Artificial Intelligence for Long-Term Monitoring of ADHD

Narin Fazlalipour[1], Anaya Sinha[2], Lucas Lam[3], Jade Zaw[4], Shatakshi Sinha[5], Adya Singh[6], Abigail Jacob[7], Kaylene Tran[8], Saukhya Sannapu[9]

[1]University of California, San Diego · [2]Princeton High School · [3]Los Altos High School · [4]Trinity School, NYC · [5]Hillcrest High School
[6]Kimball High School · [7]Hebron High School · [8]Scripps Ranch High School · [9]Vista Ridge High School

BACKGROUND

Attention-deficit/hyperactivity disorder (ADHD) is a chronic neurodevelopmental disorder characterized by persistent inattention, hyperactivity, and impulsivity that can interfere with learning, relationships, and daily functioning across childhood and adulthood [1]. Symptoms and support needs can also change with age, environment, treatment, sleep, and daily routines.

Clinical monitoring still relies largely on periodic appointments, behavioral rating scales, parent or teacher reports, and brief observations. Although these methods are essential, they capture isolated snapshots and may miss gradual changes occurring between visits or across months and years.

AI offers a possible way to analyze repeated information from electroencephalography, functional near-infrared spectroscopy, wearable sensors, sleep and activity measures, and behavior recorded during everyday tasks [2–5]. These data could complement clinical assessments by revealing patterns that are difficult to observe consistently.

Most existing research, however, evaluates diagnosis, classification, or short-term prediction. Small samples, limited diversity, inconsistent protocols, device burden, and privacy concerns continue to restrict the transition from promising AI models to validated long-term ADHD monitoring.

OBJECTIVE

This literature review and bibliometric analysis examine how AI could support long-term ADHD monitoring and identify the evidence still required for continuous clinical use.

> **How can AI be used to monitor long-term neurodevelopmental changes in individuals with ADHD?**

- Map current AI applications and major themes in ADHD research.
- Compare diagnostic performance with the requirements of longitudinal monitoring.
- Identify gaps in duration, diversity, standardization, usability, privacy, and clinical validation.

Review Scope

The bibliometric analysis maps publication patterns and research emphasis; it does not estimate treatment efficacy. Two representative studies were examined in greater detail to compare current AI capabilities with the evidence needed to measure meaningful within-person change over time.

METHODS

Literature identification: A dataset of 100 publications was collected from Web of Science using terms related to "neurodevelopment," "artificial intelligence," "ADHD," and "monitoring."

Screening: Records were reviewed for relevance and checked for duplicates, producing a final dataset of 100 unique papers.

Data extraction: Each record was tagged with author names, institutional affiliations, journal title, publication year, citation information, and author-provided keywords.

Bibliometric analysis: Metadata were organized in Microsoft Excel and imported into RStudio or Biblioshiny. Visualizations summarized annual publication patterns, research categories, and keyword co-occurrence relationships.

Focused review: Selected wearable-sensor and fNIRS studies were examined for sample characteristics, monitoring duration, model performance, and relevance to longitudinal ADHD tracking.

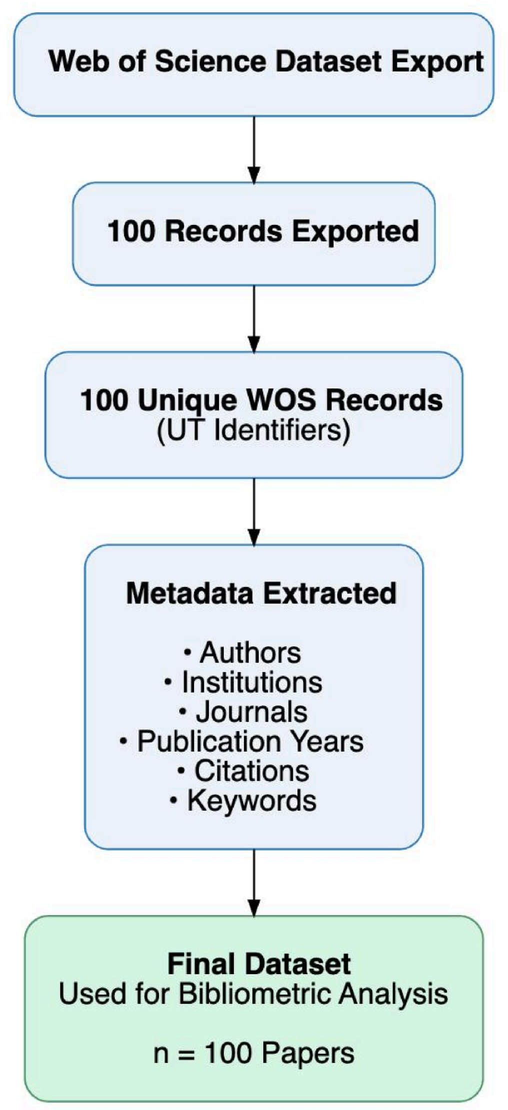

Figure 1. Web of Science methodology. One hundred unique papers were retained for bibliometric analysis. *Original figure by Adya Singh.*

