Artificial Intelligence in Pediatric Occupational Therapy Screening, Evaluation, and Intervention: A Systematic Review

Adelyne Burlas OTS, Elenor Nichols OTS,

Hanna Galan OTS, Lizbeth M. Rodríguez González OTS

Doctor of Occupational Therapy Program, Hawai'i Pacific University

OT 8520: Scholarly Practice II

Dr. Jana Cason and Dr. Michael Falcon

June 24, 2026

Abstract

Importance: Early identification of developmental challenges is associated with better long-term outcomes for children. Advances in the efficacy and accuracy of artificial intelligence (AI) may prove AI to be a great tool for support of the early identification of developmental challenges in children.

Objective: This systematic review examined the current strength of evidence regarding the use of artificial intelligence technologies in screening, evaluation, and intervention within pediatric populations, with a focus on outcomes relevant to occupational therapy practice, including developmental, functional, behavioral, and participation-related domains.

Data Sources: A literature search was completed between May 14, 2026, and May 22, 2026. Follow up searches were conducted on May 26, 2026. Databases included MEDLINE, Academic Search Complete, and PubMed using Hawai'i Pacific University's online library databases. Search terms included "Artificial Intelligence or AI or ChatGPT or Chatbot" AND "Occupational Therapy or OT" AND "Intervention or Program or Therapy" as well as combinations of these terms.

Study Selection and Data Collection: This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Published studies on use of artificial intelligence (AI) in occupational therapy screening, evaluation and intervention for the pediatric population were included in the systematic review. Data was excluded if not peer-reviewed journal articles and academic journals published in English within the last five years, lacking strict focus on AI used within occupational therapy for pediatric populations, not assessing therapy effectiveness and client outcomes, or failing to use quantitative, qualitative, or mixed methods study designs.

Findings: Five studies were included (two level 1B; three level 3B) according to the American Occupational Therapy Association's Levels of Evidence. The outcomes of three studies indicate AI may be effective in screening for developmental delays in children using everyday tasks like drawing, copying and origami. The remaining two studies demonstrated increased handwriting skills, participation, and quality of life through AI assisted interventions for children.

Conclusion and Relevance: Use of AI in screening, evaluation, and intervention in occupational therapy shows promise in early developmental delay screening and functional outcomes for children.

Key words: Artificial intelligence, evaluation, intervention, occupational therapy, pediatrics, screening

Artificial Intelligence in Pediatric Occupational Therapy Screening, Evaluation, and Intervention: A Systematic Review

Artificial intelligence (AI) is rapidly transforming healthcare by enhancing the accuracy, efficiency, and personalization of clinical services. According to Kaelin et al. (2024), AI encompasses technologies such as machine learning, predictive analytics, computer vision, and natural language processing that enable systems to analyze large amounts of data and support clinical decision-making. Within pediatric healthcare, early identification of developmental, cognitive, behavioral, and functional challenges is critical because timely intervention can improve developmental outcomes and participation in everyday activities. Occupational therapists play a central role in pediatric screening, evaluation, and intervention by supporting children's occupational performance and engagement in meaningful activities, an area highlighted by Kaelin et al. (2024). As demands for pediatric rehabilitation services continue to increase, healthcare providers are seeking innovative approaches that improve efficiency while maintaining high-quality, client-centered care, including the integration of emerging technologies (Stover & Jacobs, 2025).

Recent advances in advanced computational methods have demonstrated strong potential applications across pediatric rehabilitation settings. Rather than functioning purely as theoretical tools, AI-based systems are currently utilized to classify neurodevelopmental disorders, analyze behavioral patterns, and predict treatment outcomes. For instance, Kaelin et al. (2021) explored the use of these technologies to identify developmental delays and support participation-focused rehabilitation among children and youth with disabilities. Additional research by Kaelin et al. (2022) examined how algorithmic frameworks can capture and operationalize participation-related outcomes in pediatric rehabilitation research. Machine learning models have also shown promise for analyzing large datasets and assisting with clinical decision-making, suggesting that AI may serve as an effective screening and evaluation tool. These capabilities align closely with occupational therapy's emphasis on early identification and intervention to optimize a child's functional performance and facilitate earlier referral to essential rehabilitation services.

Beyond screening and evaluation, AI has increasingly been integrated into intervention programs designed to improve functional outcomes in children. Emerging studies have reported positive effects from AI-assisted interventions, including adaptive gaming platforms, computer vision technologies, and virtual therapeutic environments. These approaches provide opportunities for individualized treatment, continuous progress monitoring, and real-time feedback (Stover & Jacobs, 2025). In pediatric occupational therapy settings, Stover and Jacobs (2025) reported that the AI-based Korro platform demonstrated significant improvements in bilateral coordination, home-program adherence, and engagement among children with developmental challenges. These findings suggest that tailoring interventions to individual performance levels can enhance clinical effectiveness while increasing children's motivation and active participation in therapeutic activities.

