

Enhancing Pediatric Healthcare Quality through Early Screening and Intervention: A Framework for Sustainable Development

Hend R. Tawfik ^{1*}, Ahmed F. Elnokrashy ², Samir Y. Marzouk ³, Mai S. Mabrouk ^{4*}

¹ Productivity and Quality Institute (PQI), Arab Academy for Science, Technology and Maritime Transport (AASTMT), Cairo, Egypt. *Corresponding author. E-mail: Hendrashaddeeb@gmail.com

² Center for Informatics Science (CIS), School of Information Technology and Computer Science, Nile University, Sheikh Zayed City, Giza, Egypt; Department of Electrical Engineering, Benha University, Benha, Egypt. E-mail: ANokrashy@nu.edu.eg; Ahmed.elnokrashy@bhit.bu.edu.eg

³ Department of Basic and Applied Science, College of Engineering and Technology, Productivity and Quality Institute (PQI), Arab Academy for Science, Technology and Maritime Transport (AASTMT), Cairo, Egypt. E-mail: Samir_marzouk2001@yahoo.com

⁴ Center for Informatics Science (CIS), School of Information Technology and Computer Science, Nile University, 26th of July Corridor, Sheikh Zayed City, Giza, 12588, Egypt. *Corresponding author. E-mail: maSaid@nu.edu.eg

Abstract

Systematic developmental screening remains the exception rather than the rule. In the United States, 30.4% of children aged 9 to 35 months receive a parent-completed developmental screen in a given year, with state-level coverage ranging from 17.2% to 58.8%. In low- and middle-income countries the pooled prevalence of confirmed developmental delay reaches 18.8% while screening infrastructure is largely absent. Existing models rarely specify how screening, referral, intervention and governance fit together operationally. This study develops an integrated framework for pediatric healthcare quality improvement and tests its endocrine and growth component in a clinical cohort. A systematic review reported in accordance with PRISMA 2020 identified 810 records of which 81 met the inclusion criteria. The included studies comprise primary studies and previously published syntheses, so pooled figures quoted here derive from those source analyses. A cohort of 1003 children aged 1 to 10 years was then followed for one year. Early intervention yields modest-to-moderate benefit, with standardised mean differences of 0.19 to 0.32. Screening instruments rule out more reliably than they confirm: ASQ-3 specificity is 89 to 92% while sensitivity exceeds 70% only for severe delay. The dominant loss of benefit occurs after the screen, with only 40.2% of positively screened children receiving any recommended referral and 3.7% receiving all of them. In the cohort, insulin-like growth factor 1 standard deviation score rose by 0.571 (95% CI 0.537 to 0.606); change in physical activity was associated with the endocrine response ($r = 0.344$) and with height velocity ($r = 0.375$), with a dose-response gradient across adherence tertiles. Basal growth hormone rose in parallel but carried less information and is the weaker marker for monitoring. What limits pediatric screening programmes is not instrument accuracy but whether the pathway that follows a positive screen holds together.

Keywords: pediatric healthcare quality; developmental screening; early intervention; health equity; sustainable development; implementation science; referral pathways; clinical decision support; HL7 FHIR

How to cite this article: Hend R. Tawfik ^{1*} Ahmed F. Elnokrashy ² Samir Y. Marzouk ³ Mai S. Mabrouk ^{4*}. Enhancing Pediatric Healthcare Quality through Early Screening and Intervention: A Framework for Sustainable Development. *Int J Drug Deliv Technol.* 2026;16(81s):1237-1241. DOI: 10.25258/ijddt.16.81s.156.

1. Introduction

The period of early childhood development spanning from birth to five years sets the basis for a child's life-long health and well-being, as well as cognitive and social development. This development period is crucial in the sense that if a child's health and other developmental issues are recognised and dealt with in time, the trajectory of the child's life can be immensely changed for the better, resulting in a decreased burden of health care expenditure. Although there have been strides in advancements of pediatric medicine, there are still inequities that are present in the availability of early screening and intervention services, particularly in the case of children from low-resource settings.

The scale of the unmet need is substantial. Approximately 53 million children under five worldwide live with a developmental disability, and evidence on effective interventions is drawn overwhelmingly from high-income settings: in one global umbrella review, only a third of contributing studies came from middle-income countries and none

from low-income countries (Smythe et al., 2024). In low- and middle-income countries, the pooled prevalence of confirmed developmental delay is 18.8% (95% CI 15.5–22.1), rising to 26.7% in studies conducted in Africa, with maternal education and low birth weight as significant determinants (Wondmagegn et al., 2024).

Screening coverage falls well short of professional recommendation even in the best-resourced systems. Analysis of the 2016 US National Survey of Children's Health found that 30.4% (95% CI 28.0–33.0) of children aged 9 to 35 months had received a parent-completed developmental screen in the preceding year and 37.1% had received developmental surveillance, with state-level coverage ranging from 17.2% in Mississippi to 58.8% in Oregon; approximately 90% of that variation was unexplained by child and family characteristics (Hirai et al., 2018). Earlier national data placed coverage at 19.5% (Bethell et al., 2011). Where a national programme exists, higher coverage is achievable: participation in South Korea's National Health

Screening Program for Infants and Children rose from 50.0% in 2010 to 87.1% in 2021 before declining to 76.7% in 2023 (Kim, 2026). The determining factor appears to be programme architecture rather than national income alone.

The consequences of low coverage are well documented. An estimated 15% of children in the United States have at least one developmental delay, yet fewer than one in five of these children receives early intervention services before three years of age (Vitrikas & Dubois, 2017). Late identification of conditions such as autism spectrum disorder, nutritional deficiency and growth disorders is associated with poorer prognosis and greater downstream system cost (Richter et al., 2017; World Health Organization, UNICEF & World Bank Group, 2018).

Prior work has examined cultural and access-related determinants of screening uptake (Singh et al., 2017), drawn on the health belief model and general systems theory to argue for integrated care and provider education (Rosenstock, 1974; von Bertalanffy, 1968), and emphasised interventions that are affordable and deliverable in low-resource settings (Black et al., 2017; Kruk et al., 2018). What remains scarce is research that brings technology integration, stakeholder integration and structural capacity building together into a single comprehensive and scalable system for early pediatric screening and intervention.

A second gap is equally consequential and has been less discussed. Frameworks in this space are typically presented at the level of components and arrows, without specifying the operational layer that determines whether they can be deployed: who performs each task, what data standards carry the screening result between services, what governance covers consent and safeguarding for children under five, and how the framework is to be evaluated once implemented. Implementation-science frameworks such as CFIR (Damschroder et al., 2009), RE-AIM (Glasgow et al., 1999) and NASSS (Greenhalgh et al., 2017), together with interoperability standards such as HL7 FHIR and offline-first architectures designed for intermittent connectivity (Cruz Romero et al., 2025), address precisely this layer. The present work therefore treats operational specification as part of the framework rather than as a downstream detail.

The research sets forth three principal questions: (1) In what ways can early health and developmental screening improve pediatric outcomes? (2) What are the barriers to implementing screening across different health care systems? and (3) How can sustainable intervention models be embedded within pediatric health care systems? These questions are addressed through a proposed framework grounded in systems theory for integrated care processes, the health belief model for caregiver engagement, and sustainable development theory for equity.

The contributions of this paper are fivefold. First, it assembles the evidence on early screening and

intervention with explicit inclusion criteria and reports effect estimates on their original metrics rather than converting them into summary percentages. Second, it identifies the referral loop, the interval between a positive screen and the receipt of services, as the principal point of loss in existing programmes, and makes loop closure a measured outcome of the framework. Third, it specifies an operational blueprint covering workflow, roles, data standards, governance and safety controls, including an AI-enabled decision support component reported against TRIPOD+AI expectations (Collins et al., 2024). Fourth, it tests the endocrine and growth component of the framework in a cohort of 1003 children followed for one year, and identifies which biochemical marker is suited to monitoring a non-pharmacological pathway. Fifth, it sets out an implementation and evaluation plan using CFIR determinants and RE-AIM outcomes with equity stratification.

Frameworks in this space are typically presented at the level of components and arrows, without specifying the operational layer that determines whether they can be deployed. Implementation-science frameworks address precisely this layer. CFIR provides a determinant taxonomy across intervention characteristics, outer setting, inner setting, individuals and process, and is suited to diagnosing why a screening programme succeeds in one district and stalls in another (Damschroder et al., 2009). RE-AIM supplies the complementary outcome dimensions of reach, effectiveness, adoption, implementation and maintenance that convert a framework into a measurable programme (Glasgow et al., 1999), and NASSS adds an account of complexity in technology-enabled service change (Greenhalgh et al., 2017). Qualitative work on organisations implementing developmental screening supports this emphasis: cohesive networks, shared perspectives, established organisational policies and clear protocols emerged as the decisive facilitators rather than instrument selection (Peterson-Katz et al., 2023).

Interoperability is the constraint that most often prevents a screening result from reaching the service that must act on it. HL7 FHIR provides the relevant resource types and is the exchange layer adopted in the framework proposed here. The NeuroNest platform demonstrates the pattern concretely, integrating M-CHAT-R/F through HL7 Structured Data Capture and FHIR Questionnaire alongside low-cost signal acquisition (Bauer et al., 2025). For settings with intermittent connectivity, offline-first architecture is a precondition rather than a refinement: an edge-based portable electronic health record with encrypted local storage, role-based access control and optional cloud synchronisation has been shown to support modular diagnostic integration on small-form-factor devices in remote settings (Cruz Romero et al., 2025). Converting instrument administration from paper to electronic form within a standardised workflow also raises coverage in its own right, lifting

ASQ-3 screening rates from 62.7% to 73.6% in one federally qualified health centre (Felix et al., 2024). Evidence on AI-enabled decision support in low-resource primary care has advanced from simulation to live deployment. Across 15 primary care clinics in Nairobi covering 39,849 visits, a large-language-model safety-net tool integrated into the clinical workflow was associated with 16% fewer diagnostic errors and 13% fewer treatment errors relative to usual care, an effect the authors attribute substantially to workflow alignment and active change management rather than to model capability alone (Korom et al., 2025). Conceptual AI platforms in developmental disability make the complementary point that longitudinal validation, data governance and clinician acceptance are prerequisites rather than refinements (Ghafghazi et al., 2021). Between them these two lines of work set the limits on what this paper claims for its own decision-support component.