RESULTS: PUBLICATION PATTERNS

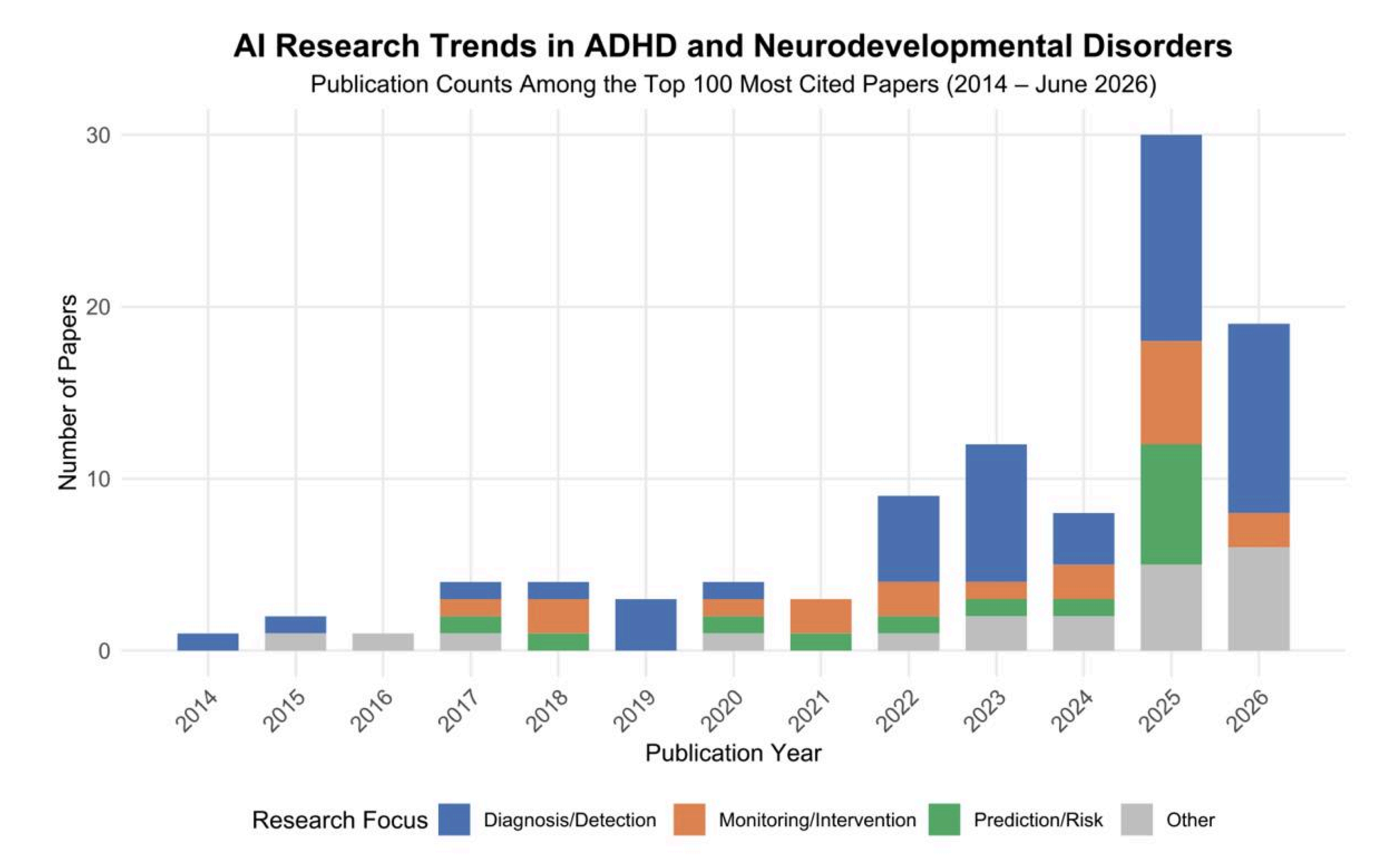

Figure 2. Publication patterns among top-cited papers. Counts rose sharply after 2021, reaching 30 papers in 2025; 2026 includes January through June only. *Original figure by Lucas Lam.*

Results at a Glance

Growth
Publication activity accelerated after 2021.

2025 peak
Thirty papers appeared in the corpus.

2026 partial year
Nineteen papers appeared through June.

Research emphasis
Diagnosis and detection remain more prominent than monitoring and intervention.

Evidence boundary
Publication growth indicates activity—not readiness for longitudinal clinical use.

RESULTS: RESEARCH THEMES

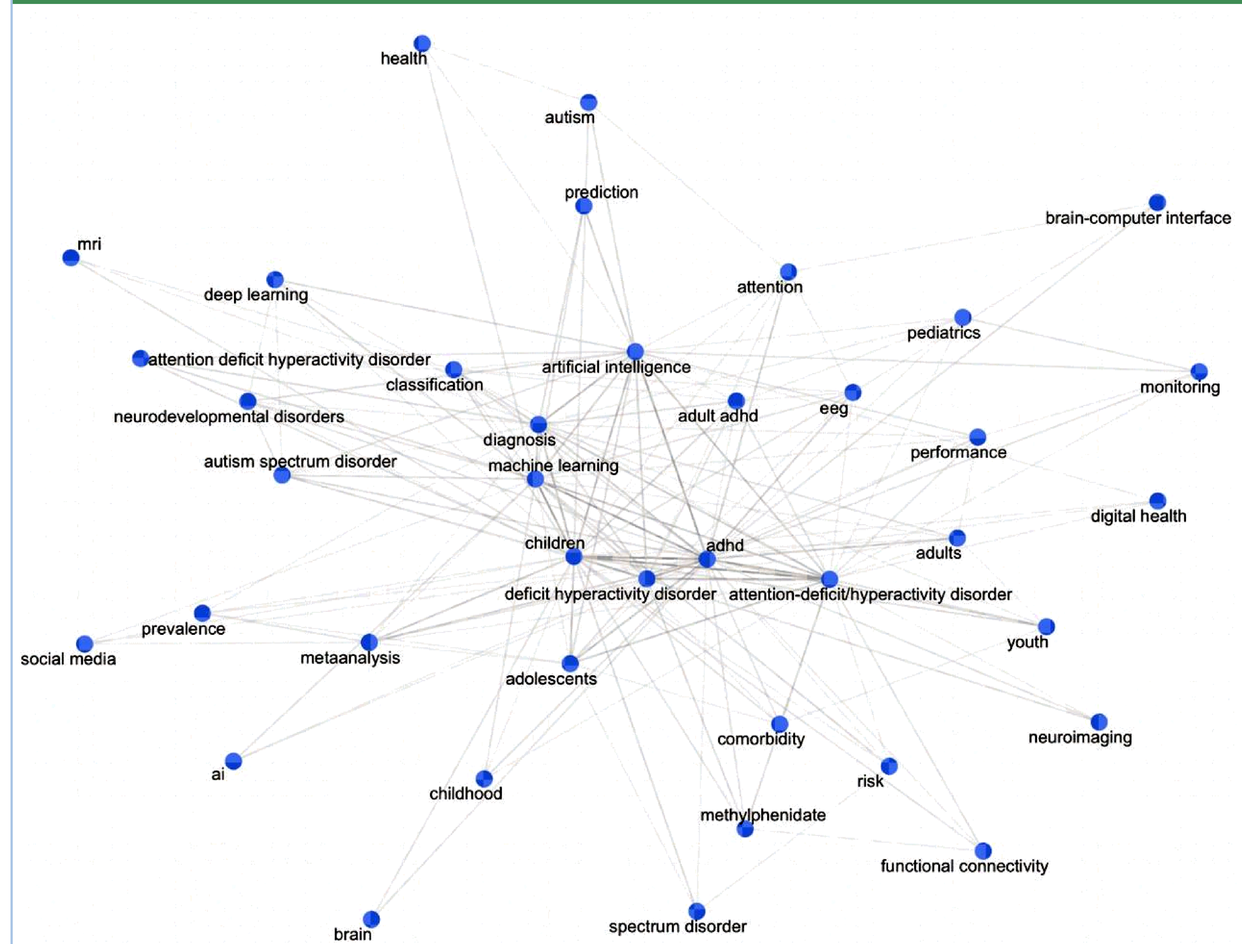

Figure 4. Keyword co-occurrence network. ADHD, AI, machine learning, diagnosis, and classification are central; monitoring and digital-health terms remain less integrated. *Original figure by Shatakshi Sinha.*

How Figure 4 Relates to the Objective

- **Central themes:** Dense links among ADHD, AI, machine learning, diagnosis, and classification show that detection remains the dominant use case.
- **Population emphasis:** Child and adolescent terms connect strongly with diagnostic research, reflecting the field's clinical focus.
- **Monitoring gap:** Monitoring and digital-health terms remain peripheral, supporting the need for research on repeated within-person change.