Despite growing interest in AI applications within pediatric rehabilitation, important questions remain regarding implementation, effectiveness, and ethical use. Occupational therapy scholars have emphasized the importance of human-centered artificial intelligence that promotes participation, occupational justice, and client-centered practice (Kaelin et al., 2024). Concerns related to algorithmic bias, transparency, accessibility, and privacy further underscore the need for careful evaluation before widespread adoption of AI technologies in clinical practice. As Kaelin et al. (2024) noted, these considerations are essential to ensuring responsible and equitable implementation. Furthermore, although individual studies have examined AI-supported screening tools, assessment methods, and intervention programs, the evidence remains dispersed across populations, conditions, and technological approaches. Therefore, a comprehensive synthesis of the literature is needed to better understand the role of AI in pediatric occupational therapy.

The purpose of this systematic review was to examine and analyze the strength of the evidence regarding the use of artificial intelligence technologies in pediatric populations. Specifically, this review evaluated the use of AI for screening, evaluation, and intervention practices. It also examined the impact of these technologies on developmental, functional, behavioral, and participation-related outcomes that are relevant to occupational therapy practice.

Method

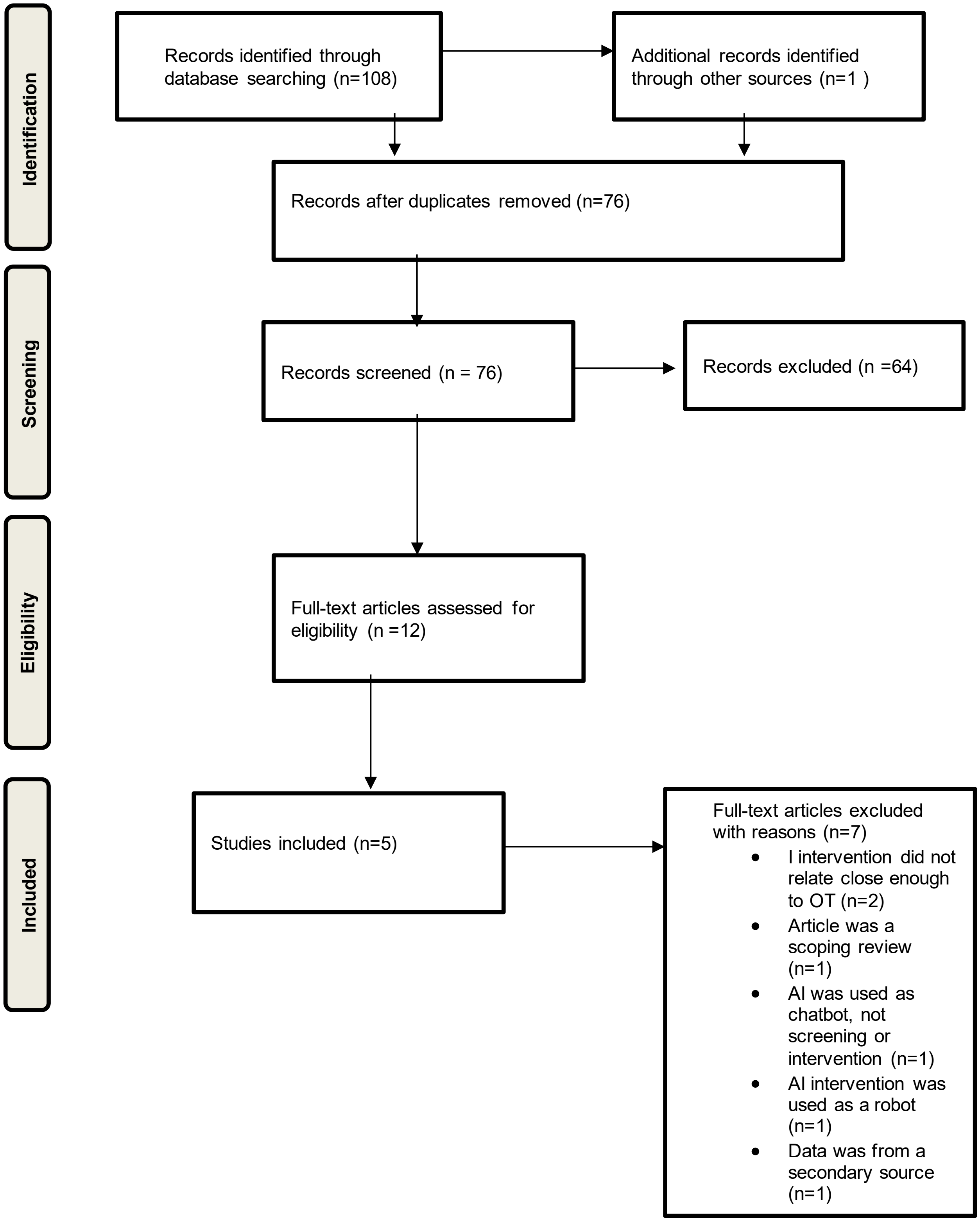
This systematic review was conducted with adherence to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). The search and collection of current literature for this systematic review was guided by the research question: How is artificial intelligence (AI) being used for screening, evaluation and intervention in pediatric occupational therapy?