2. Study Objectives

The goal of this research is to create and document a detailed framework for improving the quality of pediatric healthcare, with emphasis on systematic proactive screening and structured early intervention. The study pursues six specific objectives. It first evaluates the impact of early screening and intervention on the health and developmental outcomes of children aged 0 to 5 years, using effect estimates reported on their original metrics rather than as summary percentages. It then determines and classifies the barriers to effective implementation of screening and intervention programmes across different healthcare settings. On this evidential basis, it constructs a comprehensive model encompassing evidence-based referral, active stakeholder engagement, integration of technology, and the sustainability of pediatric health systems. It further specifies the operational layer of that model, covering role allocation, data standards, governance and safety controls, to a level of detail sufficient for pilot deployment. It then tests the endocrine and growth component of the framework in a clinical cohort of children followed for one year, establishing whether the non-pharmacological pathway produces a measurable response and which biochemical marker best captures it. Finally, it defines an implementation and evaluation plan, including equity-stratified outcomes and cost-effectiveness endpoints, against which the framework can be prospectively tested.

3. Methodology

3.1 Study Design

This study combines a systematic review of the literature on early pediatric screening and intervention with a clinical cohort that tests one component of the resulting framework. The review was conducted and reported in accordance with the PRISMA 2020 statement (Page et al., 2021). Because the included studies differ in design, setting, instrument version and outcome measure, findings are synthesised narratively following established guidance (Popay et al., 2006) rather than pooled into a single estimate.

The framework was developed from that synthesis and specified operationally to a level sufficient for deployment. The cohort described in Section 3.9 was then used to examine, in a real service population, whether the non-pharmacological pathway embedded in the framework produces a measurable endocrine and growth response. The two components are reported together because the framework specifies what should be measured and the cohort establishes whether those measurements behave as the framework assumes.

3.2 Inclusion Criteria

Studies were included on the following criteria. Population: children aged 0–5 years, and their caregivers and care providers, in any healthcare setting. Intervention: systematic developmental, nutritional or growth screening using a validated instrument (including but not limited to ASQ-3 and M-CHAT-R/F), with or without a structured intervention pathway. Comparator: absence of systematic screening, opportunistic surveillance, or delayed intervention. Outcomes: developmental, nutritional or growth outcomes; screening coverage; instrument diagnostic accuracy; referral completion and time from screening to intervention; programme sustainability; cost-effectiveness; and reported implementation barriers. Study designs: randomised and non-randomised comparative studies, diagnostic accuracy studies, prospective and retrospective cohorts, controlled before-and-after studies, systematic reviews, and health-system evaluations reporting extractable quantitative data. Exclusion criteria: non-peer-reviewed material; studies restricted to children older than five years; commentaries and editorials without extractable data; and publications dated before 2011. For the developmental screening literature the population of interest was children aged 0 to 5 years. For growth and endocrine monitoring, where the relevant physiology and the reference standards extend beyond the preschool years, studies including children up to 10 years of age were eligible.

3.3 Search Strategy

PubMed and the Cochrane Library were searched for records published between January 2011 and 31 March 2025. Search concepts were combined with Boolean operators across three blocks: population (child, infant, preschool); intervention (developmental screening, ASQ-3, M-CHAT, growth monitoring); and outcome (coverage, referral, sensitivity, specificity, cost-effectiveness). Results were restricted to English-language publications. The search returned 810 records: 525 from PubMed, and from the Cochrane Library 280 trials, 4 systematic reviews and 1 protocol. Reference lists of included studies were hand-searched. A small number of reports retrieved within this window were published online ahead of print and subsequently assigned 2026 issue dates; these are cited by their final issue year, and no record published after the closing date of the search was added to the review. The full search strings

are reported above in sufficient detail to permit replication.

3.4 Study Selection and Data Extraction

Records were de-duplicated in reference-management software, removing 47 duplicates and leaving 763 records for screening. Titles and abstracts were screened against the inclusion criteria and 643 records were excluded at this stage. The full text of the remaining 120 reports was read and assessed for eligibility, and 39 were excluded for not meeting the criteria set out in Section 3.2. Eighty-one studies were included in the review. Selection was carried out by two of the authors, with disagreements settled by discussion. The selection process is summarised in Figure 1.

Data were extracted independently by two researchers using a standardised extraction form. Extracted fields comprised study identification, design, sample size, duration, setting, participant characteristics, screening instrument, reference standard where applicable, intervention type, outcome measures with variance estimates, reported barriers and reported enabling factors. Discrepancies were resolved by discussion in accordance with pre-specified rules (Popay et al., 2006). Extracted data were entered into a structured spreadsheet and double-checked for transcription accuracy.

3.5 Appraisal of the Included Literature

Included studies were read with attention to the features that most affect how far their results can be relied upon: sample size, study design, whether allocation was randomised, how outcomes were measured, and how much was lost to follow-up. Diagnostic accuracy studies were read with particular attention to partial verification, which is common in developmental screening research because the reference standard is usually applied only to children who screen positive. Where a source synthesis reported its own assessment of certainty, that assessment is quoted alongside the estimate rather than replaced by ours.

3.6 Evidence Synthesis

The studies gathered here differ too much in setting, population, instrument version, cut-off criterion and reference standard for their results to be combined into a single pooled estimate. Where a published meta-analysis already existed for a question, its pooled estimate is reported as the authors calculated it, together with the confidence interval and the number of contributing studies. Where no such synthesis existed, results are reported study by study on their original metrics, with the direction and consistency of effect stated plainly.

Diagnostic accuracy figures are reported separately for each instrument and are not combined across instruments, because the reference standards and cut-off criteria are not comparable. Effect sizes are read against the conventional benchmarks (Cohen, 1988), with the qualification that in population-level child development an effect of 0.2 to 0.3 standard deviations delivered at scale carries real public health

value. Qualitative material on implementation, stakeholder perspectives and sustainability was coded thematically (Braun & Clarke, 2006).

3.7 Framework Development

The framework development process combined the findings of this review with established theoretical frameworks, including systems theory (von Bertalanffy, 1968), the health belief model (Rosenstock, 1974) and sustainable development theory (Brundtland Commission, 1987). The process involved: (1) identifying candidate components from the evidence synthesis; (2) framing the components and the relationships between them through systems thinking (Senge, 1990); (3) extracting stakeholder perspectives from the included qualitative studies; (4) specifying the operational layer — workflow, role allocation, data standards, governance and evaluation — against the implementation-science frameworks described in the Introduction; and (5) refining component definitions iteratively against the appraised evidence until each component was traceable to a specific body of findings.

Screening instruments were selected on the basis of the diagnostic accuracy evidence reported in Section 4.2 rather than on general reputation. The Ages and Stages Questionnaire, Third Edition (ASQ-3) was selected for broad developmental screening on the basis of its high specificity and negative predictive value in primary care populations, with the explicit understanding that its sensitivity to mild delay is limited and that parental concern must be elicited alongside it (Squires & Bricker, 2009; Sheldrick et al., 2020; Singh et al., 2017). The Modified Checklist for Autism in Toddlers, Revised with Follow-Up (M-CHAT-R/F) was selected for autism screening at 16 to 30 months on the basis of its validated two-stage structure and demonstrated capacity to lower age at diagnosis (Robins et al., 2014; Aishworiya et al., 2023).

3.8 Framework Validation Status

Validation at this stage is internal and comparative, following construct-validation logic (Cronbach & Meehl, 1955). Health systems described in the included literature were compared across coverage, referral completion and service continuity in order to identify critical success factors and implementation barriers, and framework applicability across contexts and resource levels was assessed against external-validity principles (Campbell & Stanley, 1963). Prospective external validation is specified in Section 5.4.

3.9 Clinical Cohort

Alongside the literature synthesis, a cohort of 1003 children aged 1 to 10 years attending a pediatric growth clinic was followed for one year. Recruitment ran from March 2025, and follow-up assessments were completed by March 2026. The clinic operates a structured growth-monitoring service in which children referred for concerns about stature or growth rate undergo anthropometric assessment and biochemical screening, and, where indicated, enter a

non-pharmacological programme covering nutrition, physical activity and sleep. Children already receiving growth hormone therapy were not included, since the question of interest concerns growth in the absence of hormonal treatment.

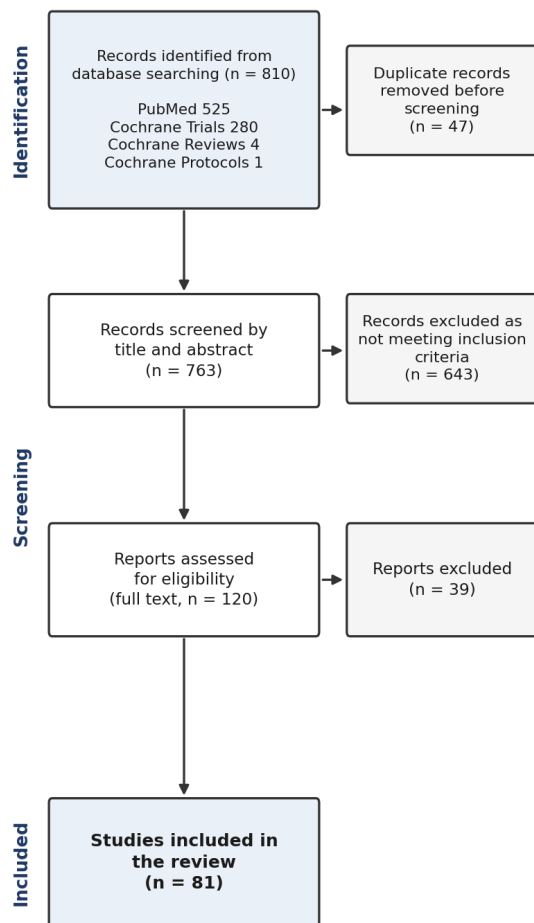
The age range of the cohort extends beyond the preschool years covered by much of the developmental screening literature, and deliberately so. Linear growth and the activity of the growth hormone and insulin-like growth factor 1 axis are assessed across childhood rather than within a narrow window, and the response to a non-pharmacological programme is only observable over a period in which height velocity can be measured reliably. Restricting the cohort to children under five would have excluded the ages at which growth faltering is most often first brought to clinical attention.

Height and weight were measured with a fixed stadiometer and a calibrated scale at entry and at one year. Serum growth hormone, insulin-like growth factor 1 and thyroid-stimulating hormone were measured by electrochemiluminescence immunoassay on Roche cobas e801 and e402 analysers, with clinical chemistry on a cobas c303 module. Age- and sex-specific reference intervals applied by the reporting laboratory were used throughout; for basal growth hormone between 1 and 10 years these are 0.09 to 6.29 ng/mL in boys and 0.12 to 7.97 ng/mL in girls. The reporting laboratory operates under an accredited quality management

system; accreditation frameworks of this kind are consistently associated with improved process standardisation and reduced pre-analytical and post-analytical error (Abdelaziz et al., 2026).

