Cureus
Part of SPRINGER NATURE

Open Access Review Article

Gross Motor Development in Institutionalized Infants and Young Toddlers: A Systematic Review

Nahla Bakri Ageeb Babekr [1], Aya Alhassan [2], Samah Mohammed Osman Mohammed Nur [3], Suzan Elhireka Koko Gadoum [4, 5], Asim Ahmed [6, 7], Mariam Z. Mubarak [8]

1. Pediatrics, Sudan Medical Specialization Board, Khartoum, SDN 2. Paediatrics, University Hospitals Birmingham NHS Foundation Trust, Birmingham, GBR 3. Dermatology, Faculty of Medicine, Al-Mughtaribeen University, Khartoum, SDN 4. School of Medicine, Ahfad University For Women, Khartoum, SDN 5. Pediatrics, Saudi Commission for Health Specialties, Dammam, SAU 6. Medicine and Surgery, University of Gezira, Wad Madani, SDN 7. Scientific Research and Editorial Affairs, SciMENA LLC, Sheridan, USA 8. Pediatric Infectious Diseases, Maternal and Child Hospital (MCH), Hail, SAU

Corresponding author: Aya Alhassan, aya.alhassan2@nhs.net

Review began 10/10/2026
Review ended 11/05/2026
Published 11/06/2026

© Copyright 2026
Babekr et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.119234

Abstract

Infants and young toddlers living in institutional care may experience limited opportunities for individualized interaction and age-appropriate movement. However, the extent of gross motor difficulties remains uncertain because available studies differ in design, assessment methods, and care settings. This systematic review examined gross motor development among children aged 0-24 months currently residing in orphanages or comparable non-family residential childcare settings. On 3 September 2026, two bibliographic databases (PubMed/MEDLINE and Scopus) were searched, supplemented by Google Scholar and the AI-assisted evidence-discovery platform Elicit. Seven peer-reviewed studies were included, comprising predominantly observational research and one non-randomized intervention study conducted across Thailand, Brazil, Sudan, and North Macedonia. The studies assessed gross motor performance, postural development, locomotion, and trunk control. Findings were synthesized narratively because clinical and methodological heterogeneity prevented meaningful meta-analysis. Common findings were low or atypical gross motor performance, postural delay, locomotor disadvantage, and trunk-control differences, although results varied across assessment instruments and outcome definitions. The intervention study reported greater improvement in motor performance following structured upright play, but non-randomized site allocation and methodological limitations prevented firm conclusions regarding effectiveness. Confidence in the findings was further limited by risk of bias, small samples, incomplete retrieval, and limited comparative evidence. Overall, the evidence suggests potential gross motor vulnerability among children in institutional care but does not establish a reliable prevalence estimate or a causal relationship with institutional residence. Stronger multicentre comparative and intervention research using standardized developmental assessments is needed.

Categories: Neurology, Pediatrics, Epidemiology/Public Health
Keywords: gross motor development, institutionalized infants, motor development, orphanages, systematic review

Introduction And Background

The first two years of life are characterized by rapid neurological, musculoskeletal, and functional development. During this period, infants progressively acquire essential motor skills, including head control, rolling, sitting, standing, and independent mobility. These skills support environmental exploration, caregiver interaction, and participation in daily activities. Although biological maturation provides the foundation for motor development, progression is also influenced by nutrition, general health, opportunities for spontaneous movement, responsive caregiving, and age-appropriate play [1-3]. As posture and mobility develop, children gain further opportunities for exploration, social interaction, and learning.

The caregiving environment may influence early motor development by providing opportunities for movement and exploration. Safe floor play, varied positioning, and caregiver-supported activities allow infants to practice emerging motor skills. Limited individualized interaction or restricted spontaneous movement may reduce these opportunities. Conversely, responsive caregiving and developmentally appropriate stimulation may encourage exploration and support postural control. The early caregiving environment should therefore be considered when interpreting gross motor performance, particularly among children with pre-existing biological or social vulnerabilities [1-3].

Infants living in orphanages and other institutional childcare settings may experience caregiving conditions that differ from those in family-based care. Some institutions have group-based routines, high child-to-caregiver ratios, rotating caregivers, and limited opportunities for individualized play. However, these conditions vary across institutions and should not be assumed to characterize every child's experience. Previous research has associated institutional care with developmental disadvantages while documenting substantial variation between children and care environments [4-6]. In this review, current institutional residence represents the exposure of interest. Biological and health-related factors, including prematurity,

low birth weight, nutritional status, and underlying medical conditions, may contribute to confounding in comparisons of developmental outcomes. Age at placement, duration of institutionalization, caregiving quality, and access to stimulation further characterize individual experiences and may influence motor performance. Consequently, observed developmental differences cannot be attributed to institutional residence alone without considering these factors.

Gross motor development is particularly relevant to developmental assessment because early motor performance can be observed directly and monitored over time. Atypical performance may involve reduced postural control, delayed acquisition of motor milestones, difficulty with locomotion, or low scores on age-referenced developmental assessments. Identifying such findings may guide developmental surveillance and indicate the need for further evaluation or supportive care. However, a single low assessment score does not necessarily establish a clinical diagnosis of developmental delay. Interpretation depends on the assessment instrument, the child's developmental age, the threshold applied, and whether findings are observed repeatedly. Biological and health-related factors, including prematurity, low birth weight, illness, and malnutrition, should also be considered when interpreting assessment results.

The Alberta Infant Motor Scale (AIMS) is a widely used observational measure of early gross motor development [7]. It assesses spontaneous motor performance, weight bearing, and postural control across four positions, providing age-referenced percentile information. Other developmental instruments assess related but distinct constructs, including locomotion, postural development, and trunk control. Consequently, results obtained using different instruments cannot automatically be considered equivalent. Even among studies using AIMS, findings may be reported as raw scores, percentiles, or study-defined delay classifications. Differences in participant age, assessment timing, and institutional context further limit direct comparisons across studies.

Previous literature, including a systematic and integrative review of institutional care, has examined developmental outcomes across multiple domains, encompassing physical, cognitive, emotional, and broader developmental functioning [4-6]. Although this literature provides valuable evidence about institutional rearing, it does not fully address the specific patterns and comparability of gross motor outcomes among children currently residing in institutional care during the first 24 months of life. This age range is clinically important because fundamental postural and locomotor milestones emerge rapidly, making age-appropriate assessment particularly relevant to identifying potential developmental difficulties. The present review therefore focuses on separately identifiable gross motor, postural, locomotor, and trunk-control outcomes rather than combining them with broader developmental measures. Specifically, it examines the reported patterns of low or atypical gross motor performance, the available comparative and intervention evidence, the methodological limitations and risk of bias of the included studies, and whether the findings are sufficiently comparable for meaningful quantitative synthesis.

Review

Materials and methods

Study Design, Reporting Framework, and Registration

This systematic review was reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement and its explanation and elaboration guidance [8,9]. Search reporting was organized with reference to PRISMA-S [10], and the search strategy was developed and audited using principles reflected in Peer Review of Electronic Search Strategies (PRESS) guidance [11]. The review was retrospectively registered in the SciMENA SciPro registry as SciPro-00000234 on 7 September 2026. An audited Version 2.0 update was subsequently submitted to the registry; at manuscript finalization, the original registration remained the current public record pending registry confirmation of the update.

The review commenced on 15 August 2026, and its core searches were completed on 3 September 2026, before the 7 September registration. No specific reason for the timing of registration was recorded.

Review Framework

The review question was structured around population, exposure or intervention, comparator, and outcome components. Studies enrolling broader age ranges were eligible only when data for children aged 0-24 months could be extracted separately. When studies reported multiple developmental domains, only separately identifiable eligible gross motor, postural, locomotor, motor-milestone, or trunk-control outcomes were considered. The review framework is summarized in Table *1*.

Component	Definition
Population	Infants and young toddlers aged 0-24 months currently living in orphanages, children's homes, institutional childcare, residential childcare, or comparable non-family residential settings.
Exposure / intervention	Primary exposure: current institutional residence. Eligible intervention studies included structured movement, upright play, massage, or comparable developmental stimulation delivered within institutional care.
Comparator	No comparator was required. Eligible comparators included family-, foster-, adoptive-, or community-reared children and usual institutional care for intervention studies.
Outcomes	Separately identifiable gross motor, postural, locomotor, motor-milestone, or trunk-control outcomes measured with standardized or validated developmental instruments. Baseline and post-intervention outcomes were kept analytically separate.