A comprehensive literature search was conducted between May 14, 2026, and May 22, 2026, using electronic databases available through Hawaii Pacific University's online library, including MEDLINE, Academic Search Complete, and PubMed. A secondary search was conducted on May 26, 2026, to identify any additional publications relevant to the research question. The inclusion criteria used to determine study eligibility were: (1) peer-reviewed journal articles published in English within the previous five years (2021-2026); (2) studies examining the use of artificial intelligence (AI) within occupational therapy for pediatric populations (children and youth); and (3) studies evaluating effectiveness and client outcomes. Quantitative, qualitative, and mixed-methods studies were included to capture a comprehensive understanding of outcomes and recurring themes related to AI applications in occupational therapy. Studies were excluded if they consisted of scoping reviews, editorials without empirical data, or opinion pieces lacking measurable intervention outcomes. Search terms included "Artificial Intelligence" OR "AI" OR "ChatGPT" OR "Chatbot" AND "Occupational Therapy" OR "OT" AND "Intervention" OR "Program" OR "Therapy," along with combinations of these terms (Appendix A).

The initial search yielded 108 articles relevant to the research question (see Figure 1). Four independent reviewers completed the screening and selection of the studies, assessed their quality and extracted data. One additional article was sourced through the American Journal of Occupational Therapy (AJOT). After duplicates were extracted 76 studies remained. The authors excluded an additional 64 records as their abstracts indicated they were scoping reviews, editorials without empirical data sets, and opinion pieces lacking clear intervention outcomes. The full text of the remaining twelve articles was assessed for eligibility. Seven articles were excluded as the AI interventions were not strictly related to occupational therapy, they were a scoping review, AI was used as a chatbot rather than for screening or intervention, or data was collected from a secondary source. A total of five studies were included in the systematic review.

Figure 1

Flow Diagram



Results

Twelve studies met the inclusion criteria. The researchers assessed each article according to its risk of bias, level of evidence, and quality. This systematic review included five studies that contained relevant information regarding the use of artificial intelligence for occupational therapy intervention and screening with pediatric patients. The authors divided the information from these articles into two distinct themes: (1) Use of AI for Pediatric Developmental Delay Screening and (2) AI-Assisted Occupational Therapy Interventions for Children. Appendix B provides a comprehensive evidence table and Appendix C details the Cochrane risk-of-bias tables used to evaluate each article.

Use of AI for Pediatric Developmental Delay Screening

Three of the five studies on the topic discussed the efficacy of AI-based screening models. The three studies were Level 3B studies (see Appendix B). Collectively, these studies suggest that artificial intelligence can serve as an effective and potentially beneficial tool for identifying developmental delays and visual-motor integration deficits in pediatric populations.

Huang et al. (2023) designed an observational, cross-sectional study in Taiwan examining the accuracy of artificial intelligence models in determining visual motor integration (VMI) skills in simple tasks such as origami, coloring, and copying. The study included 370 kindergarten children, including 182 boys and 188 girls. All participants completed the Beery-Buktenica Developmental Test of Visual-Motor Integration, Sixth Edition (VMI-6), and completed activity-based tasks that were subsequently digitized and analyzed using computer vision models. Findings demonstrated that AI-based analysis of children's task performance could predict standardized VMI outcomes, with coloring activities demonstrating the strongest predictive relationship. These results support the potential use of computer vision technologies as accessible screening tools within pediatric rehabilitation settings.

Huang et al. (2025) executed a similar pre-post observational study in Taiwan to further examine the effectiveness of using AI as a screening tool for determining if a coloring activity can identify children’s VMI developmental status accurately based on the Beery-Buktenica Developmental Test of Visual-Motor Integration, Fourth Edition (VMI-4). The study included 505 preschool children split into 404 training subjects and 101 testing subjects. Children were asked to color a provided picture of a train and were instructed to stay in the lines and not draw anything new. Different AI machine learning model approaches were used to analyze the drawings with scores determined by the gold standard of the VMI-4. Results indicated that a hybrid AI architecture achieved the strongest predictive performance, suggesting that combining multiple machine learning approaches may enhance the accuracy of developmental screening procedures.

Chen et al (2025) initiated a retrospective database cohort study in Taiwan, leveraging historical medical records to build and evaluate predictive AI machine learning models in detection of developmental delays for pediatric patients. The study included data from 2,552 outpatient pediatric clients under the age of twelve who were suspected of having developmental delays. Three machine learning models (a deep neural network, support vector machine, and decision tree model) were evaluated. Among the models tested, the decision tree demonstrated the strongest sensitivity and predictive capacity. These findings highlight the potential value of AI-driven screening systems in supporting early identification efforts while utilizing routinely collected clinical data.

Limitations of the studies on the use of AI for pediatric developmental delay screening include the use of group activities as opposed to individual assessment as children could yield different results due to interactions with others during the assessment period (Huang et al. 2023; Huang et al. 2025). The specific age range of two to twelve years old limits the results of the studies to only be applicable to these age brackets (Chen et al., 2025; Huang et al. 2023; Huang et al. 2025). Similarly, all studies were conducted in Taiwan, limiting the generalizability of findings across diverse geographic and cultural contexts (Chen, et al., 2025; Huang et al. 2023; Huang et al. 2025). According to Appendix C, the studies did not collect primary data dynamically or administer objective clinical outcome measures directly within the trial framework which increases the overall risk of bias (Chen et al., 2025; Huang et al. 2023; Huang et al. 2025).

AI-Assisted Occupational Therapy Interventions for Children

Two of the five studies on the topic discussed efficacy of AI-assisted occupational therapy programs for pediatric patients. Two of these studies were Level 1B studies (see Appendix B). All studies provided evidence that AI-assisted occupational therapy interventions may enhance functional, behavioral, and participation-related outcomes among children.

