. No pre-established cut-offs exist for the WIDEA-FS; as such, Youden's Index was used to determine cuts-off for optimal sensitivity and specificity for each domain (see Appendix O).

3. Stakeholder Focus Groups

Methods

We held 15 focus groups (of which six were online) with key stakeholders to establish priorities for, and barriers to, successful administration of a measure of child development at the 2-2½ year health and developmental review (Appendix P for participant recruitment per site). In-person groups occurred across four sites in England, chosen from within our existing networks from geographical areas with a range of levels of child deprivation (% children in absolute low-income families in 2019/20: Site 1: 20-25%; Site 2: <15%; Site 3: >25%; Site 4: 15-20%). We recruited participants purposively through charitable organisations and gatekeepers known to the research team.

Focus groups were audio recorded, transcribed verbatim and anonymised by a professional transcriber. Transcripts were analysed using Reflexive Thematic Analysis (RTA, [41]). RTA is a qualitative analysis method that enables the identification and analysis of patterns within a dataset. RTA has been identified as particularly appropriate for public health research and dissemination, as it allows for the flexible yet systematic identification and analysis of patterns that arise amongst lived experiences within public healthcare systems, from both service users' and service providers' perspectives [42–44]. RTA is a multi-stage process, during which features of a dataset are systematically coded and sorted into themes. An inductive approach was used in this analysis whereby codes and data were generated from the data through unrestricted coding. Themes were generated via a flexible, rigorous, and recursive process, during which the quality of the analysis process was ensured via data audits with members of the research team (JW, HC).

Stakeholders' views on the digitisation of a measure of child development was specified as a priority area for further investigation in this review. As such, a deductive Framework Analysis (FA), a qualitative method widely used in social policy and health research [45,46], was conducted separately, to examine stakeholders' perceptions of tool digitisation. FA allows the systematic categorisation of large volumes of qualitative data and produces highly structured outputs of summarised information, making it well-suited to dissemination to policy and lay audiences [47]. A deductive approach,

5 Agarwal et al. [35] report a specificity just below the acceptable threshold at 66%. Two further studies, Yue et al. [37,38] find poor sensitivity (2.2-50%) and moderate to good specificity (49-94%) across domains in a sample of children from rural, low-income regions of China.

where code and theme selection is driven by pre-existing theory and/or the specifics of the research question, has been identified as particularly effective in time-limited, applied research projects [48]. The sections of each transcript that related to digitisation were extracted and then deductively coded, and codes used to create an initial coding framework that was applied to the remaining transcripts. Data were charted into a framework matrix which was used to develop themes. Codes, framework, and candidate themes were discussed at regular intervals with members of the study team (JW).

Appendix Q provides details of discussion points for all focus groups.

Results

We spoke to 28 parents in our focus groups, with a broad range of self-reported ethnicity and household income, with an age range of 24 to 47 years (see Appendix R for detailed demographic data for parent participants).

Stakeholder priorities for a measure of child development at the 2-2½ year heath and developmental review

Theme 1: Just a part of the puzzle: Implementation of the tool

1.A: A useful tool in a holistic review

Summary

- Having a conversation about health and wellbeing across the family system was the **principal priority** for both parents and health visiting professionals.
- Parents, health visitors, and nursery nurses felt that a group health and developmental review format made it harder for parents and practitioners to engage in this holistic conversation.
- Professionals perceived the tool to be most useful as a conversation starter rather than a tick-list.
- Practitioners reported that how the tool is used depends on the experience and confidence level of the practitioner.
- Parents, practitioners, and DHSC policy colleagues felt strongly that a parent-reported digital tool should enhance and not replace the in-person review but also gave examples of how this might happen.

Having a conversation about health and wellbeing across the family system was the **principal priority** for both parents and health visiting professionals. They emphasised the importance of having the time and space to conduct other aspects of the assessment such as health promotion, checking other milestones (e.g., potty training) and addressing any parental concerns, including concerns about family life. Parents also expected and appreciated consideration of their own wellbeing as part of the review:

"It's about touching base with parents and saying 'how are things going? What can your child do?... kind of thing. Rather than doing a measurement on a score." [Health visiting professional group 2]

"I would expect just a warm conversation about how your family-, […] And kind of having a conversation on what's going on with you, maybe. [Parent group 1]

Parents told us that the ASQ®-3 can hinder tailored, holistic care when it was used as a "*tick-sheet"* [parent group 5] and that this detracted from their experience:

"I didn't really get much out of it. It was… just a tick box of questions, I don't think anything specifically geared to the individual child or an individual family." [Parent group 2]

Equally, concerns were raised by practitioners; although the ASQ®-3 is "just a very small part of [their] review" [health visiting professionals group 1], it is perceived by parents as a test which distracts from the broader holistic assessment. Health visiting professionals emphasised that directly observing children and using their professional judgement was crucial in assessing children's development at the 2-2½ year health and developmental review. Practitioners, policy colleagues at DHSC and parents all saw the utility of a child development measure as part of the review. For example, it can increase "parental involvement" [health visiting professionals group 3] by helping parents better understand whether their child needed any encouragement in some areas and what parents may do to support the child's development (e.g., ASQ®-3 has suggested family activities to support development in certain domains). One policy colleague described this as: "supporting them [parents] to support their child's development" [DHSC policy colleagues group]. Health visiting professionals further framed the ASQ®-3 as helpful when used as a springboard for conversations and reflections on the child's development. Parents reflected that receiving the ASQ®-3 ahead of the review appointment enabled them to engage more closely with their child's development in their own time, helping them "take stock" [parent group 5] of their child's capabilities. Some parents also indicated that having a child development *"questionnaire"* helped "*start the conversation*" [parent group 6] with the health visiting professional at the review appointment, and that the use of the ASQ®-3 within the context of the review appointment made them feel like "*I did gather some information about [child], more than if… you'd gone in blind* " [parent group 6].

Across all five health visiting professionals focus groups, professionals emphasised that how the tool is used depends on the professional's level of experience and confidence. Experienced health visitors talked about using their knowledge to judge whether a child scoring in the grey, i.e., monitoring zone on the ASQ®-3, is actually a cause for concern, whereas *"someone who's not as experienced might think, 'oh, I must follow that up' […] and probably caused concern when there's not any there"* [health visiting professionals group 3]. Trainee health visitors reinforced this perspective, explaining that as confidence and experience grow, the ASQ®-3 questions become less central to how they conduct the review. Experienced nursery nurses described learning to use their judgement to get *"the balance between what's on our agenda and what's on [the parent's] agenda…"*, whilst other nursery nurses described difficulty in navigating the tool and the holistic review:

"A parent reported tool is really tough because when you're in a visit with a child and you're looking and thinking, 'you're definitely not meeting your milestones […] but your parent's ticked 10 on everything and I have to score you 60 out of 60.'" [Nursery nurses group 1]

Parents also picked up on this, often expressing frustration when the "*health visitor just follows what's in their papers"* and expressing a desire for "*a professional to find out from that child something more, something hidden*" [parent group 2]. Both parents and health visiting professionals highlighted that how the ASQ®-3 is used can vary depending on the format of the review. Practitioners and parents considered that a 1:1 review, preferably at home, was ideal because it allows space and time for a thorough, holistic review. Whilst a minority (*n*=2) of health visiting professionals voiced conditional support for group reviews, the majority of practitioners and parents we spoke to raised concerns about the extent to which a thorough, holistic review can be achieved in a group format. Parents said they would be less likely to "*open up*" [parent group 1] about problems with other families in the room, and that *"it would raise issues for a lot of parents in terms of being able to talk about their needs"* [parent group 4].

Health visiting professionals observed that it's often "*just not appropriate*" [health visiting professionals group 2] to ask questions about parental mental health in a group setting and that the constraints of the group review setting can result in treating the ASQ®-3 as a

tick-box exercise which ultimately misses issues that may or may not be picked up by private provision, i.e., nursery or a childminder, further down the line:

"Especially with the group review setting, it's just tick, tick, tick, 'okay they're fine,' and then a few months later you get a call from nursery saying, 'oh I'm worried about this child'." [Health visiting professionals group 3]

"There should be an opportunity to speak to the health visitor. Being put in a big room, being given the questionnaire back and nobody talking to me… was kind of strange… having a conversation would have been nice." [Parent group 5]

Parents and local authority professionals were cautiously positive about how a digital offering might help make the service more efficient, emphasising that a digital offering "*should not replace the conversation with the health visitor"* [parent group 1]. A digital tool to measure child development was explicitly and consistently framed as something that could enhance, but should not replace, the 2-2½ year health and developmental review.

However, we also heard messages about how a digital tool might be used to prioritise families for the universal review (parents and health professionals) or to *"see how far [we] can stretch the mandation [sic] […], not only with skill mix, but actually with digital options, face to options, questionnaire options"*. This quote from a professional in one of our focus groups implies that a digital option could replace the in-person health and development review for some families.

The over-riding perspective across all participants was that whilst a useful tool, the ASQ®-3 "*is just part of the puzzle"* [nursery nurses group 2] and should only ever be considered part of a holistic health and development review which must involve direct observation and an in-person contact.

1.B : Parent-report versus Professional judgement

Summary

- Direct observation of the child by the professional is a priority among parents and health visiting professionals.
- The parent-reported ASQ®-3 is a way of understanding how the parent sees the child, and how much help the parent needs to support their child's development.

A related priority for practitioners and parents was that caregiver-completed ASQ®-3s should always be checked by a trained professional through observing the child, in conjunction with a discussion with the caregiver. Policy colleagues at DHSC reinforced this perspective, asserting that including parents in the assessment process is an important opportunity to help them engage in health-promoting activities, but that final judgements must always be made by "*a qualified health professional*". Although some parents felt that as they *"know [the] child best, it makes sense for me to fill it [the ASQ®-3] in"* [Parent group 1], some parents *"didn't really trust ourselves to know"* [parent group 6] and so valued professionals using their own judgement to help make an objective assessment. A few health visiting professionals acknowledged that parents' perspectives might help make a more objective assessment as "*we don't always get the true child*" [health visiting professionals group 2] at the review appointment. However, the dominant perspective was that, whilst not wishing to "*undermine the ability of the parent*" [health visiting professionals group 2], parents can't always be objective or informed reporters.

Ultimately, the consensus was that whilst the ASQ®-3 *"gives you a starting point to talk to the parents about"*, it is the conversation with the parent, combined with *"us observing that child and going through what we see at the same time"* [nursery nurses group 1] that enables health visiting professionals to make an informed assessment of the child and help the parent best support their child's development:

"If they've answered 'not yet' or 'sometimes'… you would revisit that […] if they've got any concerns written on the form, you would flag that up and discuss how to address that". [Health visiting professionals group 1]

1.C: Clarification on the purpose of the tool: "they need to know why they're doing it"

Summary

- Parents are unclear about the purpose of the 2-2½ year health and developmental review, which can lead to unrealistic expectations/ feeling disappointed.
- Parents don't necessarily understand why the ASQ®-3 questions ask what they ask, which can lead them to worry.
- Professionals emphasised the need to promote the purpose and benefits of the Healthy Child Programme to help improve attendance and manage parents' expectations.

A final priority for parents, health visiting professionals and policy colleagues at DHSC was to clarify the purpose of both the 2-2½ year health and developmental review and the ASQ®-3 within the context of the review. A pervasive experience voiced by parents was that they didn't know *"what to expect, I guess, and what to look out for before the actual appointment"* [parent group 2]. This included confusion about how the health visiting system differs from and fits in with other services, what is being assessed at the review and why, and what support families can expect to receive as a result of the review.

Health visiting professionals also recognised that many parents were unclear around the point and usefulness of both the ASQ®-3, and the 2-2½ year health and developmental review more broadly, including how it fits with nursery-led assessments:

"They need to know why they're doing it, yeah... like often we won't necessarily ring the parent before, they're just literally getting a letter with an appointment date on saying 'fill this form in'." [Health visiting professionals group 2]

"Nothing is really ever advertised about the benefits of ASQ®-3 forms and why parents need to come to all of the healthy child checks […] I think if it was advertised more as to the benefits maybe they would all be turning up." [Nursery nurses group 1]

It is worth noting, however that even parents who were motivated to attend a 2-2½ year health and developmental review reported challenges of getting an appointment in a stretched service: describing *"a constant fight"* [parent group 4], having to "*chase up the appointment because they were so overbooked*" [parent group 6] and being "*put on the back burner*" [parent group 2] because the service was short-staffed.

Parents and health visiting professionals highlighted multiple types of misunderstanding about the purpose of the ASQ®-3, from parents believing the ASQ®-3 is a diagnostic tool (e.g., thinking it's intended to diagnose Autism Spectrum Disorder), to a more generalised lack of understanding of why each question asks what it asks. Not understanding the ASQ®-3 questions could often be "*really anxiety inducing*" [parent group 5] for parents, leading some to worry about their child's development. Professionals suggested that this is indicative of a need to engage more closely with parents about child development early on, so that *"when they come to [the ASQ®-3], it isn't a surprise that your child needs to be using a pen or doing something"* [health visiting professionals group 3]. Professionals also highlighted that parents can confuse the 2-2½ year health and developmental review with the Early Years Foundation Stage (EYFS) two-year progress check, or otherwise struggle to see the value of the review when their child is also receiving assessment at nursery:

"Like nursery… if they haven't raised any concerns, [parents] say 'well… I haven't seen you in a year and a half since you done the 1-year check, so why do you need to come out and do a developmental review? They're in a provision now." [Health visiting professionals group 2]

Health visiting professionals explained to us that for the vast majority of families, the 2-2½ year health and developmental review (including the ASQ®-3) essentially *is* the intervention because it allows a conversation about what the family can do to support the child's optimal growth and development and can reassure parents. However, for many parents there was an expectation that, if concerns were to arise in the course of the review, their child would be referred to another service. When the health visiting professional suggested further monitoring of the child rather than intervention, parents could feel as if their concerns has been dismissed:

"I said… 'oh, my son is 26 months and he couldn't say a word'… and then she's only saying, 'oh, wait until he is 30 months… then let's see how it goes.' That's the way it is when I am addressing my concern and it's… like, okay, is this it, so I just have to wait?" [Parent group 4]

When they weren't referred to any service, parents described feeling like health visiting professionals *"can't help you with anything"* [parent group 2]. Policy colleagues at DHSC highlighted the need for clarification of intervention and referral pathways based on the 2-2½ year health and developmental review (including the ASQ®-3), observing that *"it's not always clear… what the next step is once [parents] have done the questionnaire."*

Theme 2: Dual use: Data versus Individual assessment

Summary

- All but two of the 24 health visiting professionals seemed unaware of ASQ®-3's intended function to provide population-level data; parents also wished for clarity on this point.
- Professionals viewed the ASQ®-3's primary function as an individual assessment tool.
- Parents and professionals demonstrated some interest in having access to regional and national-level data on child development.
- Policy colleagues at DHSC wanted a dual-purpose tool. They acknowledged the dual purpose may be difficult to achieve in practice but cautioned against separating population-level data collection from the broader developmental review.

Of the 24 professionals interviewed, only two (both nursery nurses) were aware that ASQ®-3 data are also intended to provide a population-level measure of child development, although one could not recall how she knew this: *"I don't know how I came across it"* [nursery nurses group 1]. Typically, health visiting professionals framed the ASQ®-3's primary function as a means to help detect early developmental delay in order to trigger support pathways where needed, and otherwise *"assumed [ASQ®-3 data] was just for our patient record"* [health visiting professionals group 2]. Parents also framed the ASQ®-3's main function as to trigger referrals and demonstrated a lack of clarity around whether ASQ®-3 data are used for anything other than individual-level assessment.

A minority of health visiting professionals also mentioned that the ASQ®-3 can be a helpful resource when triggering referral pathways in that it is a well-known tool recognised by other health professionals (i.e., it helped with acceptance of referrals into other services). However, professionals from another locality had opposing views, observing that *"quite often you have to explain to a paediatrician what an ASQ®-3 even is"* [nursery nurses group 1]. Parents' lack of clarity on the function of the ASQ®-3 had the potential to trigger parental anxiety:

"For me […] I felt like it was quite anxiety inducing, because I'm not entirely convinced that the point was to check if your child is meeting milestones, or if it's more of, like, a national survey." [Parent group 6]

Nevertheless, when asked by our interviewer, parents and health visiting professionals supported population-level data collection. Parents suggested that knowing how their child compared to a national average would help them contextualise their child's ASQ®-3 results, thus reducing ambiguity and legitimising practitioners' feedback. Health visiting professionals indicated that information about local and national trends would help inform their individual practice. Further, some understood that population level data could be used to evaluate intervention and services, including health visiting.

Policy colleagues at DHSC generally felt that any measure of child development needed to fulfil both purposes: first, to provide an outcome measure for early years investments that would help monitor trends and disparities across regions; and second, to enable the early detection of developmental delays across age groups through the early years. A key priority for policy colleagues at DHSC was establishing whether a single tool could fulfil these two purposes, and, if so, to establish the best way for this to be achieved. One policy colleague suggested that, as the two intended purposes of the tool are different that the assessment of the child could be separated from the data collection tool. However, several policy colleagues at DHSC emphasised the riskiness of separating population data collection from the 2-2½ year health and developmental review. It was suggested that separation of the ASQ®-3 (for data collection) from the broader developmental review would result in less data being available for population monitoring, and fewer developmental assessments being carried out. The ASQ®-3 and the 2-2½ year health and developmental review were thus viewed by some as *"mutually supportive"*, each ensuring the other's success. One policy colleague observed that there are currently "*no examples in the system of data being collected for population health monitoring purposes that is not linked to some sort of clinical activity… it would need a lot of research, I think, to go down a path which separates them."* [DHSC Policy colleagues group]

Theme 3: Barriers to universal reach

3.A: Improving the tool

Summary

- Parents and professionals suggested various aspects of the tool which could be improved.
- Parents, professionals, and policy colleagues at DHSC raised concerns about the extent to which the ASQ®-3 is appropriate for use with children from minority ethnic groups.

Health visiting professionals, policy colleagues at DHSC and parents reflected on various aspects of the ASQ®-3 that they value and which aspects that they would like to see changed (summarised in Table 4). Several professionals mentioned that they valued the free text sections of the ASQ®-3, where parents are invited to volunteer information about whether they have any concerns and what they enjoy most about their child, reflecting that these questions were best at sparking meaningful conversation. Parents and health visiting professionals highlighted broadly similar areas of concern about the questions on the ASQ®-3, indicating length, repetitiveness and outdatedness as issues that needed addressing:

"Some questions are annoyingly repetitive, and it gives the impression that they are checking that you are not lying by asking the same question over." [Parent group 5]

Participants from all groups raised concerns about the ASQ®-3's fitness for use with children from minority groups. Policy colleagues at DHSC stressed that a tool used at the 2-2½ year health and developmental review needs to *"bring equity and fairness for all families"* and expressed concerns that the ASQ®-3 may disadvantage the growing number of families in England that don't speak English as their first language. Similarly, professionals and parents highlighted the ASQ®-3's is, at present, Western-centric and may not be culturally appropriate, noting that some of the questions can subsequently prevent some children from receiving scores that are representative of their ability:

"So where you've got 'can he use cutlery?' There are some cultures that don't use cutlery, so the parents are like, 'well no', so they would fundamentally score nothing. So they might score lower when actually they shouldn't." [Health visiting professionals group 2]

3.B: ASQ®-3, the two-year health and developmental review and children with disabilities: "lost in the system"

Summary

- Parents of children with disabilities valued having a 2-2½ year health and developmental review, but often felt overlooked and neglected by the health visiting service.
- They did not see the ASQ®-3 as appropriate for their children but felt their children should be included in the population-level picture of child development and had suggestions for inclusion.

Four parents of children with disabilities and/ or specialist needs were consulted to discuss their views and experiences of the ASQ®-3 at the 2-2½ year health and developmental review.

Parents of children with disabilities value receiving a 2-2½ year health and developmental review and want their children to *"be included in some way"* [parent 4], both in terms of receiving reviews and being represented in the national data picture. Public health intelligence specialists asserted that it was important for children with disabilities to be included in population surveillance data, but acknowledged that using a universal measure to do so *"is not always going to be appropriate"* and that data need to be collected separately "*to tease out some information for that cohort separately*" [DHSC policy colleagues group].

Parents viewed the review as an important opportunity for support on *"how we're coping as a family"* [parent 1], for receiving professional advice regarding parenting, and for help managing their child's specialist care. Despite this, parents reported often feeling dismissed or overlooked, especially when their child receives care from multiple specialist teams. One mother described this as being *"lost in the system"* [parent 1] and another explained that she felt no-one wanted to take responsibility for her and her child:

"My son had been in the NICU [Neonatal Intensive Care Unit] and I was under the perinatal team so I was kind of like, who do I belong to? And the health visitors were like, 'not me, see ya!'" [Parent 2]

Table 3. Pros and cons of the ASQ®-3 by stakeholder group

	Health visiting professionals	Parents
Pros		
Free text section is useful	X	
Activity sheets that come with ASQ®-3 are useful	X	
Examples of questions (pictures) are useful	X	
Conversation + ASQ®-3 items help parents understand typical child	X	X
Cons		
American tool not appropriate for use in UK population	X	
Tool is repetitive	X	X
Questions are not age appropriate	X	X
Questions are/ wording is outdated	X	X
Questions are not culturally sensitive	X	X
Tool hard for families with English as a second language	X	X
Questions are not inclusive of multi-lingual children	X	X
Too long: socio-emotional scale needs to be integrated	X	X
Questions are complicated/ confusing	X	X

Parents emphasised that whichever tool is used to assess child development at the 2-2½ year health and developmental review should be inclusive of non-typically developing children. Parents suggested the addition of open text boxes designed to allow parents to report details of their child's disability and request specific information (e.g., support with stammering, parent 4). One parent suggested there ought to be space to report language development in non-verbal modalities i.e., sign language.

Ultimately, parents emphasised the importance of receiving contact from the health visiting team prior to the ASQ®-3 being sent out to provide an opportunity to discuss their specific circumstances. One mother described feeling hurt and disappointed when she received the ASQ®-3 as she knew it would not be relevant to her child, and felt that the health visiting team should have been more proactive in getting to know her personal circumstances and following up with her when they didn't receive her questionnaire back:

"I remember at the time […] just out of pure, just, feeling rubbish that day I just threw [ASQ®-3] in the bin. I thought, 'there's nothing to fill in, they should know this.' […] I really thought they'd be back in touch after to find out why they haven't got the review back, but they haven't. So to them, who is [child]?" [Parent 1]

Several professionals reflected on positive experiences with diligent administrative staff, who were adept at flagging specialist cases in order to *"see if it was appropriate to send ASQs"* [health visiting professionals group 2]. However, some health visiting professionals noted that there are no automated processes that add notes to clients' records indicating contextual flags and that whether an administrator would typically manually look at a client's notes to identify specialist needs varied from practice to practice. Some professionals preferred to book appointments themselves so that they could "*have that conversation with the client to see what's happening"* [health visiting professionals group 1], although acknowledged that due to limited time and resource, this wasn't always possible. Nursery nurses reflected that receiving an ASQ®-3 when a child isn't hitting typical developmental milestones *"can cause a lot of upset for parents"* and that it can lead to *"starting the whole review saying 'don't worry about that, let's put that [the ASQ®-3] away please…before you can even do anything else."* [Nursery nurses group 2]

Adopting and implementing a digital tool

Towards the end of each focus group, the interviewer asked the group about their views on a digital tool to measure child development. In addition, we spoke to five staff from two local authorities who were actively considering implementing a digital tool. Table 4 summarises the key perceived benefits of and concerns about the introduction of a digital tool for parents, health visiting professionals, policy colleagues at DHSC and local authority colleagues. Stakeholders perceived the benefits of a digital tool to be possible improvement of the user experience, modernisation of the system and improved automation and efficiency. Conversely, stakeholders expressed concerns that a digital tool may in time replace valued in-person contacts, may exclude certain groups of people, may be impractical, be incompatible with pre-existing systems, be costly, and would require thorough training and staff motivation. See Appendix Q for an in-depth description of each perceived benefit and concern.