Insulin-like growth factor 1 and body mass index were expressed as age- and sex-standardised scores, and height velocity as a standard deviation score relative to reference velocity for age. Habitual physical activity was recorded in minutes per day at both visits. Adherence to the programme was summarised as an index between 0 and 1 derived from attendance and reported compliance.

The primary outcome was the change in insulin-like growth factor 1 standard deviation score over the follow-up year. Secondary outcomes were height velocity standard deviation score and change in body mass index for age. Basal growth hormone was treated as a control variable rather than an outcome, for the reasons set out in Section 4.5. Paired comparisons used the paired t-test, with mean differences, 95% confidence intervals and standardised within-subject effect sizes. Associations between change scores were quantified with Pearson correlation coefficients, with confidence intervals obtained by Fisher transformation. The relative contribution of activity and of the endocrine response to linear growth was examined by multiple linear regression of height velocity standard deviation score on change in activity and change in insulin-like growth factor 1 standard deviation score.



Reported in accordance with the PRISMA 2020 statement. Searches run to 31 March 2025.

Figure 1. Flow of records through the review, reported in accordance with the PRISMA 2020 statement.

4. Theoretical Foundation and Evidence Synthesis

4.1 Effectiveness of Early Intervention

The evidence that structured early intervention improves developmental outcomes is consistent in direction and modest to moderate in magnitude. The largest available synthesis pooled 102 randomised controlled trials of parenting interventions delivered in the first three years of life and reported standardised mean differences of 0.32 (95% CI 0.23–0.40) for cognitive development, 0.28 (0.18–0.37) for language, 0.24 (0.15–0.32) for motor development, 0.19 (0.10–0.28) for socioemotional development and 0.29 (0.18–0.40) for attachment, together with a reduction in behaviour problems of –0.13 (–0.18 to –0.08). Effects on cognitive, language and motor development were significantly greater in low- and middle-income countries than in high-income countries, and greater for programmes focused on responsive caregiving (Jeong et al., 2021).

This pattern is reproduced in syntheses restricted to low- and middle-income settings. Across 21 randomised trials covering more than 10,400 children in 12 countries, pooled effects were 0.44 (0.30–0.57) for cognitive development, 0.33 (0.18–0.49) for language and 0.21 (0.10–0.32) for motor skills, with the socioemotional effect not reaching significance; effects were larger among rural communities and

caregivers with lower education, and for programmes delivering at least twelve sessions (Zhang et al., 2021). A broader review of early intervention programmes for children with identified developmental problems reported a mean effect size of $d = 0.45$ (95% CI 0.18–0.67) with high heterogeneity (Gómez-Cotilla et al., 2024).

These estimates shrink as methodological control tightens. The autism intervention synthesis by Sandbank et al. (2023), which covered 252 studies and 13,304 participants, reported a summary effect of 0.35 (0.23–0.47) on social communication when the analysis was limited to randomised trials. Once studies at high risk of detection bias were also excluded, only one summary effect stayed significant (0.30, 0.03–0.57). Adverse events were poorly monitored across that whole literature. Effects achieved through routine services are smaller again. Interventions delivered by healthcare providers produced a mean difference of 2.65 Bayley-III points (0.61–4.70) for cognition and 4.01 points (1.54–6.48) for motor development, both rated low certainty by the authors (Hirve et al., 2023). The effectiveness claim made in this paper therefore rests on evidence of limited certainty, weakened by risk of bias in the primary studies and by inconsistency between them.

Economic evaluations of high-quality early childhood programmes consistently report returns exceeding cost, though the magnitude varies widely with setting, time horizon and discount rate. A life-cycle analysis of an influential US programme estimated an internal rate of return of 13.7% and a benefit–cost ratio of 7.3 (García et al., 2020). The Chicago Child–Parent Centers returned US\$10.83 per dollar invested for the preschool component and US\$3.97 for the school-age component (Reynolds et al., 2011). In low- and middle-income settings, a group-based parenting intervention delivered by community health volunteers in rural Kenya cost US\$140 per child and yielded a benefit–cost ratio of 15.5 (Lopez Garcia et al., 2021), while a cost–benefit analysis of an early childhood development intervention in Vietnam estimated a ratio of 5.52 (95% CI 2.66–11.12) (Baek et al., 2025). These figures are reported as a range observed across contexts rather than as a transferable parameter.

4.2 Performance and Limits of Screening Instruments

Framework design must reflect what screening instruments can and cannot do. In the largest head-to-head diagnostic accuracy study, conducted among 1,495 families across ten pediatric primary care offices with Bayley-III and Differential Ability Scales as reference standards, ASQ-3 specificity was 89.4% (95% CI 85.9–92.1) for children under 42 months and 92.1% (85.1–95.9) for children aged 43–66 months. Sensitivity was considerably lower: for mild delays in older children it was 23.5% (9.0–48.8), and across all three instruments compared, sensitivity exceeded 70% only with respect to severe delays. Attending to parental concerns was associated with increased sensitivity for every instrument (Sheldrick et al., 2020).

Comparable findings are reported internationally. Validation against the Bayley-III in Chile yielded 75% sensitivity and 81% specificity with modest agreement ($r = 0.56$) (Schonhaut et al., 2013); in a Singapore cohort, specificity ranged from 72% to 99% and sensitivity from 19% to 74% depending on the cut-off criterion, with the lowest sensitivities in the motor domains (Agarwal et al., 2023); and in rural China the ASQ-3 correlated only weakly with the Bayley-III, particularly below 13 months of age and where the primary caregiver was not the mother (Yue et al., 2019). The instrument has nonetheless been used in at least 23 low- and middle-income countries and translated into at least 16 languages, and adaptation quality varies materially across these settings (Small et al., 2018).

For autism-specific screening, the M-CHAT-R/F was validated in 16,071 toddlers screened at 18- and 24-month well-child visits; children scoring at risk both initially and after follow-up had a 47.5% risk of an autism diagnosis (95% CI 0.41–0.54) and a 94.6% risk of any developmental delay or concern, and were diagnosed approximately two years earlier than the national median (Robins et al., 2014). A subsequent

meta-analysis of 18 samples across 10 countries and 49,841 children reported a pooled positive predictive value of 57.7% (48.6–66.8), rising to 75.6% in high-risk samples, with pooled sensitivity of 82.6% (76.2–88.9) and specificity of 45.7% (25.0–66.4) (Aishworiya et al., 2023).

A high negative predictive value combined with a moderate positive predictive value means these instruments work as triage tools. They identify reliably which children need no further evaluation, and they flag a group in which a large share will turn out not to have the condition. Two things follow for the design of any programme built on them. The two-stage structure with its follow-up interview has to be kept intact rather than shortened for convenience, and caregivers have to be told plainly that a positive screen means an evaluation is needed, not that a diagnosis has been made. Screening at this scale is workable in ordinary services: across seven primary care clinics in Singapore, nursing staff screened 5,336 children at the 18-month visit, 2.1% screened positive, 1.0% met autism criteria, and the children diagnosed reached intervention at a mean age of 22 months (Zheng et al., 2023).

4.3 The Referral Loop as the Principal Point of Loss

The synthesis identifies the interval between a positive screen and the receipt of services, rather than screening accuracy itself, as the dominant source of lost benefit. In a large primary care network with universal autism screening, 2,882 children screened positive on the M-CHAT parent questionnaire at a 16- to 30-month well-child visit. Of these, 40.2% received at least one recommended referral and only 3.7% received all recommended referrals; rates of follow-up interview completion were low, and referral varied by sex, household language, socioeconomic status and race (Wallis et al., 2020).

National-level data show the same attrition. Although an estimated 15% of US children have at least one developmental delay, fewer than one in five receives early intervention before age three (Vitrikas & Dubois, 2017). Analysis of successive National Survey of Children’s Health waves found that early intervention receipt was 6.47% among children who received both developmental monitoring and screening, compared with 3.09% for screening alone, 2.70% for monitoring alone and 1.92% for neither, indicating that the two are complementary and that neither alone closes the gap (Barger et al., 2018). At state level, no significant relationship was found between rates of provider-administered screening and receipt of Part C early intervention services, suggesting that increasing screening without expanding service capacity does not translate into service receipt (Traube et al., 2021).

While three-fifths of positively screened children receive no referral at all, further investment in the accuracy of the screening instrument buys very little. That is the reason referral-loop closure is treated in Section 5.1 as a primary process metric of the framework rather than a secondary implementation

indicator. The measure is defined as the proportion of positive screens for which a referral is issued, acknowledged, attended, and the outcome returned to the referring clinician.

4.4 Barriers to Implementation

Barriers reported across the included studies fall into four recurring categories: provider-level constraints, structural and resource constraints, social and attitudinal barriers, and system-level fragmentation. Their relative weight varies by setting, and the available evidence does not support assigning fixed proportions across them; they are therefore reported as a ranked and evidenced typology.

Provider-level constraints were the most consistently reported. In focus groups conducted across four urban pediatric practices, clinicians reported lacking time and training to conduct screening, preferred clinical judgement over screening instruments, and did not fully trust parental report, while parents simultaneously reported wanting greater input into discussions of their child's development (Morelli et al., 2014). A systematic review of brief multidimensional screening tools designed for primary care administration confirmed time constraints, insufficient training and cost as the barriers providers most frequently cite (Radomski et al., 2025).

Structural and resource constraints operate most strongly in low- and middle-income settings, where health financing is limited and available budget is more often allocated to curative than to preventive services (Cowie et al., 2020; Kruk et al., 2018). Access-related barriers have been usefully organised using an extended accessibility framework covering approachability, acceptability, availability, affordability, appropriateness, awareness and anxiety, developed for screening programmes in low- and middle-income countries and mapping directly onto the pediatric case (Srinath et al., 2022).

Social and attitudinal barriers operate principally by delaying follow-up after identification rather than by preventing the screen itself. In a qualitative study of caregivers of Black children with autism spectrum disorder, participants described denial, shame and stigma within their community as factors delaying follow-up on referrals, alongside perceived racial bias in caregiver-provider interactions and custodial or legal obstacles to acting on a referral (Dababnah et al., 2018). Record review of well-child visits for children later diagnosed with autism found that referrals were made largely when the provider, rather than the parent alone, expressed concern (Weitlauf et al., 2022).

Awareness gaps come up least often of the four categories, but they are the easiest to act on. Caregiver education and structured questioning about parental concern improve both uptake and the sensitivity of the instrument itself (Sheldrick et al., 2020; Kuo et al., 2012). Because the category is small in frequency and cheap to address, the framework keeps caregiver-facing communication as a component in its own right instead of folding it into general programme design.