Aldakhil (2024) employed a parallel-group randomized controlled trial in central Saudi Arabia examining the effectiveness of AI-based play activities on quality of life among children with ADHD. The study included 61 boys between the ages of 8 and 12 years. The intervention group participated in an AI-based play program incorporating ChatGPT, SimSimi, and Just Dance, while the control group received no intervention. Children who participated in the intervention demonstrated improvements across cognitive, emotional, social, and physical domains as measured by the PedsQL, and these improvements were maintained at follow-up. These findings suggest that AI-supported play activities may promote meaningful engagement and quality-of-life outcomes in children with ADHD.

Demirci et al. (2025) facilitated a randomized controlled trial in Ankara, Türkiye, examining the effects of an AI-assisted occupational therapy program targeting handwriting performance among children at risk for developmental coordination disorder. The study included 42 children between the ages of 8 and 12 years. Participants in the intervention group completed an AI-supported occupational therapy program based on the Model of Human Occupation. Compared with the control group, children who received the intervention demonstrated greater improvements in handwriting performance across multiple domains of the Minnesota Handwriting Assessment. These findings support the integration of AI-assisted approaches to address school-based functional skills within pediatric occupational therapy practice.

Limitations of the studies on AI-assisted occupational therapy interventions for children included sample populations that consisted exclusively of male participants (Aldakhil, 2024) and had a specific age range of 8-12 years of age (Aldakhil, 2024; Demirci et al. 2025), which limits the overall generalizability of the studies' results. Another limitation is regarding difficulty determining whether it was specifically the AI assistance that led to the improvements or another factor related to the intervention as the control groups received no intervention at all (Aldakhil, 2024; Demirci et al., 2025). According to Appendix C, the studies did not blind the evaluators during the trial (Aldakhil, 2024; Demirci et al. 2025) and did not blind self-reported outcomes (Aldakhil, 2024) which increased the overall risk of bias.

Discussion

The results of this systematic review suggest that the use of artificial intelligence technologies in pediatric occupational therapy may improve screening, the evaluation process, and intervention delivery. The five studies demonstrate how AI can enhance the efficiency and accessibility of these processes. Children showed improvements in developmental, functional, behavioral, and participation-related outcomes.

Screening is the first step in the occupational therapy process and identifies children who may benefit from further evaluation or early intervention rather than providing a formal diagnosis. The screening study used an AI model which was trained to predict whether a child will experience a developmental delay based on what services the child had received in the past (Chen et al., 2025). Following screening is the evaluation process which must be administered by licensed professionals to assess a child for therapeutic diagnosis and guide intervention planning. Two studies trained AI models to aid in the evaluation process by analyzing a coloring task or performance activities to predict children's VMI scores (Huang et al. 2023; Huang et al. 2025). Artificial intelligence was further examined for the delivery of occupational therapy interventions. AI supported interventions were used to improve handwriting skills and quality of life in children with ADHD (Aldakhil 2024, Demirci et al. 2025). The AI-driven intervention offered a holistic approach to improving the quality of life of children with ADHD by providing an alternative to the commonly used behavioral and psychological treatment approaches (Aldakhil 2024).

The researchers provided screening, evaluation, and intervention techniques led by AI technology, which expands occupational therapy services to populations that face barriers to accessing OT services, such as high cost, residing in a remote location, or lack of available services (Huang et al. 2025). The AI model utilized for screening streamlines the identification of developmental delays in children, without needing a one-on-one session with a licensed professional or paraprofessional. This makes the process more accessible to a broader range of settings. Using AI as a supplemental tool for evaluations made it possible to process the data in a way that exceeds the limitations of traditional methods, enhancing both the efficiency and accessibility of the evaluation process (Huang et al. 2023). Additionally, the coloring and origami tasks used in the interventions were simple and low cost, allowing them to be administered in any setting including remote and underserved locations (Huang et al. 2023; Huang et al. 2025). Interventions that can be completed remotely can also increase access, inclusivity, and participation, by providing a discreet option preferred by some cultures (Aldakhil, 2024).

There seems to be a relationship between play-based activities and individualized personal feedback, and children's motivation and participation. Four of the studies included in this review created activities that were fun and relatable which increased motivation and engagement in children (Huang et al., 2025). The two intervention studies used play-based interventions that provided personalized feedback in real time, which created individualized experiences which fostered greater motivation and engagement in the intervention (Aldakhil, 2024, Demirci et al., 2025). The evidence also indicated that personalized and targeted feedback can accelerate skill acquisition in children, through individualized guidance (Demirci et al., 2025).

Research suggests early intervention is vital to improving developmental outcomes in children; however, the growing demand for occupational therapy services combined with delays in screening leads to children being left undiagnosed during a critical period of development. AI provides a supplemental solution by aiding in the screening and evaluation processes, ultimately enhancing the overall efficiency and accessibility of pediatric occupational therapy services. While the five studies highlight the promising implications of AI, several methodological limitations must be acknowledged. Notably, these include small sample sizes, short-term study designs, and a general lack of geographic and cultural diversity across the existing literature on this emerging topic. Despite these research limitations, AI technologies offer a unique opportunity to address significant systemic barriers within pediatric healthcare. Obstacles such as high costs, remote locations, and the current shortage of available pediatric occupational therapists represent a critical gap in occupational justice. Implementing AI-driven systems has the potential to bridge this gap, ultimately improving clinical access, client engagement, and daily participation for underserved children.