Table 4. Perceived benefits of and concerns about introduction of a digital tool

	STAKEHOLDER GROUP			
PERCEIVED BENEFITS	**Parents**	**HV professionals**	**DHSC policy colleagues**	**Local authority colleagues**
Improving the user experience				
Constituting part of a centralised digital location for child health and development information	X	X		X
Improving accessibility by providing translation options and video/aural examples	X	X		X
Facilitate easier communication and improve relationship between parents and HVs	X	X		
Modernisation				
Maintaining service reputation and bringing service in line with national digitalisation agenda	X		X	X
Reduce cost (environmental and monetary) and risk of postage.	X	X	X	X
Automation & efficiency				
Improve service efficiency by sending scored questionnaires to HV professional before review	X	X	X	X
Improve quality of data and dataflow	X			X
Achieve greater reach with less resource				X
CONCERNS				
Replacing in-person contacts				
Digitisation in an under-resourced service - could mean that an online tool replaces the essential in-person review	X	X	X	X
Universal reach				
May exclude certain groups e.g., those experiencing digital poverty / areas with poor connectivity	X	X	X	X
Practical difficulties				
ASQ®-3 too long to comfortably use in a digital format	X	X	X	
Incompatible data management systems				
Digital tool will need to integrate with plurality of data management systems across country		X	X	X
Cost and licencing				
If the tool is licenced then negotiating a digital license might be complicated, high-risk, and costly			X	X
Staff training and motivation				
Staff will need training and motivating to use digital tool in intended way			X	X

4 Discussion

In our study, health visiting professionals and parents gave a consistent message about their principal priority for a tool to measure child development at the 2-2½ year health and developmental review: **these stakeholders valued a parent-reported child development measure that was combined with direct observation by the professional and which facilitated an in-person conversation about the child's development, and health and wellbeing across the family system.** All but two participants felt that this could not happen in a group review and practitioners and DHSC policy colleagues were clear that a digital tool, while beneficial for service efficiency, should not replace the in-person contact with direct observation of the child as part of a holistic health and development review. The prioritisation of parents and practitioners working in partnership to have a holistic conversation, facilitated by the child development tool, is strikingly consistent with the views and experiences of 85 professionals and 40 parents collected for the previous review [2] on this same topic in England. Our study and the wider evidence-base suggests that **the most appropriate tool to measure child development in the 2-2½ year health and developmental review will be one that has clinical utility and robust measurement properties for each individual child and can simultaneously collect data which can be used to analyse trends, disparities and the impact of policies and interventions.**

From our review of 34 tools available to measure child development at 0-3 years, we found two tools that met feasibility criteria for implementation at the 2-2½ year health and developmental review in England: ASQ®-3 and the CREDI Long Form (LF). Both appear to be suitable measures for collecting domain-specific population-level data: high quality studies report that they are both reliable measures and have fair to good agreement with other validated measures of child development (24 studies on ASQ®-3 and six studies on CREDI). **However, we found no research evidence reporting validation or standardisation of either tool in a UK setting, which is a key limitation.** There is also no UK-based evidence for either ASQ®-3 and CREDI-LF for use as an individual-assessment level tool, limiting conclusions that can be drawn about their utility for assessing child development in an English setting despite the fact that we know ASQ®-3 is used in this way in England [8]. Our findings are consistent with another recent review [49] that rated CREDI, ASQ®-3 and PEDS as the best tools to measure early childhood development (0-3 years) out of 27 reviewed tools, based on psychometric quality, cultural adaptability, practicality of administration and clinical utility. Although the PEDS tools met our feasibility criteria for implementation at the 2-2½ year health and developmental review, they do not provide continuous scores for each domain of child development, but rather provide categorical outcomes in domain subgroups, making them less suitable than ASQ®-3 and CREDI for assessing and collecting data on different domains of child development.

ASQ®-3 has some advantages over CREDI in that there is a far more advanced evidence-base and, as the current tool in England, has existing training and implementation materials provided by NHS England for an English context (https://portal.e-lfh.org.uk/). There are well-described challenges in health visiting services, with high demand, insufficient health visiting workforce and stretched budgets. A significant change to existing processes, such as changing a universal tool, will require resource and strategies to gain buy-in from professionals and manage the process of change. However, a key advantage of CREDI is that it is open-source and free to use.

The evidence on the performance of ASQ®-3 as an individual assessment comes from outside England and we do not know how transferable these results are to an English population. This non-UK evidence reports that ASQ®-3 is better at detecting more severe delay in groups of children at high risk of delay such as children born preterm (sensitivity 84-100%), than mild to moderate delay in general population samples (sensitivity 23-62%). In these non-UK general population samples, and using a cut-off of one standard deviation from the norm, ASQ®-3 identified 23-62 out of every 100 children who had developmental delay as determined by another gold standard reference tool. The rates of false positives (i.e., developmental delay identified in children who were in fact typically developing) amongst the general population studies ranged from 3-17% for severe delay to 11-32% for mild-moderate delay. The previous review on this topic [9] is consistent with these results, reporting that ASQ®-3

detects 86-100% of all severe developmental delay (with a false positive rate of 7-13%) and citing papers that report 39% sensitivity for mild-moderate delay in a small (n=53) general population sample (false positive rate 7%) [50].

Although the sensitivity of ASQ®-3 for detecting mild to moderate developmental delay at age 2-3 years may appear low, there are known difficulties with all efforts to detect mild to moderate delay in early childhood development, as *"enormous variability is a feature of early cognitive, language, motor and behavioural development"* [51]. The gold standard reference test used by most of our included studies was the Bayley Scales of Infant and Toddler Version 3, a clinician-administered instrument that takes approximately 70 minutes to administer [52]. Even this gold standard tool is known to underestimate mild to moderate developmental delay in children aged 1, 2 and 3 years of age [51,52], although Version 4 is reported to be more accurate [52]. The same pattern of increasing accuracy with age of child is seen for ASQ®-3 [24,29]. Moreover, the two most relevant studies on ASQ®-3 that we found (in English-speaking general population samples, Sheldrick et al. [27]; Veldhuizen et al. [28]) gave aggregate results on ASQ®-3 performance for children aged 1-42 months. The performance of the ASQ®-3 in the 2½-3 years age range is likely to differ (higher or lower) from the aggregate value reported in these studies. It is also worth noting that the ASQ®-3 may detect an even lower proportion of children with mild to moderate developmental delay than reported in our included studies, given that the reference standard is also likely to miss some cases of developmental delay.

Furthermore, it is important that any tool selected for use has been developed using modern psychometric methods. There are known issues with classical test theory, including inaccuracy, imprecision, and misleading scores [53–55]. Modern psychometric methods, based on item response theory (IRT) and/or Rasch measurement theory have been adopted for tool development in recent years. The ASQ®-3 Technical Appendix [56] states that IRT was used when ASQ was revised to its current version, but no information about this was reported in our included studies, which were more recent. We know that CREDI was developed using IRT [57]; however, as the CREDI was designed as an internationally comparable population-level measure, there is no evidence on its use at the individual level. Work is underway by the CREDI team to assess its use as an individual-level assessment and to develop cut-off scores. As developing score thresholds across cultures will necessarily be complicated, it is anticipated that this process will take some time (personal communication with the Harvard Team, October 2023). Future work is therefore needed, ensuring use of modern psychometric methods, and a representative UK-based sample, in order to determine valid and robust cut-offs for individual-level assessment of child development in England. Work is also needed to establish the whole 'package' of the child measurement tool, with appropriate materials for parents (as recommended in the NHS England training) and with agreed intervention and support pathways outlined for specific cut-off scores on the tool, taking into account the expected numbers of children who will have each score across England.

The sensitivities and specificities we report for ASQ®-3 are based on cut-offs of one standard deviation away from the mean to denote mild-moderate delay (the ASQ®-3 "monitoring zone") and > two standard deviations away from the mean to denote severe delay, in line with recommendations from the developers of ASQ®-3 and BSID-III [58,59]. However, it is possible to modify the sensitivity and specificity of a tool by using different cut-offs (scores) to identify delay. If a lower threshold for developmental delay is used, the tool will detect a higher proportion of children with delay (high sensitivity) but this will likely result in higher numbers of children with typical development being identified as delayed (i.e., higher false positives, lower specificity). A minority of studies [32,60] investigated the *optimal* balance between sensitivity and specificity of ASQ®-3 in their given population using receiver operating characteristics (ROC) and area under the curve (AUC) analyses to calculate cut-offs for indicating developmental delay, rather than using one and/ or two standard deviations below the mean. This approach is likely to be useful in further investigation of the performance of ASQ®-3 and/or CREDI in samples of children from England, in order to generate standardised scores and cut-offs based on distributions of early development among children in England. Notably, a version of the CREDI is being used in a subsample of a representative cohort of approximately 3,000 2-year-olds in England as part of a new birth cohort study, the Children of the 2020s Study [61], which provides opportunities for some validation and standardisation of the CREDI in a UK setting. These children will have the CREDI administered at age 2 (in late 2023) and 3 (in 2024) and will be administered a reference test at age 3-years and 2-months (personal communication, Pasco Fearon).

Optimal sensitivities and specificity will also depend on the intervention package that follows any given cut-off on the tool. For example, higher rates of false positives (to achieve high sensitivity) may be acceptable if the intervention pathway is a light touch and low-cost intervention with minimal anticipated harms (e.g., advice to parents or monitoring). However, even these types of intervention can cause unintended harms. Our qualitative findings indicated that the tool can trigger anxieties in parents about their child. Another study found that parents could be affronted by advice which suggested (or could be misinterpreted to imply) that they hadn't tried talking to or reading with their child to improve their speech and language [2]. This is one reason that both parents and professionals valued the conversation between a member of the health visiting team and parent, to carefully explain and make meaning of the results for the parent in the given context of their family.

The focus on a holistic review by practitioners and parents in our study highlights that contextual information and professional judgement is likely to be highly valued and informative both for identifying developmental need and in assessing which type of support might help. This echoes findings from the previous review [9] and findings from the development of the Early Language Identification Measure (ELiM), a measure designed to evaluate children's speech, language, and communication needs at the 2-2½ year health and developmental review in England [62]. Whilst there are differences between ELiM, ASQ®-3 and CREDI (specifically that ELiM is a single-domain measure focusing on early language ability whereas ASQ®-3 and CREDI measure multiple domains of early development), findings from the development of ELiM are pertinent to the present discussion. ELiM was specifically designed for use by health visiting professionals at the 2-2½ year health and developmental review, and includes both a parent-reported vocabulary checklist, and direct observation of children's language and cognitive abilities by a health visiting professional. ELiM's creators found that the combination of a parent-administered vocabulary list and professional observation resulted in the best balance of sensitivity and specificity (.98 and .63 respectively) for detecting early language delay at 2-2½ years [63]. This study concluded that *"the key is the conversation that follows ELiM, that allows the practitioner to integrate their knowledge of the child and family with the views of the parents to identify those most likely to need further engagement, and to equip parents with the skills needed to support their child's development"* [62].

The NHS England training for health visiting practitioners on using ASQ®-3 at the 2-2½ year health and developmental review is clear that it is to be used as a population measure of development, stating in highlighted text: *"The Department of Health have made the decision, based on research, to use the ASQ-3™ as the population measure of development for 2 to 2.5 year olds across England"* and there is a video explaining why the government wants a population level measure of child development at age two years (page 3, Part 1 of the training). However, despite this clear message in the NHS England training, we found a considerable lack of clarity amongst parents about the purpose of the current tool to measure child development (ASQ®-3), which is consistent with the previous review [10]. We found that all but two of the professionals that we spoke to were unaware that the intended use of ASQ®-3 was to collect population-level data for monitoring national trends and disparities.

Parents and professionals in our focus groups saw the primary purpose of the ASQ®-3 as to detect early developmental delay in the child attending the 2-2½ year health and developmental review. This has two implications. First, it suggests that there may need to be local level reinforcement or auditing of any national training provided, including checking that the messages reach parents through practitioners. Secondly, it suggests that the two potential purposes of a tool (1) individual assessment of a child to trigger support for developmental delay and (2) a population level data collection tool cannot be separated within the context of the 2-2½ year health and developmental review, as parents and practitioners are focused on benefits for the individual child.

Parents and health visiting professionals strongly felt that there needed to be more information available to explain the benefits of child development assessments in the context of the Healthy Child Programme, something also recommended to practitioners in the NHS England training. Our review suggests that there is an important need for improved and consistent training and materials for professionals and a clearer, more accessible 'package' of information to be made available to parents, regardless of which measure is used at the 2-2½ year health and developmental review. A simple and low-cost option, suggested by several professionals in our study, could be the provision of standardised, mandatory information about the ASQ®-3 in FAQ or information sheet format. This

would have the dual benefits of helping parents better understand the ASQ®-3 or other tool before filling it out, and of providing practitioners with consistent language and an interpretative framework for talking about assessing child development at age 2-2½ years.

Finally, in order for any tool to be used as a population measure of early child development, we need complete data flows from practice into local and then national information systems which is not currently the case for ASQ®-3 administered at the 2-2½ year health and developmental review. This is part of a wider issue of data completeness with the Community Services Dataset [6].

Strengths and limitations

We conducted a robust and systematic search to locate published material on tools to measure early childhood development which identified 34 tools, more than has been included in other similar reviews [49]. It is possible but unlikely that we missed any highly relevant and feasible tools. However, we were only able to review published material. We know that there is in-progress work on CREDI, and this is likely the case for other tools too, which will mean that readers need to also check for new evidence when reading this report at a later date. We recruited our professional and parent participants from across England but had difficulty recruiting parents of children with a disability (only four recruited) and foster or kinship carers (zero recruited). Our core findings from the focus groups (63 participants) are very consistent with the previous review on this topic from 2014 and other qualitative research [2], which increases our confidence about generalisability. Due to the rapid nature of our review, it was beyond the scope of the current study to complete a full psychometric evaluation following industry-standard principals (e.g., COSMIN guidelines [64]. A full-scale psychometric evaluation that considers how tools were constructed, acceptability, reliability, validity, and responsiveness [65], ideally in a UK context, would allow for more robust conclusions to be drawn. This could include reanalysis of existing datasets and new data collection.

Implications

Key finding 1: Parents and professionals valued a parent-reported child development tool combined with direct observation and which facilitated an in-person conversation about the child's development and health and wellbeing across the family system. We head that the experience and confidence of the practitioner will likely affect how far a tool can be used as a conversation starter (rather than as a tick-box).

Implication: A tool to measure child development at the 2-2½ year health and developmental review might best be embedded within an in-person holistic review of child and wider family health which includes the parent, as is currently the case in England. Stakeholders saw experience and skill of the practitioner as critical in being able to use the tool to facilitate a holistic conversation (and not as a checklist): consideration should be given to the experience and support of the practitioner using the tool within the wider health and developmental review. The impact of group reviews and digital tools should be studied carefully before implementation to ascertain that these modes do not undermine the ways in which the tool and health visiting are believed to work [1]. The most suitable tool should have both clinical utility for individual assessment and simultaneously provide meaningful data at the population level.

Key finding 2: ASQ®-3 and CREDI were tools judged as feasible to implement in the 2-2½ year health and developmental review. Evidence suggests that these are likely suitable population-level measures but there is insufficient evidence on how they would work for individual assessment of children in an English setting. This lack of evidence is particularly notable given that ASQ®-3 is already used in this way in England [1].

Implication: Evidence is needed on ASQ®-3 and CREDI to establish whether the scales are robust for use at the individual level in a UK setting, to establish population distributions of child development in this population and to establish the most appropriate cut-offs.

Commissioning analyses of CREDI data from the Children of the 2020s birth cohort study and investigating opportunities in the other planned birth cohort study in England [66] might be an efficient way to start further research on CREDI and/or ASQ®-3. In addition, fresh data collection systematically

investigating the psychometrics of these tools and testing their respective performances against gold standard of assessments, in a large and representative UK-based sample and across a range of target ages, would be highly valuable to support the robustness of the data collected at the 2-2½ year health and developmental review. There is a long-standing need for such a study, which was also recommended by the previous review on tools to measure child development at the 2-2½ year health and developmental review in 2012 [9].

Although CREDI has the advantage of being free to use, the health visiting infrastructure is under strain. Under the wrong circumstances, implementing a new tool may exacerbate strain on the service, undermine morale and lead to patchy up-take.

Key finding 3: Evidence about ASQ®-3 from English-speaking, non-UK general populations suggests that up to 62% of severe and up to 23% of mild to moderate delay might be detected by this tool using cut-offs recommended by the developers (specificity 82-89%) There is no equivalent evidence for CREDI, yet.

Implication: The low sensitivity of ASQ®-3 for detecting mild to moderate developmental delay may reflect known difficulties in identifying meaningful developmental delay before the age of four years. It is likely that any tool which is feasible to implement in the 2-2½ year health and developmental review will encounter this problem. As with ELiM, studies should consider evaluating the tool in combination with professional judgement, across different skill mix staff. The impact of training and supervision for less qualified staff should be investigated.

Key finding 4: Parents and professionals were confused about the purpose of the tool.

Implication: The provision of national level training on the selected tool is essential but not sufficient. It seems very likely that any population level data collection tool implemented in the 2-2½ year health and development review will also be used for individual assessment of children to detect developmental delay and trigger support pathways and developmental review, even if the official position is that the tool is primarily a population measure of child development. Parents and practitioners are understandably focused on benefit of the tool for helping each individual child and the most suitable tool will have clinical utility for individual assessment as well as facilitating population level data. For the parents and practitioners we spoke to, the delivery of the measure, the service pathways it triggers, were more important considerations than which specific tool is used to measure child development at the 2-2½ year health and developmental review. Ensuring the service has the capacity to cater to these priorities is therefore of particular importance, regardless of whether a different tool is introduced or the existing tool is retained.

References

1 King E, Gadsby E, Bell M, *et al.* Health visiting in the UK in light of the COVID-19 pandemic experience (RReHOPE): a realist review protocol. *BMJ*. 2023;13. doi: 10.1136/bmjopen-2022-068544.

2 McKean C, Watson, R, Charlton J, *et al.* 'Making the most of together time': development of a Health Visitor–led intervention to support children's early language and communication development at the 2–2½-year-old review. *Pilot Feasibility Stud*. 2022;8. doi: https://doi.org/10.1186/s40814-022-00978-5

3 Elearning for Healthcare. 1 to 3 years - Healthy Child Programme Schedule of Interventions Guide - DHSC. Healthy Child Programme Sched. Interv. Guide. 2022. https://www.e-lfh.org.uk/pathways-healthy-child/1-to-3-years/index.html (accessed 26 October 2023)

4 OHID. Commissioning health visitors and school nurses for public health services for children aged 0 to 19. UK Gov. 2023. https://www.gov.uk/government/publications/healthy-child-programme-0-to-19-health-visitor-and-school-nurse-commissioning/commissioning-health-visitors-and-school-nurses-for-public-health-services-for-children-aged-0-to-19 (accessed 12 October 2023)

5 Kendall S, Nash A, Braun A, *et al.* Acceptability and understanding of the Ages & Stages Questionnaires ®, Third Edition , as part of the Healthy Child Programme 2-year health and development review in England : Parent and professional perspectives. *Child Care Health Dev*. 2019;45:251–6.

6 Fraser C, Harron K, Barlow J, *et al.* How can we use the community services dataset (CSDS) for research into health visiting? Policy Research Unit Children and Families (CPRU) - NIHR 2020. https://www.ucl.ac.uk/children-policy-research/sites/children_policy_research/files/using_csds_for_research_report_08.10.20.pdf (accessed 1 December 2023)

7 Office for Health Improvement and Disparities. Health visitor service delivery metrics: annual data April 2022 to March 2023. GOV.UK. 2023. https://www.gov.uk/government/statistics/health-visitor-service-delivery-metrics-annual-data-april-2022-to-march-2023 (accessed 1 December 2023)

8 Children's Commissioner. Best beginnings in the early years: Technical report two. London: Children's Commissioner for England 2020. https://assets.childrenscommissioner.gov.uk/wpuploads/2020/07/cco-best-beginnings-in-the-early-years-tech-report-2.pdf (accessed 12 October 2023)

9 Bedford H, Walton S, Ahn J. Measures of Child Development: A review. Policy Research Unit in the Health of Children, Young People and Families 2013.

10 Kendall S, Nash A, Braun A, *et al.* Evaluating the use of a population measure of child development in the Healthy Child Programme Two Year Review. Policy Research Unit in the Health of Children, Young People and Families - UCL; Centre for Research in Primary and Community Care - University of Hertfordshire 2014.

11 Tricco AC, Lillie E, Zarin W, *et al.* PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med*. 2018;169:467–73.

12 Langlois EV, Straus SE, Antony J, *et al.* Using rapid reviews to strengthen health policy and systems and progress towards universal health coverage. *BMJ Glob Health*. 2019;4:e001178.

13 Wilson MG, Oliver S, Melendez-Torres GJ, *et al.* Paper 3: Selecting rapid review methods for complex questions related to health policy and system issues. *Syst Rev*. 2021;10:286.

14 Garritty C, Gartlehner G, Nussbaumer-Streit B, *et al.* Cochrane Rapid Reviews Methods Group offers evidence-informed guidance to conduct rapid reviews. *J Clin Epidemiol*. 2021;130:13–22.

15 Fernald LCH, Prado E, Kariger P, *et al. A Toolkit for Measuring Early Childhood Development in Low and Middle-Income Countries*. World Bank, Washington, DC 2017. https://doi.org/10.1596/29000

16 Sandler A, Brazdziunas D, Cooley, CW, *et al.* Developmental Surveillance and Screening of Infants and Young Children. *Pediatrics*. 2001;108:192–6.

17 Sheldrick RC, Garfinkel D. Is a Positive Developmental-Behavioral Screening Score Sufficient to Justify Referral? A Review of Evidence and Theory. *Acad Pediatr*. 2017;17:464–70.

18 Whiting P, Rutjes AW, Reitsma JB, *et al.* The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003;3:25.

19 da Costa A, Ferraz NN, Berezovsky A. Cognitive, motor, and visual development in healthy children in the first 42 months of life. *Arq Bras Oftamol*. 2021;84:436–41.

20 About Our Tools. Pedstest. https://pedstest.com/about-our-tools/ (accessed 27 October 2023)

21 Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;15:155–63.

22 Schonhaut L, Martinez-Nadal S, Armijo I, *et al.* Reliability and agreement of ages and stages questionnaires®: Results in late preterm and term-born infants at 24 and 48 months. *Early Hum Dev*. 2019;128:55–61.

23 Manti F, Giovannone F, Ciancaleoni M, *et al.* Psychometric Properties and Validation of the Italian Version of Ages & Stages Questionnaires Third Edition. *Int J Environ Res Public Health*. 2023;20:5014.

24 Li Y, Tang L, Bai Y, *et al.* Reliability and validity of the Caregiver Reported Early Development Instruments (CREDI) in impoverished regions of China. *BMC Pediatr*. 2020;20:475.

25 Lockhart M, Chaux R, Chevin M, *et al.* Classification Performance of the Ages and Stages Questionnaire: Influence of Maternal Education Level. *Children*. 2023;10:449.

26 Peyton C, Wroblewski K, Park J, *et al.* Validity of The Warner Initial Developmental Evaluation of Adaptive and Functional Skills (WIDEA-FS): a daily activity criterion checklist for infants and toddlers. *Pediatr Res*. 2021;90:1052–7.

27 Sheldrick RC, Marakovitz S, Garfinkel D, *et al.* Comparative Accuracy of Developmental Screening Questionnaires. *JAMA Pediatr*. 2020;174:366.

28 Veldhuizen S, Clinton J, Rodriguez C, *et al.* Concurrent Validity of the Ages and Stages Questionnaires and Bayley Developmental Scales in a General Population Sample. *Acad Pediatr*. 2015;15:231–7.

29 Schonhaut L, Armijo I, Schönstedt M, *et al.* Validity of the Ages and Stages Questionnaires in Term and Preterm Infants. *Pediatrics*. 2013;131:e1468–74.

30 Noeder MM, Logan BA, Struemph KL, *et al.* Developmental screening in children with CHD: Ages and Stages Questionnaires. *Cardiol Young*. 2017;27:1447–54.

31 Duggan C, Irvine AD, O'B Hourihane J, *et al.* ASQ-3 and BSID-III's concurrent validity and predictive ability of cognitive outcome at 5 years. *Pediatr Res*. Published Online First: 25 February 2023. doi: 10.1038/s41390-023-02528-y

32 Steenis LJP, Verhoeven M, Hessen DJ, *et al.* Parental and professional assessment of early child development: The ASQ-3 and the Bayley-III-NL. *Early Hum Dev*. 2015.

33 Agarwal PK, Xie H, Sathyapalan Rema AS, *et al.* Concurrent validity of the ages and stages questionnaires with Bayley Scales of Infant Development-III at 2 years – Singapore cohort study. *Pediatr Neonatol*. 2023;S1875957223001237.

34 Kerstjens JM, Nijhuis A, Hulzebos CV, *et al.* The Ages and Stages Questionnaire and Neurodevelopmental Impairment in Two-Year-Old Preterm-Born Children. *PLOS ONE*. 2015;10:e0133087.

35 Agarwal PK, Shi L, Daniel LM, *et al.* Prospective evaluation of the Ages and Stages Questionnaire 3rd Edition in very-low-birthweight infants. *Dev Med Child Neurol*. 2017;59:484–9.

36 Simpson S, D'Aprano A, Tayler C, *et al.* Validation of a culturally adapted developmental screening tool for Australian Aboriginal children: Early findings and next steps. *Early Hum Dev*. 2016;103:91–5.

37 Yue A, Jiang Q, Wang B, *et al.* Concurrent validity of the Ages and Stages Questionnaire and the Bayley Scales of Infant Development III in China. *PLOS ONE*. 2019;14:e0221675.

38 Yue A, Luo X, Jia M, *et al.* Concurrent validity of the MacArthur communicative development inventory, the Ages and Stages Questionnaires and the Bayley Scales of Infant and Toddler Development: A study in rural China. *Infant Child Dev*. 2021;30:e2219.

39 du Toit MN, van der Linde J, Swanepoel DW. mHealth developmental screening for preschool children in low-income communities. *J Child Health Care*. 2021;25:573–86.

40 Peyton C, Msall ME, Wroblewski K, *et al.* Concurrent validity of the Warner Initial Developmental Evaluation of Adaptive and Functional Skills and the Bayley Scales of Infant and Toddler Development, Third Edition. *Dev Med Child Neurol*. 2021;63:349–54.

41 Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qual Res Psychol*. 2021;18:328–52.

42 Campbell KA, Orr E, Durepos P, *et al.* Reflexive Thematic Analysis for Applied Qualitative Health Research. *Qual Rep*. 2021;26:2011–28.

43 Saunders CH, Sierpe A, von Plessen C, *et al.* Practical thematic analysis: a guide for multidisciplinary health services research teams engaging in qualitative analysis. *BMJ Res Methods Report*. 2023;381:e074256.

44 Clarkson C, Scott HR, Hegarty S, *et al.* 'You get looked at like you're failing': A reflexive thematic analysis of experiences of mental health and wellbeing support for NHS staff. *J Health Psychol*. 2023;28:818–31.

45 Ritchie J, Lewis J. *Qualitative research practice: a guide for social science students and researchers*. London: Sage 2003.

46 Heath G, Cameron E, Cummins C, *et al.* Paediatric 'care closer to home': stake-holder views and barriers to implementation. *Health Place*. 2012;18:1068–73.

47 Gale N, Heath G, Cameron E, *et al.* Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol*. 2013;13.

48 Hatchard J, Gillespie DOS. Using framework analysis deductively: A case study from alcohol and tobacco tax policy and modelling research. *Myths, Methods, and Messiness: Insights for Qualitative Research Analysis: Edited Proceedings of 5th Annual Qualitative Research Symposium*. University of Bath 2019:30–40.

49 Boggs D, Milner KM, Chandna J, *et al.* Rating early child development outcome measurement tools for routine health programme use. *Arch Dis Child*. 2019;104:S22–33.

50 Gollenberg AL, Lynch CD, Jackson LW, *et al.* Concurrent validity of the parent-completed Ages and Stages Questionnaires, 2nd Ed. with the Bayley Scales of Infant Development II in a low-risk sample. *Child Care Health Dev*. 2010;36:485–90.

51 Anderson PJ, Burnett A. Assessing developmental delay in early childhood — concerns with the Bayley-III scales. *Clin Neuropsychol*. 2017;31. doi: https://doi.org/10.1080/13854046.2016.1216518

52 Balasundaram P, Avulakunta I. Bayley Scales Of Infant and Toddler Development. *StatPearls [Internet]*. Treasure Island (FL): StatPearls Publishing 2022.

53 Tractenberg R. Classical and modern measurement theories, patient reports, and clinical outcomes. *Contemp Clin Trials*. 2010;31:1–3.

54 Cappelleri JC, Lundy JJ, Hays RD. Overview of Classical Test Theory and Item Response Theory for Quantitative Assessment of Items in Developing Patient-Reported Outcome Measures. *Clin Ther*. 2014;36:648–62.

55 Jabrayilov R, Emons WHM, Sijtsma K. Comparison of Classical Test Theory and Item Response Theory in Individual Change Assessment. *Appl Psychol Meas*. 2016;40:559–72.

56 Squires J, Twombly EMS, Bricker D, *et al.* ASQ-3 Technical Appendix. 2009.

57 McCoy DC, Waldman M, Fink G. Measuring early childhood development at a global scale: Evidence from the Caregiver-Reported Early Development Instruments. *Early Child Res Q*. 2018;45:58–68.

58 Brookes Publishing. ASQ-3 Technical Report. 2009. https://agesandstages.com/wp-content/uploads/2017/05/ASQ-3-Technical-Appendix_web.pdf

59 Bayley N. *Bayley Scales of infant and toddler development. 3rd edition.* San Antonio, TX, USA: The Psychological Corporation 2006.

60 Charkaluk M-L, Rousseau J, Calderon J, *et al.* Ages and Stages Questionnaire at 3 Years for Predicting IQ at 5–6 Years. *Pediatrics*. 2017;139:e20162798.

61 Centre for Longitudinal Studies, UCL. CLS | Children of the 2020s Study. https://cls.ucl.ac.uk/cls-studies/children-of-the-2020s-study/ (accessed 26 October 2023)

62 Law J, Charlton J, McKean C, *et al.* Identifying and Supporting Children's Early Language Needs. 2020.

63 Law J, Charlton J, Wilson P, *et al.* The development and productivity of a measure for identifying low language abilities in children aged 24–36 months. *BMC Pediatr*. 2023;23. doi: https://doi.org/10.1186/s12887-023-04079-x

64 Mokkink LB, Prinsen CA, Paratick DL, *et al.* COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs): User manual. 2018.

65 Smith SC, Lamping DL, Maclaine GDH. Measuring health-related quality of life in diabetic peripheral neuropathy: A systematic review. *Diabetes Res Clin Pract*. 2012;96:261–70.

66 Centre for Longitudinal Studies, UCL. CLS | Early Life Cohort Feasibility Study. https://cls.ucl.ac.uk/cls-studies/early-life-cohort-feasibility-study/ (accessed 26 October 2023)

67 McCoy DC, Cuartas J, Waldman M, *et al.* Contextual variation in young children's acquisition of social-emotional skills. *PLOS ONE*. 2019;14:e0223056.

68 Altafim ERP, McCoy DC, Brentani A, *et al.* Measuring early childhood development in Brazil: validation of the Caregiver Reported Early Development Instruments (CREDI). *J Pediatr (Rio J)*. 2020;96:66–75.

69 Waldman M, McCoy DC, Seiden J, *et al.* Validation of motor, cognitive, language, and socio-emotional subscales using the Caregiver Reported Early Development Instruments: An application of multidimensional item factor analysis. *Int J Behav Dev*. 2021;45:368–77.

70 Abdoola S. Translating tools for better parent-based assessment: An exploratory study. *S Afr J Commun Disord*. 2015;62:12 pages.

71 van Heerden A, Hsiao C, Matafwali B, *et al.* Support for the feasibility of the ages and stages questionnaire as a developmental screening tool: a cross-sectional study of South African and Zambian children aged 2-60 months. *BMC Pediatr*. 2017;17:55.

72 Rubio-Codina M, Araujo MC, Attanasio O, *et al.* Concurrent Validity and Feasibility of Short Tests Currently Used to Measure Early Childhood Development in Large Scale Studies. *PLOS ONE*. 2016;11:e0160962.

73 Rubio-Codina M, Grantham-McGregor S. Predictive validity in middle childhood of short tests of early childhood development used in large scale studies compared to the Bayley-III, the Family Care Indicators, height-for-age, and stunting: A longitudinal study in Bogota, Colombia. *PLOS ONE*. 2020;15:e0231317.

74 Schonhaut L, Armijo I, Schönstedt M, *et al.* Validity of the Ages and Stages Questionnaires in Term and Preterm Infants. *Pediatrics*. 2013;131:e1468–74.

75 Schonhaut L, Martinez-Nadal S, Armijo I, *et al.* Reliability and agreement of ages and stages questionnaires®: Results in late preterm and term-born infants at 24 and 48 months. *Early Hum Dev*. 2019;128:55–61.

76 Lockhart M, Chaux R, Chevin M, *et al.* Classification Performance of the Ages and Stages Questionnaire: Influence of Maternal Education Level. *Children*. 2023;10:449.

77 Manti F, Giovannone F, Ciancaleoni M, *et al.* Psychometric Properties and Validation of the Italian Version of Ages & Stages Questionnaires Third Edition. *Int J Environ Res Public Health*. 2023;20:5014.

78 Koushiou M, Trakoshis S, Michael N, *et al.* Exploring the Ages and Stages Questionnaire – 3 psychometric properties in Greek-Cypriot males and females during toddlerhood and preschool years: Preliminary findings. *Glob Pediatr*. 2023;4:100045.

79 Agarwal PK, Xie H, Sathyapalan Rema AS, *et al.* Concurrent validity of the ages and stages questionnaires with Bayley Scales of Infant Development-III at 2 years – Singapore cohort study. *Pediatr Neonatol*. 2023;S1875957223001237.

80 Waldman M, McCoy DC, Seiden J, *et al.* Validation of motor, cognitive, language, and socio-emotional subscales using the Caregiver Reported Early Development Instruments: An application of multidimensional item factor analysis. *Int J Behav Dev*. 2021;45:368–77.

81 Altafim ERP, McCoy DC, Brentani A, *et al.* Measuring early childhood development in Brazil: validation of the Caregiver Reported Early Development Instruments (CREDI). *J Pediatr (Rio J)*. 2020;96:66–75.

82 Li Y, Tang L, Bai Y, *et al.* Reliability and validity of the Caregiver Reported Early Development Instruments (CREDI) in impoverished regions of China. *BMC Pediatr*. 2020;20:475.

83 Alderman H, Friedman J, Ganga P, *et al.* Assessing the performance of the Caregiver Reported Early Development Instruments (CREDI) in rural India. *Ann N Y Acad Sci*. 2021;1492:58–72.

84 Gladstone M, Lancaster G, McCray G, *et al.* Validation of the Infant and Young Child Development (IYCD) Indicators in Three Countries: Brazil, Malawi and Pakistan. *Int J Environ Res Public Health*. 2021;18. doi: 10.3390/ijerph18116117

85 du Toit MN, van der Linde J, Swanepoel DW. mHealth developmental screening for preschool children in low-income communities. *J Child Health Care Prof Work Child Hosp Community*. 2021;25:573–86.

86 Peyton C, Msall ME, Wroblewski K, *et al.* Concurrent validity of the Warner Initial Developmental Evaluation of Adaptive and Functional Skills and the Bayley Scales of Infant and Toddler Development, Third Edition. *Dev Med Child Neurol*. 2021;63:349–54.

87 Peyton C, Wroblewski K, Park J, *et al.* Validity of The Warner Initial Developmental Evaluation of Adaptive and Functional Skills (WIDEA-FS): a daily activity criterion checklist for infants and toddlers. *Pediatr Res*. 2021;90:1052–7.

88 Dana Charles McCoy, Gunther Fink. CREDI User's Guide. 2017.

Appendices

Appendix A
Applying Cochrane Rapid Review Methods Recommendations

Table 5. Cochrane rapid review methods recommendations

Setting the research question and topic refinement	This RR
• Involve key stakeholders (e.g., review users such as consumers, health professionals, policymakers, decision-makers) to set and refine the review question, eligibility criteria, and the outcomes of interest. Consult with stakeholders throughout the process to ensure the research question is fit for purpose, and regarding any ad-hoc changes that may occur as the review progresses. (R1) • Develop a protocol that includes review questions, PICOS, and inclusion and exclusion criteria.	We've involved stakeholders throughout the process. This is responsive research focus, as it has been put forward by DHSC. We have maintained regular communication with DHSC and other key policy colleagues.
Setting eligibility criteria	
• Together with key stakeholders: - Clearly define the population, intervention, comparator, and outcomes. - Limit the number of interventions (R2) and comparators (R3) - Limit the number of outcomes, with a focus on those most important for decision-making. (R4) • Consider date restrictions with a clinical or methodological justification. (R5) • Setting restrictions are appropriate with justification provided. (R6) • Limit the publication language to English; add other languages only if justified. (R7) • Systematic reviews (SRs) should be considered a relevant study design for inclusion. (R8) • Place emphasis on higher quality study designs (e.g., SRs or RCTs); consider a stepwise approach to study design inclusion. (R9)	The population, intervention, comparator and outcomes was defined using the Bedford et al. [9] initial review. Date restrictions were considered based on the dates of the past review. Only English language studies were included. Systematic reviews were included in our hand searches.

Searching	
• Involve an information specialist. • Limit main database searching to CENTRAL, MEDLINE (e.g., via PubMed), and Embase (if available access). (R10) • Searching of specialized databases (e.g., PsycINFO and CINAHL) is recommended for certain topics but should be restricted to 1-2 additional sources, or omitted if time and resources are limited. (R11) • Consider peer review of at least one search strategy (e.g., MEDLINE). (R12) • Limit grey literature and supplemental searching (R13). If justified, search study registries and scan the reference lists of other SRs or included studies after screening of the abstracts and full-texts.	We used REAL as a specialist group to support with the database search strategy and the study selection and identification of studies. We limited the main database searching to Medline, PsycINFO and Web of Science. Our search strategy was peer reviewed by the specialists amongst the study team. We did not include grey literature.
Study selection	
Title and abstract screening • Using a standardized title and abstract form, conduct a pilot exercise using the same 30-50 abstracts for the entire screening team to calibrate and test the review form. • Use two reviewers for dual screen of at least 20% (ideally more) of abstracts, with conflict resolution. • Use one reviewer to screen the remaining abstracts and a second reviewer to screen all excluded abstracts, and if needed resolve conflicts. (R14) Full-text screening • Using a standardized full-text form, conduct a pilot exercise using the same 5-10 full-text articles for the entire screening team to calibrate, and test the review form. • Use one reviewer to screen all included full-text articles and a second reviewer to screen all excluded full-text articles. (R15)	For title and abstract screening, we used a standardised title and abstract screening document which was tested by the REAL team (GA, AK, and JC) and our team's reviewers (JL and RMP). This was also discussed and refined with our specialists among the study team. Conflict resolution was discussed amongst three reviewers (GA, JL and RMP). For full-text screening, we used a mapping extraction tool (Appendix I) which was tested by two reviewers (KS and RMP) on 50 articles. KS extracted 100% of the information on full-text for mapping and RMP double-checked 10%. Based on this screening, we used Table 1 to decide whether to include or exclude studies for data extraction.
Data extraction	
• Use a single reviewer to extract data using a piloted form. Use a second reviewer to check for correctness and completeness of extracted data. (R16) • Limit data extraction to a minimal set of required data items. (R17) • Consider using data from existing SRs to reduce time spent on data extraction. (R18)	After the mapping and selection of appropriate studies as described in the rapid review methods section, we tested our tool (Appendix I) for data extraction on 10 articles (RMP and KS). KS extracted data for all the articles and RMP and JL double-checked all studies.

Risk of bias assessment	
• Use a valid risk of bias tool, if available for the included study designs. • Use a single reviewer to rate risk of bias, with full verification of all judgments (and support statements) by a second reviewer. (R19) • Limit risk of bias ratings to the most important outcomes, with a focus on those most important for decision-making. (R20)	We used QUADAS-I as the risk of bias tool. RMP rated the risk of bias, verified by JL. The risk of bias tool is presented in Appendix J
Synthesis	
• Synthesize evidence narratively. • Consider a meta-analysis only if appropriate (i.e., studies are similar enough to pool). (R21) Standards for conducting a meta-analysis for an SR equally apply to an RR. • Use a single reviewer to grade the certainty of evidence, with verification of all judgments (and footnoted rationales) by a second reviewer. (R22)	We have synthesised the evidence narratively. One reviewer (JL) has written up the results with research assistance of RMP; JW then verified all judgements.
Other considerations for Cochrane RRs	
RRs should be preceded by a protocol submitted to and approved by Cochrane (R23); the protocol should be published (e.g., PROSPERO or Open Science Framework) (R24); allow for post hoc changes to the protocol (eligibility criteria etc.) as part of an efficient and iterative process (R25); document all post hoc changes; and incorporate use of online SR software (e.g., Covidence, DistillerSR, and EPPI-Reviewer) to streamline the process (R26).	

Appendix B

Search Concepts

Table 6. Search concepts for our review based on previous literature review (Bedford et al., 2013)

Concept		Related terms
Developmental	Development	Development, performance, skills, ability, disability, activity, function
	Cognitive	Cognitive, cognition, learning
	Social/emotional	Social, emotional, behaviour, socioemotional, socio-emotional
	Physical/motor	Motor skills, psychomotor, physical
	linguistics	Speech, language, linguistic, communication
Tool		Data collection, assessment, questionnaire, checklist, survey, tool, scale, inventory, diagnosis, test
Young child		Human, child, infant, preschool, early childhood, early childhood development
We combined these concepts using the AND Boolean operator Development AND tool AND young child		

Table 7. Search strategy for the three databases

Source	Version/ Platform/URL	Date of Search	RCT filter applied	Search strategy	Records retrieved
1. PubMed	https://pubmed.ncbi.nlm.nih.gov/advanced/	10/11/2022	No	(((“Assessment”[Title/Abstract] OR “questionnaire”[Title/Abstract] OR “checklist”[Title/Abstract] OR “survey”[Title/Abstract] OR “tool”[Title/Abstract] OR “scale”[Title/Abstract] OR “inventory”[Title/Abstract] OR “diagnosis”[Title/Abstract] OR “test”[Title/Abstract]) NOT (“teenager”[Title/Abstract] OR “adolescent”[Title/Abstract] OR “adolescence”[Title/Abstract] OR “young adult”[Title/Abstract] OR “adult”[Title/Abstract])) AND (“preschool”[Title/Abstract] OR “preschooler”[Title/Abstract] OR “early childhood”[Title/Abstract] OR “early years”[Title/Abstract] OR “toddler”[Title/Abstract]) AND (“Development”[Title/Abstract] OR “performance”[Title/Abstract] OR “skills”[Title/Abstract] OR “ability”[Title/Abstract] OR “disability”[Title/Abstract] OR “activity”[Title/Abstract] OR “function”[Title/Abstract] OR “cognitive”[Title/Abstract] OR “cognition”[Title/Abstract] OR “learning”[Title/Abstract] OR “social”[Title/Abstract] OR “emotional”[Title/Abstract] OR “socioemotional”[Title/Abstract] OR “socio-emotional”[Title/Abstract] OR “behaviour”[Title/Abstract] OR “motor skills”[Title/Abstract] OR “psychomotor”[Title/Abstract] OR “physical”[Title/Abstract] OR “speech”[Title/Abstract] OR “language”[Title/Abstract] OR “linguistic”[Title/Abstract] OR “communication”[Title/Abstract] OR (“developmental measures”[Title/Abstract] OR “developmental milestones”[Title/Abstract] OR “early childhood development”[Title/Abstract] OR “developmental review”[Title/Abstract]))) AND ((2012/1/1:2022/9/15[pdat]) AND (english[Filter]))	6246

2. PsycINFO	https://ovidsp.dc1.ovid.com/ovid-b/ovidweb.cgi?QS2=434f4e1a73d37e8c79e5d8c142641a5414a3561b484bc9f8caa2d4abe17c979b3c65476370e18d9ab8d8c46784a57c18233e3ff9ced33dc35a7f6494b4b2d976b3e9390782ed5672f9b744cf0a45709ed6ebc490af708363982b0995e2cf4b1bcc3ed9e937cce15ebd5af6d5c3c9c29993249177a4a464f328349769d98077910bd977dbe79faff81884193daba086fca188a5ab9954639a684137b18d1d853e12918bc4ab4a727620c373277482d4c18f65d8da0089c16dde15dcf06ffcf8888d2fceabac6a08c9	10/11/2022	No	Database: APA PsycInfo <1806 to November Week 2 2022> Search Strategy: 1 (((Assessment or questionnaire or checklist or survey or tool or scale or inventory or diagnosis or test) not (teenager or adolescent or adolescence or young adult or adult)) and (preschool or preschooler or early childhood or early years or toddler)).mp. and (Development or performance or skills or ability or disability or activity or function or cognitive or cognition or learning or social or emotional or socioemotional or socio-emotional or behaviour or motor skills or psychomotor or physical or speech or language or linguistic or communication or (Developmental measures or Developmental milestones or early childhood development or developmental review)).m_titl. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures, mesh word] (16651) 2 limit 1 to (english language and yr="2012 - 2022") (8377)	8377
3. Web of Science	https://www.webofscience.com/wos/woscc/advanced-search	10/11/2022	No	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test) NOT (teenager OR adolescent OR adolescence OR young adult OR adult) AND (preschool OR preschooler OR early childhood OR early years OR toddler) AND ((Development OR performance OR skills OR ability OR disability OR activity OR function OR cognitive OR cognition OR learning OR social OR emotional OR socioemotional OR socio-emotional OR behaviour OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication) OR (Developmental measures OR Developmental milestones OR early childhood development OR developmental review)) *(Title) * Timespan: 2012-01-01 to 2022-09-15 (Index Date) * Language: English	866
				TOTAL before de-duplication	15489
				TOTAL after de-duplication	13726

Appendix C

'Gold Standard' Papers and Test Refinement

'Gold standard' papers were identified through preliminary literature search and consultation with experts in the field (PF, SK, JW). Databases were checked for gold standard papers to see which databases are likely to contain the most relevant references. Pale green fields indicate papers that were picked up by our first test search and bright green fields indicate papers that were additionally picked up by our second test search. Neither ERIC database yielded many hits and so were excluded from the final search. In order to limit the number of results to an amount that was feasible within the confines of the study, PsycINFO, PUBMED and Web of Science were selected from the remaining databases for the final search, as between them they picked up all available references on the gold standard list apart from Black et al., 2019, which is a web article and as such was not present in academic databases.

	Test search 1 / Test search 2							Retained databases for final search strategy		
	EMBASE	ERIC	ERIC (EBSCO)	PsycINFO	Pub Med	WoS	Med line	PsycINFO	PubMed	WoS
ASQ®-3										
[60] Charkaluk et al., 2017	X			X	X	X	X			
CREDI										
[57] McCoy et al., 2018				X		X	X			
[67] McCoy et al., 2019	X			X	X		X			
[68] Altafim et al., 2020		X			X	X	X			
[69] Waldman et al., 2021		X		X		X				
ECDI2030										
Cappa et al. 2021. Identifying and minimising errors in the measurements of early childhood development: lessons learned from the cognitive testing of the ECDI2030. *International Journal of Environmental Research & Public Health,* 18(22): 11-20	X				X	X	X			
GSED										

Black et al 2019. The Global Scale for Early Development (GSED) *Early Childhood Matters* 14: 80-84										
WHO IYCD										
Lancaster et al. 2018. Creation of the WHO Indicators of Infant and Young Child Development (IYCD): metadata synthesis across 10 countries. *BMJ Global Health* 3: e000747	X				X	X				
Gladstone et al. 2021 Validation of the Infant and Young Child Development (IYCD) Indicators in Three Countries: Brazil, Malawi and Pakistan					X		X			
REVIEWS										
Richter *et al.* Early childhood development: an imperative for action and measurement at scale. *BMJ Global Health* 2019; **4**: i154-i160.	X				X	X				
Boggs *et al.* Rating early child development outcome measurement tools for routine health programme use. *Arch Dis Child* 2019; **104**: S22-S33.	X				X	X	X			
Cairney *et al*. (2021) Predictive value of universal preschool developmental assessment in identifying children with later educational difficulties: A systematic review', *PLoS ONE*, vol. 16, no. 3, pp. e0247299	X			X	X	X	X			

Table 8. Example 1 of testing search strategies

Database	Number of results	Strategy
PubMed	23.890	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement) NOT adolescent AND (preschool OR preschooler OR early childhood OR early years OR toddler OR infant) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)
	12.235 *Using the age filter: 7.646	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement OR preschool developmental assessment) AND (preschool OR preschooler OR early childhood OR early years OR toddler) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)
	245 All filters and title focused	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement) NOT adolescent AND (preschool OR preschooler OR early childhood OR early years OR toddler OR infant) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)
PsycInfo	11.195	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test) NOT adolescent AND (preschool OR preschooler OR early childhood OR early years OR toddler) AND ((Development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication) OR (Developmental measures OR Developmental milestones OR early childhood development OR developmental review))
	3.565	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test) NOT adolescent AND (preschool OR preschooler OR early childhood OR early years OR toddler) AND ((Development OR performance OR skills OR ability OR disability OR activity OR function OR cognitive OR cognition OR learning OR social OR emotional OR socioemotional OR socio-emotional OR behaviour OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication) OR (Developmental measures OR Developmental milestones OR early childhood development OR developmental review))

Database	Number of results	Strategy
Medline (OVID)	18.931	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement) NOT adolescent AND (preschool OR preschooler OR early childhood OR early years OR toddler OR infant) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)
ERIC (EBSCO)	6.067	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement) NOT adolescent AND (preschool OR preschooler OR early childhood OR early years) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)
Embase	15.963	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement) NOT adolescent AND (preschool OR preschooler OR early childhood OR early years) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)
Web of Science (WOS)	98.883	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement) NOT adolescent AND (preschool OR preschooler OR early childhood OR early years) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)

Table 9. Example 2 of testing search strategies

Database	Number of results	Strategy
PubMed	4.824	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement) NOT child AND (preschool OR preschooler OR early childhood OR early years OR toddler OR infant) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)
	4.054	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement OR preschool developmental assessment) NOT child AND (preschool OR preschooler OR early childhood OR early years OR toddler) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)
	235 All filters and title focused	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement) NOT child AND (preschool OR preschooler OR early childhood OR early years OR toddler OR infant) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)
PsycInfo	988	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test) NOT child AND (preschool OR preschooler OR early childhood OR early years OR toddler) AND ((Development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication) OR (Developmental measures OR Developmental milestones OR early childhood development OR developmental review))
	3.565	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test) NOT child AND (preschool OR preschooler OR early childhood OR early years OR toddler) AND ((Development OR performance OR skills OR ability OR disability OR activity OR function OR cognitive OR cognition OR learning OR social OR emotional OR socioemotional OR socio-emotional OR behaviour OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication) OR (Developmental measures OR Developmental milestones OR early childhood development OR developmental review))

Database	Number of results	Strategy
Medline (OVID)	4.635	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement) NOT child AND (preschool OR preschooler OR early childhood OR early years OR toddler OR infant) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)
ERIC (EBSCO)	1.304	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement) NOT child AND (preschool OR preschooler OR early childhood OR early years) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)
Embase	657	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement) NOT child AND (preschool OR preschooler OR early childhood OR early years) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)
Web of Science (WOS)	65.449	(Assessment OR questionnaire OR checklist OR survey OR tool OR scale OR inventory OR diagnosis OR test OR instrument OR measurement) NOT child AND (preschool OR preschooler OR early childhood OR early years) AND (Development* OR early child development OR early development OR performance OR skills OR ability OR disability OR activity OR function OR cogniti* OR learning OR social OR emotional OR socioemotional OR socio-emotional OR social-emotional OR behavi* OR motor skills OR psychomotor OR physical OR speech OR language OR linguistic OR communication)

Appendix D

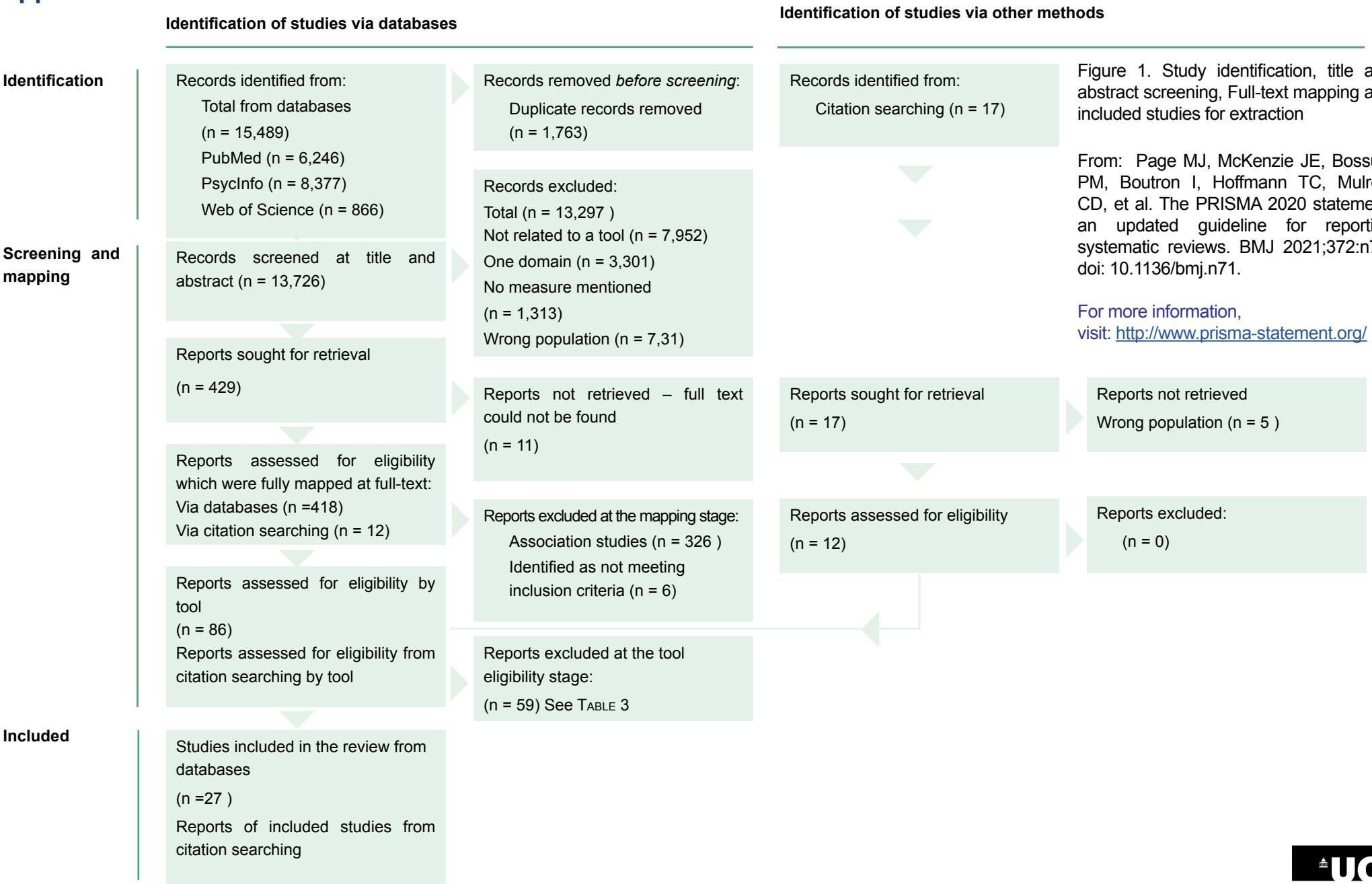

Figure 1. Study identification, title and abstract screening, Full-text mapping and included studies for extraction

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71.

For more information, visit: http://www.prisma-statement.org/

Appendix E

Mapping Extraction Tool

Table 10. Study Selection – Mapping the evidence: Extraction Tool

Extraction tool: Mapping the evidence
Reviewer
Date
Authors
Year
Study type (tool development, implementation & acceptability, reliability/validation, standardisation – see Appendix F)
Name of tool
Initialism
Age of children in study
Country study was conducted in
Available in English?
Administer in one contact?
Reviewer
Date

Appendix F

Study Type Criteria for Mapping Phase

Table 11. Definitions for criteria "study type"

Definitions for study type	Description	Number of studies from database	Number of studies from citation searching
Tool development	Study that reports the development of new tool	11	
Implementation & acceptability	Study that reports on the application of tool – how practitioners, parents, interested partners, (i.e., how easy people/users/ interested parties find to use it)	4	2
Reliability & validation	Evaluates the quality of the tool. Reliability is related to consistency of measurement. Validity relates to correctness, i.e., is the tool is measuring what it is supposed to measure.	65	10
Standardisation	Finding norms in a population (i.e., ASQ®-3 is standardised in the American population. The relevant points mentioned are where and when.)	7	
Association	Study that uses/applies a tool of interest but doesn't report on the tool. The unit of measurement is not the tool but the outcome specified in the paper.	326	
Review	Papers that provide a review / synthesis of the relevant literature	5	
Total		418	12

Appendix G

Identified Tools and Extraction from Mapping Phase

Five measures were excluded at the end of this stage due to measuring only one domain of development (Infant-Toddler Social and Emotional Assessment, socio-emotional only, three papers; Rescorla's Language Development Survey, language only, one paper; MacArthur Communicative Development Inventory, language only, one paper; Draw-a-Person Intellectual Ability Test for Children, Adolescents and Adults, cognitive only, one paper) or because they were not a measure of early childhood development (List of Capabilities Instrument, one paper).

Table 12. Results from data mapping – in-depth

	Measure	Age	English language	Time to administer	Geographical uptake (incl. UK?)	Equipment needed	Training needed	Total no. papers
1	Ages & Stages Questionnaire (ASQ®-3)	• All children were between the ages of 2 months and 5 years. • between 57 and 66 months • between the ages of 2 weeks to 42 months and 15 days • 9 months to 5.5 years • 5–24months • children between 2 and 36 months of age • 6-42 months • 8-24 months • 9-24 months • 4-40 months • Not reported (interviews with parents) • 36 months	Yes, and other languages.	• 6 studies did not report administration time. 11 out of 15 parents reported that it took them between 20–40 minutes to complete the questionnaire, and three reported that it took longer than 40 minutes. • The ASQ®-TRAK took approximately 30–60 minutes. • ASQ®-3: 19.7 (8.2) • 5-10 minutes • Not reported, however, "Some parents indicated that the questionnaire took too long to administer and although AHWs acknowledged that the tool was long and could be an issue in keeping families engaged, all reported it was necessary and did not endorse shortening it further."	• South Africa and Zambia, The Netherlands, USA, China, Australia, Colombia, China, Singapore, Australia and France. From association studies (36): Norway, USA, Rwanda, Canada, Chile, Brazil, **UK**, Netherlands, Bangladesh, China, Madagascar, Japan, Iran, Ireland, Turkey, Egypt, Uruguay	• 6 studies did not report any equipment • One study mentioned that all parents were provided with materials (e.g., blocks, crayons) to facilitate the completion of the questionnaire, as recommended in the manual. • Another mentioned that the ASQ®-TRAK includes a set of colour-illustrated flipcharts. • Another stated that the ASQ®-3 kit cost US$275. • One last study reported that a toy kit was provided, however, seems to be extra, not specifically tool related.	• 7 studies reported that parents administered the test. • Research assistants were recruited and trained in South Africa and Zambia. All were female with at least some tertiary training and experience working with children. • The ASQ®-3was administered by trained enumerators who interviewed each child's primary caregiver. Most of the ASQ®-3enumerators are from education or social work-related majors to communicate well with rural caregivers. All ASQ®-3enumerators received a weeklong formal training in the administration of the ASQ®. This training strictly followed the Ages and Stages Questionnaire Manual. It included a 2.5-day field training in which enumerators conducted ASQ®-3 interviews with the caregivers of children of similar ages and backgrounds to those in our sample. • Child health nurses employed at Congress administered the ASQ®-TRAK in the clinic. • Given the low education levels of some caregivers, items were given by interview and were only administered directly to the child if the caregiver could not provide an answer.	24

	Measure	Age	English language	Time to administer	Geographical uptake (incl. UK?)	Equipment needed	Training needed	Total no. papers
2	Parents' Evaluation of Developmental Status (PEDS)	• children ranged in age from 1 month to 6 years and 11 months • 36-83 months	Yes, in included association studies.	• The study reported that the form takes parents less than 5 minutes to complete • Another didn't report length of time	South Africa, Singapore, USA. From association studies (2): Iceland, Australia	• One study didn't report • PEDS tools were developed into a smartphone application using the same algorithm as the original paper-based tool. Paper based tool is a one-sheet questionnaire requiring no additional resources	• Parent and 1 teacher or child care worker • Caregivers	3
3	Warner Initial Developmental Evaluation of Adaptive and Functional Skills (WIDEA-FS)	10-36 months corrected age	Yes	10-15 minutes	USA	Questionnaire	Caregivers	2
4	Caregiver Reported Early Development Instruments (CREDI)	0-36 months	Yes	CREDI-SF <5 min to complete CREDI-LF ~15 min to complete Scoring time not reported, however, an app and statistical software can be provided to score	Ghana, Tanzania, Zambia, Bangladesh, Cambodia, India, Jordan, Laos, Nepal, Pakistan, Philippines, Brazil, Chile, Colombia, Guatemala, the United States, China, Lebanon, Nepal, Pakistan. From association studies (3): Tanzania, Brazil	Questionnaire (with extra materials available online)	Caregivers	6

	Measure	Age	English language	Time to administer	Geographical uptake (incl. UK?)	Equipment needed	Training needed	Total no. papers
5	Global Scales for Early Development (GSED)	0-41 months. For use with 0-3 years	Yes	Not reported	Bangladesh, Côte d'Ivoire, Pakistan, United Republic of Tanzania, Brazil, The Republic of China, the Netherlands. One paper reports data from '32 countries'	GSED-SF contains 139 caregiver-reported items using a yes/no response scale. Many items (54/139; 39%) have accompanying audio or visual clues that are presented to the caregiver while they are being interviewed and are easily incorporated into an online version of self-report	Caregiver report GSED Short Form (SF) and/or directly administered Long Form (LF).	3
6	WHO Indicators of Infant and Young Child Development (IYCD)	• 0-42 months • 0-3 years	No, but the studies translated the tool from English	Not reported	Pakistan, Brazil, Costa Rica, Nicaragua, Paraguay, Peru, Bangladesh, India, Indonesia, Kenya, Malawi, Tanzania	There is a paper version to collect data, only using the electronic sources for item demonstrations (e.g., audio sounds or photos or videos illustrating items). There is also a tablet version that is available for use and can be obtained through the IYCD research team.	Caregivers	2
7	Parent Report of Children's Abilities (PARCA-R)	• 24 months • two years • 24-27 months	Yes, and German	• Two studies did not report length of time • One study reported it typically takes 15 minutes	Australia & New Zealand, Switzerland, and UK. From association studies (1): **UK**	Questionnaire	Parents	3
8	Early Childhood Development Assessment Scale-Caregiver Survey (ECDAS-CS)	36-59 months	Yes, and other languages translated from English	Not reported	Bangladesh, India, Myanmar, China	Questionnaire	ECDAS-DA = researchers who had experience in ECE. ECDAS-CD = caregiver	1

	Measure	Age	English language	Time to administer	Geographical uptake (incl. UK?)	Equipment needed	Training needed	Total no. papers
9	Brief Early Skills & Support Index (BESSI)	2.5-5.5 and 3-5.5 years	Yes	Not reported	**UK**	Questionnaire	Nursery staff/ teachers	1
10	Early Childhood Development Index (ECDI)	• 2-4 years • 3-5 years	Yes, and other languages translated from English	9.8 minutes for completion, with a 95 per cent confidence interval of 9.5–10.1 minutes	Jamaica, India, the United States, Mexico, Bulgaria, Uganda, Philippines, Kenya, Jordan. From association studies (16): 11 national and 5 sub-national samples in the final dataset including: Central African Republic, Chad, Democratic Republic of Congo, Ghana, Kenya(Mombasa), Kenya (Nyanza Province), Madagascar (South), Malawi, Mauritania, Nigeria, Sierra Leone, Somalia (Northeast), Somalia(Somaliland), Swaziland, Togo, and Zimbabwe; Honduras, Costa Rica, Bangladesh; Vietnam, Mexico, Benin, Cote d'Ivoire, Cameroon, Chad, Congo, DR Congo, Eswatini, Gambia, Ghana, Guinea, Guinea-Bissau, Lesotho, Madagascar, Mali, Malawi, Mauritania, Nigeria, Sierra Leone, Togo, Zimbabwe, Indonesia, Uganda	Questionnaire within MICS survey	Interviewers from MICS survey to caregivers.	1

	Measure	Age	English language	Time to administer	Geographical uptake (incl. UK?)	Equipment needed	Training needed	Total no. papers
11	Early Years Toolbox (EYT)	2.5-5 years	Yes	25-30 min	Australia. From association studies (1): Australia	iPad	Child assessors trained in the use of these measures administered all tasks.	1
12	International Development and Early Learning Assessment (IDELA)	3.5-6 years	Yes	30 minutes	Afghanistan, Bolivia, Ethiopia, Uganda, Vietnam. From association studies (2): Colombia	Direct assessment requiring various low-cost stimuli (e.g. puzzle, paper, book)	The assessment was administered by a trained enumerator, usually a field officer recruited from the local population	1
13	Playful Learning Observation Tool (PLOT)	•3.5-5 years	Yes	Not reported - depends on the classroom and the time the teacher dedicates to free play time	USA	Not reported	Observer - the PLOT is a high-inference observation tool intended for use by observers very familiar with preschool classrooms (preferably Head Start classrooms) and with playful learning.	1
14	McCarthy Scales of Children's Abilities (MSCA)	2.5- 8 years	Yes, and Basque	Up to 60 minutes	Basque region of Spain	Not reported	The MSCA-E was administered individually to each child by a trained neuropsychologist, complying with the requirements for proper assessment	1
15	The Early Human Capability Index (eHCI)	2-6 years in study, though eHCI is designed for use with 3-5 year olds	translated into local languages	up to 10 minutes	Brazil, China, Kiribati, Lao People's Democratic Republic, Samoa, Tonga, Tuvalu	The eHCI requires minimal resources to be implemented; the tool is available for anyone to use free of charge,	Little enumerator training is required, and it can be completed quickly and easily by any adult who knows the child	1
16	Preschool Child Development Inventory (PCDI)	Valid and reliable for use with young children aged 3-6 years	No	Not reported	Iceland	Questionnaire	Caregiver	1
17	Mongolian Rapid Baby Scale (MORBAS)	0-42 months	No	15 min	Mongolia	None	Caregivers	1

	Measure	Age	English language	Time to administer	Geographical uptake (incl. UK?)	Equipment needed	Training needed	Total no. papers
18	Taiwan Birth Cohort Study-Developmental Instrument (TBCS-DI)	6-66 months. 6-60 months	No	Not reported	Taiwan	Not reported	Unclear	2
19	Malawi Developmental Assessment Tool (MDAT)	2 to 24 months	No, but the study translated the tool from English	Dominican Republic. From association studies (3): Malawi	A group of nine clinical psychology undergraduate evaluators conducted the assessments in three separate rooms; two evaluators assessed each child and each of them provided their own set of scores.	Not reported	Not reported	1
20	The Griffiths Developmental Scales-Chinese (GDS-C)	The average ages of the ASD and TD children were 5.51 ± 1.21 and 4.85 ± 1.06 years, respectively. For use with children 0-6 years	No, but adapted from English version	Not reported	China.	Not reported	Qualified professional assessors	1
21	The Toddler Language and Motor Questionnaire (TMLQ)	15-36 months	No	Not reported	Iceland	Questionnaire	Caregiver	1
22	Cambodian Developmental Milestone Assessment Tool (cDMAT)	0-84 months	Yes, and Khmer	Not reported	Cambodia	Not reported	The assessors were 12 nurses from local health centres and 12 preschool teachers in rural and semi-urban districts who had been engaged previously in training on disability screening and inclusive education techniques	1

	Measure	Age	English language	Time to administer	Geographical uptake (incl. UK?)	Equipment needed	Training needed	Total no. papers
23	Brigance Inventory of Early Development (IED-ii)	18-29 months	Yes	15 min for Brigance,	USA	Form/questionnaire	Parental responses + unstructured clinician observations	1
24	Mullen Scales of Early Learning (MSEL)	• 6 to 45 months at the time of the initial evaluation with the MSELM • Mean age of 3.38 years • aged 2–10 years	Yes, and Spanish	Not reported	USA. From association studies (14): USA, Norway, Australia, Bangladesh, Israel, Austria, Canada	Mainly not reported, and in one study a questionnaire and engaging materials	Not reported, and in one instance a subgroup was conducted by experienced doctoral-level clinicians, with some cognitive testing in the typical group done by trained research assistants with on-site doctoral-level supervision	3
25	Denver Developmental Screening Test (DDST-II)	6-42 months	Yes	Not reported	Colombia	Not reported	Interviewers administered the tests	2
26	Battelle Developmental Inventory (BDI-2)	6-42 months	Yes	Not reported	Non-specialised interviewers administered the tests	Not reported	Colombia. From association studies (3): USA, Colombia	2
27	Vineland Adaptive Behaviour Scales (VABS-II)	2-5 years	Yes, and other	The VABS-II was designed to be administered individually and each interview takes about an average of 30 min to complete.	India. From association studies (12): Australia, Taiwan, USA, Israel, Singapore, Austria, Canada	Questionnaire-interview	Psychologists, received training for VABS-II. The caregivers responded to semi-structured interview	1
28	Rapid Neurodevelopmental Assessment (RNDA)	2-5 years	Yes	Not reported	Bangladesh		Professionals and paraprofessionals	1
29	The Differential Ability Scales (DAS-II)	2-10 years	Yes	30-70 minutes	USA. From association studies (2): Australia.	Not reported	All cognitive and diagnostic evaluations for children in the ASD and DD groups were conducted by experienced doctoral-level clinicians, with some cognitive testing in the typical group done by trained research assistants with on-site doctoral-level supervision	1

	Measure	Age	English language	Time to administer	Geographical uptake (incl. UK?)	Equipment needed	Training needed	Total no. papers
30	Hawaii Early Learning Profile (HELP)	0-3 years	Yes	Not reported	USA	Observational measure using videos of children with familiar adults in everyday activities and in intervention sessions	Observers with a varied levels of experience and training.	1
31	The Intergrowth Neurodevelopmental Assessment (INTER-NDA)	• mean=26 months (range=22 to 30 months old) • 24 months • mean age at assessment was 24.8 months	Yes, and other languages.	• Two studies reported approximately 15 minutes • 35-45 minutes	Grenada, West Indies, Brazil, Kenya, India, Italy, and **UK**. From association studies (1): Brazil	• Care was taken to ensure that the components of the INTER-NDA's kit were familiar to Caribbean children, and commonly encountered in Caribbean households. • App and user-friendly apparatus • Electronic, tablet-based data collection and management system (the NeuroApp).	• Administered using a combination of psychometric techniques (direct administration, concurrent observation and caregiver reports). Specifically, seven non-specialist child developmental assessors. • non-specialist research staff • can be administered reliably, in the field, by trained non-specialists.	2
32	Merrill-Palmer-Revised (M-P-R)	0-78 months	Yes	30-60 minutes	USA	Stimuli book, easel book, fido book, toys and manipulatives	Trained Assessor	1
33	Bayley Scales of Infant and Toddler Development (BSID-III)	1-42 months	Yes	30-70 minutes	Various incl. **UK**	Caregiver report form; observational checklist; multiple specific stimuli	Trained clinician	23
34	Australian Developmental Screening Test (ADST)	•15-62 months	Yes	15-20 minutes	Australia	An ADST Test Kit and forms must be purchased (A$700)	NR	1
	34							**98**

Appendix H

Tools' Domains and Items Overlap

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
ASQ®-3 27 months											
Communication											
1	Without your giving him clues by pointing or using gestures, can your child carry out at least three of these kinds of directions? "Put the toy on the table", "Close the door", "Bring me a towel", "Find your coat", "Take my hand", "Get your book"		X							X	X
2	If you point to a picture of a ball/ kitty/ cup/ cat/ etc and you ask your child, "What is this?" does your child correctly name at least one picture?			X	X	X	X		X	X	X
3	When you ask her to point to her nose, eyes, hair, feet, ears and so forth, does your child correctly point to at least seven body parts? (She can point to parts of herself, you, or a doll. Mark 'sometimes' if she correctly points to at least three different body			X			X		X	X	X
4	Does your child correctly use at least two words like 'me', 'I', 'mine' and 'you'?			X	X		X				
5	Does your child make sentences that are three or four words long? Please give an example:			X	X	X				X	X
6	Without giving your child help by pointing or using gestures, ask him to "put the book on the table" and "put the shoes under the chair". Does your child carry out both of these directions correctly?			X	X	X	X				X
Gross motor											
7	Does your child walk either up or down at least two steps by himself? He may hold onto the railing or wall. (You can look for this at a store, on a playground, or at home)								X	X	

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
8	Does your child run fairly well, stopping herself without bumping into things or falling?			X		X					X
9	Does your child jump with both feet leaving the floor at the same time?			X	X	X	X				X
10	Without holding onto anything for support, does your child kick a ball by swinging his leg forward?			X		X					X
11	Does your child jump forward at least 3 inches with both feet leaving the ground at the time?					X					
12	Does your child walk up stairs, using only one foot on each stair? (The left foot is on one step, the right foot is on the next step.) She may hold onto the railing or wall.								X		
Fine motor											
13	Does your child use a turning motion with her hand while trying to turn doorknobs, wind up toys, twist tops, or screw lids on and off jars?			X		X	X				X
14	Does your child flip switches on and off?										
15	After your child watches you draw a line from the top of the paper to the bottom with a pencil, crayon, or pen, ask him to make a line like yours. Do not let your child trace your line. Does your child copy you by drawing a single line in a vertical direction?			X			X				X
16	Does your child stack seven small blocks or toys on top of each other by herself? (You could also use spools of thread, small boxes, or toys that are about 1 inch in size)			X		X	X				X
17	Can your child string small items such as beads, macaroni, or pasta wagon wheels onto a string or shoelace?										
18	After your child watches you draw a line from one side of the paper to the other side, ask her to make a line like yours. Do not let your child trace your line. Does your child copy you by drawing a single line in a horizontal direction?			X		X					X

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
Problem solving											
19	Does your child pretend objects are something else? EG does your child hold a cup to his ear, pretending it is a telephone? Does he put a box on his head, pretending it is a hat? Does he use a block or small toy to stir food?			X		X					X
20	Does your child put things away where they belong? For example, does she know her toys belong on the toy shelf, her blanket goes on her bed, and dishes go in the kitchen?			X					X		
21	When looking in the mirror, ask 'Where is [child's name]?' Does your child point to his image in the mirror?										
22	If your child wants something she cannot reach, does she find a chair or box to stand on to reach it? (EG, to get a toy on a counter or to 'help' you in the kitchen)?			X							
23	While your child watches, line up four objects like blocks or cars in a row. Does your child copy or imitate you and line up four objects in a row? You can also use spools of thread, small boxes, or other toys.					X					
24	When you point to a figure and ask your child, 'What is this?' does your child say a word that means a person or something similar? (Mark 'yes' for responses like 'snowman, 'boy', 'man', 'girl', 'daddy,' 'spaceman' and 'monkey') Please write your child's response here.										
Personal-social											
25	If you do any of the following gestures, does your child copy at least one of them?										
26	Open and close your mouth										
27	Blink your eyes										
28	Pull on your earlobe										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
29	Pat your cheek										
30	Does your child eat with a fork?									X	
31	When playing with either a stuffed animal or doll, does your child pretend to rock it, feed it, change its diapers, put it to bed, and so forth?									X	X
32	Does your child push a little wagon, stroller, or other toy on wheels, steering it around objects and backing out of corners if he cannot turn?										X
33	Does your child call herself “I” or “me” more often than her own name? EG “I do it” more often than “Juanita do it”.										
34	Does your child put on a coat, jacket, or shirt by himself?						X			X	

Overall- free text questions

35	Do you think your child hears well? If no, explain.							X			
36	Do you think your child talks like other toddlers her age? If no, explain							X			
37	Can you understand most of what your child says? If no, explain						X				X
38	Do you think your child walks, runs, and climbs like other toddlers his age? If no, explain							X			
39	Does either parent have a family history of childhood deafness or hearing impairment? If yes, explain										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
40	Do you have concerns about your child's vision? If yes, explain							X			
41	Has your child had any medical problems in the last several months? If yes, explain										
42	Do you have any concerns about your child's behaviour? If yes, explain							X			
43	Does anything about your child worry you? If yes, explain							X			

ASQ®-SE-24 months

1	Does your child look at you when you talk to him?										X
2	Does your child seem too friendly with strangers?										
3	Does your child laugh or smile when you play with her?										X
4	Is your child's body relaxed?										
5	When you leave, does your child stay upset and cry for more than an hour?										X
6	Does your child greet or say hello to familiar adults?				X					X	
7	Does your child like to be hugged or cuddled?										
8	When upset, can your child calm down within 15 minutes?										X

		AASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
9	Does your child stiffen and arch his back when he's picked up?										
10	Is your child interested in things around her, such as people, toys and food?			X							X
11	Does your child cry, scream, or have tantrums for long periods of time?										
12	Do you and your child enjoy mealtimes together?										
13	Does your child have eating problems? EG does he stuff food, vomit, eat things that are not food, or ____?										
14	Does your child sleep at least 10 hours in a 24-hour period?										
15	When you point at something, does your child look in the direction you're pointing?										
16	Does your child have trouble falling asleep at nap time or at night?										
17	Does your child get constipated or have diarrhea?										
18	Does your child follow simple directions? EG does she sit down when asked?	X									X
19	Does your child let you know how he is feeling with words or gestures? For example, does he let you know when he is hurt, hungry or tired?			X			X				
20	Does your child check to make sure you are near when exploring new places, such as a park or a friend's home?										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
21	Does your child do things over and over and get upset when you try to stop her? EG does she rock, flap her hands, spin or _____?										
22	Does your child like to hear stories or sing songs?			X		X				X	
23	Does your child hurt himself on purpose?										
24	Does your child like to be around other children? EG does she move close to or look at other children?										X
25	Does your child try to hurt other children, adults or animals? EG by kicking or biting?				X						X
26	Does your child try to show you things by pointing at them and looking back at you?										X
27	Does your child play with objects by pretending? EG does your child pretend to talk on the phone, feed a doll, or fly a toy aeroplane?	X		X		X					X
28	Does your child wake three or more times during the night?										
29	Does your child respond to his name when you call him? EG does he turn his head and look at you?										
30	Is your child too worried or fearful? If 'often' or 'sometimes', please describe										
31	Has anyone shared concerns about your child's behaviours? If 'often' or 'sometimes', please describe										
32	Overall: do you have concerns about your child's eating or sleeping behaviours? If yes, please explain							X			

		AASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
33	Overall: does anything about your child worry you? If yes, please explain							X			
34	Overall: what do you enjoy about your child?										

CREDI- LF- start for 24-29 months

1	Can the child sit or play on their own for at least 20 minute?										X
2	Can the child walk several steps on his/her own, without holding on or receiving support?									X	X
3	Can the child bend down to the ground and stand up again without falling and without holding onto a person or object?										X
4	Does the child ask you for help using signs or words when they cannot do something on their own? (EG to reach an object up high?)										
5	Does the child try to repeat sounds or words said by other people?										
6	Can the child climb onto an object such as a chair or bench?									X	X
7	Can the child figure out how to turn a spoon or object if you give it to them the wrong way round?					X					
8	Does the child stop at least briefly when told “no” or “stop that”?										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
9	Can the child kick a ball or other round object forward using their foot?	X				X					X
10	Can the child point to a person or object when asked (e.g. 'where is mama?' or 'where is the ball?')?					X				X	X
11	Can the child drink from a cup (without a lid) on their own without spilling?									X	X
12	Does the child imitate animal or other sounds (EG vroom for a car, moo for a cow)?										
13	Can the child run more than a few steps without falling or bumping into objects?	X				X					X

Start for 30-35 months

14	Can the child draw a line or shape on paper with a pen or crayon, or in the dirt with a stick?	X				X	X				X
15	Can the child answer simple questions (EG do you want water?') by saying 'yes' or 'no' rather than nodding?										
16	Can the child stack three or more small objects (EG bocks, cups, bottle caps) on top of each other?	X				X	X				X
17	Does the child imitate others' behaviours? (EG washing hands or dishes?)										X
18	Does the child sometimes share things (EG food, toys) with others without being told?										X
19	Can the child follow orders or instructions that have more than one part (EG go get water and go to bed)?					X					X
20	Can the child say five or more separate words (EG names like 'mama' or objects like 'ball')?					X					X

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
21	Is the child kind to younger children (EG speaks to them nicely and touches them gently)?										
22	Can the child walk on an uneven surface (EG a bumpy or steep road) without falling?						X				
23	Does the child listen to someone telling a story with interest?		X			X					
24	Can the child ask for something (EG food, water) by name when they want it?					X				X	X
25	Does the child involve others in play (IE play interactive games with other children)?					X				X	X
26	Can the child correctly name at least one family member other than mum and dad (EG name of brother, sister, aunt, uncle)?									X	
27	Does the child play by pretending objects are something else (EG imagining a bottle is a doll, a stone is a car, or a spoon is an airplane)?	X				X					X
28	Does the child show sympathy or look concerned when others are hurt or sad?										X
29	Can the child walk backwards?										
30	Does the child show curiosity to learn new things (EG by asking questions or exploring new areas)?		X								X
31	Can the child feed him/herself using a spoon or other utensil without spilling	X								X	
32	Can the child concentrate on one task (EG playing with friends, eating a meal) for 20 minutes?										X
33	Does the child know the name of at least two body parts (e.g., arm, eye, or nose)?	X				X			X	X	X

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
34	If you show the child an object, he/she knows well (EG a cup/ animal), can they consistently name it?	X			X	X	X		X	X	X
35	Can the child speak using short sentences of two words that go together (EG 'mama go' or 'dada eat')?					X	X			X	X
36	Can the child use a tool (EG a stick or spoon) to reach objects that are far away?	X									
37	Can the child indicate when he/she needs to go to the toilet?										X
38	Can the child say ten or more separate words (e.g., names like "Mama" or objects like "ball")?				X	X	X			X	
39	Can the child remove an item of clothing (EG take off his/her shirt)?						X			X	
40	can the child tell you when he/she is tired or hungry?		X								
41	Does the child usually finish an activity they enjoy (EG a game or book)?										X
42	Can the child easily switch back and forth between activities (EG go back to a game after being interrupted)?										
43	Can the child sing a short song or repeat parts of a rhyme from memory by his/her self?				X		X			X	
44	Can the child jump with both feet leaving the ground?	X			X	X	X				X
45	Can the child speak using sentences of three or more words that go together (e.g., "I want water" or "The house is big")?	X			X	X				X	X
46	Can the child whisper?										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
47	Does the child greet neighbours or other people he/she knows without being told (EG by saying hello or gesturing hello)?		X							X	
48	Can the child unscrew the lid from a bottle or jar?	X				X	X				X
49	Can the child correctly ask questions using any of the words “what,” “which,” “where,” or “who”?				X		X				X
50	Can the child correctly use any of the words “I,” “you,” “she,” or “he” (e.g., “I go to store,” or “He eats rice”)?	X			X		X				
51	Does the child pronounce most of his/her words correctly?	X					X	X			X
52	Can the child count up to five objects (e.g., fingers, people)?				X	X	X				X
53	Does the child ask about familiar people other than parents when they are not there (e.g., “Where is the neighbour?”)?				X						
54	If you show the child two objects or people of different size, can she/he tell you which one is the big one and which is the small one?				X	X	X				X
55	Can the child stand on one foot for several seconds without holding onto a person or object (EG wall or furniture)?					X	X				X
56	Can the child identify at least one colour (EG red, blue, yellow)?				X	X					
57	Does the child regularly use describing words such as ‘fast’, ‘short’, ‘hot’, ‘fat’, or ‘beautiful’ correctly?						X				
58	If you point to an object, can the child correctly use the words “on”, “in”, or “under” to describe where it is? (EG ‘the cup is on the table’ instead of ‘the cup is in the table’)	X			X	X	X				X
59	Can the child explain in words what common objects like a cup or chair are used for?				X	X	X				

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
60	Does the child ask 'why' questions? (EG why are you tall?)				X		X				X
61	If you ask the child to give you three objects (e.g., stones, beans), does the child give you the correct amount?				X	X	X				
62	Does the child usually put objects or toys back where they belong after using them?	X									
63	Does the child frequently act impulsively or without thinking (EG running into the street without looking)?										
64	Does the child sometimes save things like candy or new toys for the future?										
65	Can the child say what others like or dislike (e.g., "Mama doesn't like fruit," "Papa likes football")?				X		X				
66	Can the child fasten and unfasten buttons without help?						X				
67	Can the child talk about things that will happen in the future using the correct language (EG 'tomorrow he will attend school' or 'next week we will go to the market')?										
68	Can the child talk about things that have happened in the past using correct language (e.g., "Yesterday I played with my friend" or "Last week she went to the market")?				X		X				X
69	Does the child know the names of any letters (EG A, B, C)?										

CREDI-SF 24-29 months

1	If you show the child an object, he/she knows well (e.g., a cup or animal), can he/she consistently name it?	X		X		X	X		X	X	X
2	Can the child say ten or more separate words (e.g., names like "Mama" or objects like "ball")?			X		X	X			X	

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
3	Can the child sing a short song or repeat parts of a rhyme from memory by him/ herself?			X			X			X	
4	Can the child jump with both feet leaving the ground?	X		X		X	X				X
5	Can the child speak using sentences of three or more words that go together (e.g., “I want water” or “The house is big”)?	X		X		X				X	X
6	Can the child correctly ask questions using any of the words “what,” “which,” “where,” or “who”?			X			X				X
7	Can the child correctly use any of the words “I,” “you,” “she,” or “he” (e.g., “I go to store,” or “He eats rice”)?	X		X			X				
8	Does the child ask about familiar people other than parents when they are not there (e.g., “Where is the neighbour?”)?			X							
9	Can the child count up to five objects (e.g., fingers, people)?			X		X	X				X
10	Can the child identify at least one colour (e.g., red, blue, yellow)?			X		X					
11	Does the child often kick, bite, or hit other children or adults?		X								X
12	If you show the child two objects or people of different size, can he/she tell you which one is the big one and which is the small one?			X		X	X				X
13	Does the child become extremely withdrawn or shy in new situations?										
14	If you point to an object, can the child correctly use the words “on,” “in,” or “under” to describe where it is (e.g., “The cup is on the table” instead of “The cup is in the table.”).	X		X		X	X				X
15	Does the child ask “why” questions (e.g., “Why are you tall?”)?			X			X				X

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
16	If you ask the child to give you three objects (e.g., stones, beans), does the child give you the correct amount?			X		X	X				
17	Can the child explain in words what common objects like a cup or chair are used for?			X		X	X				
18	Can the child dress him/herself (e.g., put on his/her pants and shirt without help)?						X				X
19	Can the child say what others like or dislike (e.g., “Mama doesn’t like fruit,” “Papa likes football”)?			X			X				
20	Can the child talk about things that have happened in the past using correct language (e.g., “Yesterday I played with my friend” or “Last week she went to the market”)?			X			X				X

GSED-LF start for children 24-27 months until 36 months

Stream A

1	Runs well	X		X							X
2	Kicks a ball from stationary position	X		X							X
3	Runs and kicks a ball well										X
4	Kneels and then stands, without using hands										
5	hops forward on 1 foot 3 steps										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
6	Jumps with 2 feet together	X		X	X		X				X
7	Jumps over a piece of paper (widthways)	X									
8	Walks along line heel-to-toe										
9	Throws beanbag on to a cloth										X
10	Stands on 1 foot <5 seconds			X			X				X
11	Walks on tiptoes 6 or more steps										
12	Moves from sitting to standing without using hands										
13	Stands on 1 foot >5 seconds			X			X				X
14	Throws ball up into air and catches it										

Stream B

15	Points at 5 pictures in book (image set A)			X						X	X
16	Shows interest in story (set B story book)		X		X					X	
17	Follows 2 step commands			X							X

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
18	Says sentences with 2 words together			X			X			X	X
19	Names 4 pictures (image set A)	X									
20	Uses 5 clear words			X							X
21	Matches pictures (set C)										
22	Names 5 objects	X		X	X		X		X	X	X
23	Uses multiple-word utterances	X			X					X	
24	Speaks clearly in sentences										
25	Knows actions or functions of 3 or more objects			X	X		X				
26	Points to parts of whole objects (set A)										
27	Says first name										
28	Names 10 objects			X	X		X			X	
29	Understands “more” (Set H)										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
30	Identifies 2 or more colours (set E)			X	X						
31	Knows use of objects (set A)			X	X		X				
32	Names at least two colours (Set E)										
33	Identifies 5 action pictures (Set I)										
34	Identifies at least 2 shapes (set F)										
35	Talk easily about daily events										
36	Describes picture (set B)										
37	Gives logical response to a question										
38	Categorises things										
39	Matches 3 colours (set E)										
40	Understands adjective “faster” (Set G)										
41	Names actions (5) (Set I)										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
42	Taps with 2 blocks (set J)										
43	Taps with 4 blocks (set J)										
44	Taps with 8 blocks (set J)										

Stream C

45	Finds object under 2 alternating cups										
46	Inserts 2 shapes in board (3 shape block board)										
47	Inserts 3 shapes in board in 2 minutes (3 shape block board)										
48	Uses objects in play with someone	X		X							X
49	Scribbles on paper (circular scribble)								X		X
50	Builds tower of 6 blocks	X		X			X				X
51	Understands concept of “1”										
52	Inserts 3 shapes in rotated board in 2 minutes (3 shape block board)			X							
53	Builds truck/ lorry of blocks	X									

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
54	Unscrews and screws jars	X		X			X				X
55	Engages in representational play	X		X							X
56	Inserts 3 shapes in 15 seconds (3 shape block board)										
57	Copies 2-part activity										
58	Puts pegs in boards in < 30 seconds (6-hole peg board)										
59	Draws horizontal line	X									
60	Understands “more”										
61	Imitates building bridge										
62	Picks longest stick 3 of 3						X				X
63	Copies a circle						X				X
64	Builds a wall of blocks										
65	Understands concept of size			X	X		X				X
66	Understands prepositions	X		X	X		X				X

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
67	Copies a cross or plus sign										
68	Counts 3 or more objects			X	X		X				X
69	Copies a square										
70	Draws 3 or more body parts										

GSED-SF start for children 24 < 27 months

1	Can your child stack 3 or more small objects (e.g., blocks, cups, bottle caps) on top of each other?	X		X		X					X
2	Can your child walk on an uneven surface (EG a bumpy or steep road) without falling?			X							
3	Does your child usually communicate with words what he/she wants in a way that is understandable to others?	X		X				X		X	X
4	Can your child say 10 or more words in addition to ‘mama’ and ‘dada’?			X	X	X				X	
5	When looking at pictures, if you say to your child ‘what is this?’ can your child say the name of the object that you point to?	X		X	X	X			X	X	X
	GSED-SF start for children 27 < 30 months										
6	Can your child speak using short sentences of 2 words that go together (e.g., ‘mama go’ or ‘dada eat’)?			X		X				X	X

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
7	Can your child unscrew the lid from a bottle or jar?	X		X		X					X
8	Does your child help out around the house with simple chores, even if he/she doesn't do them well?									X	
9	Is your child able to go poo or pee without having accidents (wetting or soiling him-/herself)?									X	

GSED-SF start for children 30 < 33 months

10	Can your child speak using sentences of 3 or more words that go together (EG 'I want water', or 'the house is big')?	X			X	X				X	X
11	Can your child name at least two body parts (EG arm, eye, or nose)?			X					X	X	X
12	Can your child remove an item of clothing (EG take off his/her shirt)?			X						X	
13	Can your child say 15 or more separate words (EG names such as 'mama' or objects such as 'ball')?										
14	Can your child jump with both feet leaving the ground?	X		X	X	X					X
15	Can your child tell you or someone familiar his/her own name (or nickname) when asked?					X					

GSED-SF start for children 33 < 36 months

16	Can your child correctly ask questions using any of the words 'what', 'which', 'where, or 'who'?			X	X						X
17	Does your child show respect around elders?										X

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
18	Can your child correctly use any of the words “I,” “you,” “she,” or “he” (e.g., “I go to store,” or “He eats rice”)?	X		X	X						
19	Can your child sing a short song or repeat parts of a rhyme from memory by him/herself?			X	X					X	
20	Does your child know the difference between the words ‘big’ and ‘small’? For example, if you ask, ‘give me the big spoon’, can your child understand which item to give if there are 2 different sizes?			X	X	X					X
21	Does your child pronounce most of his/her words correctly?	X						X			X
22	Can your child go to the toilet by him-/herself?										
23	If you point to an object, can your child correctly use the words “on,” “in,” or “under” to describe where it is (e.g., “The cup is on the table” instead of “The cup is in the table.”).	X		X	X	X					X
24	Can your child put on at least 1 item of clothing by themselves?	X								X	
25	Can your child explain in words what common objects like a cup or chair are used for?			X	X	X					
26	Can your child draw a straight line?	X		X		X					X
27	Can your child say what he/she likes or dislikes? (EG ‘I like sweets’)?										
28	If you show your child two objects or people of different size, can he/she tell you which one is the big one and which is the smaller one?			X	X	X					X
29	Does your child regularly use describing words such as ‘fast’, ‘short’, ‘hot’, ‘fat’, or ‘beautiful’ correctly?			X							
30	Does your child know how to keep quiet when the situation requires it (EG at ceremonies, when someone is asleep)?										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
31	Does your child ask “why” questions (e.g., “Why are you tall?”)?			X	X						X
32	Can your child stand on one foot without any support for at least a few seconds?			X		X					X
33	If you ask your child to give you three objects (e.g., stones, beans), does your child give you the correct amount?			X	X	X					
34	Does your child understand the term ‘longest’? For example, if you ask your child to choose ‘which is the longest of 3 objects (EG 3 spoons or sticks), would he/she be able to choose the longest?					X					X
35	Can your child talk about things that have happened in the past using correct language (e.g., “Yesterday I played with my friend” or “Last week she went to the market”)?			X	X						X
36	Can your child tell a story?										X
37	Can your child tell you when he/she is happy, angry or sad?		X								
38	Can your child name at least 1 colour (EG red, blue, yellow)?					X					
39	Can your child count up to 5 objects (EG fingers, people)?			X	X	X					X
40	If you draw a circle, can your child do it just as you did?					X					X
41	Can your child tell you when others are happy, angry or sad?										
42	Can your child talk about things that will happen in the future using correct language (e.g., “Tomorrow he will attend school” or “Next week we will go to the market”)?			X	X						
43	Can your child fasten and unfasten buttons without help?			X							

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
44	Can your child dress him-/herself completely (except for shoelaces, buttons and zippers)?				X						X
45	Can the child say what others like or dislike (e.g., “Mama doesn’t like fruit,” “Papa likes football”)?			X	X						

PEDS:R 30 months

1	Please list any concerns about your child’s learning, development, and behaviour.										
2	Do you have any concerns about how your child talks and makes speech sounds?	X		X			X				X
3	Do you have any concerns about how your child understands what you say?										X
4	Do you have any concerns about how your child uses his or her hands and fingers to do things?										
5	Do you have any concerns about how your child uses his or her arms and legs?	X									
6	Do you have any concerns about how your child behaves?	X									
7	Do you have any concerns about how your child gets along with others?										
8	Do you have any concerns about how your child is learning to do things for himself/herself?										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
9	Do you have any concerns about how your child is learning preschool or school skills?										
10	Do you have any concerns about how your child is behind others or can't do what other kids can?										
11	Do you have any concerns about your child's health or about how he or she sees, hears, eats, or sleeps?	X	X								
12	Please list any other concerns.										

PEDS:DM (developmental milestones) 30 months

1	Does your child try to scribble with crayons or markers?					X					X
2	How many of these body parts can your child point to if you say, "Where are your eyes? "..."Where is your nose?"..."feet?"..."hair?"..."mouth?"..."ears?"	X		X			X			X	X
3	If you ask your child ***"What's this?"*** and showed things like a spoon, cup, doll, truck, box, crayon, cookie, chair, or light, how many names for these or other things does he or she say?	X		X	X	X	X			X	X
4	How does your child get up a set of stairs?	X								X	
5	Does your child try to help when it is time to put things away?	X									
6	Does your child use two things together such as taking a doll for a ride in a toy car, having a truck carry things, or having a doll take a pet for a walk?										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
WIDEA-FS											
1	Easily drinks formula or breast milk										
2	Easily swallows baby food										
3	Chews solid food										
4	Finger feeds										
5	Eats using a spoon			X							
6	Drinks from cup without a lid			X							X
7	Eats using a fork	X									
8	Holds arms up so you can put shirt on										
9	Removes socks			X			X				
10	Pulls pants down			X			X				
11	Pulls up a zipper once it is started										
12	Puts on t-shirt	X					X				

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
13	Removes all clothes										
14	Indicates a wet diaper										
15	Indicates a soiled diaper										
16	Voids into potty chair or toilet						X				
17	Sits on potty chair and has bowel movement						X				
18	Rolls both ways										
19	Maintains sitting without support										X
20	Crawls short distance										
21	Walks few feet with assistance (cruises)										X
22	Scoots/ crawls 10 feet or moves his/her wheelchair 10 feet										
23	Walks 10 feet independently			X							X
24	Crawls up stairs								X		

		AASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
25	Gets on and off a chair			X							X
26	Walks up stairs with hand held	X							X		
27	Understands words for people in immediate family (mommy, daddy) (R)			X							
28	Demonstrates 2 syllable babbling (baba)(E)										X
29	Understands words for some common objects (R)			X		X					X
30	Gestures a social greeting (wave, blow a kiss) (E)		X	X							
31	Carries out a 1 step oral request with gesture (pick up toy, cup) (R)										
32	Uses single words or signs to request or communicate (E)			X			X				X
33	Carries out a 1 step oral request without gesture (R)	X									X
34	Identifies one body part (R)	X		X			X		X		X
35	Identifies three or more body parts (R)	X		X			X		X		X
36	Points at pictures (R)					X					
37	Has at least 10 words or 10 signs (E)			X	X	X	X				

		AASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
38	Combines words or signs to make needs known (E)	X		X	X	X	X				X
39	Names pictures (E)	X		X	X	X	X		X		X
40	Plays peekaboo, patty cake or so big										X
41	Looks for object dropped out of sight										X
42	Initiates social contacts with peers										
43	Takes turns rolling a ball										X
44	Imitates another child										
45	Recognises familiar song		X	X	X		X				
46	Starts mechanical toy or computer										
47	Can pretend play with a doll or toy	X									X
48	Turns pages in a book										
49	Points at pictures when you read a story					X					

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
50	Helps with simple household tasks						X				

WHO IYCD

Motor

1	Does your child try to move his/her head (or eyes) to follow an object or person?										
2	While your child is on his/her back, does he/she bring his/her hands together?										
3	Does your child put objects or hands to his/her mouth?										
4	Does your child hold his/her head steady for at least a few seconds, without it flop-ping to the side?										
5	When he/she is on his/her tummy, does your child hold his/her head straight up, looking around for more than a few seconds?										
6	Does your child try to reach for objects that are in front of him/her by extending one or both arms?										
7	Does your child reach for AND HOLD an object, at least for a few seconds?										
8	Does your child sit WITH support, either leaning against something (furniture or person), or by leaning forward on his/ her hands?										
9	If an object falls to the grounds, does your child look for it?									X	

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
10	Does your child drink from an OPEN cup while you hold it?										
11	Does your child bang objects together, or bang an object on the table or on the ground?										
12	Does your child pass an object from one hand to the other?										
13	Does your child sit upright, with fairly straight back and neck control, without holding on to you, an object, or resting hands on the floor?										
14	Does your child pick up small items, like pieces of food USING THUMB AND FINGER TIPS?										
15	Does your child intentionally SQUEEZE or PAT/HIT an object that makes a noise in order to hear it?										
16	Does your child pull him/herself up from the floor while holding onto something for support?										
17	Does your child pick up small items, like pieces of food using thumb and JUST ONE FINGER TIP?										
18	When someone holds one or both hands (just to balance), or when your child is holding furniture or other object WITH JUST ONE HAND, does he/she make a few steps forward, without tripping?									X	
19	Can your child push a toy car intention- ally, all four wheels on the ground?	X									
20	Does your child stand up WITHOUT holding onto anything, even if just for a few										
21	Does your child make any marks on paper or on the ground with a crayon/ pencil/ pen or a stick?				X				X		

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
22	Does your child climb onto an object (rock, porch, step, chair, bed, low table etc.)?			X						X	
23	Does your child bend down or squat to pick up an object from the floor and then stand up again, without help from a per- son or object?			X							
24	Does your child take several steps (3–5) forward without holding onto any person or object for support, even if he/she falls down immediately?			X						X	
25	Does your child make a scribble on paper, or on the ground, in a BACK AND FORTH manner?								X		
	For example, can he/she move the crayon/ pencil/ pen or stick back and forth?										
26	Does your child drink WITHOUT HELP from an OPEN cup?			X						X	
27	Does your child walk well, with coordination without falling down often?									X	
	With one foot in front of the other (rather than shifting weight side to side, stiff-legged)?										
28	While standing, does your child purpose- fully THROW the ball and not just drop it?					X					
29	Does your child unscrew a lid, at least turning the lid one rotation?	X		X		X	X				
30	Does your child stack at least two objects on top of each other, such as bottle tops, blocks, stones, etc.?	X		X		X	X				
31	While standing, does your child kick a ball by swinging his/her leg forward?	X		X		X					

		AASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
32	Does your child run well, without falling or bumping into objects?	X		X		X					
33	Does your child stand on one foot WITH SUPPORT by holding onto something or somebody?										
34	Does your child walk sideways, two or more steps WITHOUT any support?										
35	Does your child run and kick a ball, and do so successfully?					X					
36	Does your child jump WITH BOTH FEET LEAVING THE GROUND AT ABOUT THE SAME TIME?	X		X	X	X	X				
37	If you draw a straight line does your child do it, just as you did?	X		X		X	X				
38	If you draw a circle does your child do it, just as you did?					X	X				
39	Does your child stand on one foot WITH- OUT any support for at least a few seconds?			X		X	X				
40	Does your child dress him/herself completely (except for shoelaces, buttons and zippers)?				X	X	X				
	Language and cognitive										
41	Does your child turn his/her head towards your voice or some noise?										
42	When you TALK to your child, does he/she smile, make noises, or move arms, legs or trunk in response?										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
43	Does your child make sounds when LOOKING at toys or people (not crying)?										
44	Does your child make single sounds like “buh” or “duh” or “muh”?										
45	Does your child make noise or gesture TO GET YOUR ATTENTION?		X								
46	Does your child make two similar sounds TOGETHER like baba, mumu, pepe (single consonant vowel combinations)?									X	
47	Does your child point, shake head, or otherwise gesture to let you know WHAT HE/ SHE WANTs?										
48	Does your child say at least one word?										
49	Does your child try to SHOW you some- thing by reaching at you, pulling on your hand or clothes, using gestures, making noises (not words)?		X								
50	Does your child put together noises that sound sentence-like, even if they are not words?										
51	If you or someone else invites your child to play a familiar game (peek-a-boo, wave bye-bye, etc.), does your child know what to do without you showing them?										
	Does he/she understand, even if you don’t show him/her what to do? Does he/she try to participate, even if it’s not exactly correct? Does he/she attempt the movement,										
52	Does your child identify at least one object? For example, when you ask “where is the ball/spoon/cup/cloth/door/plate/bucket etc.” does your child look at or point to (or even name) the object?	X		X	X	X	X		X	X	
53	Without you making a gesture, does your child follow a simple spoken command or direction?	X	X							X	
	Does the child understand when you ask some- thing? Will he/she do at least one thing you ask the child to do? Can your child do it without you using a gesture?										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
54	Does your child ask for something (e.g., food, water) by name when he/she wants it?			X			X			X	
55	Does your child usually understand what you say to him/her when you say it only once? For example, you don't have to repeat what you said to be understood.							X			
56	Does your child identify at least one body part? For example, if you ask, "where's your eyes?" can he/she point to his/her eyes?	X		X			X		X	X	
57	Does your child follow directions with more than one step? For example, "Go to the kitchen and bring me a spoon"?			X		X					
58	Does your child use two WORDS together in a meaningful phrase/speak in short two-word sentences? For example, "mama go," "give mama," "daddy gone," etc.			X		X	X			X	
59	Does your child say at least six words?			X	X	X	X		X	X	
60	Does your child understand the words over/ on top of/ above, next to, under and in front of? If you ask your child to put something under the table or on top of the chair, does he/she understand?	X		X	X	X	X				
61	When looking at pictures, if you say to your child 'what is this?', does he/she SAY the name of the object that you point to?	X		X	X	X	X		X	X	
62	Does your child identify at least seven objects? For example, when you ask "where is the ball/spoon/cup/cloth/door/plate/bucket etc." does your child look at or point to (or even name) the objects?	X									

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
63	Does your child usually communicate with words what he/she wants in a way that is understandable to others?	X		X			X	X			
64	Does your child ask questions using the words what, words what, which, where, why, or who?			X	X		X				
65	Does your child speak in sentences that are 3–4 words long?	X			X	X	X			X	
66	Does your child know the difference be- tween the words “big” and “small”? For example, if you ask, “Give me the big spoon” can your child understand which one to give if there are two different sizes?			X	X	X	X				
67	Does your child talk about or explain things that have happened in the past using the past tense? For example, “I went to the market yesterday” (NOT “I go to the market yesterday”). Does your child tell you about something that happened on a last special event? Does he/she talk about anything that happened in the past?			X	X		X				
68	Does your child count up to 5?			X	X	X	X				
69	Does your child understand the term “longest”? For example, if you ask him/her to choose “which is the longest of 3 objects?” e.g., 3 spoons or sticks – would he/she be able to choose the longest?					X	X				
70	Does your child tell a story?						X				
	Socio-emotional										
71	Does your child stop crying or calm down when you come to the room after being out of sight, or when you pick him/her up?		X								
72	When you are about to pick up your child, does he/she act happy or excited?										

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
73	If you play a game with your child, does he/she respond with interest?		X							X	
	For example, if you play peek-a-boo, pat-a-cake, wave bye-bye etc. does your child smile, widen their eyes, kick or move arms or vocalize.										
74	Does your child smile when you smile or talk with him/her?										
75	Does your child smile, move excitedly, kick legs, move arms or trunk, or make coo noises when a known person enters the room or speaks to him/her?										
76	Is your child interested when he/she sees other children playing?		X								
	Does he/she watch, smile, or look excited?										
77	Is your child very interested to play with a new toy or object?		X			X					
78	Does your child usually get along with other children he/she plays with such as siblings or other family members?										
	When your child spends time with other children, does he/she mostly agree on things and not quarrel?										
79	When your child sees other children playing, does your child play next to them without joining in?		X								
	Does he/she play alongside?										
80	Does your child play together with other children not just by sitting next to each other and doing the same thing, but actively interacting?			X		X				X	
81	When your child needs to use the toilet, does he/she show you by pulling on his/her clothes, holding him/herself, crying, or some other way?			X							

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
82	Does your child share things with other children or family members without being told to do so?			X							
83	Does your child show sympathy or look concerned?			X							
	If your child's friend or sibling is sad, what does your child do? Does your child try to com- fort the other?										
84	If your child has a problem with another child, fighting over a toy, or something else, does your child come to you or an- other adult and ask for help?										
85	Does the child usually show respect for authority figures (e.g., teacher, grandparent, elder)?				X						
86	Does your child worry if a known person is sick or hurt?										
	Is child interested in the wellbeing/health of others?										
87	Does your child follow through on instructions, for example, does he/she finish his/her chores?										
88	Does your child keep working on something until he/she is finished?			X							
89	Does your child use "imaginary" or not visible objects in play?	X								X	
	For example, pretending to talk on a cell phone or pretending to give food to a child (even though they have nothing to give)										
90	Does your child have difficulty taking turns when playing with others?									X	

		ASQ®--3 27 months	ASQ®--SE	CREDI-LF 24-36 months	CREDI-SF 24-29 months	GSED-LF 24-36 months	GSED-SF 24-36 months	PEDS:R	PEDS:DM	WIDEA-FS	WHO IYCD
	General behaviour										
91	Does your child pay attention when some- one is talking to him/her?		X								
	Does he/she look and respond to you when you are talking or asking child to do something?										
92	Does your child sit or play quietly on his/her own for at least several minutes?			X							
93	Does your child hold on to your and has a hard time leaving your side?		X								
94	Does your child get upset when the care- giver leaves the room?		X								
95	Does your child get upset when the care- giver leaves him/her with someone else?										
96	Does your child take an object and pretend it is something else?	X	X	X		X					
	For example, does your child use something to pretend it is a phone?										
97	When a known adult asks your child to do something, does he/she comply?										
98	Does your child ever try to imitate your actions around the house?			X						X	
99	Does your child play well in a group of children?										
100	Does your child kick, bite, or hit other children or adults?		X		X						

Appendix I

In-Depth Extraction tool

Domain	Extraction
Authors	
Name of study	
Year published	
Study methods - Study aim:	
Study methods - Population (n and characteristics):	
Study methods - Data collection dates:	
Notes on study methods:	
Reviewer's initials - deep extraction:	
Date reviewed - deep extraction:	
Tool design - age range of tool:	
Tool design - scoring system:	
Tool design - digital version? (Yes/No)	
Tool design - Administered by? (carer/ healthcare professional/ other)	
Tool design - training required?	
Tool design - Equipment required?	
Notes on tool design:	
Reliability - Inter-rater reliability (carers):	
Reliability - Inter-rater reliability (parents and professionals):	
Reliability - Test-retest:	
Reliability - Parallel forms:	
Reliability - Split-half:	
Notes on reliability:	
Validity - Gold standard measure:	
Validity - Threshold/ cut-off:	
Validity - Sensitivity:	
Validity - Specificity:	
Validity - PPV	
Validity - NPV:	
Validity - Percentage agreement:	
Notes on validity:	
Standardisation - Normative scores:	
Standardisation - Country of standardisation:	
Notes on standardisation:	
NOTES	

Appendix J

Risk of Bias

Domains for risk of bias and individual items assessed per domain (extracted from QUADAS-I [18]

Applicability or generalisability

A1. Was the spectrum of patients representative of the patients who will receive the test in practice?

A2. Were selection criteria clearly described?

Bias

B1. Is the reference standard likely to correctly classify the target condition?

B2. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?

B3. Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?

B4. Did patients receive the same reference standard regardless of the index test result?

B5. Was the reference standard independent of the index test (i.e., the index test did not form part of the reference standard)?

B6. Were the index test results interpreted without knowledge of the results of the reference standard?

B7. Were the reference standard results interpreted without knowledge of the results of the index test?

B8. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?

B9. Were withdrawals from the study explained?

Reporting

R1. Was the execution of the index test described in sufficient detail to permit replication of the test?

R2. Was the execution of the reference standard described in sufficient detail to permit its replication?

R3. Were uninterpretable/ intermediate test results reported?

Table 13. Risk of bias tool – QUADAS-I

		Applicability or generalisability		Bias								Reporting			
Tool	**Study**	A1	A2	B1	B2	B3	B4	B5	B6	B7	B8	B9	R1	R2	R3	**Study type**
ASQ®	[70]	Unclear	Yes	Yes	NA	No	NA	NA	NA	NA	No	No	Yes	NA	No	implementation & acceptability
ASQ®	[71]	Unclear	Yes	NA	NA	No	NA	NA	NA	NA	Yes	No	Yes	NA	No	reliability
ASQ®	[32]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	validation
ASQ®	[34]	No	Yes	Yes	Unclear	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	validation
ASQ®	[35]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	No	Yes	Yes	Yes	Unclear	validation
ASQ®	[30]	Unclear	Yes	Yes	Unclear	Yes	Yes	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	No	validation
ASQ®	[60]	No	Yes	Yes	Yes	No	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	No	validation
ASQ®	[37]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	validation
ASQ®	[72]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	No	validation
ASQ® AND PEDS	[27]	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	No	validation
ASQ®	[36]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	No	reliability/ validation
ASQ®	[73]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	No	validation
ASQ®	[38]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	Yes	validation
ASQ®	[31]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	validation

ASQ®	[28]	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	validation
ASQ®	[74	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	validation
ASQ®	[75]	No	Yes	Yes	NA	NA	NA	No	Unclear	Unclear	Yes	Yes	Yes	Unclear	No	reliability
ASQ®	[76]	No	Yes	Unclear	No	Yes	Yes	Yes	Unclear	Unclear	No	Yes	Yes	Unclear	No	validation
ASQ®	[77]	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	No	validation
ASQ®	[78]	No	Yes	NA	NA	Yes	Yes	NA	NA	NA	Yes	Yes	Yes	NA	No	reliability
ASQ®	[79]	Unclear	Unclear	Unclear	Unclear	Yes	Yes	Yes	Unclear	Unclear	Yes	No	Yes	Yes	No	validation
CREDI	[80]	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Unclear	Unclear	Yes		Yes	Yes	No	reliability/ validation
CREDI	[81]	No	Yes	Yes	Yes	No	Yes	Yes	Unclear	Unclear	Yes	No	Yes	Yes	No	validation
CREDI	[82]	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	No	validation
CREDI	[83]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	Unclear	validation
IYCD	[84]	Yes	Yes	NA	NA	No	NA	NA	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	reliability/ validation
PEDS	[85]	No	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	validation
WIDEA	[86]	No	Yes	Yes	Unclear	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	Unclear	validation
WIDEA	[87]	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Unclear	Yes	Yes	Yes	No	validation

The study number refers to the number in the references.

Appendix K

Tool Overviews

AGES AND STAGES QUESTIONNAIRE (ASQ®-3)

The ASQ®-3 is a tool designed for use with 0–66-month-olds and is designed to screen for developmental delay. There are separate questionnaires available for 21 age categories; questionnaires can be completed by parents/carers either in hard copy or online. Little training is required for paraprofessionals to score the questionnaire; training materials and a User's Guide are available for a fee on the ASQ®-3 website. The ASQ®-3 provides a score out of 60 for each developmental domain which are compared to cutoff scores that are presented in the ASQ®-3's technical appendices. Scores of ≥2SD below the mean in any domain are indicative of a need for follow-up assessment; scores of ≥1SD below the mean score indicate that the child should be monitored and given follow-up activities for practicing skills in these areas. The ASQ website provides cost information for individual purchase of ASQ®-3. A starter kit, which includes 21 paper ASQs, a user guide, and a CD with printable PDFs is costed at USD295.

PARENTS' EVALUATION OF DEVELOPMENTAL STATUS (PEDS)

Two versions of the PEDS featured in the papers in our review: the revised PEDS (PEDS-R) and the PEDS: Developmental Milestones (PEDS:DM).

The PEDS-R is a surveillance tool and screening test for use with 0–8-year-old children and is designed to elicit parent/caregivers' concerns about their child's development and health. It features twelve items which focus on whether parents/carers have concerns about various aspects of their child's development (e.g., "*do you have any concerns about how your child behaves?*"). PEDS-R can be completed either in hard copy or online and no specific equipment is required. On-demand training is available for a fee via the PEDS website. PEDS-R is scored to identify mental health, social-emotional, and behavioural (MEB) problems, and developmental delays/disorders (DD). Scores result in one of seven 'paths' depending on prevalence and intensity of MEB and/or DD (path A1= high MEBDD risk – path D/E = low MEB risk and low DD risk).

The PEDS:DM Is a brief surveillance and screening tool for developmental delay in children aged 0-8 years and intended to replace use of informal developmental milestone checklists. It comprises 6-8 items, each of which directly corresponds to a specific developmental domain. PEDS:DM can be used as a stand-alone test, or alongside PEDS-R to provide a more comprehensive picture of early child development. A scoring template is used to score parents' responses, which are then transferred to a longitudinal developmental 'growth' chart. Failure on any item suggests probable difficulties in its associated domain and performance below the 16th percentile. The PEDS website provide cost information for individual purchase. A pack of 50 PEDS-R booklets, a response form, a scoring/interpretation form, directions and action steps is costed at USD52. PEDS:DM materials start at USD69.

THE WARNER INITIAL DEVELOPMENTAL EVALUATION OF ADAPTIVE AND FUNCTIONAL SKILLS (WIDEA-FS)

The WIDEA-FS is a 50-item check list designed to describe the emerging functional skills in 0–36-month-olds. Each item consists of an explicit operationally defined task that is part of an everyday activity and is rated on a scale of one (never performs task) to four (always performs task). The WIDEA-FS can be completed directly by the parents/caregiver; there does not appear to be an online version available. The total score ranges from 50 to 200 points; once a child achieves the maximum score in each domain, one can be assured that basic skills have been achieved (Peyton et al., 201b). No information could be located about training, necessary equipment or licensing costs.

THE CAREGIVER REPORTED EARLY DEVELOPMENT INDEX (CREDI)

The CREDI tools are for use with 0-36-month-olds and comprise two forms: a long form (LF) and a short form (SF). Although both forms are reported by caregivers using a simple yes/no/don't know response scale, each serves a different purpose and produces different scores. The CREDI-LF was designed for large-scale research and evaluation projects to provide domain-specific information on children's ECD. As such, it is intended to be sensitive to the impacts of interventions, policies, and other developmental inputs and provides

both domain-specific and global development scores. The CREDI-SF was designed to provide a brief "snapshot" of children's overall ECD for national monitoring, household surveys, or other large-scale data collection efforts. CREDI is a relatively new tool with first user guidance issued in 2017[88] (last updated Oct 2023) and with the first paper on tool development published in 2018[57]. Training materials and assessors' guides are available on the CREDI website, and it is recommended that assessors receive at least one day's training on the CREDI. The CREDI-LF provides both an overall score and domain scores whereas the CREDI-SF includes 20 items in each of six 6-month age brackets and produces a single Developmental Score (*D*-score) that reflects children's overall development across domains. Two included papers [57,68] indicate the CREDI can be completed online but it is unclear exactly how the digital tool is accessed. Scores can be generated using an app or a statistical software package provided via the CREDI website. All materials are freely available via the CREDI website.

THE GLOBAL SCALES FOR EARLY DEVELOPMENT (GSED)

The GSED tools are for use with 0-36-month-olds and comprise two forms: a long form (LF) and a short for (SF). The GSED tools are designed to be used for large-scale data collection and monitoring efforts and for research and programmatic evaluation and are therefore not designed to be used for screening individual children for developmental delays or impairments. Whilst GSED-SF is caregiver- reported and typically administered via in-person interview with the child's primary carer, GSED-LF is directly administered by a trained professional. Training materials are in development, though it is unclear whether these have been released at time of writing (September 2023). Detailed item guides and user guides are available on the GSED website and a GSED Training Manual is also available. To be certified to administer the GSED-SF and GSED-LF, assessors must complete a GSED training and pass required knowledge assessments. Both GSED-SF and GSED-LF produce a single, holistic score of child development (*D*-score*)*. The *D*-score is calculated from the yes/no responses to an age-appropriate set of items. The *D*-score can also be transformed into a Development-for-Age z-score (*DAZ* score), an age-independent score that allows for easier comparison between samples from different ages or countries. A comprehensive scoring guide is provided and includes links to either an online calculator or a downloadable statistical package. All materials are freely available via the GSED website.

THE WHO INDICATORS OF INFANT AND YOUNG CHILD DEVELOPMENT

The IYCD is a tool designed to measure development for children ages 0-3 years across cultures. The IYCD is parent-carer reported that is administered by a trained, non-specialist individual. IYCD was created for use on a tablet-based system using open-source software (Open Data Kit). The tool consists of 100 items which produce age-standardised development scores (*DAZ* score). Training is indicated to take 2-3 day, conducted in-person; training materials are available through the IYCD website, although the research team have been unable to successfully access the website.

Appendix L

Developmental Domains Covered by Tool

Table 14. Developmental domains covered by each tool

	Domains						
Individual-level measures							
ASQ®-3	Communication	Motor	Problem solving	Personal/ social			
PEDS-R	Language	Motor	Behaviour	Self-help	School and social skills	Global/ cognitive	Health
PEDS:DM	Language	Motor	Socio-emotional	Self-help			
WIDEA-FS	Communication	Mobility	Social cognition	Self-care			
CREDI	Language	Motor	Cognition	Social-Emotional	Mental health		
GSED	Language	Motor	Cognition	Social-Emotional	Adaptive		
IYCD	Language	Motor	Socio-emotional and behaviour				

Appendix M

Tool Characteristics

Table 15. Characteristics of tools: individual-level measures

Tool items	Parent reported?	Training for professional	Administration time (minutes)	Scoring and cut-offs	Culturally adaptable?	Age range (months)	Digital? y/n	Licensing costs?
ASQ®-3 43 items	Y, with professional input	Minimal training. Training DVDs and seminars available through ASQ®-3 website. Indicative costs $49.95. E-learning modules available for practitioners via NHS England.	10-15	Domain-specific and global scores out of 60 ≥2SD below the mean → follow-up assessment of ≥1SD below mean → monitor + practice activities	N*	0-66	Y but possible GDPR issue- Brookes servers located in USA	Y
PEDS-R 12 items	Y, with professional input	Minimal training. 1.5-hour training with test and certificate stating completion of PEDS training, indicative cost $125.	5 5	Scores indicate presence of mental health, social-emotional, and behavioral (MEB) problems, and developmental delays (DD). Scores = one of seven 'paths' depending on prevalence and intensity of MEB and/or DD (path A1= high MEBDD risk to paths D/E = low MEB & low DD risk). Each pathway has a specific referral recommendation.	N**	0-96	Y	Y
PEDS:DM 6-8 items	Y	NR		Scoring template is placed over parent's responses to score answers. Failure on any item suggests probable difficulties in that domain and performance below the 16th percentile.	N**	0-96	Y	Y
WIDEA-FS 50 items	Y	No training is reported.	10	No scoring protocol is provided.	NR	0-36	N	NR

Table 16. Characteristics of tools: population-level measures

Tool items	Parent reported?	Training for professional	Administration time (minutes)	Scoring and cut-offs	Culturally adaptable?	Age range (months)	Digital? y/n	Licensing costs?
CREDI LF 69 items	Y, with professional input	It is recommended that assessors receive at least one day of training on the CREDI.	15	Scoring time NR. Scoring through custom app. LF: domain-specific and a global scores.	Y	0-36	Unclear	N- freely available
SF 20 items	Y, with professional input	Training material freely available online.	<5	SF: global score only.	Y	0-36		N- freely available
GSED LF 70 items	N- direct observation	5-7 days 2-3 days	30-75	LF and SF: global scores only. Can be transformed into age-adjusted Development-for-Age scores (DAZ). Scoring through custom app.	Y	0-36	Y	N- freely available
SF 45 items	Y, with professional input	Online training courses in development though not released yet. To be certified to administer the SF and LF, assessors must complete a GSED training and pass required knowledge tests	15-25		Y	0-36	Y	N- freely available
WHO IYCD 100 items	Y, with professional input	Package of training materials available through IYCD website. Training takes 2-3 days, in-person.	NR	NR	Y	0-36	Y	N- freely available

* ASQ®-3 Arabic, Chinese, French, Spanish, and Vietnamese language translations available.

**PEDS Spanish language translation available.

Appendix N

Reliability and Validity Data

Table 17. Reliability and validity data by measure

	Reliability			Convergent validity		Discriminant validity	Known groups validity
Measure	**Internal consistency**	**Test-retest**	**Inter-rater**				
ASQ®-3	.55 - .88	.53-.94	.80	**Reference standard**	**General population, non-English-speaking, *n* studies= 2** Agarwal et al., 2023 Schonhaut et al., 2013		
				BSID-III	**By domain:** Problem solving: .22 Communication: .44-.59 Gross motor: .36 Fine motor: .13 Personal social: .16-.37 **Total score:** .75	Low correlations between conceptually divergent domains apart from ASQ®-3 fine motor which had a higher correlation with BSID-III receptive, expressive language and cognitive (.26, .22, .24) than with BSID-III gross motor domains (.17)	NR
				BSID-III	**At-risk sub-populations (e.g., preterm birth, children from deprived areas), non-English speaking, *n* studies= 4** Rubio-Codina et al., 2016 Simpson et al., 2016 Yue et al., 2019 Yue et al., 2021		
					By domain: Problem solving: .12-.60 Communication: .23-.67 Gross motor: .18-.49 Fine motor: .13-.46 Personal social: NR	NR	NR NR

PEDS-R	NR	NR	NR		**At-risk sub-populations (e.g., preterm birth, children from deprived areas), English speaking, *n* studies= 1** duToit et al., 2021		
				VABS-3	**By domain:** Receptive language: .34 Expressive language: .40 Fine motor:.26 Gross motor: .36 Adaptive behaviour: .59 Social-emotional: .37 Self-help: .27 Literacy: .27	NR	NR
PEDS:DM	NR	NR	NR	VABS-3	**By domain:** Receptive language: .29 Expressive language: .22 Fine motor: .33 Gross motor: .35 Adaptive behaviour: .48 Social-emotional: .29 Self-help: NR Literacy: .44	NR	NR
WIDEA-FS	NR	NR	NR		**At-risk sub-populations (e.g., preterm birth, children from deprived areas), English speaking, *n* studies= 2** Peyton et al., 2021a Peyton et al, 2021b		
				BSID-III	Social cognition .26 Communication .65 Mobility .41	NR	Children with special health needs demonstrated significantly lower WIDEA-FS scores than typically developing children.

CREDI	.66-.91	.70 -.81	NR	BSID-III INTER-NDA ASQ®-SE CDI BOI PRIDI	**General population, mixed English and non-English speaking, *n* studies = 6** Alderman et al., 2021 Altafim et al., 2020 Li et al., 2020 McCoy et al., 2017 McCoy et al., 2018 Waldman et al., 2021		
					By domain: Language: .26-.90 Motor: .18-.90 Cognition: .12-.93 Socio-emotional: .15-.76	Convergent and discriminant relationships found for both CREDI language and motor subscales, which correlated most strongly with their respective counterparts and from alternative ECD measures. CREDI cognitive subscale scores correlated most strongly with BSID and INTER-NDA expressive language subscales CREDI socio-emotional scores correlated most strongly with BSID language subscales. ASQ®:SE correlated most strongly with CREDI cognitive subscale score, rather than CREDI socio-emotional scores as predicted.	CREDI scores correlate with: Height-for-age (HAZ): .16-.22. 1 *SD* increment HAZ = .12 *SD* increase in CREDI scores Home stimulation: .16-.25 Highest home learning environment scores = CREDI scores .95 *SD* higher Household wealth: CREDI scores .20 *SD* higher in top than bottom quintile Caregiver education: CREDI scores .20 *SD* higher when caregiver education is >12 years compared to no formal education. Disability Higher CREDI scores amongst children who are non-disabled (> .50 *SD*).

GSED	≥ .98	≥ .98	≥ .98	BSID-III	**General population, non-English speaking, *n* studies = 1** Gladstone et al., 2023		
					By domain >0.88 with all domains of the BSID-III (LF and SF)	NR	Age-adjusted GSEDs correlate in the hypothesized directions with all posited convergent measures (e.g., height-for-age *r*= 0.19.)
					Total score >0.97		
WHO IYCD	NR	0.84-0.96	0.78-0.95	-	**General population, non-English speaking, *n* studies = 1** Gladstone et al., 2021		
					NR	NR	Age-standardized IYCD scores correlates with: height for age (HAZ): .25 Weight for age (WAZ): .25 Maternal education: .37 SES: .36 Home family environment: .22

Appendix O

Sensitivity and Specificity Data

Table 18. Sensitivity and specificity by tool

	Population	Gold standard test (for a positive result in the gold standard test, we assume true presence of delay)	**Mild delay (cut off < -1SD on both measures)** Tables show number of children who pass (+) and fail (-) on the tool of interest in comparison to the reference measure. A fail is indicative of possible developmental delay.	**Moderate-severe delay (cut off < -2SD on both measures)**
ASQ®				
General population, English-speaking, *n*= 2				
Sheldrick et al. 2020	USA, convenience sample, general population, 9–42-month-olds	Bayley Scales of Infant and Toddler Development (BSID-III)	Interim data not presented Prevalence based on BSID-III: NR Sensitivity: 23.1% Specificity: 89.4%	Interim data not presented Prevalence based on BSID-III: NR Sensitivity: 60% Specificity: 89.4%
Veldhuizen et al., 2015	Canada, general population, 1–36-month-olds	Bayley Scales of Infant and Toddler Development	NR	<table><tr><td></td><td>BSID-III +</td><td>BSID-III -</td><td>Total</td></tr><tr><td>ASQ®-3 +</td><td>470</td><td>10</td><td>480</td></tr><tr><td>ASQ®-3 -</td><td>100</td><td>7</td><td>107</td></tr><tr><td>Total</td><td>570</td><td>17</td><td>587</td></tr></table>Prevalence based on BSID-III: 2.9% (17/587) Sensitivity: 41.2% (7/17) Specificity: 82.5% (470/570)

<table>
<tr><td colspan="5">General population, non-English speaking, n= 3</td></tr>
<tr><td>Agarwal et al., 2023</td><td>Singapore, population cohort, 23–25-month-olds, ASQ®-3 in Chinese, Malay, Tamil and English languages</td><td>Bayley Scales of Infant and Toddler Development</td><td>NR</td><td>Interim data not presented

Prevalence based on BSID-III: NR
Sensitivity: 21.1%
Specificity: 97.4%</td></tr>
<tr><td>Charkaluk et al., 2017</td><td>France, population cohort, 36-month-olds, ASQ®-3 in French language</td><td>Wechsler Preschool and Primary Scale of Intelligence (WPPSI), at age 5-6</td><td>Interim data not presented

Prevalence based on WPPSI: 7.3% (69/939)
Predictive value[a]:
Sensitivity: 77%
Specificity: 68%</td><td>NR</td></tr>
<tr><td>Steenis et al., 2015</td><td>Netherlands, general population, 18–24-month-olds, ASQ®-3in Dutch language</td><td>Bayley Scales of Infant and Toddler Development</td><td><table><tr><td></td><td>BSID-III +</td><td>BSID-III -</td><td>Total</td></tr><tr><td></td><td>292</td><td>57</td><td>349</td></tr><tr><td></td><td>123</td><td>93</td><td>216</td></tr><tr><td>Total</td><td>415</td><td>150</td><td>565</td></tr></table>Prevalence based on BSID-III: 26.5% (150/565)
Sensitivity: 62%
Specificity: 70.4%</td><td><table><tr><td></td><td>BSID-III +</td><td>BSID-III -</td><td>Total</td></tr><tr><td></td><td>477</td><td>5</td><td>482</td></tr><tr><td></td><td>75</td><td>8</td><td>83</td></tr><tr><td>Total</td><td>552</td><td>13</td><td>565</td></tr></table>Prevalence based on BSID-III: 2.3% (13/565)
Sensitivity: 61.5%
Specificity: 84.4%</td></tr>
</table>

<table>
<tr><td colspan="5">At-risk sub-population, English speaking, n= 2</td></tr>
<tr>
<td>Duggan et al., 2023</td>
<td>Ireland, children with low birthweight, 24-month-olds</td>
<td>Bayley Scales of Infant and Toddler Development

Kaufmann Brief Intelligence Test (age 5-6)</td>
<td>
<table>
<tr><td></td><td>BSID-III +</td><td>BSID-III -</td><td>Total</td></tr>
<tr><td>ASQ®-3 +</td><td>166</td><td>30</td><td>196</td></tr>
<tr><td>ASQ®-3 -</td><td>57</td><td>25</td><td>82</td></tr>
<tr><td>Total</td><td>223</td><td>55</td><td>278</td></tr>
</table>
Prevalence based on BSID-III: 19.8% (55/278)
Sensitivity: 40% (25/55)
Specificity: 74.4% (166/223)

By domain:
Problem-solving: sensitivity: 20.7% (6/29) specificity: 98.8% (241/249)
Communication: sensitivity: 25% (11/44) specificity: 99.5% (233/234)
Motor: sensitivity: 50% (20/40) specificity: 76.1% (181/238)
</td>
<td>
<table>
<tr><td></td><td>BSID-III +</td><td>BSID-III -</td><td>Total</td></tr>
<tr><td>ASQ®-3 +</td><td>194</td><td>2</td><td>196</td></tr>
<tr><td>ASQ®-3 -</td><td>71</td><td>11</td><td>82</td></tr>
<tr><td>Total</td><td>265</td><td>13</td><td>278</td></tr>
</table>
Prevalence based on BSID-III: 4.7% (13/278)
Sensitivity: 84.6% (11/13)
Specificity: 73.2% (194/265)
</td>
</tr>
</table>

<table>
<tr><td>Noeder et al., 2017</td><td>USA, children with coronary heart disease, 6–36-month-olds</td><td>Bayley Scales of Infant and Toddler Development</td><td>Interim data not presented

Prevalence based on BSID-III: NR
By domain:
Problem-solving: sensitivity: 67%
specificity: 87%
Communication: sensitivity: 90%
specificity: 84%
Fine motor: sensitivity: 65%
specificity: 84%
Gross motor: sensitivity: 77%
specificity: 92%</td><td>NR</td></tr>
<tr><td colspan="5">At-risk sub-population, non-English speaking, n= 6</td></tr>
<tr><td>Agarwal et al., 2016</td><td>Singapore, children born prematurely with very low birthweight, 24-month-olds, ASQ®-3-3 language not specified other than “parents had mixed language and literacy backgrounds”.</td><td>Bayley Scales of Infant and Toddler Development</td><td>NR</td><td><table><tr><td></td><td>BSID-III +</td><td>BSID-III -</td><td>Total</td></tr><tr><td>ASQ®-3 +</td><td>66</td><td>4</td><td>70</td></tr><tr><td>ASQ®-3 -</td><td>34</td><td>21</td><td>55</td></tr><tr><td>Total</td><td>100</td><td>25</td><td>125</td></tr></table>Prevalence based on BSID-III: 20% (25/125)
Sensitivity: 84% (21/25)
Specificity: 66% (66/100)</td></tr>
</table>

<table>
<tr>
<td>Kerstjens et al., 2015</td>
<td>Netherlands, children born prematurely, 22–26-month-olds, ASQ®-3 in Dutch language</td>
<td>Bayley Scales of Infant and Toddler Development (without language subscale)</td>
<td>
<table>
<tr><th></th><th>BSID-III +</th><th>BSID-III -</th><th>Total</th></tr>
<tr><td>ASQ®-3 +</td><td>161</td><td>2</td><td>163</td></tr>
<tr><td>ASQ®-3 -</td><td>48</td><td>13</td><td>61</td></tr>
<tr><td>Total</td><td>209</td><td>15</td><td>224</td></tr>
</table>
Prevalence based on BSID-III: 5.8% (13/224)
Sensitivity: 87% (13/15)
Specificity: 99% (161/163)
</td>
<td>
<table>
<tr><th></th><th>BSID-III +</th><th>BSID-III -</th><th>Total</th></tr>
<tr><td>ASQ®-3 +</td><td>163</td><td>0</td><td>163</td></tr>
<tr><td>ASQ®-3 -</td><td>51</td><td>10</td><td>61</td></tr>
<tr><td>Total</td><td>214</td><td>10</td><td>224</td></tr>
</table>
Prevalence based on BSID-III: 4.5% (10/224)
Sensitivity: 100% (10/10)
Specificity: 76% (163/214)
</td>
</tr>
</table>

<table>
<tr>
<td>Schonhaut et al., 2013</td>
<td>Chile, mixed population (term, preterm, and very preterm), 8-30-month-olds, ASQ®-3 in Spanish language</td>
<td>Bayley Scales of Infant and Toddler Development</td>
<td>Interim data not presented

<u>30 months old, whole sample:</u>
Prevalence based on BSID-III: 18.3% (56/306)
Sensitivity: 82%
Specificity: 84%

<u>Term 8–30-month-olds:</u>
Sensitivity: 59%
Specificity: 87%

<u>Late preterm 8–30-month-olds:</u>
Sensitivity: 80%
Specificity: 73%

<u>Extreme preterm 8–30-month-olds:</u>
Sensitivity: 86%
Specificity: 86%</td>
<td>NR</td>
</tr>
<tr>
<td>Simpson et al., 2016</td>
<td>Australia, indigenous children, 2–36-month-olds, ASQ®-3</td>
<td>Bayley Scales of Infant and Toddler Development</td>
<td>NR</td>
<td>
<table>
<tr><th></th><th>BSID-III +</th><th>BSID-III -</th><th>Total</th></tr>
<tr><td>ASQ®-3 +</td><td>55</td><td>2</td><td>57</td></tr>
<tr><td>ASQ®-3 -</td><td>5</td><td>5</td><td>10</td></tr>
<tr><td>Total</td><td>60</td><td>7</td><td>67</td></tr>
</table>
Prevalence based on BSID-III: 10.5% (7/67)
Sensitivity: 71.4% (5/7)
Specificity: 91.7% (55/60)
</td>
</tr>
</table>

Yue et al., 2019	Rural China, at-risk, poverty, 17–24-month-olds, ASQ®-3 in Chinese language	Bayley Scales of Infant and Toddler Development	Interim data not presented Prevalence based on BSID-III: NR By domain: Problem-solving: sensitivity: 18.7% specificity: 84.4% Communication: sensitivity: 28% specificity: 90% Motor: sensitivity: 50% specificity: 67.9%	Interim data not presented Prevalence based on BSID-III: NR By domain: Problem-solving: sensitivity: 18.6% specificity: 93.8% Communication: sensitivity: 28.6% specificity: 83.1% Motor: sensitivity: 41.7% specificity: 94.3%
Yue et al., 2021	Rural China, at-risk, poverty, 17–24-month-olds, ASQ®-3 in Chinese language	Bayley Scales of Infant and Toddler Development (Communication subscale only)	Prevalence based on BSID-III: 44.9% Communication domain sensitivity: 14.7% Specificity: 49%	Prevalence based on BSID-III: 21.1% Sensitivity: 2.2% Specificity: 87.3%
PEDS				
General population, English-speaking, *n*= 1				
Sheldrick et al., 2020	USA, convenience sample, general population, 9–42-month-olds	Bayley Scales of Infant and Toddler Development	Interim data not presented <u>PEDS-R</u> Prevalence based on BSID-III: NR Sensitivity: 28% Specificity: 79.6% <u>PEDS:DM</u> Prevalence based on BSID-III: NR Sensitivity: 67.2% Specificity: 42.7% <u>PEDS-R and PEDS:DM together</u> Prevalence based on BSID-III: NR Sensitivity: 22.7% Specificity: 83.9%	Interim data not presented <u>PEDS-R</u> Prevalence based on BSID-III: NR Sensitivity: 78.9% Specificity: 79.6% <u>PEDS:DM</u> Prevalence based on BSID-III: NR Sensitivity: 60.8% Specificity: 42.7% <u>PEDS-R and PEDS:DM together</u> Prevalence based on BSID-III: NR Sensitivity: 78.9% Specificity: 83.9%

At-risk sub-population, non-English context				
DuToit et al., 2021	South Africa, at-risk, poverty, 36–83-month-olds, PEDS-R & PEDS:DM, language unclear	Vineland Adaptive Behaviour Scales (VABS-3)	Interim data not presented <u>PEDS-R and PEDS:DM together</u> Prevalence based on VABS-3: 80.1% (221/276) Sensitivity: 92.6% Specificity: 22.5%	Interim data not presented
WIDEA-FS				
At-risk sub-population, English-speaking, *n*= 1				
Peyton et al., 2021a	USA, children at risk of neurodevelopmental delay, 10-30-month-olds	Bayley Scales of Infant and Toddler Development	Prevalence based on BSID-III: NR By domain[b]: Mobility: Cut-off: 36 Sensitivity: 88% Specificity: 87% Communication: Cut-off: 52 Sensitivity: 100% Specificity: 51% Social cognition: Cut-off: 38 Sensitivity: 60% Specificity: 77% Self-care: Cut-off: 42 Sensitivity: 60% Specificity: 87%	NR

[a]WPPSI scores of <85 (-1SD below mean) indicated low IQ at age 5-6.
[b]Peyton et al. use ROC analysis and Youden's Index to establish the optimal WIDEA-FS cut-offs for detecting developmental delay as measured by BSID-III.

Appendix P

Focus Group Participant Numbers by Data Collection Sites

Table 19. Recruitment of focus group participants by data collection site

Group		Site						
		1	2	3	4	Virtual	Total groups	Total participants
Parents of children aged 2-3 years	N groups	2	3	-	1	1	7	
	N participants	8	10	-	10	1		29
Health visiting teams (full skill mix)	N groups	-	2	1	-	2	5	
	N participants	-	9	8	-	7		24
Local authorities with digital ASQ®-3	N groups	-	-	-	-	2	2	
	N participants	-	-	-	-	5		5
DHSC policy colleagues including public health intelligence specialists	N groups					1	1	
	N participants					5		5
Total groups		2	5	1	1	6	**15**	
Total participants per site		8	19	8	10	18		**63**

Appendix Q

Focus Group Topic Guides

Parents

- Could you tell us about what happened or the process during your child's 2-2½ year review? This includes if you've just booked it but haven't had it yet, or even if you were never offered it. Tell us a bit about your experience with this so far.
  - Prompts: Where it happened?
  - Who was there (practitioners)?
  - How long it was?
  - Was it hard/ easy to get to?
  - 1:1 or in a group?
  - Who completed the form, and did you see it?
  - If you completed the form yourself, did you feel confident answering the questions?

- What was your child/ the child in your care doing whilst the form was filled in?
  - Prompt: Was the child there?
  - Prompt: Were you involved at all?

- Did you find out what the review said about your child?
  - Prompt: If yes, did you understand what the review said about your child?
  - Prompt: Did you have questions about what the review found? Did someone answer your questions?

- If you've had your 2-2½ year review, is there anything that you would have liked to be different/ that you particularly liked/ that worked well for you?
  - Prompt: what were the most important considerations when planning to attend your child's/ the child in your care's 2-2½ year review?
  - Prompt: Did anything make it easy/ difficult to attend?

- Tell me a little bit about filling out the form. Would you have liked more support with filling out the form?

- Do you think the form was useful and appropriate for your child? If yes why/ if no, why not?

- There has been some suggestion of making the form digital instead- what do you think of this? Would you prefer it/ not prefer it to stay in paper form? Why?

- What did you think about the feedback you received after filling the form, if any?
  - What kind of feedback would you have liked?
  - Prompt: For example, specific feedback about your child's behaviour, or anything that the results showed, and you wanted to talk/ask about.

- Is there anything specific to you, your child/ or your family that made your experience more complex or difficult?

- What are the three main priorities that could make the 2-2½ year review even better for your child/family?

Parents of children with disability

- Could you tell me how was the review offered?
  - (Prompts: was a letter sent to your house/ phone call from the health visiting service/ other? Any explanation of what the review was for?)

- As part of your invitation to the 2-2½ year review, were you also sent a questionnaire to fill out about your child's development?

- If yes, how did you feel about the questionnaire?

- Do you think the form was useful and appropriate for your child? If yes why/ if no, why not?

- Were you able to fill it out? Did you?

- If no, what did you receive when you received your invitation letter through the post? Just the letter, or was there anything else?

- Could you tell me a bit more about what the review itself was like? Prompts: Things like where it happened and who conducted it. What sort of thing the practitioner asked etc.

- How did you feel about the review?
  - Prompt: Did you find it useful?
  - Prompt: Did you have questions about what the review found? Did someone answer your questions?

- Was there anything that you would have liked to be different/ that you particularly liked/ that worked well for you?
  - Prompt: what were the most important considerations when planning to attend your child's/ the child in your care's 2-2½ year review?

If you haven't had time to review:

- Were you aware of/ expecting an invitation to the 2-2½ year review?

- Did you have a 9-12 month review?
  - If yes: can you remember any details of what it was like? Were you given a questionnaire to fill out at this review?

- Is it important to you to be offered a review with a member of the health visiting team (or

- If yes, why? What about the 2-2½ year review would be valuable to you? If no, why not?

For all:

- Is your child under the care/ supervision of any other specialist teams? If yes, could you tell me a bit about the kind of care they receive?
  - Prompt: Do you feel like your child has enough/ the right contact time with the other specialist teams they see?
  - Prompt: How do the developmental reviews fit around this additional care?

- Do you think data should be collected on children with disabilities so that they're a part of this national picture of child development, or do you think it's inappropriate to measure development in this way for these children?

- What are the three main priorities that could make the 2-2½ year review better for your child/family?

Health visiting professionals

- What is the protocol for the 2-2½ year review in your locality as you understand it?
  - Who makes initial contact with parents when the review is due? How is contact made? What If you don't hear back from a family?

- In your experience, what typically happens at the 2-2½ year review?
  - Prompt: Where does it typically take place? (At home/ in a clinic/on-line, varies according to families' needs?)
    - Prompt: Who is usually there?
  - Prompt: who usually administers the measure of child development? Parent/ you/ nursery nurses/ other?
  - Prompt: what is the method of delivery for review where English is not the first language?

- Do you ever use the ASQ®-3-SE in addition to ASQ®-3 to review social and emotional development?

- Do you ever adapt the ASQ®-3 ?
  - Prompt: EG skipping questions, doing them out of order, using own judgements…

- Do parents understand the results? Do you ever talk them through the results? If yes- when, in what way?

- What do you do next? What do you do with the data?

- (For nursery nurses) I'd be really interested to know how you feel about your role within the health visiting skill mix team or system. So how you interface with other members of the team, so health visitors, staff nurses – how do you feel it works?

- Do you ever have parent saying they've had this review at nursery? Are nurseries every using ASQ®-3 as far as you're aware?

- How confident are you that the ASQ®-3 gives an accurate (full) picture of the 2-year-old?
  - Prompt: Does it pick up on things that need attention?
  - Prompt: If it doesn't meet expectations/not confident, how do you think it could be improved?

- What do you use the ASQ®-3 for within [local area] or within your own practice?
  - Prompt: what's its purpose in your practice?
  - Prompt: how does the ASQ®-3 work within the wider 2-2½ year review in your experience?

- Do you think it's appropriate/ feasible for the government to use the data gathered using the ASQ®-3 to monitor disparities and trends in child development across the country?

- How would you feel if the ASQ®-3 were to be changed for another tool designed to measure child development at age 2-2½?

- What kind of training did you have specifically for the ASQ®-3?
  - Have you ever completed the on-line training module?

- Could you tell me about how you record children's ASQ®-3 results in your practice?
  - Prompt: Who inputs the ASQ®-3 data into the electronic records for the child?
  - Prompt: Is it easy to find afterwards in the electronic record?
  - Prompt: Is it shared with or available to any other professionals?

- Is there every any information sharing of these results? Would you ever share children's ASQ®-3 data with any other professionals or teams and if so, why?
  - Prompt: Would you ever communicate this data to other teams?

- In your experience, do you know of any professionals/ local authorities using children's ASQ®-3 results to get a broad picture of what is happening with children's development in their local authority?

Local authority colleagues

- Could you each introduce yourselves and briefly summarise where you're at in terms of using a digital tool at the 2-2½ year review? i.e., are some of your localities already using it or is it something you're looking into but haven't adopted yet?
  - If a digital tool, including the digital ASQ, is in use in your region, how long has it been in use?

- Now could you talk a bit to why you became interested in a digital tool?
  - How did a digital tool get onto your local authority's agenda?
  - Why did you want a digital tool in your area?
  - What do you hope to achieve with a digital tool?

- Was there a business case to make for moving to a digital tool and if so, how was it made?
  - How was the justification for resource made?
  - Ballpark costs?

Adapt the following depending on whether digital tool is in use, or is intended to be implemented soon:

- In terms of practical implementation – what's it been like using the digital tool? How's it going?
  - Who is the digital questionnaire available to (/who would it be available to…)– when and how is it distributed to respondents?
  - What format does the tool take? On a bespoke app/ web link/ other?
  - Who can access the digitised responses?
  - At what point do the responses become available?
  - How are the digitised responses used by healthcare practitioners?
  - What benefits/ disadvantages have you seen?
  - Are these pros/ cons due to the tool itself, or the way it's been digitised?
  - Any feedback from either service users or service providers on implementation of the digital tool?
  - Have you implemented any kind of evaluation yet?
  - Are there any characteristics of your local system that have facilitated/ made digital implementation difficult?
  - In terms of technical implementation – one of the issues that the DHSC has flagged with us in the past is that some or most of Brookes' servers are based in America. I'm wondering what the case is with your digital provision?
  - Do you know where your data are stored?
  - How does/ do you think it could be made to work within the existing systems you work with?
  - If you were to give advice to other local areas considering moving to a digital tool, what would you say?

- What's the most important thing you think we should know about implementing a digital tool?
  - Any key take-home messages?

Policy colleagues

- Starting broad, why have a universal measure of child development age 2?
  - Prompt: What about uses beyond your own team? Are there any? If so, are they equally important?

- How do you think LAs should be using the results of the measure, both for each child and for their whole population?

- In an ideal world, what would the universal measure allow you and your team to do or achieve?

- Are there any problems or major challenges with achieving these purposes through a universal measure of child development age 2?

- Thinking broadly, what are the important aspects of any measure of child development for you and your team?
  - Prompt: what about the scope of a measure in terms of the domains it assesses?
  - Prompt: what about how it can be used in practice?
  - Further prompt: Who, when, where and format e.g., on paper or other
  - Prompt: What about from the user perspective? Any important considerations about its use by parents/carers?
  - Prompt: what about the data / data flow?
  - Prompt: what about resource and costs?
  - Prompt, do any of these aspects stand out as more important to policy teams?

- Now thinking more narrowly, there are currently multiple tools in use to measure development age 2: ASQ®-3, ELiM and WellComm. Could you tell me about this?

- Do you see any of these tools or approaches having advantages over others?
  - Prompt: ELiM– can you tell me more about why & how this was developed?
  - Prompt: What about other established measures of child development e.g. the Early Years Foundation Stage profile?

- Is there anything else you'd like to tell us – either about what you think works best or what needs to be avoided?

Appendix R

Parent's Sociodemographic Data – Focus Group

Table 20. Sociodemographic data for parent focus group participants (total n= 28)

	Mean (range)
Parent age (years)	34.4 (24-47)
No. children	*N* (%)
1	11 (39.3%)
2	13 (46.4%)
3	4 (14.3%)
Ethnicity (self-identified)	*N* (%)
British	7 (25%)
White	5 (17.9%)
Bangladeshi	3 (10.7%)
Black	2 (7.1%)
White British	2 (7.1%)
Afro-Caribbean	1(3.6%)
Albanian	1 (3.6%)
Arab	1 (3.6%)
Asian	1 (3.6%)
Black African	1 (3.6%)
Bulgarian	1 (3.6%)
Indonesian Japanese	1 (3.6%)
Latvian	1 (3.6%)
Not reported	1 (3.6%)
Pre-tax household income	*N (%)*
<£10,000	1 (3.6%)
£10,001-£20,000	3 (10.7%)
£20,001-£30,000	5 (17.9%)
£30,001-£40,000	1 (3.6%)
£40,001-£50,000	4 (14.3%)
≥£50,001	13 (46.4%)
Not reported	1 (3.6%)

Appendix S

Digital tool – full analysis

Perceived benefits

IMPROVING THE USER EXPERIENCE

A digital offering was perceived by all groups as having the potential to improve the user experience by constituting part of a broader, centralised online hub wherein all information about children's health and development, including access to developmental review materials, would be kept. LA colleagues emphasised the importance of digitisation efforts to "*really be a system-wide piece of work rather than, you know, a standalone digital ASQ®-3".* Stakeholders suggested that a digital offering may make the tool more accessible, reflecting on a digital offering's potential to provide text, video, and aural examples of the questions to aid parents' understanding. They also highlighted the importance of being able to administer the tool in the parent/carer's primary languages, and hoped that a digital offering might make this a possibility. Parents and health visiting professionals also noted that a digital hub could provide instant messaging channels that would improve parent/carer engagement with the health visiting service, facilitating easier, more personalised communication between parent/carers and the health visiting service.

MODERNISATION: A DIGITAL TOOL FOR A DIGITAL WORLD

All stakeholder groups highlighted the need for modernisation of service offerings. From policy and local authority colleagues' perspectives, a key consideration was the reputational impact of retaining analogue tools/ services; participants emphasised that providing an online service would fit with government objectives to digitise across all national services. Participants across all groups noted that current reliance on the postal service to distribute key information and materials for the 2-2½ year health and development review is a risky and costly strategy, both in terms of monetary and environmental cost. From a commissioning perspective, local authority, and policy colleagues at DHSC emphasised that it would be prudent to ensure that any services commissioned in the future are as up-to-date and 'future-proof' as possible: *"I think it's unrealistic that we can have no digital system for another 10-15 years. It already feels outdated."* [Policy colleague]

AUTOMATION AND EFFICIENCY

A digital tool was perceived as having potential to automate processes, ultimately making the service more efficient. Policy and local authority colleagues emphasised a digital offering has potential to save time and money on administration costs and would represent *"an alternative to stuffing envelopes with ASQ®-3 questionnaires and posting them out… we're looking at the capacity of staff it takes to actually send them out."* [Local authority colleague]. We understood from professionals in our focus groups and our wider work outside this study with stakeholders that it is common practice to send out *ASQ*®-3 in the post.

Parents felt that a digital offering complete with video examples of each ASQ®-3 item may help them provide more objective responses, subsequently leading to better quality data. Policy colleagues at DHSC emphasised the urgent need for automated data management such that parents' inputted ASQ®-3 responses should populate automatically into local authorities' data management systems so that they can then be "*automatically flowing to the back end for any statistical purposes."* One local authority currently trialling a digital ASQ®-3 confirmed this as a priority as, due to system incompatibility and cost of licencing, their staff must manually input ASQ®-3data from the location to which the web-based ASQ®-3 exports data, to their local data management system, thus effectively doubling the workload associated with this task.

Concerns

DIGITAL MUST NOT REPLACE IN-PERSON CONTACTS

A final crucial consideration is that parents, DHSC policy colleagues and local authority professionals took care, when discussing ways in which a digital offering may help make the service more efficient, to emphasise that digital offering "*should not replace the conversation with the health visitor"* [parent group 1]. A digital option was explicitly framed as something that could enhance, not replace, the 2-2½ year health and development review as detailed above, and that time and energy saved on sending and scoring hard copies of the ASQ®-3 could be spent on other crucial aspects of the review, such as health promotion:

"This isn't a plan to just say 'just do it online'… even for those families where everything is tickety-boo, there will still be a contact because there are still messages and information to be given at that contact." [LA colleague]

However, we also heard messages about how a digital tool might be used to prioritise families for the universal review (parents and health professionals) or to *"see how far [we] can stretch the mandation […], not only with skill mix, but actually with digital options, face to options, questionnaire options",* implying that a digital option would replace the in-person health and development review for some people.

UNIVERSAL REACH

Participants from all groups raised concerns about the extent to which a digital tool may have universal reach. Most commonly, concerns were raised that digital poverty and national inequalities in internet connectivity may exclude certain groups of families. Local authority and policy colleagues at DHSC acknowledged this and emphasised that *"we wouldn't ever want a digital-only system so as to be able to include families who are not able to get online."* [DHSC Policy colleagues group]

PRACTICALITIES OF USING A DIGITAL TOOL

Parents, health visiting professionals observed that the ASQ®-3, particularly in combination with the ASQ®-SE, is very long and so may prove difficult to complete on a laptop/phone in one sitting. A priority for policy colleagues at DHSC was that any digital tool be designed carefully to ensure its fitness for purpose as *"if you're having to print it out and go in, do it and upload it… that removes the whole purpose of the digital side of it."* [DHSC Policy colleagues group]

DATA MANAGEMENT AND SYSTEMS

Policy and local authority colleagues raised concerns about the need for a digital tool to integrate effectively with local authorities' current data management systems. Both acknowledged the potential difficulty of devising a digital format compatible with the various data management systems used across the country (e.g., EMIS, Rio, SystmOne). It was important to health visiting professionals that automatically imported parent-reported data be made available in a way *"that we can edit it, rather than it just being what the parent said… that goes on their notes"* [nursery nurses group 1] so that health visiting professionals can provide their professional assessment. Health visiting professionals also raised pragmatic concerns that automatically imported ASQ®-3 scores may become out of date by the time of their health and development review appointment, as *"quite often parents will rebook their appointments"* [Nursery nurses group 1].

COST AND LICENCING

Policy colleagues at DHSC noted that a lack of unified data management system could make offering a digital option complicated and costly for local authorities. One local authority team that had begun the process of negotiating a licence with Brookes Publishing for a digital ASQ®-3 found the process complicated and associated with considerable financial risk and ultimately concluded that the risks outweighed the benefits of having a digitally available tool, suggesting that negotiating a digital licence "*should be offered nationally"* [LA colleague group].

STAFF TRAINING AND MOTIVATING PRACTITIONERS

Policy colleagues at DHSC recognised that "*a lot of training, a lot of guidance, a lot of encouragement",* and reinforcement of existing infrastructure would be needed to ensure health visiting teams deliver and record the results of a digital tool in the intended way. Similarly, local authority colleagues stressed that staff will *"need to be taught and coached".* A key consideration raised in this group was how, due to a variety of factors, the nature of the health visiting role has changed over time, and how a shift to digitalisation would represent further change. Staff buy-in, by providing a clear message on the purpose and correct usage of the digital tool in the context of the health and development reviews, was deemed to be critical:

"There are skills with using this kind of technology that they need to adapt to… their course doesn't really prep them for working in this way. We really need to bring [health visitors] them with us, because we can't lose any more." [Local authority colleagues group]

NIHR | Policy Research Unit Children and Families

The NIHR Children and Families Policy Research Unit is hosted by UCL Great Ormond Street Institute of Child Health

UCL