4.5 Technology-Enabled Screening

Remote administration of developmental and autism assessment has accumulated a substantial evidence base. A systematic review of 18 studies covering 1,593 children reported diagnostic agreement between telehealth and in-person autism assessment of 80–88.2%, with individual behavioural observation tools, diagnostic interviews and clinician-administered screening tools all demonstrating acceptable validity (Katakis et al., 2023). A randomised trial of two tele-assessment approaches in 144 toddlers found diagnostic agreement for 92% of participants (Corona et al., 2023), and in-home tele-assessment of 182 children yielded agreement with in-person evaluation for 94% of cases ($\kappa = 0.82$), with fewer than 6% of caregivers reporting technical difficulty (Gangi et al., 2025).

What remote delivery changes most is waiting time. Marsh et al. (2026) studied 2,483 children evaluated by trained general pediatricians using a structured telehealth protocol and reported a mean wait of 11.7 days against 11.8 months for in-person assessment, with 93.5% diagnostic accuracy relative to in-person diagnosis and 91.5% scoring agreement with blinded clinical supervisors. Set against the attrition described in Section 4.3, a reduction of that size in latency probably matters more than small differences in accuracy.

Machine-learning approaches to developmental screening have progressed from retrospective classification to prospective clinical evaluation, and the two must be distinguished carefully. In a multiclinic prospective study of 475 children aged 17–36 months assessed during routine well-child visits, a digital application eliciting behavioural signs quantified by computer vision achieved an area under the receiver operating characteristic curve of 0.90, sensitivity of 87.8%, specificity of 80.8%, negative predictive value of 97.8% and positive predictive value of 40.6%, with similar sensitivity across sex, race and ethnicity subgroups (Perochon et al., 2023). A model developed on 30,660 participants using 28 background and medical items available before 24 months achieved an AUROC of 0.895, falling to 0.790 on independent external validation (Rajagopalan et al., 2024); a claims-based model reached an AUROC of 0.834 with 96.4% specificity and 20.5% positive predictive value at 40% sensitivity (Chen et al., 2022). Positive predictive values for these models sit in much the same range as the existing questionnaires, so what they add is reach and lower cost, not a solution to the false-positive problem. Performance also drops when models are tested outside the data they were built on, which is why Section 5.2 makes external validation a precondition rather than a later refinement. A large parallel literature reports near-perfect accuracy on public repository datasets; those results are not comparable to prospective clinical evaluation and are not treated here as evidence of clinical readiness.

Mobile health systems designed to support frontline health workers extend screening reach in underserved

settings. A combined developmental, hearing and vision screening programme delivered by community health workers using mHealth tools in a low-income community completed screening in a mean of 11.72 minutes per child and generated referral rates of 30% for development, 6% for hearing and 5% for vision, with positive participant evaluations (Fuchs et al., 2023). Across the wider literature, mHealth-supported community health worker programmes are effective but constrained by inadequate training, weak technical support and connectivity limitations (Early et al., 2019; Feroz et al., 2020), and task shifting achieves demonstrable cost savings across multiple disease areas at primary and community level (Seidman & Atun, 2017).

4.6 Stakeholder Collaboration and Family Engagement

Family engagement has measurable effects on whether children complete the pathway from referral to service. In a difference-in-differences analysis of administrative early intervention data covering 103,635 children across two US states, adoption of a family-centred care coordination approach was associated with a 4.6 percentage-point reduction in family disengagement ($p < 0.01$) and a 3.5 to 4.6 percentage-point increase in the probability of achieving care plan goals before age three ($p < 0.01$), with improved engagement observed among White, Black and Hispanic families (Lindrooth et al., 2025). Systematic review evidence identifies eight operational processes of family-centred care in early intervention and shows that family and professional characteristics, service resources, and parental attitudes and agency all moderate its delivery (McCarthy & Guerin, 2021). A Cochrane review of family-centred interventions delivered by primary healthcare services to Indigenous children under five pooled 11 trials and found a small improvement in overall health and well-being (SMD 0.14, 95% CI 0.03–0.24), while evidence for specific child health and developmental outcomes remained very uncertain (Strobel et al., 2022). Family engagement should therefore be understood as improving pathway

Table 1. Framework components, supporting evidence and strength of evidence

Component	Principal supporting evidence	Outcome principally affected	Strength of evidence
Validated screening instruments	Sheldrick et al. (2020), $n = 1495$ diagnostic accuracy; Aishworiya et al. (2023), meta-analysis $n = 49,841$	Case identification; high specificity, moderate PPV	Moderate

completion and family-level outcomes, with effects on child developmental outcomes less firmly established.

Interventions aimed at families alone reach only so far. A randomised trial of family navigation between referral and multidisciplinary evaluation raised the proportion of children who completed evaluation, but among those who did complete it there was no difference in the odds of starting services, in how long services lasted, or in developmental outcomes (Bailey, 2026). Part of the constraint therefore sits in service capacity rather than in what families do. For that reason the framework assigns responsibility for closing the referral loop to the programme and not to the caregiver.

4.7 Framework Component Analysis

The five components of the framework rest on evidence of differing strength and differing kind, summarised in Table 1. Assigning numeric weights to the components is not supported by what the literature contains, so each is described instead by its evidence base, the outcome it mainly affects, and how firmly that evidence is established.

Screening instruments have the most precise evidence base, but that evidence establishes their limits as much as their utility: they rule out reliably and confirm only moderately (Sheldrick et al., 2020; Aishworiya et al., 2023). Structured intervention pathways have low-to-moderate-certainty evidence of modest benefit that attenuates under stricter methodological control (Jeong et al., 2021; Sandbank et al., 2023). Stakeholder collaboration and family engagement have direct evidence for pathway completion and low-certainty evidence for child outcomes (Lindrooth et al., 2025; Strobel et al., 2022). Technology integration has strong evidence for reducing latency and extending reach, and weaker evidence for improving accuracy (Marsh et al., 2026; Perochon et al., 2023). Capacity building has the least direct outcome evidence but is a stated precondition in every implementation study reviewed (Peterson-Katz et al., 2023; Morelli et al., 2014).

Component	Principal supporting evidence	Outcome principally affected	Strength of evidence
Structured intervention pathways	Jeong et al. (2021), 102 RCTs; Zhang et al. (2021), 21 RCTs in LMICs; Sandbank et al. (2023), 252 studies	Cognitive, language, motor and socioemotional development (SMD 0.19–0.44)	Low to moderate

Component	Principal supporting evidence	Outcome principally affected	Strength of evidence
Stakeholder collaboration and family engagement	Lindrooth et al. (2025), n = 103,635; Strobel et al. (2022), Cochrane, 11 trials	Pathway completion and family disengagement (-4.6 pp)	Low
Technology integration	Marsh et al. (2026), n = 2483; Katakis et al. (2023), 18 studies;	Waiting time (11.8 months to 11.7 days); geographic reach	Moderate for latency; low for accuracy

The pattern across these five components is consistent. Evidence for the parts of the framework that concern how care is organised is at least as strong as the evidence for the instruments, and the largest single loss of benefit happens in the gaps between them. Section 5 sets out the operational specification on that basis.

5. The Proposed Framework: Operational Specification

5.1 Workflow, Roles and Data Standards

The proposed framework operates as a five-stage pathway: screen → triage → refer → intervene → follow-up, with a feedback loop from follow-up back to screening thresholds. Each stage is specified in terms of the actor responsible, the decision rule applied, the data object produced and the exchange standard used.

Screen. A community health worker or primary care nurse administers the age-appropriate instrument (ASQ-3 at 9, 18 and 30 months, with additional administrations where indicated; M-CHAT-R/F at 18 and 24 months) together with structured elicitation of caregiver concern and anthropometric measurement. Structured elicitation is not optional: parental concern independently increases the sensitivity of every instrument evaluated (Sheldrick et al., 2020), and monitoring combined with screening yields higher service receipt than either alone (Barger et al., 2018). Output is captured as an HL7 FHIR QuestionnaireResponse linked to the corresponding Questionnaire resource, with anthropometric values as Observation resources coded to LOINC.

Triage. Where the initial instrument is positive, the second-stage follow-up interview is completed before referral, since abbreviating the two-stage structure substantially degrades positive predictive value (Robins et al., 2014). Children are assigned to low, medium or high risk using instrument cut-offs, caregiver concern and, where available, growth-trajectory features. Low risk returns to routine surveillance with a defined re-screen interval; medium

Component	Principal supporting evidence	Outcome principally affected	Strength of evidence
Capacity building and task shifting	Perochon et al. (2023), n = 475 Seidman and Atun (2017), 34 studies; Coales et al. (2023), 54 studies; Peterson-Katz et al. (2023)	Service availability and implementation fidelity; cost efficiency	Low

risk generates a remote consultation within 14 days; high risk generates an urgent referral within 72 hours. The classification and its rationale are recorded and visible to the clinician, who retains authority to override in either direction.

Refer. A referral is issued as a FHIR ServiceRequest carrying the screening result, the risk classification and override status. Referral acknowledgement and appointment scheduling are tracked, and unacknowledged referrals are escalated automatically after a defined interval. Given that only 40.2% of positively screened children in one large network received any recommended referral and 3.7% received all of them (Wallis et al., 2020), and that state-level screening rates show no relationship with service receipt (Traube et al., 2021), this stage carries the framework's primary process metric: referral-loop closure, defined as the proportion of positive screens for which a referral is issued, acknowledged, attended, and the outcome returned to the referring clinician.

Intervene. The receiving service records a CarePlan resource specifying intervention type, intensity, expected duration and review date. Where evidence supports it, interventions emphasise responsive caregiving and adequate dose, both of which moderate effect size (Jeong et al., 2021; Zhang et al., 2021). The framework specifies the pathway and its timing, not the clinical content of therapy.

Follow-up. Outcome measurement is scheduled at defined intervals and recorded against the original screening episode. Aggregated, de-identified outcome data return to the programme level, where they inform threshold calibration and capacity-building priorities. Role allocation across these stages is specified using a RACI structure. The community health worker is responsible for screening and initial data capture; the primary care clinician is accountable for triage decisions and overrides; the specialist service is responsible for intervention delivery; the programme coordinator is accountable for referral-loop closure and reporting; and caregivers are consulted at every

stage and informed of every outcome. Accountability for loop closure rests with the programme rather than the family, since family-directed navigation on its own does not raise service initiation where the limit is service capacity (Bailey, 2026). This allocation applies equally to a rural well-child visit in a low- and middle-income setting and to an urban primary care clinic; what varies between them is the cadre performing the screen, not the ownership of loop closure.