Strengths and Limitations

This systematic review demonstrated several methodological strengths that support the credibility and transparency of the findings while also presenting limitations that should be considered when interpreting the results. Strengths of this systematic review include the use of a multi-person research team to reach consensus in article selection, adherence to PRISMA guidelines, and the development of a flow diagram to transparently document the screening and selection process. On the other hand, this systematic review encountered several limitations including challenges in locating studies that explicitly addressed the primary research question. This difficulty was compounded by an unusual external factor: publishers had recently retracted several prominent articles exploring AI within occupational therapy interventions. Consequently, a relatively small number of studies ultimately fulfilled the final inclusion criteria. Additional limitations include the possibility that relevant articles may not have been identified through the databases and search terms employed, and the inherent subjectivity involved in organizing articles into thematic categories.

Implications for Occupational Therapy Practice

The findings of this systematic review suggest that artificial intelligence has the potential to enhance pediatric occupational therapy practice across screening, evaluation, and intervention. AI-based tools may support occupational therapists in identifying developmental concerns earlier, improving assessment efficiency, and delivering individualized interventions tailored to each child's specific needs. Importantly, AI should be regarded as a complementary tool that augments, rather than replaces, occupational therapists' clinical reasoning, professional judgment, and family-centered care. Continued research is needed to establish the effectiveness, feasibility, and ethical implementation of AI technologies within pediatric occupational therapy settings.

Key Takeaways for Occupational Therapy Practice:

- AI-based screening and predictive models may assist occupational therapists in the early identification of developmental delays and functional challenges, supporting timely referral and intervention.
- AI-enhanced assessment tools can provide objective data to complement traditional evaluation methods and inform clinical decision-making.
- Personalized AI-driven interventions, including adaptive games, virtual reality, and cognitive training programs, may improve engagement, participation, and functional outcomes in pediatric populations.

- Occupational therapists should critically evaluate the evidence, accessibility, privacy considerations, and clinical applicability of AI technologies before integrating them into practice.

Conclusion

Studies included within this systematic review provide evidence on the effectiveness of artificial intelligence technologies in supporting pediatric occupational therapy screening, evaluation, and intervention practices. Additional research is necessary to determine the long-term effectiveness, clinical application, and integration of AI-based tools within pediatric occupational therapy while maintaining client-centered and evidence-based care. Current findings suggest that AI has the potential to enhance occupational therapy practice by supporting individualized interventions, improving clinical decision-making, and promoting meaningful participation and functional outcomes among pediatric populations. As AI technologies continue to advance, occupational therapists may utilize these tools as complementary resources that enhance, rather than replace, the clinical expertise and therapeutic relationship central to occupational therapy.

AI Disclosure Statement: During the preparation of this systematic review, the authors used Claude AI to improve grammar and readability. After using this tool, the authors reviewed and edited the content. The authors take full responsibility for the final content.

References

Aldakhil, A. F. (2024). Investigating the impact of an AI-based play activities intervention on the quality of life of school-aged children with ADHD. *Research in Developmental Disabilities*, *154*, 104858. https://doi.org/10.1016/j.ridd.2024.104858

Chen, S., Huang, C., Weng, S., & Oyang, Y. (2025). Detection of pediatric developmental delay with machine learning technologies. *PLOS ONE*, *20*(5), e0324204. https://doi.org/10.1371/journal.pone.0324204

Demirci, O., Yilmaz, G. G., & Köse, B. (2025). Artificial intelligence-supported occupational therapy program on handwriting skills in children at risk for developmental coordination disorder: Randomized controlled trial. *Research in Developmental Disabilities*, *161*, 105009. https://doi.org/10.1016/j.ridd.2025.105009

Huang, C., Yu, Y., Chen, K., Lin, G., & Hsieh, C. (2023). Using artificial intelligence to identify the associations of children's performance of coloring, origami, and copying activities with visual–motor integration. *American Journal of Occupational Therapy*, *77*(5), 7705205110. https://doi.org/10.5014/ajot.2023.050210

Huang, T., Chen, K., Lin, G., & Huang, C. (2025). Using a coloring activity to identify children's development of visual–motor integration: An application of artificial intelligence. *Annals of Medicine*, *57*(1), 2578725. https://doi.org/10.1080/07853890.2025.2578725

Kaelin, V.C., Nilsson, I., & Lindgren, H. (2024). Occupational therapy in the space of artificial intelligence: Ethical considerations and human-centered efforts. *Scandinavian Journal of Occupational Therapy*, *31*(1), 2421355. https://doi.org/10.1080/07853890.2025.2578725

Stover, A.D., & Jacobs, K. (2025). Embracing artificial intelligence (AI) in occupational therapy practice: Bridging workforce gaps and redefining care. *Work*, *80(3)*, 1021–1028. https://doi.org/10.1177/10519815241312447

Appendix A

Search Terms

Artificial Intelligence or AI or ChatGPT or Chatbox
AND
Occupational Therapy or OT
AND
Intervention or Program or Therapy