TABLE 1: Review framework

AIMS, Alberta Infant Motor Scale [7]. The framework allowed single-group institutional studies as well as comparative and intervention designs; a comparator was not mandatory.

Eligibility Criteria

Eligibility was assessed during title and abstract screening and again at full-text assessment. Studies were not excluded based on methodological quality or risk-of-bias assessment. Risk of bias was evaluated after inclusion and used to inform interpretation. The inclusion and exclusion criteria are summarized in Table *2*.

Criterion	Decision	Operational rule
Population	Include	Current residents of institutional or non-family residential childcare within the eligible age range.
Age	Include	Participants aged 0-24 months, or separately extractable data for that age range from a broader sample.
Age	Exclude	Broader age samples without separately extractable data for participants aged 0-24 months.
Setting	Include	Orphanages, children's homes, institutional childcare centres, residential childcare institutions, or comparable non-family residential care.
Setting	Exclude	Hospital, rehabilitation, family-home, post-adoption, foster-placement, or other settings that did not represent current institutional residence.
Outcome	Include	Separately identifiable gross motor, postural, locomotor, motor-milestone, or trunk-control outcome.
Outcome	Exclude	No eligible gross-motor outcome, including studies limited to anthropometry, cognition, language, social, behavioural, emotional, or fine-motor outcomes.
Outcome extractability	Exclude	Combined or overall developmental/motor scores when a gross-motor-specific result could not be separately extracted.
Assessment	Include	Standardized or validated developmental assessment with an eligible motor/postural domain.
Design	Include	Cross-sectional, cohort, longitudinal observational, case-control, comparative observational, randomized controlled, non-randomized, or quasi-experimental studies with eligible outcome data.
Publication/report type	Exclude	Reviews, meta-analyses, editorials, commentaries, case reports, protocols without completed results, duplicate/overlapping reports, and reports without sufficient original extractable data.
Conference/grey literature	Conditional	Conference abstracts or grey literature were logged during the evidence audit but were not included in the primary synthesis unless a sufficiently complete report supported full eligibility assessment and extractable outcome data.
Condition-specific populations	Exclude	Studies focused exclusively on major neurological, genetic, congenital, or neuromuscular disorders unless eligible institutionalized participants and gross-motor data were separately analyzable.

TABLE 2: Inclusion and exclusion criteria

Eligibility decisions considered population, age, current institutional setting, study design, report type, and whether an eligible motor outcome could be extracted separately. This outcome requirement ensured that combined developmental measures were not interpreted as direct gross motor findings when no motor-specific result was available. High risk of bias was not an exclusion criterion.

Information Sources and Search Strategy

The main search was conducted on 3 September 2026 using two bibliographic databases, PubMed/MEDLINE and Scopus, supplemented by Google Scholar and the AI-assisted discovery platform Elicit. The database searches covered records from inception without publication-date, language, geographical, publication-type, subject-area, or source filters. Google Scholar was searched in relevance order, and the first 100 results were selected for screening, which was the planned screening limit. This restriction may have missed eligible studies ranked beyond the first 100 results. Elicit was used to identify additional candidate studies for bibliographic and eligibility checking; its generated summaries were not used as article abstracts or primary evidence. The search syntax was adapted to each source: PubMed used MeSH and title/abstract terms, Scopus used TITLE-ABS-KEY fields, and Google Scholar and Elicit used source-appropriate queries. ASTA and ALEX were planned supplementary sources but contributed no imported records. Reference-list checking, citation searching, targeted web searches, and repository auditing continued through 7 September 2026. The formal PRISMA accounting covers the 652 records entering the documented screening workflow; other audit leads remained separately documented. The exact core search strategies are presented in Table *3*.

Source	Search string / query	Core search date
PubMed/MEDLINE	("Child, Institutionalized"[Mesh] OR "Orphanages"[Mesh] OR institutionalized[Title/Abstract] OR institutionalised[Title/Abstract] OR "institution-reared"[Title/Abstract] OR "institutionally reared"[Title/Abstract] OR orphanage*[Title/Abstract] OR orphaned[Title/Abstract] OR orphans[Title/Abstract] OR "institutional care"[Title/Abstract] OR "residential care"[Title/Abstract] OR "residential childcare"[Title/Abstract] OR "residential child care"[Title/Abstract] OR "residential institution"[Title/Abstract] OR "residential institutions"[Title/Abstract] OR "child care institution"[Title/Abstract] OR "child care institutions"[Title/Abstract] OR "children's home"[Title/Abstract] OR "children's homes"[Title/Abstract] OR "baby home"[Title/Abstract] OR "baby homes"[Title/Abstract]) AND ("Infant"[Mesh] OR "Infant, Newborn"[Mesh] OR "Child, Preschool"[Mesh] OR infant*[Title/Abstract] OR baby[Title/Abstract] OR babies[Title/Abstract] OR toddler*[Title/Abstract] OR child*[Title/Abstract]) AND ("Motor Skills"[Mesh] OR "Psychomotor Performance"[Mesh] OR "gross motor"[Title/Abstract] OR "motor development"[Title/Abstract] OR "motor developmental"[Title/Abstract] OR "motor performance"[Title/Abstract] OR "motor function"[Title/Abstract] OR "motor functions"[Title/Abstract] OR "motor skill"[Title/Abstract] OR "motor skills"[Title/Abstract] OR "motor delay"[Title/Abstract] OR "motor delays"[Title/Abstract] OR "motor milestone"[Title/Abstract] OR "motor milestones"[Title/Abstract] OR "psychomotor development"[Title/Abstract] OR "psychomotor developmental"[Title/Abstract] OR neuromotor[Title/Abstract] OR locomotor[Title/Abstract] OR "Alberta Infant Motor Scale"[Title/Abstract] OR Bayley[Title/Abstract] OR "Peabody Developmental Motor Scale"[Title/Abstract] OR "Peabody Developmental Motor Scales"[Title/Abstract] OR PDMS[Title/Abstract] OR Griffiths[Title/Abstract])	3 Sep 2026
Scopus	TITLE-ABS-KEY (institutionalized OR institutionalised OR institutionalization OR institutionalisation OR "institution-reared" OR "institutionally reared" OR orphanage* OR orphaned OR orphans OR "institutional care" OR "residential care" OR "residential childcare" OR "residential child care" OR "residential institution" OR "residential institutions" OR "child care institution" OR "child care institutions" OR "children's home" OR "children's homes" OR "baby home" OR "baby homes") AND TITLE-ABS-KEY (infant* OR newborn* OR neonat* OR baby OR babies OR toddler* OR child*) AND TITLE-ABS-KEY ("gross motor" OR "motor development" OR "motor developmental" OR "motor performance" OR "motor function" OR "motor functions" OR "motor skill" OR "motor skills" OR "motor delay" OR "motor delays" OR "motor milestone" OR "motor milestones" OR "psychomotor development" OR "psychomotor developmental" OR neuromotor OR locomotor OR "Alberta Infant Motor Scale" OR Bayley OR "Peabody Developmental Motor Scale" OR "Peabody Developmental Motor Scales" OR PDMS OR Griffiths)	3 Sep 2026
Google Scholar	("gross motor" OR "motor development" OR "psychomotor development" OR "Alberta Infant Motor Scale") (institutionalized OR institutionalised OR orphanage OR "institutional care" OR "residential care") (infant OR toddler OR child)	3 Sep 2026
Elicit	What studies have assessed gross motor development in infants and young toddlers aged 0-24 months currently living in orphanages, institutional care, or residential childcare settings, using the Alberta Infant Motor Scale or another standardized gross motor assessment?	3 Sep 2026

TABLE 3: Exact core search strategies

Supplementary leads that did not complete the formal screening and full-text eligibility pathway remained in the audit log and were not counted as included evidence.