RESULTS: SELECTED STUDIES

Wearable monitoring
A 12-week study collected movement, heart rate, sleep, and questionnaire data from 30 Chinese adolescents. Objective measures classified ADHD with an AUC of 0.844; combined objective and subjective measures reached 0.933 [3]. The small, narrow sample limits generalizability.

fNIRS functional connectivity
Machine-learning analysis compared 75 medication-naïve adults with ADHD and 75 controls. A linear support vector machine achieved 0.800 accuracy and 0.799 F1-score [4]. This supports classification, not long-term within-person tracking.

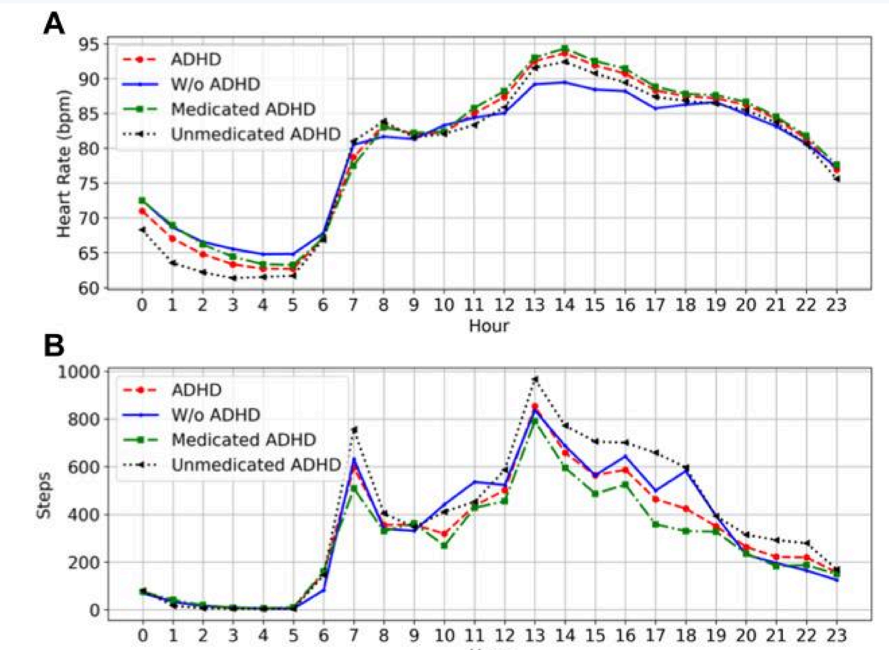

Figure 3. Mean Values of 24-Hour Patterns of Different Groups. ADHD and without ADHD (W/o ADHD) lines are the mean of all participants in the ADHD and without ADHD groups, respectively. The medicated and unmedicated lines are the mean of medicated ADHD and unmedicated ADHD participants, respectively. (A) Heart-rate patterns. (B) Step patterns. bpm = beats per minute. *Reproduced from Jiang et al. [3].*

Current systems detect ADHD-related patterns; evidence that they can track clinically meaningful change over years remains limited.

RESULTS: DIAGNOSTIC VS. LONGITUDINAL EVIDENCE

Evidence	What it establishes	What remains missing
12-week wearables [3]	Passive signals can supplement assessment.	Larger, more diverse cohorts followed longer.
24-hour group patterns [3]	Heart-rate and step profiles vary across comparison groups.	Individual trajectories measured across months or years.
Task-state fNIRS [4]	Connectivity features support classification.	Repeated within-person measurement.
Bibliometric trends	AI-related research is expanding.	Evidence that monitoring improves care.

Results synthesis: The field is moving toward repeated real-world sensing, but long-term within-person validity remains the central evidence gap.

What is established
AI can identify group-level patterns and extract informative signals over defined observation windows.

What is not established
Current studies do not show that those signals reliably measure clinically meaningful change within one person over years.

What must come next
Prospective studies must connect individual trajectories with symptoms, functioning, treatment response, and improved care.

DISCUSSION

Publication growth and strong diagnostic results show momentum, but not readiness for long-term monitoring. Cross-sectional models classify groups at one time point; longitudinal systems must detect meaningful change within the same person while accounting for medication, sleep, context, adherence, and missing data.

Interpretation

- Diagnosis and classification remain central; monitoring-related terms are less integrated.
- Wearables and smart environments demonstrate feasibility, not multi-year clinical validity.
- Repeated multimodal data could complement—but should not replace—clinical assessment.

Clinical and Ethical Implications

- Outputs must communicate uncertainty and preserve clinician oversight.
- Continuous monitoring requires consent, data minimization, secure storage, and child-specific safeguards.
- Models must be tested across diverse populations to reduce bias and avoid harmful surveillance.

Limitations

- The top-100 cited-paper sample may not represent the full field.
- Representative studies used limited samples and follow-up periods.
- Sensors, protocols, outcomes, and evaluation methods remain inconsistent.

Future Directions

- Build diverse longitudinal cohorts anchored to individual baselines.
- Standardize repeated measurement, adherence, and missing-data reporting.
- Validate models against clinically meaningful within-person change.
- Develop low-burden, interpretable, privacy-preserving systems.

EVIDENCE VERDICT

Established: AI can classify ADHD-related patterns and analyze multimodal signals.

Promising: Wearables and smart environments can collect repeated real-world information.

Not established: Reliable multi-year tracking of clinically meaningful neurodevelopmental change.

EVIDENCE-TO-PRACTICE PATH

1. **Diverse cohorts** — Represent age, sex, race, comorbidity, and treatment context.
2. **Repeated multimodal data** — Measure behavior, activity, sleep, and neurophysiology over time.
3. **Individual trajectories** — Compare each person with their own baseline, not only a control group.
4. **Transparent governance** — Protect privacy, explain uncertainty, and retain clinician oversight.
5. **Prospective validation** — Test whether alerts correspond to meaningful clinical change.

CONCLUSION

AI's most defensible role in ADHD care is not continuous diagnosis, but longitudinal change detection. A monitoring system would establish a personal baseline, distinguish sustained deviation from ordinary day-to-day variability, and present clinicians with interpretable trajectories rather than a single label or score. Behavioral, physiological, neurophysiological, and contextual signals could be combined, but any alert should state its confidence, data sufficiency, and the observations that produced it.

The central challenge is therefore a mismatch between what current models optimize and what long-term care needs. Classification studies are designed to separate ADHD from comparison groups; monitoring must recognize meaningful change within the same person. Group-level accuracy cannot establish that capability. The decisive benchmark is whether a model's trend or alert corresponds prospectively to changes in symptoms, daily functioning, treatment response, or quality of life—and whether acting on it improves care.

Future research should be organized around that benchmark. Studies should predefine clinically meaningful change, collect repeated measurements across diverse cohorts, document adherence and missingness, and compare algorithmic trajectories with patient-, caregiver-, and clinician-reported outcomes. Medication, development, sleep, routines, and environment should be modeled as explanatory context rather than noise. Deployment should require informed consent, data minimization, interpretable uncertainty, and human review. Success would mean detecting important change earlier without replacing clinical judgment or turning everyday life into unaccountable surveillance.

REFERENCES

[1] Thapar, A., & Cooper, M. (2016). Attention deficit hyperactivity disorder. *The Lancet, 387*(10024), 1240–1250. https://doi.org/10.1016/S0140-6736(15)00238-X

[2] Bongurala, A. R., Save, D., & Virmani, A. (2024). Revolutionizing attention deficit hyperactivity disorder with artificial intelligence. *PLOS Mental Health, 1*(6), e0000181. https://doi.org/10.1371/journal.pmen.0000181

[3] Jiang, Z., Chan, A. Y. L., Lum, D., Wong, K. H. T. W., Leung, J. C. N., Ip, P., Coghill, D., Wong, R. S., Ngai, E. C. H., & Wong, I. C. K. (2025). Wearable signals for diagnosing ADHD in adolescents: A feasibility study. *JAACAP Open, 3*(4), 875–889. https://doi.org/10.1016/j.jaacop.2024.11.003

[4] Lim, S., Dong, S.-Y., McIntyre, R. S., Chiang, S. K., & Ho, R. (2025). Exploring functional connectivity in ADHD: An fNIRS study with machine learning analysis. *IEEE Journal of Biomedical and Health Informatics, 29*(10), 7473–7485. https://doi.org/10.1109/JBHI.2025.3564487

[5] Berrezueta-Guzman, J., Pau, I., Martin-Ruiz, M.-L., & Maximo-Bocanegra, N. (2021). Assessment of a robotic assistant for supporting homework activities of children with ADHD. *IEEE Access, 9*, 93450–93465. https://doi.org/10.1109/ACCESS.2021.3093233

Data and figure notes: Bibliometric data were drawn from Web of Science. Figures 1, 2, and 4 are original team figures. Figure 3 is reproduced from Jiang et al. [3].