Appendix B

Evidence Table for Artificial Intelligence in Pediatric Occupational Therapy Screening, Evaluation, and Intervention

Author/Year	Level of Evidence Study Design Risk of Bias	Participants Inclusion Criteria Study Setting	Intervention and Control Groups	Outcome Measures	Results
Aldakhil (2024) https://doi.org/10.1016/j.ridd.2024.104858	**Level of Evidence**: Level 1B **Study Design:** Randomized controlled design (Parallel-group RCT). **Risk of Bias**: Low to Moderate risk. **Strengths/Limitations** Utilized computer-based randomization and confirmed baseline homogeneity; limited by a male-only sample and reliance on self/proxy-report questionnaires.	**Participants:** *N*= 61 school-aged boys (*M*= 10.0 years *SD*= 1.4; 0% female/ 100% male) **Inclusion Criteria:** Age 8–12 years; formal ADHD clinical diagnosis (DSM-5); symptoms > 6 months; onset < 6years old; active parental involvement; constant psychotropic medication dosage; capable of following AI tool instructions. **Study Setting:** Public elementary school therapeutic environments in central Saudi Arabia.	**Intervention (Experimental) Group:** *n*= 30 Received an AI-based play activity program consisting of twelve 45-minute sessions over 4 weeks (3 sessions/week) using tools like ChatGPT, SimSimi, and Just Dance AI to address cognitive, emotional, social, and physical domains. **Control Group**: *n*= 31 Passive control group that maintained standard school routines and traditional activities without AI exposure.	**Standardized Questionnaires:** ● Pediatric Quality of Life Inventory (PedsQL) Arabic version (Child Self-Report and Parent Proxy-Report) **Evaluated Dimensions:** ● Physical Functioning ● Emotional Functioning ● Social Functioning ● School Functioning	Repeated-measures ANOVA showed large, statistically significant improvements in total child self-report scores ($F(2, 131) = 435.47$, $p < .001$, $n^2 = .875$) and parent proxy-report scores ($F(2, 131) = 378.53$, $p < .001$, \n$^2 = .857$) from pre-test to post-test, which were sustained at 7-week follow-up **Nonsignificant Findings:** No primary quality-of-life dimensions or baseline-to-follow-up metrics demonstrated non-significant changes for the intervention cohort.

Chen et al. (2025) https://doi.org/10.1371/journal.pone.0324204	**Level of Evidence:** Level 3B **Study Design:** Retrospective database cohort study leveraging historical medical records to build and evaluate predictive machine learning models. **Risk of Bias:** Moderate risk of selection/spectrum bias due to an outpatient-only sample with high male gender imbalance (69.6% male). **Strengths/Limitations:** Large sample size utilizing real-world clinical records; limited by high male imbalance (69.6% male) and single-center hospital data constraints lacking external geographic validation.	**Participants**: *N* = 2,552 outpatients *(M* = 72.34 months; 30.4% female / 69.6% male) ● Cases form the outpatient department (OPD) visits database reported between January 1, 2012, and December 31, 2016. ● Outpatients under 12 years of age who made one or more Outpatient Department (OPD) visits. **Inclusion Criteria:** Pediatric patients visiting the hospital clinic suspected of developmental delay (DD) between January 1, 2012, and December 31, 2016, with formal DSM-5-TR diagnostic verification. **Study Setting**: The outpatient department (OPD) of the rehabilitation clinic at En Chu Kong Hospital in Taiwan.	As a non-interventional predictive modeling study, participants were divided into two analytic groups based on diagnostic confirmation: ● Analytic *Developmental Delay (DD)* Group (*n* = 1,719): Outpatients verified to have clinical developmental delays based on formal multidisciplinary evaluations. ● *Non-Developmental Delay (Non-DD Group) (Control Group):* Outpatients presenting with clinical suspicion or psychiatric co-morbidities who did not meet standard clinical diagnostic thresholds for	**Input Feature Medical Service Predictors:** ● Occupational Therapy Service (OTS) frequency. ● Physical Therapy Service (PTS) frequency. ● Speech Therapy Service (STS) frequency. **Algorithmic Frameworks Evaluated:** ● Decision Tree (DT). ● Deep Neural Network (DNN). ● Support Vector Machine (SVM).	**Significant Findings:** Machine learning algorithms successfully parsed therapy utilization patterns to detect diagnostic category boundaries. Decision Tree (DT) models achieved superior clinical classification performance when prioritizing high sensitivity, delivering a peak sensitivity of 0.902 and a positive predictive value (PPV) of 0.723. **Nonsignificant Findings:** Stratified sampling protocols based on gender were not implemented due to sample trait boundaries. Advanced complex DNN and