Record Management and Study Selection

All retrieved records were merged into a spreadsheet-based review database and deduplicated using PMID, DOI, other bibliographic identifiers, and normalized title matching, followed by verification of uncertain matches. Screening proceeded in two stages: title and abstract review, followed by full-text eligibility assessment. Study screening, full-text eligibility assessment, and data extraction were conducted by Nahla Bakri Ageeb Babekr, Aya Alhassan, and Asim Ahmed. The workflow was AI-assisted and source-audited; the title/abstract decisions were recorded as provisional in the review log, and potentially eligible or unclear records were advanced to retrieval. Record-level decisions and reasons for exclusion or non-retrieval were retained. Available full texts and bibliographic sources informed subsequent eligibility and source-verification checks; generated discovery summaries were not substituted for source reports. Reports removed during final bibliographic source verification were recorded separately from full-text eligibility exclusions. The audit does not document independent duplicate human screening, inter-reviewer agreement, or a formal adjudication process, and these procedures are not reported as completed.

Data Extraction

Data were extracted into a standardized, source-linked spreadsheet. Study-level variables included author, publication year, country, care setting, institutional-residence status, design, total and eligible sample sizes, participant age, sex, birth and nutritional characteristics, and duration of institutionalization when available. Comparator and intervention details were recorded separately. Outcome-level fields included the assessment instrument, motor construct or domain, assessment time point, and extractable outcome estimates. For dichotomous outcomes, numerators and denominators were retained; continuous and comparative findings were recorded as reported, including means, standard deviations, medians, ranges or interquartile ranges, effect measures, confidence intervals, and p values. Baseline data and post-intervention findings were kept analytically separate. The audit spreadsheet links extracted outcomes to available source files or URLs and, where supplied, tables or page locations, enabling comparison with the original reports. Extraction entries were reconciled against available sources during the final audit. The records do not establish a formal independent duplicate-extraction or discrepancy-adjudication procedure, and neither is reported as completed.

Risk-of-Bias Assessment

Risk of bias was assessed according to study design. The non-randomized quasi-experimental study and observational studies were appraised using design-appropriate Joanna Briggs Institute (JBI) critical-appraisal tools, including the quasi-experimental, prevalence, cohort, and analytical cross-sectional checklists as appropriate [12-14]. The study-level audit records the responses to applicable JBI questions (Yes, No, Unclear, or Not applicable), together with evidence-based rationales and source locations. For synthesis, appraisal findings were summarized across selection, confounding or allocation, missing data or follow-up, outcome measurement, and analysis or reporting domains. Detailed study-level critical-appraisal assessments are provided in the Appendix. Overall labels of low risk, some concerns, or high risk represented review-defined qualitative judgments based on the pattern and significance of domain-level concerns. These categories were not derived from JBI numeric scores or cutoffs. Risk-of-bias findings were considered when interpreting study estimates, especially where sampling, confounding, site allocation, or attrition limited inference; they were not used as eligibility-based exclusions. Independent duplicate appraisal and formal adjudication were not documented.

Data Synthesis and Evidence Confidence

A structured narrative synthesis was undertaken using established narrative-synthesis methodology, with reporting additionally guided by the Synthesis Without Meta-analysis (SWiM) guideline [15,16]. Findings were grouped by study design, instrument, and motor construct; age, outcome definitions, thresholds, and timing guided comparisons. AIMS results were compared descriptively where their reporting definitions were sufficiently similar; postural, locomotor, developmental-quotient, motor-milestone, and trunk-control outcomes were discussed separately because these instruments measure non-equivalent constructs. Single-group institutional estimates, comparisons with family-reared children, and intervention outcomes were kept distinct; baseline estimates were not combined with post-intervention effects. Study-level estimates were interpreted in relation to the available denominators, thresholds, and assessment times. Meta-analysis was not performed because outcome constructs, definitions, and assessment time points were insufficiently comparable for meaningful pooling. Even the AIMS studies differed in age composition, assessment timing, and operational definitions. A formal Grading of Recommendations Assessment, Development and Evaluation (GRADE) rating was not undertaken. Overall confidence in the narrative findings was considered qualitatively in relation to study design, risk of bias, consistency, precision, directness, and completeness of available evidence. In particular, the interpretation distinguished more directly reported findings from those limited by small samples, attrition, confounding, or nonstandard outcome definitions.

Results

Study Selection

The audited screening workflow identified 652 records from the original database and discovery searches. After removal of 181 duplicates, 471 unique records were screened, and 390 were excluded before retrieval. Full-text reports were sought for 81 records; 46 were not retrieved, leaving 35 assessed for eligibility. Of these, 26 were excluded after full-text assessment, and nine met the eligibility criteria. Two otherwise eligible reports were not retained during the final bibliographic source-verification audit because stable, authoritative article-level publisher sources could not be verified despite repeated attempts. Their removal concerned verification of the publication source rather than failure to meet the clinical eligibility criteria. These removals excluded one reported randomized massage intervention study and one descriptive AIMS study involving 17 children, potentially narrowing the available intervention and observational evidence. Seven peer-reviewed studies therefore remained in the final narrative synthesis. The 46 unretrieved reports and the two subsequent source-verification removals may have limited the completeness of the evidence captured. The study selection process is summarized in Figure *1*.

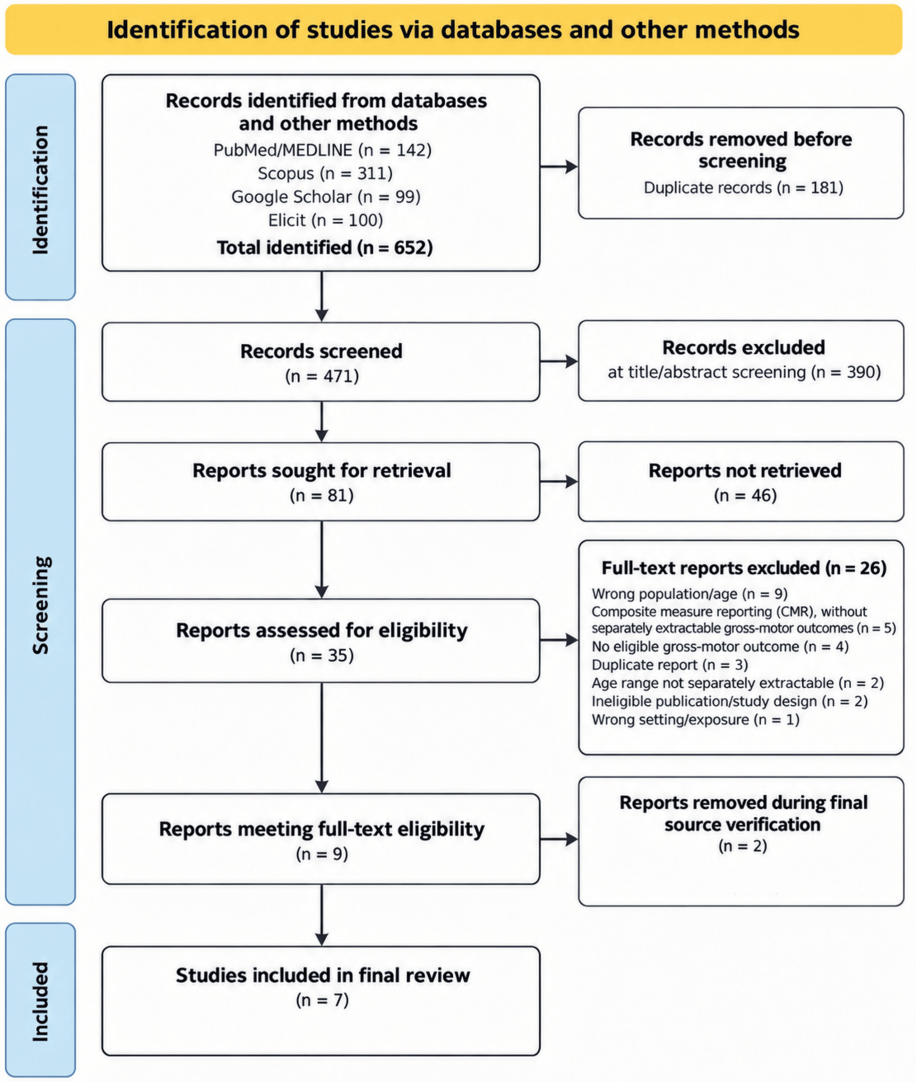


FIGURE 1: PRISMA Flow Diagram of Study Selection

The PRISMA flow diagram summarizes identification, screening, retrieval, eligibility assessment, final source verification, and inclusion of studies. The search conducted on September 3, 2026 identified 652 records: 142 from PubMed, 311 from Scopus, 99 from Google Scholar, and 100 from Elicit. After removal of 181 duplicates, 471 unique records were screened, of which 390 were excluded before retrieval. Eighty-one reports were sought for retrieval; 46 could not be retrieved, leaving 35 reports for full-text eligibility assessment. Twenty-six reports were excluded after full-text assessment. Nine reports met full-text eligibility criteria; two were subsequently removed during final source verification, resulting in seven included peer-reviewed studies. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; CMR, composite measure reporting.

Study Characteristics

Seven peer-reviewed studies were included [17-23], comprising 361 participants: 322 children in institutional care and 39 family-reared controls. Studies were conducted in Thailand, Brazil, Sudan, and North Macedonia. Four used the AIMS [7]; the remaining studies reported Brunet-Lézine postural development [24], Griffiths locomotion [25], or Segmental Assessment of Trunk Control (SATCo) outcomes [26]. The Daily Activities of Infants Scale (DAIS) [27] was retained as a contextual activity measure rather than a direct motor-development score. Their designs, populations, assessment instruments, and age ranges are summarized in Table *4*.

Study / country	Design	Sample	Age	Motor assessment
Castanho AAG et al., 2004 [17] Brazil	Descriptive cross-sectional	n = 30 institutionalized	0-18 mo	AIMS
Espié E et al., 2011 [18] Sudan	Prospective longitudinal cohort	n = 151 institutionalized	0-24 mo	Brunet-Lézine postural domain
Prommin S et al., 2022 [19] Thailand	Purposive cluster quasi-experimental, non-randomized	n = 30 analyzed	6-10 mo	AIMS
Prommin S et al., 2020 [20] Thailand	Longitudinal observational cohort	n = 11 institutionalized	1.5-11.5 mo	AIMS
Rajchanovska D et al., 2008 [21] North Macedonia	Comparative cross-sectional	n = 78 total; 39 institutionalized + 39 controls	6-12 mo	Griffiths locomotion
Singer Brugiolo AS et al., 2024 [22] Brazil	Longitudinal observational cohort	n = 28 institutionalized	0-18 mo	AIMS
Arifin NF et al., 2025 [23] Thailand	Comparative cross-sectional observational	n = 33 institutionalized	3-8 mo	SATCo; DAIS contextual

TABLE 4: Characteristics of included studies

AIMS, Alberta Infant Motor Scale [7]; DAIS, Daily Activities of Infants Scale [27]; SATCo, Segmental Assessment of Trunk Control [26]. Additional assessment instruments were the Brunet-Lézine scale [24] and the Griffiths scales [25]. DAIS was retained as a contextual activity measure rather than a direct motor-development score.

Additional participant and setting characteristics were available for some studies. All seven studies recruited children living in orphanages or comparable residential childcare. Castanho and Blascovi-Assis reported 20 boys and 10 girls, all residents for more than 10 days [17]. Espié et al. reported 88 boys and 63 girls, including 57 children with low birth weight [18]. The upright-play study compared 16 intervention participants with 14 institutional-care controls [19]. The family-reared comparison included 39 institutionalized and 39 age-matched family-reared infants [21]. Singer Brugiolo et al. assessed 17 children longitudinally and 11 only once [22]. Arifin et al. compared 16 full-term with 17 preterm or low-birth-weight infants within institutional care [23]. Comparable durations of institutionalization were not consistently available across the studies.

Key Gross Motor Findings

AIMS-based results indicated different patterns of motor performance across the included studies [17,19,20,22]. Castanho and Blascovi-Assis reported a shift in percentile distribution, but a score below the 50th percentile indicates performance below the reference median and is not, by itself, a clinical diagnosis of developmental delay [17]. Prommin et al. reported variation in AIMS percentiles across repeated assessments [20], while Brugiolo et al. reported study-defined delay classifications at admission [22]. Other instruments identified postural delay, locomotor disadvantage, or trunk-control differences [18,21,23]. Study-specific results and their interpretive limitations are summarized in Table *5*.

Study	Outcome	Main result	Interpretive caution
Castanho AAG et al. [17]	AIMS percentile distribution	29/30 below the 50th percentile; five in the 5th-10th band and four in the 10th-25th band.	Below the 50th percentile is a distributional finding, not a validated clinical delay threshold.
Espié E et al. [18]	Brunet-Lézine postural DQ <70	20.0% (18/90), 10.8% (4/37), 8.3% (2/24), 30.4% (7/23), and 71.4% (5/7) at 2, 4, 9, 12, and 18 months, respectively.	Follow-up denominators declined from 90 to seven; estimates are time-specific available-case proportions.
Prommin S et al. [19]	AIMS during upright-play intervention	Baseline percentiles 36.6 vs 42.5; 12-week values 64.6 vs 45.5; reported change favored intervention, d = 0.74, p = 0.042.	Purposive site assignment and site-care differences limit causal interpretation.
Prommin S et al. [20]	Repeated AIMS percentiles	Mean percentile 55.4; 7/11 had at least one score below the 10th percentile and 3/11 at or below the 5th percentile.	Very small cohort with missing later observations.
Rajchanovska D et al. [21]	Griffiths locomotion	Locomotor stagnation in 46%; developmental quotient 77.69 in institutionalized infants vs 111.24 in controls.	Unadjusted comparative findings with selection and confounding concerns.
Singer Brugiolo AS et al. [22]	AIMS delay at admission	12/28 classified as delayed at the first assessment.	Admission estimate is more interpretable than variable follow-up after physiotherapy exposure.
Arifin NF et al. [23]	SATCo trunk control	Median total SATCo 13 in full-term vs 10 in preterm/low-birth-weight infants; U = 81.5, p = 0.045.	Single-site purposive sample; DAIS was contextual rather than a direct motor outcome.

TABLE 5: Principal gross motor findings

AIMS, Alberta Infant Motor Scale [7]; DQ, developmental quotient; SATCo, Segmental Assessment of Trunk Control [26]; DAIS, Daily Activities of Infants Scale [27]. Additional assessment instruments were the Brunet-Lézine scale [24] and the Griffiths scales [25]. Numerical findings are reported as presented in the source studies.

Structured comparison of the four AIMS studies showed why their estimates were not interchangeable. Castanho and Blascovi-Assis reported that 29 of 30 children scored below the 50th percentile at a single assessment, a distributional result rather than a validated delay proportion [17]. In the longitudinal study by Prommin et al., seven of 11 children had at least one score below the 10th percentile across repeated observations [20]. Brugiolo et al. classified 12 of 28 children as delayed at the first assessment [22]. In the upright-play study, group mean AIMS percentiles were reported at baseline and after 12 weeks, and the intervention comparison was analyzed separately from descriptive prevalence findings [19]. These results differ in thresholds, observation periods, and outcome definitions and were therefore compared descriptively rather than pooled.

Intervention Findings

One included study evaluated structured upright play and was considered separately from observational findings [19]. AIMS percentile improvement was greater in the intervention group, but allocation was by institution or site rather than individual randomization. Consequently, intervention exposure could not be separated confidently from site-level and caregiving differences, and the finding does not establish intervention efficacy. These intervention results were not combined with baseline or observational estimates.

Quantitative Synthesis

Meta-analysis was not performed because the studies differed in motor constructs, outcome thresholds, assessment times, and designs. Their measures included AIMS percentile classifications, Brunet-Lézine developmental quotients, Griffiths locomotion, and SATCo trunk-control scores. These outcomes were not sufficiently equivalent for a clinically meaningful pooled estimate. The four AIMS studies were compared descriptively by population, threshold, and assessment timing (Table *5*), rather than pooled, because they reported different age ranges, one-time versus repeated measurements, intervention versus observational outcomes, and non-equivalent definitions [17,19,20,22].

Risk of Bias

No included study was judged at overall low risk of bias. Of seven studies, four were classified as high risk

and three as having some concerns. Confounding, site-level intervention allocation, and attrition were particularly relevant to internal validity and causal interpretation. Small or purposively selected samples and single-institution recruitment primarily limited generalizability, while differing outcome thresholds and assessment schedules restricted comparability. Accordingly, findings from studies at high risk of bias were interpreted more cautiously than findings with fewer concerns, rather than treated as equally strong evidence. The design-specific JBI assessments are summarized in Table *6*.

Study	JBI checklist	Selection	Confounding / allocation	Missing data / follow-up	Outcome measurement	Analysis / reporting	Overall
Castanho AAG et al. [17]	JBI Prevalence	High concern	High concern	Unclear	Some concerns	Some concerns	High
Espié E et al. [18]	JBI Cohort	Some concerns	High concern	High concern	Low concern	High concern	High
Prommin S et al. [19]	JBI Quasi-Experimental	High concern	High concern	Some concerns	Low concern	High concern	High
Prommin S et al. [20]	JBI Cohort	Some concerns	Some concerns	Some concerns	Low concern	Low concern	Some concerns
Rajchanovska D et al. [21]	JBI Analytical Cross-Sectional	High concern	High concern	Low concern	Some concerns	Some concerns	High
Singer Brugiolo AS et al. [22]	JBI Cohort	Some concerns	High concern	High concern	Low concern	High concern	Some concerns
Arifin NF et al. [23]	JBI Analytical Cross-Sectional	Some concerns	Some concerns	Low concern	Low concern	Low concern	Some concerns

TABLE 6: JBI design-specific risk-of-bias summary

JBI, Joanna Briggs Institute [12-14].

Overall categories are review-defined synthesis labels derived from domain-level judgments and are not JBI numeric cutoffs. The study-level risk-of-bias judgments were informed by current JBI risk-of-bias methodology [12-14] and summarize the seven included evidence reports [17-23]. These judgments are presented in Figure *2*.

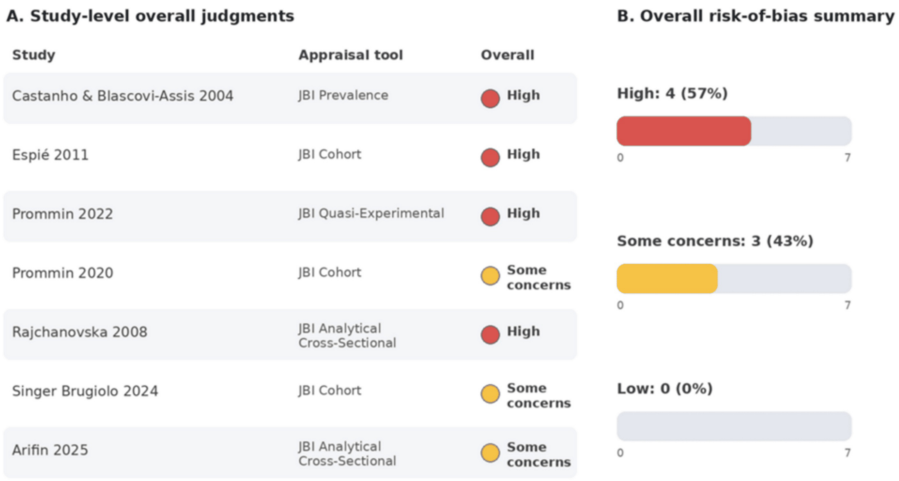


FIGURE 2: Risk of bias assessment of the included reports.

The figure presents the seven included studies: Castanho and Blascovi-Assis [17], Espié et al. [18], Prommin et al. (2022) [19], Prommin et al. (2020) [20], Rajchanovska et al. [21], Singer Brugiolo et al. [22], and Arifin et al. [23]. All studies were assessed using design-appropriate JBI critical-appraisal tools [12-14].

Overall risk-of-bias judgments were high risk in four studies and some concerns in three studies; no study was judged overall low risk. Overall categories are review-defined synthesis labels and are not JBI numeric cutoffs.

Discussion

This systematic review identified recurring indications of gross motor vulnerability among infants and young toddlers living in institutional care. Its main contribution is not a single prevalence estimate or evidence of a causal effect of institutional residence. Instead, the review brings together and critically appraises seven studies conducted across different settings and using different measures of early motor development. The findings concerned global gross motor performance, postural development, locomotion, and segmental trunk control, which are related but distinct outcomes. Differences in sampling, thresholds, assessment timing, and risk of bias prevent direct numerical comparisons and a single pooled estimate.

The size and methodological limitations of the evidence base are important findings. Four of the seven studies were judged to be at high overall risk of bias and three had some concerns; none was judged to be at low risk. The studies suggest possible motor-development difficulties, but do not establish their population prevalence, the magnitude of any association with institutional residence, or a causal effect. These limitations also mean that findings from studies with greater risks of confounding, selection bias, or attrition warrant less interpretive weight than more directly supported observations.

The findings are broadly consistent with earlier literature on institutional care and developmental disadvantage [4-6], but the scope and strength of the evidence differ. The systematic and integrative review by van IJzendoorn et al. [4] addressed development across multiple domains and care transitions, whereas the present review focuses on motor-specific outcomes during current institutional residence at 0-24 months. Its concerns are reflected in the low AIMS performance reported by Castanho and Blascovi-Assis [17] and the locomotor comparison reported by Rajchanovska et al. [21]. However, longitudinal fluctuations in AIMS performance reported by Prommin et al. [20] and differences in trunk control by birth status within institutional care reported by Arifin et al. [23] indicate that neither the pattern nor the possible explanation is uniform. Broader accounts of responsive caregiving, health, nutrition, and stimulation [5,6] offer context rather than evidence that these factors caused the findings observed here. The present synthesis adds an outcome-specific comparison without combining motor measures with cognitive or psychosocial development.

The AIMS-based studies illustrate why low performance and clinically defined delay must be distinguished. Castanho and Blascovi-Assis reported a downward shift in percentile distribution, but performance below the 50th percentile is not itself a validated diagnosis of developmental delay [17]. Prommin et al. observed substantial within-child variability in repeated AIMS measurements [20], whereas Singer Brugiolo et al. reported delay classifications at admission [22]. A separate Prommin intervention study reported changes in AIMS percentiles after upright play and was interpreted as intervention evidence rather than as an estimate of baseline delay [19]. The observational findings suggest a similar pattern of concern, but they reflect different outcome definitions and assessment time points. Their reported proportions and percentiles should therefore not be interpreted as equivalent estimates of developmental delay.

Studies using instruments other than AIMS addressed different aspects of motor development. Espié et al. reported age-specific postural delay using Brunet-Lézine developmental quotients, but substantial attrition limits interpretation of later follow-up estimates [18]. Rajchanovska et al. observed lower Griffiths locomotor performance in institutionalized infants than in family-reared controls; the unadjusted comparison remains susceptible to selection and residual confounding [21]. Arifin et al. reported differences in segmental trunk control between full-term and preterm or low-birth-weight infants within an orphanage population, but the study did not establish that birth characteristics caused the observed differences [23]. Prematurity, low birth weight, nutritional status, duration of institutionalization, and caregiving conditions are plausible contributors to developmental variation; however, their individual effects could not be established consistently from these studies. Postural delay, locomotor differences, and trunk-control scores should therefore be interpreted as separate constructs rather than numerically interchangeable expressions of one outcome.

Intervention evidence was limited to one non-randomized upright-play study [19]. AIMS percentile improvement favoured the intervention group. However, allocation occurred by institution or site rather than by individual randomization. The observed difference therefore could not be separated confidently from caregiving, environmental, or other site-specific characteristics. These findings justify further investigation of structured movement and caregiver-supported play, but do not establish efficacy or support a particular programme, intensity, or duration.

Overall, the findings suggest possible motor vulnerability rather than a reliably quantified effect of institutional residence. The small evidence base, heterogeneous outcomes, and absence of studies at low overall risk of bias limit clinical inference. Better comparative and intervention research is needed before specific developmental risks or treatment effects can be estimated reliably.

Strengths and Limitations

A major strength of this review is its focused synthesis of gross motor development in children aged 0-24 months currently living in institutional care. The database-specific search strategies and retrieval outcomes were documented, with supplementary discovery and source-verification activities recorded. Observational and intervention findings were considered separately, and design-appropriate risk-of-bias tools were applied. The final manuscript audit also checked study design, age range, sample size, instrument terminology, eligibility decisions, and PRISMA accounting.

The small evidence base also reflects limitations of the available literature. Only seven studies remained in the final synthesis, and no study was judged to be at low overall risk of bias: four were at high risk and three had some concerns. Small or purposively selected samples, single-institution recruitment, attrition, incomplete reporting, and limited adjustment for confounding restrict internal validity and, in some studies, generalisability. These factors limit confidence in precise estimates of developmental difficulty, between-group differences, and causal interpretation.

The review process had additional limitations. Google Scholar screening was restricted to the first 100 relevance-ranked results, potentially missing eligible reports beyond this limit. Of 81 reports sought for retrieval, 46 were unavailable; this substantial loss may have affected evidence completeness and introduced selection bias. Two otherwise eligible reports were not retained after final bibliographic source verification because stable, authoritative article-level publisher sources could not be verified despite repeated attempts. Their omission may also have affected the representation of available evidence. The screening workflow was AI-assisted and source-audited, while independent duplicate screening, duplicate extraction, inter-reviewer agreement, and formal adjudication were not documented; these are limitations of the review process rather than characteristics of the included studies. The protocol was registered retrospectively, and the reported audited update was awaiting registry confirmation at manuscript finalization.

These review-process limitations are distinct from weaknesses in the primary studies. Across the included reports, different assessment instruments, percentile thresholds, age distributions, and follow-up times limited comparability. A formal GRADE assessment was not performed, and meta-analysis was not undertaken because the available outcomes were not sufficiently equivalent for meaningful pooling. Consequently, confidence is greater in the observation that several studies reported motor-development concerns than in any estimate of their frequency, magnitude, or cause.

Clinical and Research Implications

The findings support attention to early gross motor development among infants in institutional care, but do not establish a screening interval or a particular treatment protocol. Repeated, age-appropriate motor assessment may be considered as part of broader developmental surveillance. Such an approach is informed by developmental practice and the variability reported in the included studies, rather than by direct evidence in this review of an effective assessment schedule. Interpretation should consider prematurity, birth weight, nutritional and medical status, opportunities for movement, and the caregiving context; the included studies did not establish independent causal effects for these factors.

Future studies should use prospective, adequately sized, multicentre designs to examine whether reported motor differences occur across a wider range of institutional settings. Investigators should characterize age at assessment, duration and quality of institutional care, nutritional and birth characteristics, and caregiving conditions. Standardized, age-appropriate motor instruments and prespecified thresholds are needed. Studies should report instrument-specific scores and relevant follow-up denominators separately, allowing results to be compared without conflating distinct developmental outcomes.

Comparative observational studies should recruit appropriate family- or community-reared controls. Analyses should account for prematurity, low birth weight, malnutrition, chronic illness, and socioeconomic conditions. Intervention studies require robust allocation methods, adequate statistical power, standardized co-interventions, prespecified outcomes, prospective registration, and complete follow-up. These approaches are needed to improve estimates of developmental differences and to evaluate intervention effectiveness more reliably.

Conclusions

Across seven heterogeneous studies, findings suggest a concerning pattern of low or atypical gross motor performance, postural delay, locomotor differences, and trunk-control difficulties among infants and young toddlers living in institutional care. The frequency and magnitude of these findings remain uncertain because of differences in assessment methods and study quality. The review does not establish a causal effect of institutional care or a reliable prevalence estimate of motor delay.

For this reason, repeated, age-appropriate standardized motor assessments may be considered as part of broader developmental surveillance, although the review does not establish an optimal screening schedule

or intervention. Stronger prospective multicentre studies with appropriate family-reared comparison groups and adjustment for key confounders, together with adequately powered intervention studies using robust allocation methods and complete follow-up, are needed before specific clinical recommendations can be established.

Appendices

Item	JBI appraisal question	Response	Evidence / rationale	Source location
1	Castanho and Blascovi-Assis (2004) [17]	JBI Prevalence	High	
Q1	Was the sample frame appropriate to address the target population?	No	One selected institution is not a representative sampling frame for institutionalized infants generally.	Methods/Results pp.439–441
Q2	Were study participants sampled appropriately?	Unclear	The article states that 30 children were included but does not clearly describe census/consecutive sampling or eligibility flow.	Methods/Results pp.439–441
Q3	Was the sample size adequate?	No	Small n=30 and no sample-size calculation.	Methods/Results pp.439–441
Q4	Were the study subjects and setting described in detail?	Yes	Age, sex, institutional setting, nursery distribution, reasons for placement and visits were described.	Methods/Results pp.439–441
Q5	Was data analysis conducted with sufficient coverage of the identified sample?	Unclear	Coverage of all eligible infants during the study period was not reported.	Methods/Results pp.439–441
Q6	Were valid methods used for identifying the condition/outcome?	Yes	AIMS is a validated standardized motor assessment, although <50th percentile is not a clinical-delay threshold.	Methods/Results pp.439–441
Q7	Was the outcome measured in a standard, reliable way for all participants?	Unclear	Same AIMS procedure was used, but assessor training, blinding and study-specific reliability were not reported.	Methods/Results pp.439–441
Q8	Was there appropriate statistical analysis?	Yes	Descriptive counts and percentages were appropriate for the stated characterization objective.	Methods/Results pp.439–441
Q9	Was the response rate adequate, and if not, was the low response rate managed appropriately?	Unclear	Eligible denominator and nonparticipation were not reported.	Methods/Results pp.439–441
2	Espié et al. (2011) [18]	JBI Cohort	High	
Q1	Were comparison groups similar and recruited from the same population?	N/A	Single-cohort study; no separate unexposed comparison group.	Methods p.2; flow/Results pp.3–4
Q2	Were exposures measured similarly to assign participants to exposed and unexposed groups?	N/A	Single-cohort study; all participants were living in the same orphanage.	Methods p.2; flow/Results pp.3–4
Q3	Was the exposure measured in a valid and reliable way?	Yes	Residence/admission in Mygoma orphanage was documented prospectively.	Methods p.2; flow/Results pp.3–4
Q4	Were confounding factors identified?	Yes	Birth weight, medical pathology, origin, transfers and several child factors were recorded.	Methods p.2; flow/Results pp.3–4
Q5	Were strategies to deal with confounding factors stated?	No	Only simple risk-factor comparisons were reported; no multivariable or time-varying confounder control.	Methods p.2; flow/Results pp.3–4
Q6	Were participants free of the outcome at the start of the study?	N/A	The review estimates developmental status rather than incident disease in an outcome-free	Methods p.2; flow/Results

			cohort.	pp.3–4
Q7	Were outcomes measured in a valid and reliable way?	Yes	Brunet–Lézine postural domain was applied with a defined DQ<70 delay threshold.	Methods p.2; flow/Results pp.3–4
Q8	Was follow-up time reported and sufficient for outcomes to occur?	Yes	Assessments at 2, 4, 9, 12 and 18 months were prespecified and sufficient to observe developmental change.	Methods p.2; flow/Results pp.3–4
Q9	Was follow-up complete, and were reasons for loss described?	No	Follow-up fell from 151 enrolled to 90, 37, 24, 23 and 7; deaths, discharge and illness changed the observed sample.	Methods p.2; flow/Results pp.3–4
Q10	Were strategies used to address incomplete follow-up?	No	Available-case results were presented without attrition weighting, imputation or sensitivity analysis.	Methods p.2; flow/Results pp.3–4
Q11	Was appropriate statistical analysis used?	No	Time-point proportions did not account for repeated measures, changing denominators or informative attrition.	Methods p.2; flow/Results pp.3–4
3	Prommin et al. (2022) [19]	JBI Quasi-Experimental	High	
Q1	Is it clear in the study what is the cause and what is the effect?	Yes	The upright-play intervention preceded repeated AIMS follow-up assessments.	Methods pp.4–8
Q2	Were the participants included in comparisons similar?	No	Experimental and control infants came from different orphanage sites and site-level caregiving differences were present.	Methods pp.4–8; Table 1
Q3	Were participants included in comparisons receiving similar treatment/care other than the exposure or intervention of interest?	No	Site and caregiver environments differed, so the intervention was confounded with institutional setting.	Methods pp.4–8
Q4	Was there a control group?	Yes	A contemporaneous institutionalized control group was used.	Methods; Figure 1
Q5	Were there multiple measurements of the outcome both pre and post the intervention?	Yes	AIMS was assessed at baseline and 4, 8 and 12 weeks.	Methods; Tables 2–4
Q6	Was follow-up complete and, if not, were differences adequately described and analyzed?	Unclear	Four of 34 recruited infants were excluded and no intention-to-treat or sensitivity analysis was reported.	Figure 1; Methods
Q7	Were outcomes measured in the same way for participants included in comparisons?	Yes	The same AIMS procedure and blinded trained assessors were used across groups.	Methods pp.6–7
Q8	Were outcomes measured in a reliable way?	Yes	AIMS reliability was reported with high inter-rater/intra-rater coefficients.	Methods pp.6–7
Q9	Was appropriate statistical analysis used?	Unclear	Repeated analyses were reported, but purposive site assignment and site-level confounding limit causal interpretation.	Results; Tables 2–4
4	Prommin et al. (2020) [20]	JBI Cohort	Some concerns	
Q1	Were comparison groups similar and recruited from the same population?	N/A	Single orphanage cohort with no separate unexposed comparison group.	Methods pp.2–3; Results p.4
Q2	Were exposures measured similarly to assign participants to exposed and unexposed groups?	N/A	All participants shared the institutional exposure.	Methods pp.2–3; Results p.4
Q3	Was the exposure measured in a valid and reliable way?	Yes	All infants newly admitted at newborn to <1.5 months were recruited from one defined orphanage.	Methods pp.2–3; Results p.4
Q4	Were confounding factors identified?	Yes	Biological factors, illness, admission age and growth were recorded.	Methods pp.2–3; Results p.4
			Corrected age was used for preterm infants, but	

Q5	Were strategies to deal with confounding factors stated?	No	no broader confounder-adjusted analysis was performed.	Methods pp.2–3; Results p.4
Q6	Were participants free of the outcome at the start of the study?	N/A	The objective was developmental trajectory/variability, not incident delay among outcome-free infants.	Methods pp.2–3; Results p.4
Q7	Were outcomes measured in a valid and reliable way?	Yes	AIMS is validated; one trained physical therapist assessed all infants and intra-rater ICC was 0.99.	Methods pp.2–3; Results p.4
Q8	Was follow-up time reported and sufficient for outcomes to occur?	Yes	Monthly assessments from 1.5 to 11.5 months were prespecified and developmentally appropriate.	Methods pp.2–3; Results p.4
Q9	Was follow-up complete, and were reasons for loss described?	No	Two of 13 infants were adopted and excluded; two preterm infants lacked the last two assessments.	Methods pp.2–3; Results p.4
Q10	Were strategies used to address incomplete follow-up?	No	No imputation or sensitivity analysis for missing visits/adoptions was reported.	Methods pp.2–3; Results p.4
Q11	Was appropriate statistical analysis used?	Yes	Descriptive trajectory and variability analyses were aligned with the exploratory study objective, with limitations from small n/missingness.	Methods pp.2–3; Results p.4
5	Rajchanovska et al. (2008) [21]	JBI Analytical Cross-Sectional	High	
Q1	Were the criteria for inclusion in the sample clearly defined?	Yes	Institutionalized and family-reared infants aged 6–12 months were defined with eligibility criteria.	Methods pp.185–188; Tables 1–3
Q2	Were study subjects and the setting described in detail?	Yes	Institutional home, age range, group sizes and selected biological criteria were described.	Methods pp.185–188; Tables 1–3
Q3	Was the exposure measured in a valid and reliable way?	Yes	Institutional versus biological-family residence was clearly classified.	Methods pp.185–188; Tables 1–3
Q4	Were objective, standard criteria used to measure the condition/outcome?	Yes	Griffiths scale domains, mental age and developmental quotient were used as standardized criteria.	Methods pp.185–188; Tables 1–3
Q5	Were confounding factors identified?	Yes	Age matching and perinatal eligibility were considered, but anthropometric differences remained.	Methods pp.185–188; Tables 1–3
Q6	Were strategies to deal with confounding factors stated?	No	No multivariable adjustment or other strategy addressed residual confounding/anthropometric imbalance.	Methods pp.185–188; Tables 1–3
Q7	Were outcomes measured in a valid and reliable way?	Unclear	Griffiths was standardized, but assessor training, blinding and study-specific reliability were not reported.	Methods pp.185–188; Tables 1–3
Q8	Was appropriate statistical analysis used?	Yes	Independent-group tests were reported, but interpretation remains limited by unadjusted confounding.	Methods pp.185–188; Tables 1–3
6	Singer Brugiolo et al. (2024) [22]	JBI Cohort	Some concerns	
Q1	Were comparison groups similar and recruited from the same population?	N/A	Single institutional cohort with no unexposed comparison group.	Methods pp.2–4; Results pp.4–6
Q2	Were exposures measured similarly to assign participants to exposed and unexposed groups?	N/A	All included participants were living in the same institution.	Methods pp.2–4; Results pp.4–6

Q3	Was the exposure measured in a valid and reliable way?	Yes	Institutional residence and admission timing were documented from institutional records.	Methods pp.2–4; Results pp.4–6
Q4	Were confounding factors identified?	Yes	Birth, health and institutional factors were collected, although many birth variables were missing.	Methods pp.2–4; Results pp.4–6
Q5	Were strategies to deal with confounding factors stated?	No	No adjusted analysis was used to address confounding by age, biological risk, duration in care or physiotherapy exposure.	Methods pp.2–4; Results pp.4–6
Q6	Were participants free of the outcome at the start of the study?	N/A	Baseline developmental delay was an outcome of interest at admission.	Methods pp.2–4; Results pp.4–6
Q7	Were outcomes measured in a valid and reliable way?	Yes	AIMS was validated for Brazil; assessors were trained and intra-/inter-rater ICC values exceeded 0.80.	Methods pp.2–4; Results pp.4–6
Q8	Was follow-up time reported and sufficient for outcomes to occur?	Unclear	Assessment every ~2 months was planned, but actual follow-up varied and 11/28 had only one assessment.	Methods pp.2–4; Results pp.4–6
Q9	Was follow-up complete, and were reasons for loss described?	No	Only 17/28 had repeated assessments; reasons/timing for incomplete follow-up were not fully analyzed.	Methods pp.2–4; Results pp.4–6
Q10	Were strategies used to address incomplete follow-up?	No	No statistical strategy addressed unequal follow-up or missing longitudinal observations.	Methods pp.2–4; Results pp.4–6
Q11	Was appropriate statistical analysis used?	No	Descriptive percentiles did not model repeated measures, and all infants received individualized physiotherapy during follow-up.	Methods pp.2–4; Results pp.4–6
7	Arifin et al. (2025) [23]	JBI Analytical Cross-Sectional	Some concerns	
Q1	Were the criteria for inclusion in the sample clearly defined?	Yes	Full-term and preterm/LBW inclusion and exclusion criteria were explicitly defined for corrected age 3–8 months.	Methods pp.2–6; Tables 3–6 pp.7–9
Q2	Were study subjects and the setting described in detail?	Yes	Single orphanage setting, caregiving routine, group sizes and demographic characteristics were described.	Methods pp.2–6; Tables 3–6 pp.7–9
Q3	Was the exposure measured in a valid and reliable way?	Yes	Gestational age and birth weight were taken from medical/history records using explicit thresholds.	Methods pp.2–6; Tables 3–6 pp.7–9
Q4	Were objective, standard criteria used to measure the condition/outcome?	Yes	SATCo used standardized static, active and reactive scoring; DAIS procedures were described.	Methods pp.2–6; Tables 3–6 pp.7–9
Q5	Were confounding factors identified?	Yes	Age/corrected age, gestational age, birth weight and maternal conditions were recognized as relevant factors.	Methods pp.2–6; Tables 3–6 pp.7–9
Q6	Were strategies to deal with confounding factors stated?	No	Groups were compared without multivariable adjustment or matching beyond corrected age.	Methods pp.2–6; Tables 3–6 pp.7–9
Q7	Were outcomes measured in a valid and reliable way?	Yes	SATCo assessor training and high intra-/inter-rater ICCs were reported; DAIS study-specific reliability was not tested.	Methods pp.2–6; Tables 3–6 pp.7–9
Q8	Was appropriate statistical analysis used?	Yes	Mann–Whitney U, effect size, power and Spearman correlations were appropriate for the reported non-normal scores.	Methods pp.2–6; Tables 3–6 pp.7–9

TABLE 7: Detailed study-level JBI critical-appraisal assessments

Response options: Yes; No; Unclear; N/A (not applicable). Overall judgments are review-defined qualitative interpretations of appraisal concerns and were not computed from numerical JBI cutoffs.

JBI, Joanna Briggs Institute [12-14]; AIMS, Alberta Infant Motor Scale [7]; DQ, developmental quotient; SATCo, Segmental Assessment of Trunk Control [26]; DAIS, Daily Activities of Infants Scale [27]. Additional assessment instruments referenced in the appraisal rationales were the Brunet-Lézine scale [24] and the Griffiths scales [25].

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Asim Ahmed, Nahla Bakri Ageeb Babekr, Aya Alhassan, Suzan Elhireka Koko Gadoum, Samah Mohammed Osman Mohammed Nur, Mariam Z. Mubarak

Acquisition, analysis, or interpretation of data: Asim Ahmed, Nahla Bakri Ageeb Babekr, Aya Alhassan, Suzan Elhireka Koko Gadoum, Samah Mohammed Osman Mohammed Nur, Mariam Z. Mubarak

Drafting of the manuscript: Asim Ahmed, Nahla Bakri Ageeb Babekr, Mariam Z. Mubarak

Critical review of the manuscript for important intellectual content: Asim Ahmed, Nahla Bakri Ageeb Babekr, Aya Alhassan, Suzan Elhireka Koko Gadoum, Samah Mohammed Osman Mohammed Nur, Mariam Z. Mubarak

Supervision: Nahla Bakri Ageeb Babekr, Aya Alhassan, Mariam Z. Mubarak

Disclosures

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** The study received financial support of USD 1,400 from 10 Scholars Academy to support research. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

The authors acknowledge the use of Zettelly (Stockholm, Sweden) and the 10 Scholars App (Honolulu, HI, USA) to support record organization, screening, source-linked data extraction and auditing, reference checking, DOI verification, and manuscript formatting. All study-selection decisions, extracted data, scientific interpretations, and final manuscript content were independently reviewed and remain the responsibility of the authors.

References

1. Nurturing care for early childhood development: a framework for helping children survive and thrive to transform health and human potential. (2018). Accessed: September 28, 2026: https://www.who.int/publications/i/item/9789241514064.
2. Black MM, Walker SP, Fernald LC, et al.: Early childhood development coming of age: science through the life course. Lancet. 2017, 389:77-90. 10.1016/S0140-6736(16)31389-7
3. Britto PR, Lye SJ, Proulx K, et al.: Nurturing care: promoting early childhood development . Lancet. 2017, 389:91-102. 10.1016/S0140-6736(16)31390-3
4. van IJzendoorn MH, Bakermans-Kranenburg MJ, Duschinsky R, et al.: Institutionalisation and deinstitutionalisation of children 1: a systematic and integrative review of evidence regarding effects on development. Lancet Psychiatry. 2020, 7:703-20. 10.1016/S2215-0366(19)30399-2
5. Berens AE, Nelson CA: The science of early adversity: is there a role for large institutions in the care of vulnerable children?. Lancet. 2015, 386:388-98. 10.1016/S0140-6736(14)61131-4
6. van IJzendoorn MH, Palacios J, Sonuga-Barke EJ, et al.: Children in institutional care: delayed development and resilience. Monogr Soc Res Child Dev. 2011, 76:8-30. 10.1111/j.1540-5834.2011.00626.x
7. Eliks M, Gajewska E: The Alberta Infant Motor Scale: a tool for the assessment of motor aspects of neurodevelopment in infancy and early childhood. Front Neurol. 2022, 13:927502. 10.3389/fneur.2022.927502
8. Page MJ, McKenzie JE, Bossuyt PM, et al.: The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021, 372:n71. 10.1136/bmj.n71
9. Page MJ, Moher D, Bossuyt PM, et al.: PRISMA 2020 explanation and elaboration: updated guidance and

exemplars for reporting systematic reviews. BMJ. 2021, 372:n160. 10.1136/bmj.n160
10. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, Koffel JB: PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev. 2021, 10:39. 10.1186/s13643-020-01542-z
11. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C: PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol. 2016, 75:40-6. 10.1016/j.jclinepi.2016.01.021
12. Stone JC, Barker TH, Aromataris E, et al.: From critical appraisal to risk of bias assessment: clarifying the terminology for study evaluation in JBI systematic reviews. JBI Evid Synth. 2023, 21:472-7. 10.11124/JBIES-22-00434
13. Munn Z, Stone JC, Aromataris E, Klugar M, Sears K, Leonardi-Bee J, Barker TH: Assessing the risk of bias of quantitative analytical studies: introducing the vision for critical appraisal within JBI systematic reviews. JBI Evid Synth. 2023, 21:467-71. 10.11124/JBIES-22-00224
14. Barker TH, Stone JC, Sears K, et al.: Revising the JBI quantitative critical appraisal tools to improve their applicability: an overview of methods and the development process. JBI Evid Synth. 2023, 21:478-93. 10.11124/JBIES-22-00125
15. Rodgers M, Sowden A, Petticrew M, et al.: Testing methodological guidance on the conduct of narrative synthesis in systematic reviews: effectiveness of interventions to promote smoke alarm ownership and function. Evaluation. 2009, 15:49-73. 10.1177/1356389008097871
16. Campbell M, McKenzie JE, Sowden A, et al.: Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ. 2020, 368:l6890. 10.1136/bmj.l6890
17. Castanho AA, Blascovi-Assis SM: Caracterização do desenvolvimento motor da criança institucionalizada. Fisioter Bras. 2004, 5:437-42. 10.33233/fb.v5i6.3185
18. Espié E, Ouss L, Gaboulaud V, et al.: Against the odds: psychomotor development of children under 2 years in a Sudanese orphanage. J Trop Pediatr. 2011, 57:412-7. 10.1093/tropej/fmq117
19. Prommin S, Siritaratiwat W, Bennett S, Mato L, Keeratisiroj O, Kamruecha W: Effects of play in an upright position on intra-individual variability of gross motor and language development in institutionalized infants. Int J Environ Res Public Health. 2022, 19:11804. 10.3390/ijerph191811804
20. Prommin S, Bennett S, Keeratisiroj O, et al.: Instability of gross motor development during the first year in orphaned infants: a longitudinal observation study. Early Child Dev Care. 2020, 190:2041-9. 10.1080/03004430.2018.1555825
21. Rajchanovska D, Todorovski T, Kochankovska L, et al.: Psychomotor development of the institutionalized children at the age of 6-12 months. Medicus. 2008, 9:185-94.
22. Brugiolo AS, Pacheco Martins K, Palermo Ferreira MC, et al.: Motor development of infants in a children's shelter institution: a longitudinal study. Movimenta. 2024, 17:1-11.
23. Fauziah Arifin N, Siritaratiwat W, Yonglitthipagon P, Mato L, Chatprem T, Sangkarit N: Segmental trunk control in sitting between full-term and preterm infants raised in the orphanage setting. Int J Environ Res Public Health. 2025, 22:1824. 10.3390/ijerph22121824
24. Josse D: [Revised Brunet-Lézine: Psychomotor Development Scale of Early Childhood]. Éditions et Applications Psychologiques, Issy-les-Moulineaux; 1997.
25. Griffiths R: The Abilities of Babies: A Study in Mental Measurement . University of London Press, London; 1954.
26. Butler PB, Saavedra S, Sofranac M, et al.: Refinement, reliability, and validity of the segmental assessment of trunk control. Pediatr Phys Ther. 2010, 22: 246-57. 10.1097/PEP.0b013e3181e69490
27. Bartlett DJ, Fanning JK, Miller L, et al.: Development of the Daily Activities of Infants Scale: a measure supporting early motor development. Dev Med Child Neurol. 2008, 50:613-7. 10.1111/j.1469-8749.2008.03007.x