JOUR UMRANIYE PEDIATR
2026;6(1):1–9 doi: 10.14744/upd.2026.84755

REVIEW
*Pediatrics*

# Artificial intelligence for the busy paediatrician

**Doğan Şengül**

**Department of Software Engineering, İstanbul Sabahattin Zaim University, Faculty of Engineering and Natural Sciences, İstanbul, Türkiye**

## ABSTRACT

Artificial intelligence (AI) is rapidly entering paediatric practice, yet clinicians often lack a concise, practical orientation. We summarise what paediatricians should know about contemporary AI in a compact and clinically useful context. We performed a narrative state-of-the-art review of peer-reviewed literature published between January 2017 and April 2026 (last search date: 30.04.2026) using PubMed, Scopus, and special issues on AI in paediatrics. Search terms combined "artificial intelligence," "machine learning," "deep learning," "large language models," and "natural language processing" with "paediatric/pediatric," "child," "neonatal," and paediatric subspecialty keywords. Priority was given to peer-reviewed studies, regulatory and ethics guidance, and recent reviews addressing clinical implementation, validation, bias, explainability, and governance. Because the review was narrative rather than systematic, no PRISMA flow diagram was prepared. Findings were organised around five clinical domains, with cross-cutting pitfalls synthesised. Mature applications include image-based diagnostics (radiology, ophthalmology, and dysmorphology), deterioration and sepsis prediction in the paediatric intensive care unit, asthma exacerbation forecasting, autism screening, accelerated rare-disease diagnosis from whole-genome data, precision dosing, and LLM-based clinical documentation and parental education tools. In selected narrow tasks under controlled test conditions, reported performance may approach, and in some studies overlap with, expert-level performance; however, generalisability is highly dependent on dataset quality, population similarity, and external validation. Conversely, paediatric AI faces unique threats: small and developmentally heterogeneous datasets, label scarcity, racial and gender bias, dataset shift, opaque "black box" reasoning, and complex consent for minors. Robust external validation in diverse paediatric populations remains the exception rather than the rule. AI is best understood as augmented intelligence that supports, rather than replaces, the paediatrician. Practical adoption requires AI literacy, structured evaluation of tools against age-stratified evidence, vigilance for bias, and explicit local governance. Paediatricians who engage early will help shape safe, equitable, child-specific AI.

**Keywords:** Artificial intelligence; clinical decision support systems; deep learning; diagnostic imaging; machine learning; natural language processing; paediatrics.

**Cite this article as:** Şengül D. Artificial intelligence for the busy paediatrician. Jour Umraniye Pediatr 2026;6(1):1–9.

**Received (Başvuru):** 12.05.2026 **Revised (Revizyon):** 15.06.2026 **Accepted (Kabul):** 16.06.2026 **Online (Online yayınlanma):** 10.07.2026

**Correspondence (İletişim):** Dr. Doğan Şengül. İstanbul Sabahattin Zaim Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Yazılım Mühendisliği Bölümü, İstanbul, Türkiye.
**Phone (Tel):** +90 505 929 39 81 **e-mail (e-posta):** dogan.sengul@izu.edu.tr

# Yoğun pediatristler için yapay zekâ

**ÖZET**

Yapay zekâ (YZ), günlük pediatri pratiğine hızla girmektedir; ancak yoğun rutinde klinisyenler kısa ve uygulanabilir bir kaynak bulmakta güçlük çekebilmektedir. Bu derleme, her pediatristin güncel YZ uygulamaları hakkında bilmek isteyebileceği kısa, klinik açıdan kullanışlı ve pratik bilgileri sunmayı amaçlamaktadır. Ocak 2017-Nisan 2026 (son tarama tarihi: 30 Nisan 2026) arasında yayımlanan hakemli literatürün anlatısal nitelikte güncel bir derlemesi yapıldı. PubMed, Scopus ve pediatrik YZ'ye ayrılmış güncel özel sayılar tarandı. Arama terimleri (İngilizce) "yapay zekâ", "makine öğrenmesi", "derin öğrenme", "büyük dil modelleri" ve "doğal dil işleme" kavramları ile "pediatri", "çocuk", "neonatal" ve pediatrik yan dal anahtar sözcükleri birleştirilerek oluşturuldu. Klinik uygulama, doğrulama, yanlılık, açıklanabilirlik ve yönetişimi ele alan hakemli çalışmalara, düzenleyici ve etik kılavuzlara ve güncel derlemelere öncelik verildi. Sistematik olmaktan çok anlatısal bir derleme amaçlandığından PRISMA akış diyagramı hazırlanmadı. Bulgular beş klinik alanda toplandı ve sentezlendi. Olgunluğa ulaşmış uygulamalar arasında görüntü temelli tanı (radyoloji, oftalmoloji, dismorfoloji), pediatrik yoğun bakımda kötüleşme ve sepsis öngörüsü, astım atağı tahmini, otizm taraması, tüm genom verisinden hızlandırılmış nadir hastalık tanısı, hassas dozlama ve büyük dil modeli tabanlı klinik dokümantasyon ile ebeveyn eğitimi araçları yer almaktadır. Dar tanımlı seçili görevlerde bildirilen performans, uzman düzeyine yaklaşabilmekte ve bazı çalışmalarda bu düzeyle örtüşebilmektedir; ancak genellenebilirlik veri kalitesine, popülasyon benzerliğine ve doğrulama (validasyon) süreçlerine büyük ölçüde bağlıdır. Buna karşın pediatrik YZ; küçük ve gelişimsel açıdan heterojen veri kümeleri, ırk ve cinsiyet yanlılıkları, veri kayması, şeffaflıktan yoksun "kara kutu" modeller ve reşit olmayanlarda karmaşık onam süreçleri gibi kendine özgü tehditlerle karşı karşıyadır. Farklılık ve çeşitlilik gösteren pediatrik popülasyonlarda dış doğrulama hâlâ kuraldan çok istisnadır. YZ, pediatristin yerine geçen değil, onu destekleyen artırılmış zekâ olarak anlaşılmalıdır. Pratik benimseme; YZ okuryazarlığı, araçların yaşa göre tabakalandırılmış kanıtlarla yapılandırılmış değerlendirilmesi, yanlılığa karşı dikkat ve duruma özgü yönetişim gerektirmektedir. YZ süreçlerine katılmak için gecikmeksizin çaba gösteren pediatristler, çocuklara özgü güvenli ve adil bir YZ'nin biçimlenmesine katkı sağlayacaktır.

**Anahtar Kelimeler:** Derin öğrenme; doğal dil işleme; klinik karar destek sistemleri; makine öğrenmesi; pediatri; tanısal görüntüleme; yapay zekâ.

**ORCID ID**

DŞ: 0000-0002-2285-3907

İstanbul Sabahattin Zaim Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Yazılım Mühendisliği Bölümü, İstanbul, Türkiye

## INTRODUCTION

Few topics have generated as much excitement, anxiety, and clinical inertia within paediatrics as artificial intelligence (AI). The general paediatrician is increasingly expected not only to interpret clinical data but also to understand rapidly evolving AI-related claims, tools, and risks: a regulator clears another algorithm; a national academy issues guidance on chatbots in clinic; a hospital pilots a sepsis predictor on the wards; a parent arrives with a printed exchange from a large language model (LLM) and a list of pointed questions about a rash. Within roughly a decade, AI has moved from the periphery of medicine to a position where it touches triage, imaging interpretation, electronic health record (EHR) documentation, dosing, and even how parents prepare for the visit (1–3).

Paediatrics carries unique features that intensify both the promise and the risk of AI. Children are not small adults: physiology, dosing, image appearance, disease epidemiology, and reference ranges all change with age and developmental stage. Paediatric datasets are smaller, rare outcomes are common, and the data labels available to train models often reflect the clinical world of adults rather than that of children (4, 5). Add to this the special obligations of paediatric care: guardianship, assent, lifelong consequences of misclassification, and the digital literacy of the family. It becomes clear that paediatricians cannot delegate the AI strategy.

This narrative state-of-the-art review is organised as a series of compact "AI bits": short, practical takeaways for the busy clinician. We first summarise the core terminology so that the rest of the article can be read fluently (Table 1 and Fig. 1). We then survey contemporary applications across five clinical domains, with selected high-quality examples. We close with the cross-cutting pitfalls of paediatric AI, an evaluation framework that fits into a clinic day (Fig. 2), and the educational and governance implications for the next five years. The aim of this work is orientation rather than exhaustive coverage. Once these pages have been read, the paediatrician should be in a position to scrutinise an AI claim, to put sharper questions to vendors and editors, and to judge whether a given tool earns its place in the daily workflow.

## SEARCH STRATEGY AND SCOPE

Consistent with a narrative-review design, sources were identified through PubMed, Scopus, and recent special issues dedicated to AI in paediatrics, with hand-searching of the reference lists of pivotal reviews. The search window ran from January 2017 to April 2026, with the final pass completed on April 30, 2026. Within this window, the terms "artificial intelligence," "machine learning," "deep learning," "large language models," and "natural language processing" were paired with "paediatric/pediatric," "child," "children," "neonatal," and selected paediatric subspecialty keywords. Peer-reviewed paediatric studies, regulatory and

**Table 1. Glossary of essential AI terminology for the practising paediatrician**

| Term | Plain-language definition | Paediatric example |
|---|---|---|
| Artificial intelligence (AI) | Umbrella term for computer systems that perform tasks usually requiring human cognition. | AI-enabled symptom checkers in primary care. |
| Machine learning (ML) | A subset of AI in which algorithms improve at a task through exposure to data. | Predicting hospitalisation risk in paediatric asthma. |
| Deep learning (DL) | ML using multilayer neural networks, especially powerful with images, signals and text. | Detecting fractures and pneumonia on chest radiographs. |
| Convolutional neural network (CNN) | DL architecture optimised for spatial pattern recognition in images. | Scoliosis quantification; retinopathy of prematurity grading. |
| Recurrent / transformer network | DL architectures for sequences and language; transformers underlie modern LLMs. | PICU vital-sign streams; ambient clinical scribes. |
| Natural language processing (NLP) | Techniques for understanding and generating human language. | Extracting child-abuse cues from unstructured paediatric notes. |
| Large language model (LLM) | Generative model trained on massive text; produces fluent answers and summaries. | Parental education on developmental dysplasia of the hip. |
| Generative AI | AI that creates new content (text, images, audio). | Drafting discharge summaries; synthetic paediatric data for teaching. |
| Reinforcement learning | Learning by trial-and-error rewarded for desired outcomes. | Closed-loop ventilator weaning in neonatal care. |
| Federated learning | Multi-site training where data stay local and only model updates are shared. | Pooling rare-disease cohorts across paediatric centres. |
| Explainable AI (XAI) | Methods that make model reasoning interpretable to humans. | Saliency maps highlighting suspicious regions on a chest film. |
| Augmented intelligence | Framing AI as a clinician's assistant rather than a replacement. | Decision support that flags abnormal newborn screen patterns. |

ethics guidance, and recent reviews addressing clinical implementation, validation, bias, explainability, and governance were given priority. We used non-paediatric or non-peer-reviewed sources for definitional or contextual purposes.

The review is narrative rather than systematic; therefore, no formal PRISMA flow diagram was prepared, and we did not attempt a meta-analytic synthesis. Limitations of this approach are acknowledged: narrative reviews are subject to selection bias by design. The rapidly moving field implies that some emerging tools may not yet be reflected in the indexed literature at the time of writing.

## CLINICAL AND RESEARCH IMPLICATIONS

### AI Bits, Defined: A 60-Second Taxonomy

Artificial intelligence is a broad term for computer systems that perform tasks usually tied to human cognition, including pattern recognition, language understanding, and decision support (1, 6). Within AI, machine learning (ML) denotes algorithms that improve at a task by learning from data rather than explicit rules. Deep learning (DL) is a subset of ML that uses multilayer artificial neural networks; convolutional neural networks (CNNs) handle images well, while transformer-based networks underpin modern LLMs and increasingly handle a growing share of EHR time series (1, 5, 7). Natural language processing (NLP) is the family of techniques that enables computers to understand and generate human language, and it is the motor of clinical scribes, document summarisation, and chatbots (1, 8).

Two more terms now circulate in everyday clinical talk among researchers. Generative AI describes systems that produce new content (text, images, and code) and includes ChatGPT-class LLMs. Federated learning describes training methods in which models are updated across several hospitals without raw patient data leaving each site, a particularly attractive approach for rare paediatric diseases (4, 9, 10). Figure 1 places these terms in their nested relationship; Table 1 provides a glossary intended for daily use, and Figure 2 maps how each of these technologies enters and exits a paediatric clinical workflow.

Two practical clarifications matter for clinicians. First, what is often called "AI in medicine" today is, more precisely, augmented intelligence. As Bhargava and Knake have argued, this software serves as an assistant to the physician and not as a replacement for the physician (1, 2). Second, AI tools are narrow in scope. A model that detects retinopathy of prematurity is not the same as a model that predicts sepsis, and validation in one task or population does not carry over by default to another. This tight scope is a strength when kept in mind and a danger when forgotten.

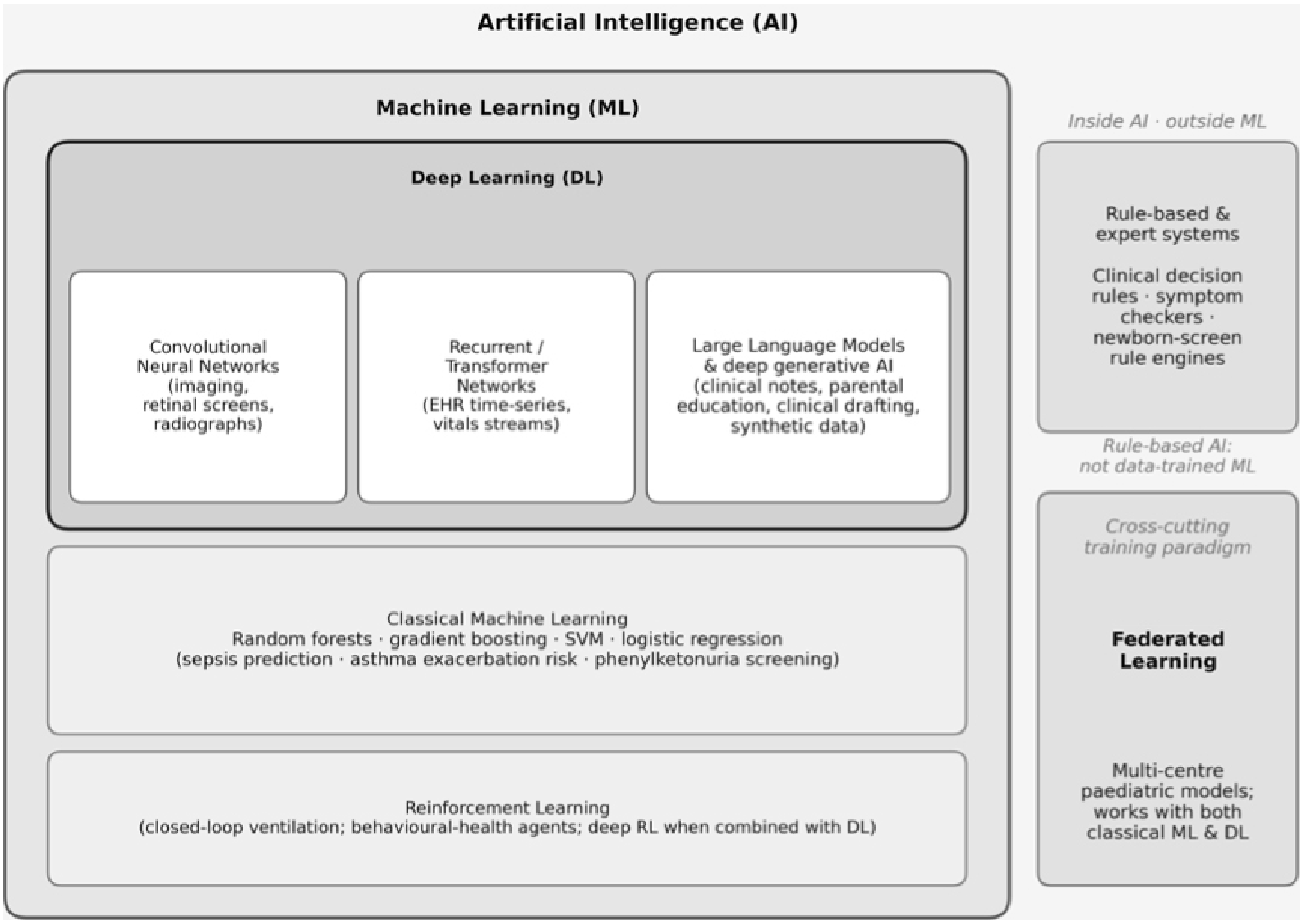


**Figure 1.** Conceptual taxonomy of artificial intelligence with paediatric examples.

Nested rectangles show that machine learning is a subset of artificial intelligence and deep learning is a subset of machine learning. Selected paediatric exemplars are listed for each level, including convolutional neural networks for imaging, transformer networks for electronic health record time series and large language models for parental education and clinical documentation.

## Where AI Is Already Useful in Paediatric Practice

### *Imaging and Dysmorphology*

Image interpretation is reported to be the most mature application of AI in paediatric medicine (Table 2). In selected narrow tasks, such as scoliosis quantification, fracture and physical injury detection, cystic fibrosis scoring on chest radiographs, identification of tuberculosis and bacterial pneumonia patterns, and automated grading in retinopathy of prematurity and paediatric cataract, deep CNNs have reported performance approaching that of paediatric subspecialists, with occasional reports of overlap under controlled test conditions; such comparisons depend strongly on the test set, scanner, and population used (3, 7, 11, 12). In dysmorphology, facial-image classifiers can suggest candidate genetic syndromes, such as trisomy 21, DiGeorge, and Goldenhar syndromes, from a single photograph, supporting earlier referral in regions with limited access to medical genetics (4, 7). The major caveats are that paediatric-specific training data remain scarce relative to adult datasets and that performance characteristics often degrade when the AI is exposed to portable radiographs, different scanners, or a different ethnic distribution than that on which it was trained (3, 4, 13).

### *Risk Prediction and the Paediatric Intensive Care Unit*

AI-based early warning systems are advancing rapidly in critical care. Kamaleswaran and colleagues, working with nearly 500 children admitted to a paediatric intensive care unit (PICU), trained an ML algorithm that flagged severe sepsis up to eight hours earlier than the conventional electronic health record screen (3, 14). Even a modest lead time of this size may translate into earlier antibiotic administration and a measurable reduction in organ failure. Models for clinical deterioration, bronchopulmonary dysplasia (BPD) in preterm infants, and paediatric urinary tract infection (UTI) in febrile children are emerging from international consortia (5, 15). Closed-loop technology has also reached the bedside: technology-driven automated weaning of mechanical ventilation continuously titrates pressure to a clinician-chosen tidal volume target, illustrating how

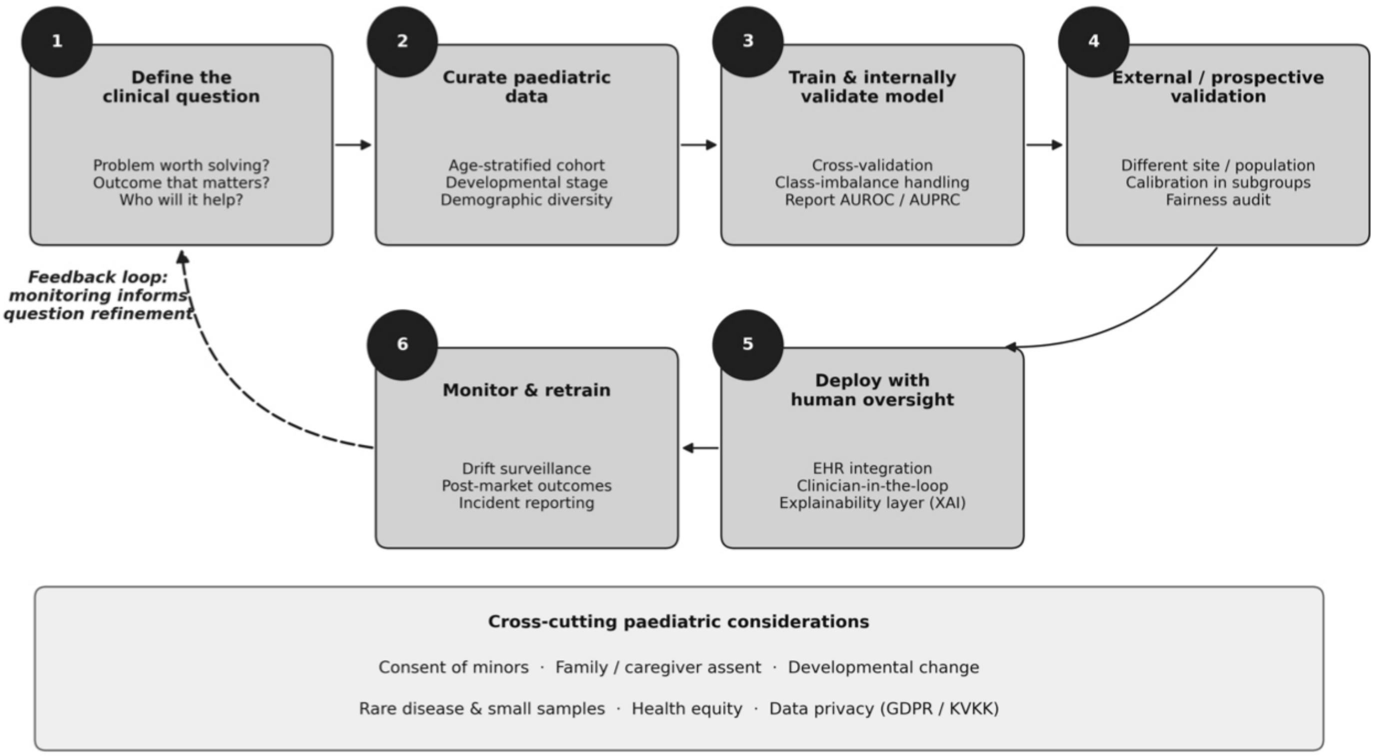


**Figure 2.** The paediatric AI lifecycle: a practical six-step framework for the paediatrician.

The cycle moves from defining the clinical question, through data curation, training and internal validation, external or prospective validation, deployment with human oversight and continuous post-market monitoring. A feedback loop returns lessons to the question-formulation step. Cross-cutting paediatric considerations should apply at every stage: consent of minors, family or caregiver assent, developmental change, rare diseases and small samples, equity and data privacy under GDPR and the Turkish Personal Data Protection Law (KVKK).

an algorithm can perform a structured task tirelessly when supervised (2). Paediatricians should ask whether such tools have been prospectively validated in a population resembling their own and how alerts are routed, displayed, and acted upon.

*Chronic Disease and Behavioural Health*

In asthma, ML has been used to predict hospitalisation risk from EHRs combined with weather, neighbourhood characteristics, and community viral load and to forecast the persistence of childhood asthma into later years (3). Smartphone microphones can substitute for office spirometry, and inhaler-mounted sensors fed into algorithms can identify environmental triggers and adherence gaps (3, 16). For autism spectrum disorder, the device evaluated by Megerian and colleagues, authorised by the United States Food and Drug Administration, combines caregiver questionnaires, short home videos, and primary-care input to assist diagnosis in children aged 18-72 months (3, 17). Such a tool may shorten a wait that, in many regions, still exceeds four years from the first parental concern and that falls hardest on girls and on Black, Latino, or rural children. For attention-deficit/hyperactivity disorder (ADHD), conduct disorder, and post-traumatic stress disorder, ML models have shown promise for early screening, severity quantification, and identification of clinically meaningful subtypes (5, 18, 19). Finally, AI-driven cognitive behavioural therapy chatbots reduced depressive symptoms in a randomised trial in young adults, and their paediatric counterparts are being studied as bridges across mental health waitlists (3, 20).

*Genetics, Rare Disease, and Metabolic Screening*

Paediatric AI shines where the data are at their most challenging: in rare diseases. AI-augmented rapid whole-genome sequencing has been used to deliver provisional molecular diagnoses to critically ill neonates within hours rather than weeks, with measurable changes in management (3, 21). Pattern-recognition systems now help screen for phenylketonuria, mitochondrial disorders, and other inborn errors of metabolism without a false-positive penalty, and rare-disease auxiliary diagnosis platforms can shortlist candidate disorders from clinical phenotype data alone (5, 22). A large subset of rare diseases is hereditary and affects children; for these families, even a modest reduction in the diagnostic odyssey is meaningful.

*Therapeutics, Dosing, and the Operating Room*

Paediatric pharmacokinetics are a paradigm case for personalisation, since size and maturity drive dose far more than they do in adults. ML-based dosing models have been shown to outperform conventional approaches for selected antibiotics in children, including ceftaroline, by integrating weight, renal function, and organism susceptibility (5). In peri-operative care, an AI workflow tool was associated with a reduction in

**Table 2. Selected paediatric AI applications by clinical domain**

| Domain | Representative tasks | Illustrative tools/studies | Maturity* |
|---|---|---|---|
| Imaging & dysmorphology | Fracture and physical injury detection; pneumonia and tuberculosis triage; scoliosis grading; retinopathy of prematurity; congenital cataract; syndrome facial recognition. | Paediatric chest-radiograph CNNs; retinal screening systems; FDNA Face2Gene-type platforms (3, 7, 11, 12). | Mature in narrow tasks; limited paediatric external validation. |
| Critical care & deterioration | Sepsis, septic shock, BPD, UTI, deterioration and mortality prediction; closed-loop ventilation. | PICU sepsis ML models with 8-h lead time; automated ventilator weaning algorithms (2, 3, 14, 15). | Several FDA-cleared adult tools; paediatric-specific evidence emerging. |
| Chronic disease & behavioural health | Asthma exacerbation and persistence prediction; smartphone spirometry; autism, ADHD, depression and PTSD support. | Smartphone-based spirometry; FDA-authorised AI device for autism (Canvas Dx); cognitive behavioural therapy chatbots (3, 17, 18, 20). | Selected products marketed; broader evidence still maturing. |
| Rare disease & genetics | Rapid whole-genome diagnosis in NICU/PICU; phenylketonuria and metabolic screening; phenotype-based rare-disease shortlisting. | Rapid whole-genome sequencing pipelines; RDAD systems; ML-based metabolic newborn-screening models (3, 21, 22). | Rapid clinical translation in tertiary centres. |
| Therapeutics & workflow | Precision dosing; perioperative opioid stewardship; ambient documentation; LLM-based parental education. | ML-based ceftaroline dosing; OR Advisor for opioid reduction; ChatGPT-class LLMs for parental information (3, 5, 23, 24). | Rapidly evolving; close human supervision required. |

*: Maturity descriptions are qualitative; AUROC: Area under the receiver operating characteristic curve; BPD: Bronchopulmonary dysplasia; CNN: Convolutional neural network; FDA: Food and Drug Administration; LLM: Large language model; ML: Machine learning; NICU: Neonatal intensive care unit; PICU: Paediatric intensive care unit; RDAD: Rare Disease Auxiliary Diagnosis system; UTI: Urinary tract infection.

opioid administration from 85% of surgeries to <1% in a paediatric surgery centre, with implications for short-term recovery and long-term opioid exposure (3, 23). Across these examples, the algorithm augments the prescriber rather than replacing pharmacological judgment.

*Documentation, Communication, and Parental Education*

Generative AI—specifically LLMs such as ChatGPT and clinical equivalents—has shifted from novelty to daily use in only two years (24, 25). For paediatricians, three categories deserve attention. Ambient clinical documentation tools transcribe and summarise encounters, returning structured notes for human review and reducing keyboard time. Patient-facing chatbots can deliver explanations of paediatric conditions at an acceptable level of quality. In the comparative study by Li and colleagues, three contemporary LLMs answered standardised questions on developmental dysplasia of the hip with accuracy rates of 78-83% on independent specialist scoring, a result that places these tools as a supplement to, and never a substitute for, professional counselling (5, 24). Parents themselves now enter the consultation having already "consulted" a chatbot; mixed-methods studies of paediatric chronic-disease cohorts show that families show greater intention to use clinically curated chatbots than to use general-purpose LLMs, citing trust, credibility, and clinician involvement as decisive factors (5, 26). The principal risks are confident hallucinations, embedded bias, lack of paediatric calibration, and misinformation about vaccines and other public-health-sensitive topics (8, 25).

### AI Bits for Paediatricians

*Small Samples, Rare Outcomes, and Developmental Change*

The defining mathematical fact about paediatric AI is data scarcity. Rare diseases, rare emergencies, and even common diseases, when stratified by narrow age bands, quickly produce datasets too small for confident generalisation (4, 5, 9, 11). Models trained on adolescents may underperform in toddlers, and a model trained on toddlers in 2018 may underperform in toddlers in 2026 because of changes in clinical practice, vaccination patterns, or coding habits, a phenomenon known as dataset shift (5, 9). Multicentre and federated learning approaches help, but no algorithm escapes the requirement for ongoing surveillance after deployment.

*Bias, Equity, and "Algorithmic Ageism"*

AI faithfully reproduces the patterns on which it is trained, including inequities (3, 9, 27). Demonstrated examples in paediatrics include diagnostic delays in autism for Black, Latino, female, and rural children embedded in screening tools and pulse oximetry algorithms that perform less accurately in children with darker skin (3). The remedy is not to refuse AI but to demand that vendors document the demographic composition of training and validation cohorts, that developers explicitly test for equity gaps, and that local users monitor performance by subgroup. Children additionally need protection from "algorithmic ageism": tools developed primarily for adults and silently extrapolated downward without revalidation.

**Table 3. A five-minute appraisal checklist for an AI tool offered to a paediatric practice**

| Domain | Question to ask | Red flag if... |
|---|---|---|
| Clinical question | What exact decision will this AI improve, in which age group and at what stage of care? | Vendor describes a generic "platform" with no specific decision target. |
| Training data | How many paediatric cases were used? What is the age, sex, ethnicity, comorbidity and site distribution? | Training was adult-only or single-centre; no demographic breakdown is provided. |
| Validation | Was external or prospective paediatric validation performed? What are AUROC, calibration and subgroup performance? | Performance reported only on the training/test split from one institution. |
| Bias and equity | Has fairness been audited across race, sex, language, disability and socioeconomic strata? | No equity analysis; minority subgroups under-represented or unreported. |
| Workflow integration | How is the output displayed, who can override it and what is the expected alert burden? | "Always-on" alerts with no override pathway and unclear ownership. |
| Explainability | Can you understand, at the case level, why the model produced a given output? | Pure black box with no XAI layer or rationale. |
| Privacy and consent | Where do paediatric data flow? How is consent of the minor and family handled? | Data leave the institution to a vendor cloud without clear safeguards. |
| Post-deployment governance | Who monitors performance drift, equity and adverse events after go-live? | No defined owner, dashboard, or audit cadence. |
| Evidence base | Are peer-reviewed trials available with patient-oriented outcomes, ideally aligned with CONSORT-AI / TRIPOD-AI? | Only conference abstracts, vendor white papers, or in-silico claims. |
| Regulatory status | Is the tool authorised for the intended paediatric use by FDA, MHRA, EMA, or the Turkish Medicines and Medical Devices Agency? | Marketed as "CE-marked" only for adults but offered for paediatric care. |

AI: Artificial intelligence; AUROC: Area under the receiver operating characteristic curve; CE: Conformité Européenne; CONSORT-AI: Consolidated Standards of Reporting Trials extension for AI; EMA: European Medicines Agency; FDA: Food and Drug Administration; MHRA: Medicines and Healthcare products Regulatory Agency; TRIPOD-AI: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis extension for AI; XAI: Explainable AI.

*The Black Box, Hallucinations, and Explainability*

Many high-performing models, especially deep neural networks and LLMs, do not transparently disclose how they reach a conclusion. Saliency maps, counterfactual explanations, and rule extraction, collectively known as explainable AI (XAI), partially address this, but a clinician should never feel compelled to act on an output whose logic is opaque and that contradicts their clinical reasoning (5, 9). LLMs additionally produce hallucinations: smooth-sounding statements that are factually wrong, at times including fabricated references. Therefore, the clinical use of LLMs requires both human verification of factual claims and institutional policies that limit their unsupervised role in patient care (8, 25).

*Privacy, Consent, and the Special Status of the Child*

Children cannot meaningfully consent to the long-term reuse of their data, but the consequences of such reuse may follow them into adulthood. Several frameworks set out the basic rights that apply here: the European Union General Data Protection Regulation (GDPR), the EU AI Act, and Türkiye's Personal Data Protection Law (KVKK) (4, 28, 29). Among these rights, the right to erasure of personal data carries even greater weight when the data belong to a child. In Türkiye, the secondary regulation on the deletion, destruction, and anonymisation of personal data operationalises these rights and is directly relevant to paediatric data stewardship (29). Genetic, behavioural, and image data deserve heightened protection. Paediatricians act as custodians of paediatric data until the child grows into adult autonomy; this responsibility must be reflected in informed consent processes for any AI-related research or clinical use.

*Regulatory and Clinical Maturity*

More than 1000 AI-enabled medical devices have been authorised by the FDA according to the agency's public list, but the great majority were cleared for adult use, and few have undergone prospective paediatric trials with patient-oriented outcomes (3, 4, 30). Existing regulatory pathways were not built for self-updating algorithms; the FDA's proposed predetermined change control plans illustrate one direction, but local hospitals will still need clear governance structures, including model registries, performance dashboards, and incident reporting. The cautionary tale of an early symptom-checker chatbot that overstated its diagnostic accuracy in paediatric scenarios underlines the cost of premature commercialisation (5, 31).

## HOW TO READ AN AI CLAIM IN FIVE STEPS

When confronted with a new AI tool, the paediatrician can run a five-step check that follows the same line of thinking used for any new test. To begin with, what is the exact clinical question and target population, including age stratum? Second, on what data was the model trained, and is the demographic and developmental composition close to that of the local population? Third, has external or prospective validation been carried out in a paediatric cohort? If yes, how is performance reported: sensitivity, specificity, area under the receiver operating characteristic curve, calibration, and, where relevant, the results of fairness audits? Fourth, how does the tool integrate into the EHR and workflow, who acts on its outputs, and what is the alert burden? Fifth, what is the post-deployment monitoring plan for drift and harm? These five questions, summarised in Table 3, are deliberately compatible with a busy clinic schedule and align with the paediatric AI lifecycle shown in Figure 2.

In clinical research, the same logic applies during peer review and editorial work. Reporting standards such as STARD-AI, CONSORT-AI, SPIRIT-AI, TRIPOD-AI, and DECIDE-AI offer structured checklists; familiarity with these instruments allows paediatricians to comment on AI manuscripts even without a computer science background (4, 9, 31).

## EDUCATION, GOVERNANCE, AND THE NEXT FIVE YEARS

Surveys consistently find that paediatric trainees and faculty report low confidence in AI literacy, despite high enthusiasm for its potential (3, 4, 32). Curriculum proposals such as MEd2030 explicitly add modules on biomedical informatics and AI, transdisciplinary collaboration, and the ethics of digital therapeutics; incremental local steps include adding AI critical appraisal sessions to journal clubs, requiring AI tool review as part of new-technology committees, and pairing each AI deployment with an outcome audit at six and twelve months (3, 32). At the institutional level, governance committees should include not only clinicians and informaticians but also patient and family representatives, ethicists, and equity champions (4, 5).

In Türkiye and the broader region, several near-term priorities are visible. Paediatric-specific reference datasets are scarce in Turkish-speaking populations and ought to be developed through multicentre collaboration. Local validation of imported algorithms should become the rule rather than the exception. The case is sharpest for radiology, ophthalmology, and growth-assessment tools, whose performance shifts with the acquisition hardware and the reference norms in use. Family-facing LLM use in vaccine counselling, infant nutrition, and adolescent mental health needs Turkish-language quality benchmarks. Finally, residency training programmes would benefit from a defined AI core competency, articulated within existing paediatrics curricula and assessed through structured cases.

## CONCLUSIONS

AI is no longer a future technology in paediatrics; it is a present, uneven, and consequential reality. The most important shift required of the paediatrician is conceptual rather than technical: to read AI as augmented intelligence, narrow in scope and dependent on the quality of the data and labels behind it. The next shift is practical. We must apply the same disciplined questions to an algorithm that we already apply to a new diagnostic test or a new therapeutic approach: paediatric validation, equity audits, transparent reasoning, and clear post-deployment governance. The final shift is collective: to engage early with the design, evaluation, and regulation of paediatric AI so that algorithms come to reflect the realities of children's development, family context, and lifelong stakes. Paediatricians who do this will not merely use AI; they will shape it.

**Ethics Committee Approval:** In this study, no patient data or human/animal subjects were involved and no ethics committee approval was required.

**Conflict of Interest:** No conflict of interest was declared by the author.

**Financial Disclosure:** The author declared that this study has received no financial support.

**Use of AI for Writing Assistance:** Not declared.

**Acknowledgements:** The author thanks Associate Prof. Fatma Sevinç Şengül, MD, for providing informal feedback before preparation of the manuscript.

**Peer-review:** Externally peer-reviewed.

**Etik Kurul Onayı:** Bu çalışmada herhangi bir hasta verisi veya insan/hayvan denek kullanılmamıştır ve etik kurul onayı gerekmemiştir.

**Çıkar Çatışması:** Yazar çıkar çatışması bildirmemiştir.

**Mali Destek:** Yazar bu çalışma için mali destek almadıklarını beyan etmişlerdir.

**Yazma Yardımı için Yapay Zeka Kullanımı:** Beyan edilmedi.

**Teşekkür:** Yazar, makalenin hazırlanmasından önce görüş ve önerilerini paylaşan Doç. Dr. Fatma Sevinç Şengül'e teşekkür eder.

**Hakemli inceleme:** Harici olarak hakemli.

## REFERENCES

1. Bhargava H, Salomon C, Suresh S, Chang A, Kilian R, van Stijn D, et al. Promises, pitfalls, and clinical applications of artificial intelligence in pediatrics. J Med Internet Res 2024;26:e49022.
2. Knake LA. Artificial intelligence in pediatrics: the future is now. Pediatr Res 2023;93:445–6.
3. Khalatkar V. Artificial intelligence in pediatrics. Indian Pediatr 2025;62:869.
4. Kuru K, Caswell N, Hart GR, Quon J, Ehwerhemuepha L. Editorial: Artificial intelligence and machine learning in pediatrics. Front Artif Intell 2026;9:1814141.
5. Balla Y, Tirunagari S, Windridge D. Pediatrics in artificial intelligence era: a systematic review on challenges, opportunities, and explainability. Indian Pediatr 2023;60:561–9.
6. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng 2018;2:719–31.
7. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25:44–56.
8. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit Health 2021;3:e745–50.
9. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med 2019;17:195.

10. Sheller MJ, Edwards B, Reina GA, Martin J, Pati S, Kotrotsou A, et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. Sci Rep 2020;10:12598.
11. Otjen JP, Moore MM, Romberg EK, Perez FA, Iyer RS. The current and future roles of artificial intelligence in pediatric radiology. Pediatr Radiol 2022;52:2065–73.
12. Reid JE, Eaton E. Artificial intelligence for pediatric ophthalmology. Curr Opin Ophthalmol 2019;30:337–46.
13. Davendralingam N, Sebire NJ, Arthurs OJ, Shelmerdine SC. Artificial intelligence in paediatric radiology: Future opportunities. Br J Radiol 2021;94:20200975.
14. Kamaleswaran R, Akbilgic O, Hallman MA, West AN, Davis RL, Shah SH. Applying artificial intelligence to identify physiomarkers predicting severe sepsis in the PICU. Pediatr Crit Care Med 2018;19:e495–503.
15. Filipow N, Main E, Sebire NJ, Booth J, Taylor AM, Davies G, et al. Implementation of prognostic machine learning algorithms in paediatric chronic respiratory conditions: a scoping review. BMJ Open Respir Res 2022;9:e001165.
16. Larson EC, Goel M, Boriello G, Heltshe S, Rosenfeld M, Patel SN. SpiroSmart: using a microphone to measure lung function on a mobile phone. In: Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp ’12); 2012 Sep 5-8; Pittsburgh (PA), USA. New York (NY): ACM; 2012. p.280–9.
17. Megerian JT, Dey S, Melmed RD, Coury DL, Lerner M, Nicholls CJ, et al. Evaluation of an artificial intelligence-based medical device for diagnosis of autism spectrum disorder. NPJ Digit Med 2022;5:57.
18. Aylward BS, Abbas H, Taraman S, Salomon C, Gal-Szabo D, Kraft C, et al. An introduction to artificial intelligence in developmental and behavioral pediatrics. J Dev Behav Pediatr 2023;44:e126–34.
19. Saxe GN, Ma S, Ren J, Aliferis C. Machine learning methods to predict child posttraumatic stress: a proof of concept study. BMC Psychiatry 2017;17:223.
20. Prochaska JJ, Vogel EA, Chieng A, Kendra M, Baiocchi M, Pajarito S, et al. A therapeutic relational agent for reducing problematic substance use (Woebot): Development and usability study. J Med Internet Res 2021;23:e24850.
21. Clark MM, Hildreth A, Batalov S, Ding Y, Chowdhury S, Watkins K, et al. Diagnosis of genetic diseases in seriously ill children by rapid whole-genome sequencing and automated phenotyping and interpretation. Sci Transl Med 2019;11:eaat6177.
22. Zhu Z, Gu J, Genchev GZ, Cai X, Wang Y, Guo J, et al. Improving the diagnosis of phenylketonuria by using a machine learning-based screening model of neonatal MRM data. Front Mol Biosci 2020;7:115.
23. Low D. Closing the gateway from surgery to persistent opioid use. Boston (MA): Institute for Healthcare Improvement; 2019; republished by Catalysis (Appleton, WI). Available from: https://createvalue.org/articles_and_news/closing-gateway-surgery-persistent-opioid-use/. Accessed 30 Apr, 2026.
24. Li S, Jiang J, Yang X. Preliminary assessment of large language models’ performance in answering questions on developmental dysplasia of the hip. J Child Orthop 2025;19:207–12.
25. Cascella M, Montomoli J, Bellini V, Bignami E. Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research Scenarios. J Med Syst 2023;47:33.
26. Lau CWY, Kupiec K, Livermore P. Exploring the acceptance and opportunities of using a specific generative AI chatbot to assist parents in managing pediatric rheumatological chronic health conditions: mixed methods study. JMIR Pediatr Parent 2025;8:e70409.
27. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019;366:447–53.
28. European Parliament and Council of the European Union. Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Off J Eur Union 2024; L, 2024/1689 (12 July 2024). Available from: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng. Accessed 30 Apr, 2026.
29. Kişisel Verileri Koruma Kurumu. Kişisel verilerin silinmesi, yok edilmesi veya anonim hale getirilmesi hakkında yönetmelik. Resmî Gazete 28 Ekim 2017; sayı: 30224. Ankara: KVKK; 2017. Available from: https://www.kvkk.gov.tr/Icerik/2053/Kisisel-Verilerin-Silinmesi-Yok-Edilmesi-veya-Anonim-Hale-Getirilmesi-Hakkinda-Yonetmelik Accessed 30 Apr, 2026. [In Turkish]
30. U.S. Food and Drug Administration. Artificial intelligence-enabled medical devices. Silver Spring (MD): FDA; updated 2025. Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-enabled-medical-devices. Accessed 30 Apr, 2026.
31. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med 2020;26:1364–74.
32. Paranjape K, Schinkel M, Nannan Panday R, Car J, Nanayakkara P. Introducing artificial intelligence training in medical education. JMIR Med Educ 2019;5:e16048.