			Developmental Delay (DD).		SVM configurations did not show statistically superior F1-scores compared to simpler DT variants under acute sensitivity constraints.
Demirci et al. (2025) https://doi.org/10.1016/j.ridd.2025.105009	**Level of Evidence:** Level 1B **Study design:** Randomized control trial **Risk of Bias:** Low **Strengths/Limitation** Strong methodological rigor utilizing a true experimental control; limited by a localized sample and lack of longitudinal long-term follow-up beyond post-intervention.	**Participants:** *N*= 42 (Children 8-12 y/o) **Inclusion criteria:** Children aged 8–12 years identified as "at risk" or presenting with coordination challenges meeting explicit diagnostic standard thresholds for DCD. **Study Setting:** Clinic of the Department of Occupational Therapy: Gulhane Health Science University Ankara, Turkiye.	**Intervention group:** *n*=21 Completed an AI-supported occupational therapy handwriting program utilizing computer vision or adaptive machine learning feedback systems to shape kinematic patterns and execution mechanics. **Control group:** *n*=21 Received conventional, standard-of-care occupational therapy handwriting interventions focusing	**Standardized Clinical Measures:** ● Evaluation Tool of Children's Handwriting (ETCH) indices. ● Movement Assessment Battery for Children (MABC) screening indicators. **Kinematic Performance Variables:** ● Handwriting speed. ● Line legibility. ● Range of pen pressure spatial adherence.	**Significant Findings:** The intervention group showed significant improvements in writing speed, range, legibility, shape, alignment, and size outcome measures. Using AI-supported handwriting intervention administered by an occupational therapist more than likely contributed to these improvements due to the feedback that is provided in real time from the AI and the increase in opportunity for practicing repetitively.

			on traditional motor practice without AI integration.		**Nonsignificant Findings:** Baseline demographic arrays and pre-test motor scores demonstrated non-significant differences between the experimental and control cohorts ($p > .05$), establishing proper pre-intervention homogeneity.
Huang, C. et al. (2023) https://doi.org/10.5014/ajot.2023.050210	**Level of Evidence:** Level 3B **Study Design:** Observational, cross-sectional study Examining the accuracy of AI models in determining VMI skills in simple tasks like origami, coloring and copying. **Risk of Bias:** Low to Moderate risk. ● Nonrandomized ● Group testing environment	**Participants:** Validation cohort of preschool/ kindergarten children (*N*=370; *M* age 5.2 years, *SD*= 0.63; age range 40-77 months; *n*= 182 boys [49.2 %] *n*= 188 girls [50.8%] **Inclusion Criteria:** ● Children in 2nd and 3rd years of kindergarten. ● Did not attend special schools. ● Liven in Kaoshing or Tainan City. ● Able to participate in group activities and follow instructions. **Study Setting:**	**Single Observation Cohort** (*N* = 370): Non-experimental design; all participants completed identical drawing tasks. Activity products (origami, coloring, and copying) were digitized via photograph and analyzed through automated computer vision models to predict gold-standard motor testing results.	**Gold Standard Comparison Metric:** ● Beery-Buktenica Developmental Test of Visual-Motor Integration, 6th Edition (Beery VMI-6) containing 30 items (3 scribbling, 27 shapes. copying). ○ **VMI-6** was used for standard scoring	**Significant Findings:** Computer vision successfully extracted and quantified physical paper properties to predict gold-standard motor testing results. Coloring showed the strongest individual task correlation ($R^2 = 0.355 - 0.473$) Followed by Copying ($R^2 = 0.316 - 0.444$) While Origami had the lowest individual predictive capacity

	Strengths/Limitations: Innovatively automates standard paper-and-pencil performance analysis through computer vision; limited by convenience sampling in a single geographic zone, which might affect generalizability.	Early childhood educational kindergarten settings in southern Taiwan.		compared with AI scoring. ◦ Measured visual motor integration (VMI). ◦ 30 items (3 for scribbling and 27 where they copied a shape. **Machine Learning Frameworks:** ● Image feature extraction via Deep Convolutional Neural Networks (EfficientNetB7 processing 2,560 features). ● Predictive scoring via Support Vector Machine (SVM) Regressor and Random Forest algorithms.	$(R^2 = 0.269 - 0.340)$ The duo combination of Origami + Copying delivered the optimal balance, achieving a high coefficient of determination $(R^2 = 0.390\ to\ 0.577)$ **Nonsignificant Findings:** Combining all three tasks together did not yield a statistically significant improvement in classification variance $(R^2 = 0.400\ to\ 0.550)$ compared to the streamlined dual-task Origami + Copying matrix. Therefore, incorporating the time-consuming coloring task provided no significant added statistical benefit.

				AI models analyzed the children's activities and predicted their VMI scores. ● **Coloring:** colored black/white train pic and asked to stay in lines ● **Origami:** made a dog following 4 steps, that was then photographed 8 times ● **Copying:** asked to copy a drawing of a person in a black box using the same shapes and scale Both scores were compared to see if the activities could accurately be used to show VMI skills and if AI model scoring was accurate to the VMI-6.	