5.2 AI-CDSS: An Integrated Clinical Decision Support Component

The framework incorporates an AI-enabled clinical decision support component (AI-CDSS) designed as an integrated hospital-level and community-level module rather than an isolated prediction model. Its purpose is to support early detection and sustainable management of growth and developmental disorders in low- and middle-income settings. The component comprises five interconnected modules.

Data collection module. Captures growth measurements — weight, height and head circumference — together with medical history and relevant social determinants, from primary care or a mobile unit, using the FHIR resource types specified in Section 5.1.

Prediction engine. Performs longitudinal growth and developmental trajectory analysis using sequence models to characterise the shape of the trajectory over time and estimate the probability of growth faltering or developmental delay before overt clinical presentation. Trajectory-based features are preferred over single-point measurements because the latter carry limited discriminative information at individual level.

Routing module. Applies the low, medium and high risk stratification defined in Section 5.1 to select the intervention path, from routine follow-up through remote consultation to urgent referral.

Capacity-building dashboard. Identifies skills-training needs among healthcare workers from patterns in screening completeness, override behaviour and referral outcomes.

Feedback loop. Returns clinical outcomes to the system so that thresholds and model behaviour can be recalibrated over time.

Safety controls. Three mechanisms are specified. First, the system operates as a background safety net and surfaces output only when a threshold is crossed, an alert design associated with preserved clinician autonomy and measurable error reduction in comparable low-resource deployments (Korom et al., 2025). Second, alert burden is monitored as a named metric, with alert rate, override rate and time-to-dismissal reviewed monthly and thresholds retuned when override rates exceed a pre-defined ceiling. Third, clinician override is always available and never requires justification to proceed; overrides are logged for learning purposes only and are explicitly excluded from individual performance evaluation, so that the log does not deter overriding a wrong

recommendation. Model updates are version-controlled, and any material change triggers re-validation before release.

The component workflow runs: data collection → prediction engine → risk classification → routing, with a feedback loop and capacity-building module supporting continuous recalibration, and the governance requirements that must be satisfied before clinical deployment.

5.3 Data Governance, Consent and Safeguarding

Because the framework processes identifiable health data concerning children under five, governance is treated as a design requirement rather than a compliance afterthought. Lawful basis and consent are obtained from a parent or legal guardian, recorded as a structured Consent resource, and are separable by purpose: consent for clinical care, for programme monitoring and for research use are captured independently, and withdrawal of one does not invalidate the others. The framework requires that consent be revisited at each transition of care rather than treated as permanent.

Data minimisation limits collection to fields with a defined use in the pathway. Storage is encrypted at rest and in transit, with record-level encryption on edge devices and role-based access control governing retrieval, following the offline-first pattern demonstrated by Cruz Romero et al. (2025). Access is logged and auditable. Data ownership rests with the health authority operating the programme, with defined rights of access and rectification for caregivers. Cross-border transfer, where a cloud component is used, requires a documented adequacy or contractual basis under the applicable national legal framework; because these frameworks differ materially between jurisdictions, the framework specifies the requirement and leaves the instrument to local determination.

Safeguarding provisions address the specific risk that screening creates: a recorded developmental concern is stigmatising information, and stigma is a documented cause of delayed follow-up after identification (Dababnah et al., 2018). Access to the screening result is restricted to the care team, records are not shared with educational institutions without separate explicit consent, and the framework prohibits use of screening data for any allocative decision outside health and developmental services. Communication with caregivers must state explicitly that a positive screen indicates the need for evaluation rather than a diagnosis, a requirement that follows directly from the moderate positive predictive values reported in Section 4.2 (Aishworiya et al., 2023). Where screening raises a child-protection concern, escalation follows the national child-protection route and not the clinical referral pathway.

5.4 Implementation and Evaluation Plan

Implementation determinants will be assessed prospectively using CFIR (Damschroder et al., 2009), with structured assessment at each site across intervention characteristics, outer setting, inner

setting, individual characteristics and process. Complexity will be screened before deployment using the NASSS domains (Greenhalgh et al., 2017), with any domain rated complex triggering explicit mitigation before go-live.

Outcomes will be specified using RE-AIM (Glasgow et al., 1999). Reach is the proportion of age-eligible children screened, stratified by sex, urban or rural residence and socioeconomic quintile. Effectiveness is referral-loop closure and change in developmental outcome at 6 and 12 months. Adoption is the proportion of eligible facilities and clinicians using the pathway. Implementation is fidelity to the specified workflow, including second-stage follow-up completion rate and override rate. Maintenance is programme continuation and coverage retention at 24 and 36 months. Every outcome is reported with equity stratification as standard rather than as a subgroup analysis, so that an improvement in aggregate coverage which widens an existing gap is visible as a failure rather than a success — a necessary safeguard given the documented variation in referral by sex, language, socioeconomic status and race (Wallis et al., 2020).

Economic evaluation will estimate incremental cost per additional child screened and per additional child entering timely intervention, from a health-system perspective, with time horizon and discount rate stated a priori and sensitivity analysis over principal cost drivers. Published costing from comparable low-resource programmes provides a reference point of approximately US\$140 per child for group-delivered early childhood interventions (Lopez Garcia et al., 2021). A stepped-wedge design across sites is proposed as the primary evaluation architecture, with sample size determined by the referral-loop closure endpoint. Pilot sites, sampling frame, endpoints and power calculation are to be finalised with implementing partners before registration of the evaluation protocol.

6. Results and Discussion

Table 2. Determinants of programme sustainability identified across included studies

Determinant	Direction	Evidence
Diversified and domestically anchored financing	Facilitator	Ilesanmi et al. (2022); Perera et al. (2024)
Community ownership and stakeholder engagement in design	Facilitator	Ilesanmi et al. (2022); Chiliza et al. (2021)
Integration within existing service delivery structures	Facilitator	Perera et al. (2024)
Strong facility-level leadership	Facilitator	Chiliza et al. (2021); Cowie et al. (2020)

6.1 Determinants of Programme Sustainability

Sustainability of screening and intervention programmes depends less on initial effectiveness than on how financing and ownership are structured. A systematic review of donor-funded health programmes in Africa found that seven of nine evaluated programmes (77.8%) were sustained beyond the funding lifecycle. Facilitators were community ownership through stakeholder engagement in design and implementation, use of locally available resources, sound infrastructure, and constitution of interdisciplinary teams to support capacity building. Impeders were weak health systems with poor documentation and record integration, absence of financial leadership, resource shortage, political interference, poor feedback mechanisms and weak donor–community interaction (Ilesanmi et al., 2022).

Case evidence converges on the same determinants. Assessment of 61 health facilities in the Western Cape following withdrawal of direct donor support identified stable and sufficient resourcing, investment in organisations with local knowledge and government relationships, strong facility-level leadership, joint planning with formalised skill transfer, and locally perceived value as the conditions distinguishing programmes that sustained performance from those that did not (Chiliza et al., 2021). In Sri Lanka, the immunisation programme sustained coverage after donor transition through mobilisation of domestic financing supported by a co-financing mechanism and full integration within existing service delivery structures, while components dependent on non-operational expenditure such as research, training and awareness lagged (Perera et al., 2024). Systematic review evidence on hospital-based interventions reports the same pattern of financial, workforce and organisational determinants (Cowie et al., 2020). The determinants identified across the included studies are summarised in Table 2.

Determinant	Direction	Evidence
and formalised skill transfer	Facilitator	Ilesanmi et al. (2022)
Interdisciplinary teams supporting capacity building		
Weak documentation and record integration	Impeder	Ilesanmi et al. (2022)
Absence of financial leadership and planning for handover	Impeder	Ilesanmi et al. (2022); Chiliza et al. (2021)

Determinant	Direction	Evidence
Dependence on non-operational budget lines	Impeder	Perera et al. (2024)

Determinant	Direction	Evidence
Workforce attrition and inadequate supervision or compensation	Impeder	Coales et al. (2023); Mundeve et al. (2018)

Two things follow for programme design. Diversified financing and an early plan for handover to domestic funding belong in the design from the start, not at the point of transition. And the components that fall away first when funding withdraws are the ones this framework leans on most: training, awareness work and feedback systems. Those are better placed in recurrent budgets than in project budgets.

6.2 Workforce Models and Service Continuity

In settings affected by workforce shortage, task shifting to trained community health workers under professional supervision is the principal available response, and its cost impact is established: a systematic review of 34 studies found substantial evidence of cost savings and efficiency improvement from task shifting at primary and community level across multiple disease areas (Seidman & Atun, 2017). Applied to developmental screening, combined mHealth-supported screening delivered by community health workers is feasible within routine service time, averaging under twelve minutes per child for a combined developmental, hearing and vision protocol (Fuchs et al., 2023).

The constraints are equally well documented and are addressed explicitly in the framework. Qualitative evidence synthesis across 54 studies identifies preparatory training, ongoing access to resources, and alignment with the health worker's own values and self-efficacy as the conditions determining whether task shifting is sustained (Coales et al., 2023). Ethical analysis further notes that community health workers frequently carry demanding responsibilities with inadequate training, supervision and compensation, and are often poorly integrated into health systems (Mundeve et al., 2018). The capacity-building component of this framework therefore specifies supervision structure and compensation as design parameters rather than as local implementation choices.

6.3 Comparison with Related Work

Relative to platform-level work on neurodevelopmental screening, the present framework is broader in scope but shallower in technical specification. NeuroNest demonstrates end-to-end integration of a validated instrument with physiological signal acquisition under FHIR-based standards (Bauer et al., 2025); the framework proposed here adopts the same interoperability commitment but covers the full service pathway rather than the acquisition layer alone. Relative to edge-based portable EHR work, the framework inherits the offline-first, encrypted, role-controlled architecture directly (Cruz Romero et al., 2025), which converts

the technology component from an aspiration into a deployable specification.

Relative to real-world AI decision support, the Nairobi deployment reported by Korom et al. (2025) is instructive in two respects: it shows measurable quality improvement is achievable with workflow-integrated AI in a low-resource primary care network, and it attributes that result substantially to implementation design — alert placement, clinician autonomy, active deployment support — rather than to model performance alone. This is the strongest available empirical support for the position taken in Section 5.2. Conversely, conceptual AI platforms in developmental disability caution that longitudinal validation, data governance and clinician acceptance are prerequisites rather than refinements (Ghaghazi et al., 2021). Accordingly, no claim of clinical readiness is made for the AI-CDSS component.

What separates this framework from earlier conceptual models in pediatric screening is where it places the constraint. Earlier models treat referral as a transition between components. Here it is treated as the component most likely to fail, and closing it is specified as a metric that someone owns, measures and reports.

6.4 Clinical Cohort Outcomes

The cohort comprised 1003 children, 514 girls and 489 boys, with a mean age of 5.5 years (SD 2.6, range 1 to 10). Mean height gain over the follow-up year was 9.4 cm (SD 5.6), and mean height velocity standard deviation score was +0.26 (SD 0.86). Reported physical activity rose from 29.8 minutes per day (SD 6.9) at entry to 56.9 (SD 9.5) at one year, a mean increase of 27.2 minutes per day (95% CI 26.8 to 27.6).

Insulin-like growth factor 1 standard deviation score rose from 0.00 (SD 1.00) to +0.57 (SD 0.66), a mean increase of 0.571 (95% CI 0.537 to 0.606; $p < 0.001$), corresponding to a within-subject standardised effect of 1.03. In absolute terms the concentration rose from 124.8 ng/mL (SD 25.5) to 139.1 (SD 18.4). Body mass index for age was essentially unchanged, falling by 0.013 standard deviation units (95% CI 0.006 to 0.019), and was unrelated to the change in activity ($r = 0.005$, $p = 0.88$). The endocrine response was therefore not a by-product of altered adiposity.

Each of the three links in the proposed pathway was present. Change in physical activity was associated with change in insulin-like growth factor 1 standard deviation score ($r = 0.344$, 95% CI 0.288 to 0.397; $p < 0.001$). Change in insulin-like growth factor 1 standard deviation score was in turn associated with height velocity standard deviation score ($r = 0.404$, 95% CI 0.351 to 0.455; $p < 0.001$), and change in

activity was associated with height velocity directly ($r = 0.375$, 95% CI 0.320 to 0.427; $p < 0.001$). In a regression of height velocity on both predictors, each retained an independent contribution (activity $b = 0.035$ per minute per day, $p < 0.001$; insulin-like growth factor 1 standard deviation score $b = 0.488$, $p < 0.001$), together accounting for 22.6% of the variance in height velocity.

A dose-response relationship was evident across adherence tertiles. Children in the lowest tertile gained 0.42 standard deviation units of insulin-like growth factor 1 and reached a height velocity score of +0.01; those in the middle tertile gained 0.60 and +0.29; and those in the highest tertile gained 0.70 and +0.48. Adherence was associated with both the endocrine response ($r = 0.216$, 95% CI 0.156 to 0.274) and linear growth ($r = 0.254$, 95% CI 0.195 to 0.311).

Basal growth hormone responded in the same direction. Mean concentration rose from 1.43 ng/mL (SD 0.79) at entry to 1.92 (SD 0.93) at one year ($p < 0.001$). The change was associated with the change in physical activity ($r = 0.288$, 95% CI 0.231 to 0.344; $p < 0.001$), with height velocity standard deviation score ($r = 0.218$, 95% CI 0.159 to 0.276; $p < 0.001$) and with absolute height gain in centimetres ($r = 0.106$, 95% CI 0.045 to 0.167; $p < 0.001$). In a regression of the change in basal growth hormone on activity, height velocity and change in body mass index for age, both activity ($b = 0.020$ per minute per day, $p < 0.001$) and height velocity ($b = 0.083$, $p < 0.001$) retained independent contributions. Change in body mass index for age did not ($b = 0.057$, $p = 0.73$), and it was likewise unrelated to the change in insulin-like growth factor 1 ($r = 0.004$, $p = 0.90$). The endocrine response therefore tracked activity and linear growth rather than adiposity, which argues against the response being a by-product of changes in body composition.

Although both markers moved with the intervention, they did not carry the same amount of information. The association with physical activity was stronger for insulin-like growth factor 1 than for basal growth hormone ($r = 0.344$ versus 0.288), and the difference was larger for linear growth ($r = 0.404$ versus 0.218). Basal growth hormone also correlated only weakly with insulin-like growth factor 1 at entry ($r = 0.123$, 95% CI 0.062 to 0.184). Taken across the cohort, activity, height velocity and body mass index together explained 9.7% of the variance in the change in basal growth hormone, rising to 16.3% once the change in insulin-like growth factor 1 was added. The explanation is physiological: growth hormone is secreted in pulses, so a single basal sample captures one point on a fluctuating profile, whereas insulin-like growth factor 1 is stable across the day and integrates the activity of the axis over weeks. Both measures confirm that the axis responded, but insulin-like growth factor 1 measures that response with less noise, and it is therefore the more suitable marker for monitoring a non-pharmacological programme.

These findings support the position that short stature of constitutional or nutritional origin responds to

change in nutrition, activity and sleep, and that the response is measurable through the growth hormone and insulin-like growth factor 1 axis without recourse to hormonal treatment. That is a different clinical situation from established growth hormone deficiency, in which replacement therapy is the only effective option, and the distinction matters for how referral pathways are designed.

The pattern is consistent with the wider literature. Exercise stimulates growth hormone secretion (Birzniece, 2019), and sustained exercise raises circulating insulin-like growth factor 1 (Kwon et al., 2025). Intervention studies in children have nonetheless found repeatedly that basal or morning growth hormone does not move even when height and insulin-like growth factor 1 improve: a 24-week jumping programme in children with short stature raised height and insulin-like growth factor 1 while leaving basal growth hormone unchanged (Wang et al., 2025), and among 250 children, insulin-like growth factor 1 was higher in the physically active group while growth hormone showed no difference (Uvaidillaeva, 2025). The present cohort reproduces that dissociation.

Two limitations bear directly on interpretation. The cohort was observed rather than randomised, so secular growth, regression to the mean and family-level confounding cannot be separated from the effect of the programme. The dose-response gradient across adherence tertiles is suggestive rather than decisive, since adherence is itself associated with family circumstances that independently affect growth.

7. Limitations and Future Research Directions

Several limitations qualify these findings. The review relied on published literature and may under-represent unpublished and grey literature, particularly from low- and middle-income settings where publication rates are lower. Restriction to English-language publications is likely to have compounded this under-representation in precisely the regions the framework is intended to serve.

Heterogeneity across included studies in setting, population, instrument version, cut-off criterion and reference standard was substantial, and for several outcomes precluded quantitative pooling. Diagnostic accuracy estimates are particularly vulnerable to partial verification bias, since studies rarely administer the reference standard to screen-negative children; reported sensitivity in this literature should therefore be regarded as an upper bound. The effectiveness claims made for this framework should therefore be read as indicative rather than precise.

The effect sizes reported for early intervention derive predominantly from parenting and developmental interventions delivered under trial conditions. Effects achieved in routine service delivery are consistently smaller, and the framework should not be expected to reproduce trial-level effects at scale.

Framework validation relied on comparative analysis of existing literature rather than prospective implementation trials, which limits conclusions about

real-world effectiveness. The operational specification in Section 5 is internally coherent and evidence-aligned but untested; its central claim, that specifying and measuring referral-loop closure improves service receipt, is falsifiable and constitutes the primary hypothesis for prospective evaluation.

Cultural adaptation requirements for framework implementation across diverse settings require further investigation. While the analysis identified cultural and attitudinal barriers, detailed exploration of successful adaptation strategies remains limited and represents an important area for future research.

Priority directions for future work are: (1) prospective evaluation of framework implementation across contrasting healthcare settings using the stepped-wedge design specified in Section 5.4, with referral-loop closure as the primary endpoint; (2) cost-effectiveness analysis of framework implementation versus usual care across different economic settings; (3) longitudinal studies tracking child outcomes under framework-based versus usual care; (4) development and external validation of the AI-CDSS prediction engine against TRIPOD+AI (Collins et al., 2024) and PROBAST (Wolff et al., 2019) standards, including pre-specified fairness assessment; and (5) formal cultural adaptation studies establishing which framework elements are core and which are locally modifiable.

8. Conclusion

This study sets out an actionable strategy for improving pediatric healthcare quality through systematic early screening and structured intervention, integrating five components: validated screening, structured intervention pathways, stakeholder engagement, standards-based technology integration and capacity building. The strategy is grounded in synthesised evidence across diverse healthcare settings and is directed at improving outcomes and equity in pursuit of the health targets of SDG 3.

The synthesised evidence supports a specific and comparatively modest claim. Structured early intervention delivered in the first three years improves developmental outcomes by approximately 0.2 to 0.4 standard deviations depending on domain and setting, with larger effects in low- and middle-income contexts and smaller effects when delivered through routine services. Effects of this magnitude are meaningful at population scale but do not support claims of transformation at individual level, and the certainty of the underlying evidence is low to moderate.

Abbreviations

AAP	American Academy of Pediatrics
AI-CDSS	Artificial-intelligence-enabled clinical decision support system
ASQ-3	Ages and Stages Questionnaires, Third Edition
AUROC	Area under the receiver operating characteristic curve

Coverage remains far below professional recommendation even in high-income systems, with US developmental screening reaching roughly three in ten eligible children and varying more than threefold between states, while national programme architecture elsewhere has achieved coverage above 75%. The determining factor is how services are organised, financed and reimbursed rather than national income alone, which is why the framework treats programme architecture as its primary object.

The most consequential finding of this synthesis concerns what happens after a positive screen. Where fewer than half of positively screened children receive any recommended referral, and where state-level screening rates show no relationship with service receipt, additional investment in screening coverage or instrument accuracy yields limited return. The framework therefore gives the referral loop an owner, a tracking mechanism and an escalation rule, and reports closure as an outcome. That is where it departs from earlier models.

Technology contributes chiefly by reducing latency and extending reach rather than by improving diagnostic accuracy, with remote assessment reducing waiting from months to days at diagnostic agreement of 80–94%. Machine-learning approaches show prospective discrimination around 0.90 in well-conducted studies but retain positive predictive values comparable to existing questionnaires and lose performance on external validation; they are accordingly specified with mandatory external validation and fairness assessment before deployment. Sustained commitment from multiple stakeholders alongside diversified financing will be essential to achieving the intended outcomes. Because the framework rests on multiple theoretical foundations and on synthesised empirical evidence, management approaches will need to permit context-specific adaptation without altering its core elements. The next stage of this work is prospective implementation evaluation and external validation of the decision-support component.

This research adds to the evidence for integrated, systems-based approaches to improving pediatric healthcare. By establishing an integrated framework from a systematic assessment of the literature, it provides actionable recommendations for healthcare staff, policymakers and implementing bodies concerning the improvement of pediatric healthcare quality on the path towards health equity.

CFIR	Consolidated Framework for Implementation Research
CHW	Community health worker
CI	Confidence interval
FHIR	Fast Healthcare Interoperability Resources
HL7	Health Level Seven International
LMIC	Low- and middle-income country

LOINC	Logical Observation Identifiers Names and Codes	PROBAST	Prediction model Risk Of Bias ASsessment Tool
M-CHAT-R/F	Modified Checklist for Autism in Toddlers, Revised with Follow-Up	RACI	Responsible, Accountable, Consulted, Informed
NASSS	Non-adoption, Abandonment, Scale-up, Spread and Sustainability	RCT	Randomised controlled trial
NPV	Negative predictive value	RE-AIM	Reach, Effectiveness, Adoption, Implementation, Maintenance
PICOS	Population, Intervention, Comparator, Outcomes, Study designs	SDG	Sustainable Development Goal
PPV	Positive predictive value	TRIPOD+AI	Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis, artificial intelligence extension
		TSH	Thyroid-stimulating hormone

CRedit authorship contribution statement

Hend R. Tawfik: Conceptualization, Methodology, Investigation, Writing – original draft. Ahmed F. Elnokrashy: Methodology, Software, Visualization, Writing – review & editing. Samir Y. Marzouk: Formal analysis, Validation, Writing – review & editing. Mai S. Mabrouk: Supervision, Formal analysis, Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethics approval

The review component of this study drew on published literature and did not require ethical approval. The clinical cohort was based on records generated during routine clinical care in a licensed laboratory and clinic setting. All records were de-identified before analysis, no additional procedures were performed for research purposes, and the work was conducted in accordance with the principles of the Declaration of Helsinki.

Data availability

The literature synthesised in this study is published and is cited in the reference list. The individual-level data supporting the clinical cohort are not publicly available because they derive from clinical records, but a de-identified analysis dataset is available from the corresponding author on reasonable request and subject to the applicable institutional data-sharing conditions.

Funding

No funding was received to assist with the preparation of this manuscript.

References

- Abdelaziz, A. S., Marzouk, S. Y., Mabrouk, M. S., 2026. Medical laboratory accreditation programs: a comparative review of international and national frameworks. *eJIFCC* vol. 37 (5), 960-971.
- Agarwal, P. K., Xie, H., Sathyapalan Rema, A. S., Rajadurai, V. S., Lim, S. B., Meaney, M., Daniel, L. M., 2023. Concurrent validity of the Ages and Stages Questionnaires with Bayley Scales of Infant Development-III at 2 years – Singapore cohort study. *Pediatrics and Neonatology* vol. 65 (3), 236–242.

- Aishworiya, R., Ma, V. K., Stewart, S., Hagerman, R., Feldman, H. M., 2023. Meta-analysis of the Modified Checklist for Autism in Toddlers, Revised/Follow-up for screening. *Pediatrics* vol. 151 (6), e2022059393.
- Baek, Y., Tran, T., Fisher, J., Ademi, Z., 2025. Economic returns on investing in early childhood development in Vietnam: A cost-benefit analysis. *BMC Health Services Research* vol. 25, 1–12.
- Bailey, D. B., 2026. Supporting family engagement in early intervention. *Pediatrics*. Advance online publication.
- Barger, B., Rice, C., Wolf, R., Roach, A., 2018. Better together: Developmental screening and monitoring best identify children who need early intervention. *Disability and Health Journal* vol. 11 (3), 420–426.
- Bauer, A., Bosl, W., Aalami, O., Schmiedmayer, P., 2025. Toward scalable access to neurodevelopmental screening: Insights, implementation, and challenges. *arXiv:2503.13472*.
- Bethell, C., Reuland, C., Schor, E., Abrahms, M., Halfon, N., 2011. Rates of parent-centered developmental screening: Disparities and links to services access. *Pediatrics* vol. 128 (1), 146–155.
- Birzniece, V., 2019. Exercise and the growth hormone-insulin-like growth factor axis. *Current Opinion in Endocrine and Metabolic Research* vol. 9, 1-6.
- Black, M. M., Walker, S. P., Fernald, L. C. H., Andersen, C. T., DiGirolamo, A. M., Lu, C., ..., Grantham-McGregor, S., 2017. Early childhood development coming of age: Science through the life course. *The Lancet* vol. 389 (10064), 77–90.
- Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* vol. 3 (2), 77–101.
- Brundtland Commission, 1987. Our common future: report of the World Commission on Environment and Development. Oxford University Press, Oxford.
- Campbell, D. T., Stanley, J. C., 1963. Experimental and quasi-experimental designs for research. Houghton Mifflin.
- Chen, Y.-H., Chen, Q., Kong, L., Liu, G., 2022. Early detection of autism spectrum disorder in young

- children with machine learning using medical claims data. *BMJ Health & Care Informatics* vol. 29 (1), e100544.
- Chiliza, J., Laing, R., Feeley, F. G., Borba, C. P. C., 2021. Program sustainability post PEPFAR direct service support in the Western Cape, South Africa. *PLoS ONE* vol. 16 (5), e0251230.
- Coales, K., Marsden, D., Merzoug, K., 2023. Perspectives of health workers engaging in task shifting to deliver health care in low- and middle-income countries: A qualitative evidence synthesis. *Global Health Action* vol. 16 (1), 2228112.
- Cohen, J., 1988. *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., ..., Logullo, P., 2024. TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* vol. 385, e078378.
- Corona, L. L., Weitlauf, A. S., Hine, J., Berman, A., Miceli, A., Nicholson, A., ..., Warren, Z., 2023. A randomized trial of the accuracy of novel telehealth instruments for the assessment of autism in toddlers. *Journal of Autism and Developmental Disorders* vol. 53 (9), 3618–3629.
- Cowie, J., Nicoll, A., Dimova, E. D., Campbell, P., Duncan, E. A., 2020. The barriers and facilitators influencing the sustainability of hospital-based interventions: A systematic review. *BMC Health Services Research* vol. 20 (1), 588.
- Cronbach, L. J., Meehl, P. E., 1955. Construct validity in psychological tests. *Psychological Bulletin* vol. 52 (4), 281–302.
- Cruz Romero, S. A., Mercado Hernandez, M. J., Ali Rivera, S. Y., Santiago Fernandez, J. A., Lugo Beauchamp, W. E., 2025. Design of an edge-based portable EHR system for anemia screening in remote health applications. *arXiv:2507.15146*.
- Dababnah, S., Shaia, W. E., Champion, K., Nichols, H. M., 2018. “We had to keep pushing”: Caregivers’ perspectives on autism screening and referral practices of Black children in primary care. *Intellectual and Developmental Disabilities* vol. 56 (5), 321–336.
- Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., Lowery, J. C., 2009. Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science* vol. 4, 50.
- Early, J., Gonzalez, C., Gordon-Dseagu, V., Robles-Calderon, L., 2019. Use of mobile health (mHealth) technologies and interventions among community health workers globally: A systematic review. *Health Promotion Practice* vol. 20 (6), 805–817.
- Felix, G., Kuhn, M., Bloomfield, B., Bacon, E., Wallis, K. E., 2024. Standardizing and improving primary care-based electronic developmental screening for young children in federally qualified health center clinics. *Maternal and Child Health Journal* vol. 28 (6), 1067–1076.
- Feroz, A., Jabeen, R., Saleem, S., 2020. Using mobile phones to improve community health workers performance in low- and middle-income countries. *BMC Public Health* vol. 20, 49.
- Fuchs, K., Eksteen, S., Swanepoel, D. W., 2023. Combined screening of early childhood development, hearing and vision by community health workers using mHealth tools in a low-income community. *Early Child Development and Care* vol. 193 (11–12), 1273–1285.
- Gangi, D. N., Corona, L. L., Weitlauf, A. S., Hine, J., Nicholson, A., Juarez, A. P., ..., Warren, Z., 2025. In-home tele-assessment for autism in toddlers: Validity, reliability, and caregiver satisfaction with the TELE-ASD-PEDS. *Journal of Developmental and Behavioral Pediatrics* vol. 46 (2), e112–e120.
- García, J. L., Heckman, J. J., Leaf, D. E., Prados, M. J., 2020. Quantifying the life-cycle benefits of an influential early-childhood program. *Journal of Political Economy* vol. 128 (7), 2502–2541.
- Ghafehazai, S., Carnett, A., Neely, L., Das, A., Rad, P., 2021. AI-augmented behavior analysis for children with developmental disabilities: Building towards precision treatment. *arXiv:2102.10635*.
- Glasgow, R. E., Vogt, T. M., Boles, S. M., 1999. Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *American Journal of Public Health* vol. 89 (9), 1322–1327.
- Greenhalgh, T., Wherton, J., Papoutsis, C., Lynch, J., Hughes, G., A’Court, C., ..., Shaw, S., 2017. Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *Journal of Medical Internet Research* vol. 19 (11), e367.
- Gómez-Cotilla, R., Herruzo-Cabrera, J., Trenados-Barrios, Y., Pino-Osuna, M. J., 2024. Efficacy of early intervention programmes: Systematic review and meta-analysis. *Psicología Educativa* vol. 30 (1), 45–56.
- Hirai, A. H., Kogan, M. D., Kandasamy, V., Reuland, C., Bethell, C., 2018. Prevalence and variation of developmental screening and surveillance in early childhood. *JAMA Pediatrics* vol. 172 (9), 857–866.
- Hirve, R., Adams, C., Kelly, C. B., McAullay, D., Hurt, L., Edmond, K. M., Strobel, N., 2023. Effect of early childhood development interventions delivered by healthcare providers to improve cognitive outcomes in children at 0–36 months: A systematic review and meta-

- analysis. *Archives of Disease in Childhood* vol. 108 (4), 247–257.
- Ilesanmi, O. S., Bello, A. E., Afolabi, A. A., 2022. Sustainability of donor-funded health-related programs beyond the funding lifecycle in Africa: A systematic review. *Cureus* vol. 14 (5), e24643.
- Jeong, J., Franchett, E. E., Ramos de Oliveira, C. V., Rehmani, K., Yousafzai, A. K., 2021. Parenting interventions to promote early child development in the first three years of life: A global systematic review and meta-analysis. *PLoS Medicine* vol. 18 (5), e1003602.
- Katakis, P., Schlieff, M., Barnett, P., Rains, L. S., Rowe, S., Pilling, S., Johnson, S., 2023. Diagnostic assessment of autism in children using telehealth in a global context: A systematic review. *Review Journal of Autism and Developmental Disorders* vol. 10, 1–19.
- Kim, J. S., 2026. Trends in participation and abnormal screening results in the National Health Screening Program for Infants and Children in South Korea, 2010–2023. *Kosin Medical Journal*. Advance online publication.
- Korom, R., Kiptinness, S., Adan, N., Ahmed, K., Alam, N., Alam, S., ..., Singhal, K., 2025. AI-based clinical decision support for primary care: A real-world study. *arXiv:2507.16947*.
- Kruk, M. E., Gage, A. D., Arsenaault, C., Jordan, K., Leslie, H. H., Roder-DeWan, S., ..., Pate, M., 2018. High-quality health systems in the Sustainable Development Goals era: Time for a revolution. *The Lancet Global Health* vol. 6 (11), e1196–e1252.
- Kuo, D. Z., Houtrow, A. J., Arango, P., Kuhlthau, K. A., Simmons, J. M., Neff, J. M., 2012. Family-centered care: Current applications and future directions in pediatric health care. *Maternal and Child Health Journal* vol. 16 (2), 297–305.
- Kwon, J., Kim, S., Lee, D., 2025. Effects of exercise interventions on IGF-1 in healthy populations: a systematic review and meta-analysis. *Physical Activity and Nutrition* vol. 29 (1), 1–11.
- Lindrooth, R. C., Perrillon, M. C., Hardy, R., Rosenberg, S. A., 2025. Does family-centered care coordination improve outcomes for children enrolled in early intervention services? An analysis of Colorado's GO4IT initiative. *Health Services Research* vol. 60 (1), e14392.
- Lopez Garcia, I., Saya, U. Y., Luoto, J. E., 2021. Cost-effectiveness and economic returns of group-based parenting interventions to promote early childhood development: Results from a randomized controlled trial in rural Kenya. *PLoS Medicine* vol. 18 (9), e1003746.
- Marsh, K., Coulter, K., Baker, C., Fowler, K., Kwon, S., Warren, Z., 2026. Reliability, accuracy, and timeliness of autism telehealth evaluations by pediatricians. *Telemedicine and e-Health*. Advance online publication.
- McCarthy, E., Guerin, S., 2021. Family-centred care in early intervention: A systematic review of the processes and outcomes of family-centred care and impacting factors. *Child: Care, Health and Development*, 48(1), 1–32.
- Morelli, D. L., Pati, S., Butler, A., Blum, N. J., Gerdes, M., Pinto-Martin, J., Guevara, J. P., 2014. Challenges to implementation of developmental screening in urban primary care: A mixed methods study. *BMC Pediatrics* vol. 14, 16.
- Mundeva, H., Snyder, J., Ngilangwa, D. P., Kaida, A., 2018. Ethics of task shifting in the health workforce: Exploring the role of community health workers in HIV service delivery in low- and middle-income countries. *BMC Medical Ethics* vol. 19, 71.
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., et al., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* vol. 372, n71.
- Perera, P. N., Dayaratne, T. T., Rannan-Eliya, R. P., Wickramasinghe, R., 2024. Factors impacting sustained coverage in the context of donor transitions: Experience from Sri Lanka. *Health Policy and Planning*, 39(Suppl 1), i65–i77.
- Perochon, S., Di Martino, J. M., Carpenter, K. L. H., Compton, S., Davis, N., Eichner, B., ..., Dawson, G., 2023. Early detection of autism using digital behavioral phenotyping. *Nature Medicine* vol. 29 (10), 2489–2497.
- Peterson-Katz, K., Piccone, G., Tuzi, A., Kulkarni, T., Reynolds, J. N., 2023. Barriers and facilitators to implementation of developmental screening and early intervention in Canadian organizations following completion of a training and coaching model: A thematic analysis. *Frontiers in Health Services* vol. 3, 1155568.
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., ..., Duffy, S., 2006. Guidance on the conduct of narrative synthesis in systematic reviews. *ESRC Methods Programme*.
- Radomski, A. D., Bergmann, C., Alfaro, A., Chan, E., Newton, A. S., 2025. Brief multidimensional screening tools for young children's mental health and development for administration by primary care providers: A systematic review. *BMC Primary Care* vol. 26, 88.
- Rajagopalan, S. S., Zhang, Y., Yahia, A., Tammimies, K., 2024. Machine learning prediction of autism spectrum disorder from a minimal set of medical and background information. *JAMA Network Open* vol. 7 (8), e2429229.
- Reynolds, A. J., Temple, J. A., White, B. A. B., Ou, S.-R., Robertson, D. L., 2011. Age-26 cost-benefit analysis of the Child-Parent Center early education program. *Child Development* vol. 82 (1), 379–404.
- Richter, L. M., Daelmans, B., Lombardi, J., Heymann, J., Boo, F. L., Behrman, J. R., ..., Darmstadt, G. L., 2017. Investing in the foundation of sustainable development: Pathways to scale up

- for early childhood development. *The Lancet* vol. 389 (10064), 103–118.
- Robins, D. L., Casagrande, K., Barton, M., Chen, C. M. A., Dumont-Mathieu, T., Fein, D., 2014. Validation of the Modified Checklist for Autism in Toddlers, Revised with Follow-up (M-CHAT-R/F). *Pediatrics* vol. 133 (1), 37–45.
- Rosenstock, I. M., 1974. The health belief model and preventive health behavior. *Health Education Monographs* vol. 2 (4), 354–386.
- Sandbank, M., Bottema-Beutel, K., Crowley LaPoint, S., Feldman, J. I., Barrett, D. J., Caldwell, N., ..., Woynaroski, T., 2023. Autism intervention meta-analysis of early childhood studies (Project AIM): Updated systematic review and secondary analysis. *BMJ* vol. 383, e076733.
- Schonhaut, L., Armijo, I., Schonstedt, M., Alvarez, J., Cordero, M., 2013. Validity of the Ages and Stages Questionnaires in term and preterm infants. *Pediatrics* vol. 131 (5), e1468–e1474.
- Seidman, G., Atun, R., 2017. Does task shifting yield cost savings and improve efficiency for health systems? A systematic review of evidence from low-income and middle-income countries. *Human Resources for Health* vol. 15, 29.
- Senge, P. M., 1990. *The fifth discipline: The art and practice of the learning organization*. Doubleday.
- Sheldrick, R. C., Marakovitz, S., Garfinkel, D., Carter, A. S., Perrin, E. C., 2020. Comparative accuracy of developmental screening questionnaires. *JAMA Pediatrics* vol. 174 (4), 366–374.
- Singh, A., Yeh, C. J., Blanchard, S. B., 2017. Ages and stages questionnaire: A global screening scale. *Boletín Médico del Hospital Infantil de México (English Edition)* vol. 74 (1), 5–12.
- Small, J. W., Hix-Small, H., Vargas-Baron, E., Marks, K. P., 2018. Comparative use of the Ages and Stages Questionnaires in low- and middle-income countries. *Developmental Medicine & Child Neurology* vol. 61 (4), 431–443.
- Smythe, T., Scherer, N., Nanyunja, C., Tann, C. J., Olusanya, B. O., 2024. Strategies for addressing the needs of children with or at risk of developmental disabilities in early childhood by 2030: A systematic umbrella review. *BMC Medicine* vol. 22, 51.
- Squires, J., Bricker, D., 2009. *Ages & stages questionnaires: A parent-completed child monitoring system* (3rd ed.). Paul H. Brookes Publishing.
- Srinath, A., van Merode, F., Rao, S. V., Pavlova, M., 2022. Barriers to cervical cancer and breast cancer screening uptake in low- and middle-income countries: A systematic review. *Health Policy and Planning* vol. 38 (4), 509–527.
- Strobel, N. A., Adams, C., McAullay, D. R., Edmond, K. M., 2022. Family-centred interventions for Indigenous early childhood well-being by primary healthcare services. *Cochrane Database of Systematic Reviews* vol. 12, CD012463.
- Traube, D. E., Reynolds, S. M., Cederbaum, J. A., Sinatra, R., 2021. Relationship between state-level developmental screening and IDEA Part C early intervention rates. *Journal of Early Intervention* vol. 43 (4), 297–310.
- Uvaidillaeva, N., 2025. Physical activity and growth hormone axis markers in schoolchildren: a cross-sectional study. *Central Asian Journal of Medical Sciences* vol. 11 (2), 88–96.
- Vitrikas, K., Dubois, D., 2017. Developmental delay: When and how to screen. *American Family Physician* vol. 96 (1), 36–43.
- Wallis, K. E., Guthrie, W., Bennett, A. E., Gerdes, M., Levy, S. E., Mandell, D. S., Miller, J. S., 2020. Adherence to screening and referral guidelines for autism spectrum disorder in toddlers in pediatric primary care. *PLoS ONE* vol. 15 (5), e0232335.
- Wang, L., Chen, Y., Liu, H., 2025. Influence of a 24-week jumping exercise programme on growth speed and the GH-IGF-1-IGFBP-3 axis in children with short stature. *BMC Pediatrics* vol. 25, 1–11.
- Weitlauf, A. S., Miceli, A., Vehorn, A., Dada, Y., Pinkham, A., Ezell, K., Warren, Z., 2022. Black families' experiences of developmental screening: Review of well child visits to inform enhanced ASD risk assessment. *Journal of Developmental and Behavioral Pediatrics* vol. 43 (9), 503–510.
- Wolff, R. F., Moons, K. G. M., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., ..., Mallett, S., 2019. PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. *Annals of Internal Medicine* vol. 170 (1), 51–58.
- Wondmagegn, T., Girma, B., Habtemariam, Y., 2024. Prevalence and determinants of developmental delay among children in low- and middle-income countries: A systematic review and meta-analysis. *Frontiers in Public Health* vol. 12, 1301524.
- World Health Organization, United Nations Children's Fund, World Bank Group, 2018. *Nurturing care for early childhood development: A framework for helping children survive and thrive to transform health and human potential*. World Health Organization.
- Yue, A., Jiang, Q., Wang, B., Abbey, C., Medina, A., Shi, Y., Rozelle, S., 2019. Concurrent validity of the Ages and Stages Questionnaire and the Bayley Scales of Infant Development III in China. *PLoS ONE* vol. 14 (9), e0221675.
- Zhang, L., Ssewanyana, D., Martin, M.-C., Lye, S., Moran, G., Abubakar, A., ..., Malti, T., 2021. Supporting child development through parenting interventions in low- to middle-income countries: An updated systematic review. *Frontiers in Public Health* vol. 9, 671988.
- Zheng, R. M., Chan, Y. H., Aishworiya, R., Choo, C. C., Kang, Y. Q., Sung, M., ..., Kwan, Y. H.,

2023. Validity and feasibility of using the Modified Checklist for Autism in Toddlers, Revised with Follow-Up (M-CHAT-R/F) in primary care clinics in Singapore. *Autism* vol. 28 (4), 878–890.

von Bertalanffy, L., 1968. General system theory: Foundations, development, applications. George Braziller.