Huang, T. et al.(2025) https://doi.org/10.1080/07853890.2025.2578725	**Level of Evidence:** Level 3B **Study Design:** ● Pre-post observational study. ● No control group. ● Nonrandomized. **Risk of Bias:** Low to Moderate risk. ● No control group and randomization. ● Group testing environment. ● But had a large sample size and clear procedure. **Strengths/Limitations** Large sample size optimized with upsampling techniques; limited by data collection in group kindergarten classes where peers may cause distractions, and centering on a single train drawing layout.	**Participants:** *N* = 505 preschool children (*M* = 57.64 months, *SD* = 11.10; split into 404 training and 101 testing subjects; majority sample presented with typical development patterns). **Inclusion Criteria:** Typically developing children and cohorts presenting with fine motor/coordination vulnerabilities capable of finishing structured drawing, coloring, or folding manual dexterity tasks. **Study Setting:** Kindergarten class activity environments and pediatric research settings in Taiwan.	**Single Model Development Dataset** (*N* = 505): Non-experimental study design. Imbalanced baseline training distributions corrected using the Synthetic Minority Oversampling Technique (SMOTE) to evenly assign typical vs. delayed cohorts (*n*=326 each, totaling 652 unsampled subjects).	**Automated Computer Vision Engine:** ● Image feature extraction via Deep Convolutional Neural Networks (EfficientNetB7 processing 2,560 distinct shape markers). **Gold Standard Assessment Tool:** ● Beery-Buktenica Developmental Test of Visual-Motor Integration, Fourth Edition (VMI-4) age-standardized scores. ○ **VMI-4** used as standard measured developmental status (typical/delay) 27 items where children are	**Significant Findings:** The integrated hybrid AI architecture (Approach 3) combining SVM regression with an XGBoost classifier delivered optimal clinical predictive capacity. On the independent testing dataset, the model successfully identified VMI delays with an overall accuracy of 80.20%, sensitivity of 73.68%, and specificity of 81.71%. SMOTE up sampling significantly enhanced structural model sensitivity to prevent diagnostic misclassification. **Nonsignificant Findings:** There were no significant demographic variations in mean age or raw baseline total VMI-4 scores between the split training and

				asked to copy geometric forms. **Coloring: Children were asked to color** a provided picture of a train. They were told to stay in the lines and to not draw other things. **AI** analyzing drawing with scores of gold standard VMI-4. **Algorithmic Frameworks:** ● Approach 1–2: Direct linear regressors. ● Approach 3: Layered Support Vector Machine (SVM) regression coupled with an eXtreme Gradient Boosting (XGBoost) classifier.	testing subsets ($p > .05$). Standalone regression protocols (Approach 1) focusing on simple raw score conversion demonstrated inadequate predictive power ($R^2 = 0.58$ to 0.65), proving basic regression lines are clinically insufficient without secondary classifier layering.

Appendix C

Risk-of-Bias Table

Risk-of-Bias Table for Randomized Controlled Trial (RCT) and Non-RCT (Two or More Group Design)										
	Selection Bias (Risk of bias arising from randomization process)			Performance Bias (effect of assignment to intervention)		Detection Bias		Attrition Bias	Reporting Bias	Overall risk-of-bias (low, moderate, high
Citation	Random Sequence Generation	Allocation Concealment (until participants enrolled and assigned)	Baseline difference between intervention groups	Blinding of Participants During the Trial	Blinding of Study Personnel During the Trial	Blinding of Outcome Assessment: Self-reported outcomes	Blinding of Outcome Assessment: Objective Outcomes (assessors aware of intervention received?)	Incomplete Outcome Data (data for all or nearly all participants	Selective Reporting (results being reported selected on basis of the results?)	
Aldakhi (2024)	+	?	+	–	–	–	+	+	+	L
Demirci et al. (2025)	+	+	+	–	–	+	+	+	+	L

Note. Categories for risk of bias are as follows: Low risk of bias (+), unclear risk of bias (?), high risk of bias (–). Scoring for overall risk of bias assessment is as follows: 0–3 minuses, low risk of bias (L); 4–6 minuses, moderate risk of bias (M); 7–9 minuses, high risk of bias (H).

Citation. Table format adapted from Higgins, J. P. T., Sterne, J. A. C., Savović, J., Page, M. J., Hróbjartsson, A., Boutron, I., . . . Eldridge, S. (2016). A revised tool for assessing risk of bias in randomized trials. Cochrane Database of Systematic Reviews 2016, Issue 10 (Suppl. 1), 29–31. https://doi.org//10.1002/14651858.CD201601

Risk of Bias for Before-After (Pre-Post) Studies with No Control Group (One Group Design)												
Citation	Study question or objective clear	Eligibility or selection criteria clearly described	Participants representative of real-world patients	All eligible participants enrolled	Sample size appropriate for confidence in findings	Intervention clearly described and delivered consistently	Outcome measures pre-specified, defined, valid/reliable, and assessed consistently	Assessors blinded to participant exposure to intervention	Loss to follow-up after baseline 20% or less	Statistical methods examine changes in outcome measures from before to after intervention	Outcome measures were collected multiple times before and after intervention	Overall risk of bias assessment (low, moderate, high risk)
Chen et al. (2025)	Y	Y	Y	Y	Y	N	Y	N	Y	N	N	M
Huang et al. (2023)	Y	Y	Y	Y	Y	N	Y	NR	Y	N	N	L
Huang et al. (2025)	Y	Y	Y	Y	Y	N	Y	NR	Y	N	N	L

Note. Y = yes; N = no; NR = not reported. Scoring for overall risk of bias assessment is as follows: 0–3 N, Low risk of bias (L); 4–8 N, Moderate risk of bias (M); 9–11 N, High risk of bias (H).

Citation. Table format adapted from National Heart Lung and Blood Institute. (2014). Quality assessment tool for before–after (pre–post) studies with no control group. Retrieved from https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools