Indian Pediatrics (2023) 60:561-569
https://doi.org/10.1007/s13312-023-2936-8

SYSTEMATIC REVIEW

Pediatrics in Artificial Intelligence Era: A Systematic Review on Challenges, Opportunities, and Explainability

YASHASWINI BALLA,[1] SANTOSH TIRUNAGARI,[2] DAVID WINDRIDGE[3]

[1]*Neurosciences Department, Alder Hey Children's NHS Foundation Trust, Liverpool, United Kingdom.*
[2]*Department of Psychology, Middlesex University, London, United Kingdom.*
[3]*Department of Computer Science, Middlesex University, London, United Kingdom.*
Correspondence to: Dr Santosh Tirunagari, Department of Psychology, Middlesex University, London, United Kingdom.
s.tirunagari@mdx.ac.uk

Background: The emergence of artificial intelligence (AI) tools such as ChatGPT and Bard is disrupting a broad swathe of fields, including medicine. In pediatric medicine, AI is also increasingly being used across multiple subspecialties. However, the practical application of AI still faces a number of key challenges. Consequently, there is a requirement for a concise overview of the roles of AI across the multiple domains of pediatric medicine, which the current study seeks to address.

Aim: To systematically assess the challenges, opportunities, and explainability of AI in pediatric medicine.

Methodology: A systematic search was carried out on peer-reviewed databases, PubMed Central, Europe PubMed Central, and grey literature using search terms related to machine learning (ML) and AI for the years 2016 to 2022 in the English language. A total of 210 articles were retrieved that were screened with PRISMA for abstract, year, language, context, and proximal relevance to research aims. A thematic analysis was carried out to extract findings from the included studies.

Results: Twenty articles were selected for data abstraction and analysis, with three consistent themes emerging from these articles. In particular, eleven articles address the current state-of-the-art application of AI in diagnosing and predicting health conditions such as behavioral and mental health, cancer, syndromic and metabolic diseases. Five articles highlight the specific challenges of AI deployment in pediatric medicines: data security, handling, authentication, and validation. Four articles set out future opportunities for AI to be adapted: the incorporation of Big Data, cloud computing, precision medicine, and clinical decision support systems. These studies collectively critically evaluate the potential of AI in overcoming current barriers to adoption.

Conclusion: AI is proving disruptive within pediatric medicine and is presently associated with challenges, opportunities, and the need for explainability. AI should be viewed as a tool to enhance and support clinical decision-making rather than a substitute for human judgement and expertise. Future research should consequently focus on obtaining comprehensive data to ensure the generalizability of research findings.

Key words: *Artificial intelligence, Data science, Deep learning, Large language model.*

Protocol Registration: International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY): INPLASY202350045.

Science and technology have made significant advancements with the introduction of artificial intelligence (AI), and machine learning (ML) has been a game-changer. ML has enabled computers to learn without explicit programming by combining computer science and statistics [1]. ML has gained momentum in many fields, including healthcare, thanks to emerging tools like ChatGPT, Bard and Glass AI 2.0. These tools are transforming industries by enabling conversations between humans and machines. ChatGPT, a large language model (LLM), has immense potential to assist in healthcare, including helping patients with mental health issues and aiding healthcare providers in decision-making [2,3]. Recently, Glass Health introduced Glass AI 2.0, a similar LLM to ChatGPT, but with a clinical knowledge database created and maintained by clinicians to generate differential diagnoses and clinical plan outputs [4].

The integration and scope of such AI tools in healthcare are growing rapidly. Pediatrics is a field with practical challenges like complex comorbidities, increasing emergency admissions, and a lack of access to pediatric care providers, which could hinder the provision of quality and timely care [5]. ML implementations can streamline the pediatric work-force and assist clinical decision-making by enabling physicians to focus on patient-centered care plans by making better use of their clinical knowledge and time [1]. ML techniques can analyze vast datasets and create predictive models that go beyond human cognitive capabilities.

Even though technological advancements are expanding the integration and scope of ML in pediatrics, there are challenges associated with the implications of AI, such as unintentional bias from data, like racial segregation and

The original version of this article was revised due to a retrospective Open Access order.

under-performing algorithms, which could jeopardize patient care [6]. To mitigate these issues, it may be better to focus on collective human-support AI systems instead of complete automation. Additionally, ML could facilitate medical training or provide evidence-based care to patients using AI-based web or mobile applications with the help of human-in-the-loop systems. Therefore, it is crucial to evaluate the explainability of AI models, potential opportunities, and challenges when integrating ML in healthcare, especially for the pediatric population.

METHODS

The field of artificial intelligence (AI) is rapidly growing and evolving, with a variety of buzzwords and terms that can be confusing to navigate. In order to clarify these terms and their relationship to one another, this section provides definitions of key buzzwords related to AI in **Box I**. Additionally, a visual representation of the relationship among these buzzwords is presented in **Fig. 1**. By establishing a clear understanding of these terms, the subsequent methodology can be more easily understood and applied.

Since this study focuses on ML in pediatrics; for this purpose, we employ a qualitative approach, expressing research findings and interpretations in terms of non-numeric data. The research thus uses secondary analysis to constitute an evidence-based literature review and associated analysis.

Search Strategy and Keywords

In this review, the focus is on assessing the challenges and opportunities of ML in pediatrics. The search was performed through authentic databases to obtain the relevant information. The search strategy was based on using a set of keywords and Boolean operators.

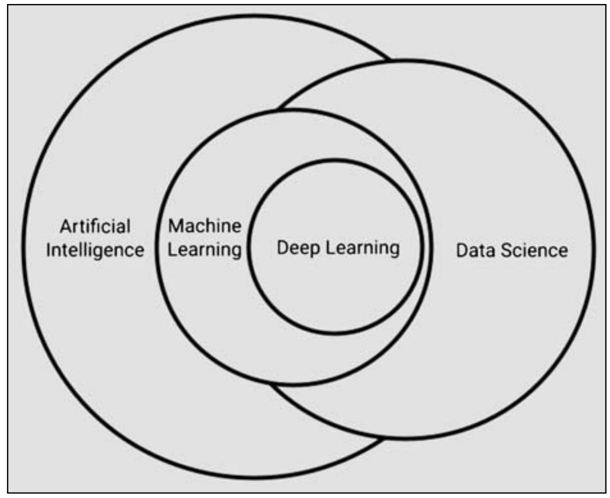

Fig. 1 Relationship between artificial intelligence (AI), machine learning (ML), deep learning and data science. AI is the creation of computer systems that can perform tasks that typically require human intelligence. Machine learning is a branch of AI that enables computers to learn from data without being programmed explicitly. Deep learning is a subset of machine learning that focuses on neural network algorithms. Data science deals with everything related to data, including collecting, cleaning, analyzing and interpreting it.

Box I Simplified Definitions of the Commonly Used Artificial Intelligence Buzzwords

Artificial intelligence (AI): This term refers to the creation of intelligent systems that simulate human thinking and behavior. AI systems can be designed to perform tasks such as speech recognition, decision-making, and problem-solving.

Statistics: Involves making inferences about a population based on a sample of data. It can be used to make predictions and identify patterns within data.

Machine learning: A subset of AI, machine learning involves designing algorithms that can learn from data without being explicitly programmed. It involves finding patterns within data that can be used to make predictions.

Deep learning: A subset of ML, deep learning uses neural network-based methods to generate repeatable predictions by finding patterns within data. It is particularly useful for complex data sets and can be used for tasks such as language, image and speech recognition.

Data science: This is the study of data, which involves data preparation, transformation, and analytics. Data scientists use a variety of techniques and tools to make sense of large data sets and to identify patterns that can be used for decision-making.

Keywords that were used during the search for the desired topic that is challenges and opportunities of machine learning in pediatrics are: “machine learning AND pediatrics”, “machine learning”, “challenges faced during paediatrics care AND technology”, and “significance of machine learning AND paediatrics care”, “paediatrics AND machine learning history”, “machine learning AND future in paediatrics care.”

Databases, Data Extraction and Selection Criteria

The databases that were used during the searching process include PubMed Central and Europe PubMed Central [7]. PRISMA guidelines for systematic reviews were used [8] for the data extraction process. The selection criterion defining the inclusion of relevant information is peer-reviewed journals from scholarly databases, the English publication language, the timeline of the last 7 years (2016-2022), and the information relevant to the research objectives. The aforementioned criteria are considered while searching for relevant data.

Data Collection

For data collection, Preferred Reporting Items for Systematic Reviews and Meta Analysis (PRISMA) guidelines were adopted as shown in **Fig. 2**. The data was collected as per the inclusion and exclusion criteria guidelines in Section 2.2. Initial search results retrieved were *N*=210 searches (this included scholarly articles and, initially, grey literature). Later, primary screening was performed, and duplicated results were removed, leaving 86 articles. These were then passed down to secondary screening, where exclusion based on title/abstracts, language, and timeline was performed, resulting in 39 articles. Finally, the content analysis was performed for quality check, and 20 articles were then selected for the review discussion.

RESULTS

Three themes have been designed based on the achievable results viz., *i*) the current state-of-the-art functioning of ML algorithms in pediatric medicine, *ii*) the challenges of ML algorithm deployment in pediatric medicine, and *iii*) the future outlook of ML in pediatric medicine. The data analysis from this review can be found at: *https://github.com/tsantosh7/supplementary_material/blob/main/review_data_analysis.pdf*

The Current State-of-the-Art

There are many sub-specialty areas in pediatrics, including neonatology, pediatric endocrinology, pediatric emergency, nephrology, neurology, rheumatology, ophthalmology, behavioral medicine, respiratory medicine and many more. ML has integrated into areas where some showed profound outcomes, and others needed improvements. The section will detail the current perspectives on ML application in pediatric medicine.

PTSD diagnosis: Several studies have investigated the use of deep learning models for neuroimaging to classify PTSD in children who have experienced natural disasters [9,10]. Ge and colleagues [9] found that property loss and lifestyle deterioration were the most probable variables for predicting PTSD using ML algorithms. PTSD is a lasting dysfunctional condition in children, and the ability of ML to perform predictive classification has important implications for early intervention and treatment.

Cancer diagnosis: ML has also been applied to cancer diagnosis in pediatric care. Fathi, et al. [11] used the neuro-fuzzy inference system (NFIS) as a deep learning (DL) model for diagnosing pediatric leukemia patients, with the system predicting cases by extracting information from a patient's neutrophil count from blood test records. NFIS has been investigated as a prognostic tool for cancer detection in children, where the prognosis is particularly important, and is therefore highly demanded.

Metabolic condition diagnosis: ML has shown promise in detecting genetic metabolic conditions. Zhu, et al. [12] discussed that even though metabolic data of children are noisy (complex) in a clinical setting, the ML significantly extracts some useful findings from metabolic data to screen phenylketonuria (PKU) in children without a false-positive diagnosis. Moreover, PKU detection in newborns and susceptibility was also reported to be diagnosed with ML models. Rare Disease Auxiliary Diagnosis system (RDAD) is one of the examples of detecting rare phenotype-based metabolic disorders in children [13].

Eye disorder diagnosis: ML has also made contributions to ophthalmic medicine, such as predicting myopia

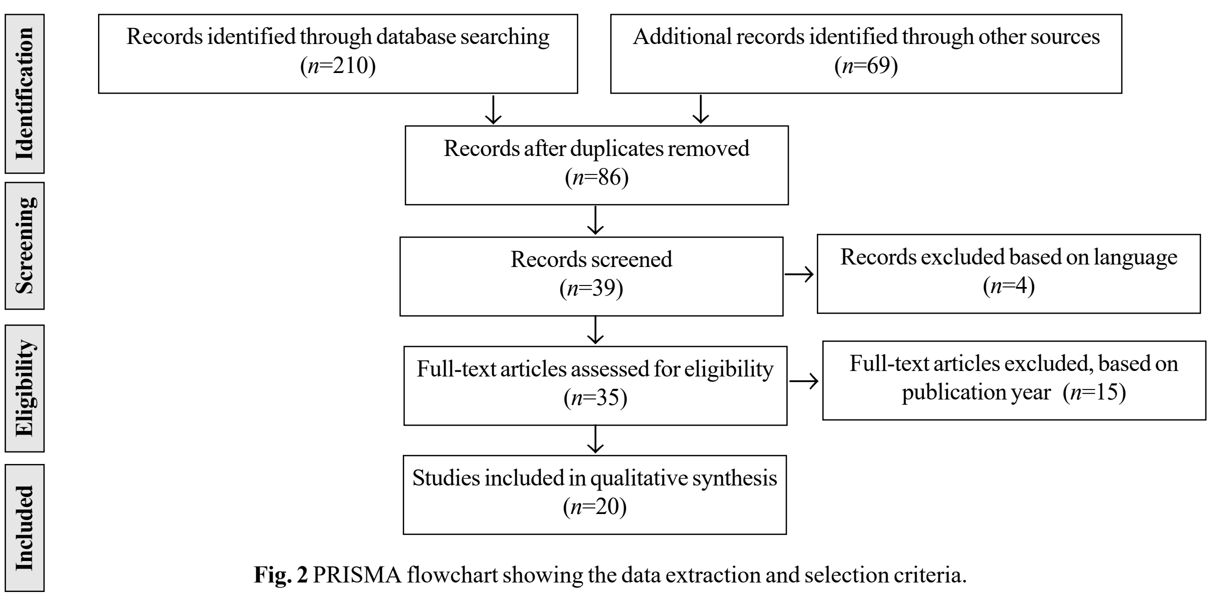

Fig. 2 PRISMA flowchart showing the data extraction and selection criteria.

development in school-aged children [14,15]. Lin and colleagues [14] observed the real-time clinical refraction data by applying ML models to predict myopia development, where acceptable prediction among children was found. Another study [15] used a regression model to predict childhood myopia in Chinese children. At a 95% confidence interval, a suitable diagnostic accuracy was reported. In addition, the study also analyzed the impact of factors on atropine-treated myopia that control intraocular pressure in children [15]. Different AI models have been used to detect factors that can enhance myopia control by optimizing atropine use [16].

Behavioral disorders: AI and ML have the potential to revolutionize analysis of behavioral problems in pediatrics, including autism spectrum disorder (ASD), conduct disorder (CD) and attention deficit hyperactive disorder (ADHD) [17,18]. These conditions lead to disruptive behaviors in children. By analyzing disruptive behaviors, AI and ML can identify patterns and correlations, which can lead to a more accurate diagnosis of behavioral problems and help in the development of more effective treatment plans [18]. Moreover, AI can also analyze speech patterns and behavior in children with ASD and identify specific markers that are indicative of ASD. This information can then be used to develop more personalized treatment plans that are tailored to the individual child's needs [19].

Abuse analysis: Child abuse is an unaddressed public health challenge but causes mental disturbances in children and post-traumatic stress disorder (PTSD) [20]. A study proposed the concept that the development of a convolutional neural network can facilitate the detection of childhood sexual abuse. ML can perform self-figure drawings to be used as a comparative figure with drawings of non-victims to identify cases of maltreatment in children [21].

Amrit, et al. [22] applied ML models to cases of abuse in children in the Netherlands using the data extracted from child specialties. They characteristically converted unstructured clinical notes into structured data and used a classification algorithm to characterize the abuse cases. Literature also studied the effects of ML on child abuse detection. ML has progressed in the predictive analysis of childhood abuse cases.

Improving PICU efficiency: In recent years, the use of AI/ML technologies in pediatric intensive care units (PICU) has greatly advanced patient outcomes for children with severe illnesses. As a result, PICU death rates have significantly decreased, with some studies reporting rates as low as 1-2% [23].

Challenges of ML Deployment in Pediatrics

The key challenges to deploying AI in clinical practice are the barriers preventing AI's clinical translation. The safety and timely transformation into practice, accuracy, algorithmic biases, data biases, brittleness, and irregular interpretability are some barriers that affect the AI-based service delivery in clinical settings [24].

Irregularity in temporality: Clinical patient data are often not recorded at regular intervals causing irregularity. For example, blood pressure measurements are collected only when a patient visits, as and when necessary, or during regular appointments. Irregularities or missing samples are inevitable because a patient can be absent from the appointment, or a clinician may cancel or reschedule the appointment [25]. Xiao, et al. in [26], proposed two important challenges of temporality and irregularity. According to their findings, clinical electronic health records (EHRs) comprise short- and long-term records for patient health trajectories. However, clinical complexity is present in long-term cases, which may interfere with the interpretation and ML-based structuring of clinical data. Furthermore, their analysis showed that the data density of clinical records varies among patients and can produce irregular samples for testing.

Biases in the diagnostic outcomes: The data used to train AI/ML models must be diverse and inclusive of different racial and ethnic groups to avoid biases in the diagnostic outcomes [24]. The limited amount of data available for conditions, such as chromosomal anomalies and genetic malformations, can affect the accuracy of AI/ML models. Approximately 80% of rare diseases (RD) are hereditary in nature, and 75% of them afflict children. Although each case is uncommon, RDs are thought to impact 350 million individuals worldwide cumulatively. Furthermore, the complexity of some syndromic conditions, such as trisomy 21 and Turner syndrome and the variability in their presentations can make it difficult for AI/ML models to diagnose these conditions accurately. So, the lack of standardization in the data collected for diagnosis and the use of different diagnostic tools by different healthcare providers can make it challenging for algorithmic models to diagnose syndromic conditions effectively. Moreover, the interpretation and implementation of AI/ML diagnostic results in the clinical setting require careful consideration of ethical and legal considerations, such as patient privacy and informed consent.

Bias and quality of data: There is also the risk of bias in AI and ML algorithms. If the algorithms are trained on biased data, they can perpetuate that bias in their predictions and decisions. Especially, as PICUs often collect large amounts of patient data, which may be fragmented, inconsistent, or of poor quality, and can impact the accuracy of AI and ML algorithms, which can have serious consequences for critically ill children. Integrating AI and ML into PICUs also requires collaboration between healthcare providers, AI and ML experts, and patient representatives to ensure that the technology is used ethically and meets the needs of patients and families [6].

Lack of variety in data: Understanding behavioral problems in pediatrics is complex, and AI/machine learning can aid in this process. However, several challenges need to be addressed. Firstly, the data used to train AI models is limited, as it is difficult to collect large amounts of data on children's disruptive behavior, such as ADHD and CD. This can result in models that are not representative of the population and may lead to incorrect predictions. Moreover, children's behavior constantly changes, and AI models must be updated frequently to reflect these changes . Likewise, children's behavior is influenced by many factors, including their environment, genetics, and development stage, making it difficult to predict their behavior accurately. In addition, there is risk of AI models reinforcing existing biases, which can lead to unfair treatment of children [18].

So, there is a need for interdisciplinary collaboration between AI researchers, pediatricians, and child psychologists to ensure that AI models are developed and used to benefit children. This requires a deep understanding of the complexities of child behavior and the development of ethical and transparent AI models. Although AI has the potential to play a significant role in understanding behavioral problems in pediatrics, it is important to address the challenges mentioned above. With the right approach, AI can help pediatricians and child psychologists to better understand and treat children with behavioral problems of ADHD, CD and ASD, leading to improved outcomes for children and their families.

Record duplication: Another significant barrier is record duplication, which limits the use of AI in practice. Vogl [27] discussed this barrier and reported that record duplications are prevalent, especially in children's welfare and protection services, and prevent processing structured data into AI-based DL methods. AI lacks unique identifiers that may differentiate duplicated records. AI itself merges duplicated data across large datasets where challenges for individual-level detection can be incurred. In addition, where records are inaccurately merged, additional challenges arise as a result. The risk of misattributed information is another risk of AI.

Interpretability of the ML models: Another challenge is the lack of interpretability of AI and ML algorithms. These algorithms can make predictions based on vast amounts of data, but it can be difficult to understand how they arrived at a particular conclusion, making it difficult for healthcare providers to trust and use the technology effectively [24]. The interpretability and explainability challenges are discussed further in the Section on explainable AI (XAI).

A recent study by McCartney, et al. [28] commented on the practical challenges associated with AI. The example of the Babylon app was an important predictor of the problems. There are still some challenges in the NHS, which first developed the app for the purpose of a symptoms checker in pediatric patients. The problems are mainly associated with the ineffective evaluation of these clinical apps before commercializing. Babylon was not tested for the safest care and treatment. The Babylon app serves as a notable example illustrating these problems, particularly in the context of the NHS. Originally developed as a symptoms checker for pediatric patients, the app encountered several challenges due to inadequate evaluation before commercialization. Notably, Baby-lon had not undergone sufficient testing to ensure the provision of safe and effective care and treatment. Therefore, it is noticeable that technological advancements, especially AI-based models, should be tested for their proper accuracy before introducing them in clinical practices. The Babylon app, an AI driven diagnostic and triage system, initially claimed 100% accuracy in symptom checking and result evaluation. However, this claim was later proven to be incorrect. Such type of political and legal challenges further complicates the introduction of AI-based technologies into healthcare settings.

Additionally, using AI and ML raise ethical concerns about patient privacy and data security. Healthcare providers must ensure that patient data is protected and used only for the purpose of providing care. Furthermore, Davendralingam, et al. [29] highlighted challenges of data security issues, legal and ethical considerations, and issues with standardizing clinical terms. All these challenges must be addressed to improve AI implications in pediatrics.

Opportunities for ML algorithms in Pediatric Medicine

Clarke, et al. [1] found that AI has improved the precision in diagnosis, with ML models able to identify abnormal findings from normal clinical radiographs of pediatric patients [30]. Pediatric ophthalmology could also benefit from AI in detecting retinopathy of prematurity (ROP) and congenital cataracts, taking into account unique aspects of designing AI applications that differ from adults [31]. Diseases such as asthma require stratification due to heterogeneity in disease severity and response to clinical treatment and trajectories, and ML can improve the classification of such diseases [32]. AI-based algorithms, such as Natural Language Processing (NLP), are also effective in drug development, identifying the most specific druggable targets [33].

One potential application of AI/ML is the prediction of patient deterioration, which can help clinicians respond quickly to this potentially life-threatening condition. AI/ML can analyze large amounts of patient data, including vital signs, laboratory results, and medical history, and use this information to predict which patients are at risk of worsening [34]. Another opportunity for AI and ML in PICUs is in the personalized treatment of patients by analyzing patients' data, medical history and current condition to develop individualized treatment plans. For example, AI and ML can be used to determine the most appropriate medications and doses based on a child's age, weight, and medical history.

Computational advancement helps track the progress of children with behavioral problems over time. AI can identify changes in behavior and determine the effectiveness of different treatments by analyzing data from multiple sources, such as medical records, behavioral assessments, and patient reports. This information can then be used to make informed decisions about future treatment plans and to adjust them as needed [35]. The sophisticated DL algorithms can analyze data from various sources, including genetic data, family history, and environmental factors; AI can therefore identify children who are at risk of developing behavioral problems of ADHD, CD and ASD. This information can then be used to provide early interventions and support.

AI/ML also has enormous potential in diagnosing syndromic conditions such as Prader-Willi syndrome, Angelman syndrome and Huntington disease in pediatrics. AI can process vast amounts of genetic information about the aforementioned conditions and identify patterns that may not be immediately apparent to the human eye [36]. For example, AI can help identify specific patterns in genetic data that may indicate a particular syndrome, such as Prader-Willi syndrome, allowing for a faster and more accurate diagnosis. This can be particularly useful in diagnosing rare or complex syndromes where traditional methods may be insufficient and give false negative or false positive results. By using AI, pediatricians can make more informed decisions about testing and treatment, leading to improved outcomes for children and their families. Another opportunity is using AI for image analysis, such as in diagnosing craniofacial syndromes such as Goldenhar syndrome, trisomy 21 and DiGeorge syndromes. AI algorithms can analyze facial images and identify specific features that are indicative of a particular syndrome, allowing for a more accurate and objective diagnosis. This can be particularly useful in detecting syndromes early and improving treatment outcomes. In addition, AI can help reduce the time and resources required for diagnosis, allowing pediatricians to focus on providing care to children in need. By automating routine tasks, such as data collection and analysis, pediatricians can spend more time with patients and families, improving the overall quality of care.

Filipow [37] reported that ML can potentially diagnose chronic respiratory conditions such as chronic obstructive pulmonary disease and chronic airway obstruction. Shu, et al. [38] discussed several applications of AI; however, it is crucial to improve the predictive architecture of algorithms to improve diagnostic outcomes and precision. Furthermore, it is widely acknowledged in the literature that the introduction of an AI model into the clinical setting should be based on a thorough assessment of its precision, diagnostic efficacy, and reliability. This evaluation process can serve as a valuable lesson for future iterations of apps such as Babylon, encouraging the implementation of extensive validation practices to enhance their utility and effectiveness.

Explainable AI (XAI)

AI can provide a great benefit to pediatric medicine by assisting in the diagnosis, treatment, and management of diseases. However, it is also essential to explain the decision-making process to both doctors and patients to build trust and provide a better understanding of the reasoning behind a particular diagnosis or treatment recommendation. This need led to the development of explainable artificial intelligence (XAI), which aims to make AI processes more transparent and interpretable. Deep learning algorithms, in particular, can benefit from XAI, as they learn every aspect of the decision-making process on their own, known as "neural weights" [39,40].

There are two main approaches to XAI [40]: model-based and post-hoc explanation-based. Model-based methods make the ML model directly interpretable to medical practitioners, while post-hoc explanation-based methods translate the model into a more understandable format. The counterfactual explanation is an extension of XAI that can help identify and supplement diagnostic indicators. It allows medical professionals to ask, "what if" questions, such as "how would this disease look in an adult?" or "how likely is this diagnosis if the patient were 10 years older?" These methods can be particularly useful in pediatrics, as they can help make connections to general medicine and supplement traditional DL methods.

Limitations

Despite the potential benefits of AI/ML in pediatric medicine and with the recent emergence of AI 2.0 tools such as ChatGPT, Bard, and GLASS A.I 2.0 potentially offering benefits for diagnosing and managing diseases; deploying these tools in practice has several limitations. One significant limitation is the context-specific nature of ML. For example, when developing an ML model to predict hospital mortality in children admitted to the ICU, it is important to ensure that the model performs well on a validation cohort from a different ICU to demonstrate its ability to generalize in different contexts. Similarly, an ML model designed to analyze as-thma in a specific population e.g., ethnic group, should be able to generalize effectively to other diverse populations as well. These examples highlight the importance of testing and validating ML models across different populations and contexts to ensure their reliability and applicability in diverse settings.

Another limitation in using ML in pediatric medicine is the lack of experiential learning with these models. Experiential learning allows clinicians to learn from their experiences and improve their decision-making over time. However, ML models are pre-determined and do not have the ability to learn from new experiences unless specifically retrained, which can lead to inaccurate predictions or an inability to adapt to changes in the patient's condition over time. Additionally, when there are limited training samples available, statistical noise may arise in the ML predictions. These challenges can be compounded by other biases, such as those related to gender or demographics [41].

To address these limitations, researchers have developed various strategies, including the use of data augmentation to address category imbalance, interpretation and explanation-based ML methods, and monitoring the deployment context [42]. Additionally, collecting demographic data is crucial for mitigating biases in the data, especially given the potential for compounding biases related to gender and demographics in pediatric medicine [41]. Overall, while ML tools offer tremendous potential for improving pediatric medicine, their deployment comes with significant limitations that need to be addressed. Context-specific testing, validation, and monitoring are crucial to ensure that these tools are effective and accurate in different patient populations. Additionally, researchers must find ways to incorporate experiential learning into these models to ensure that they adapt and improve over time (**Fig. 3**). This process starts with defining the problem to be solved using ML, followed by gathering and cleaning relevant data. Exploratory analysis provides a basic understanding of the data and helps identify potential biases in the dataset. The next step is to build a model by training and evaluating the dataset, which is a critical phase in the cycle. As new data becomes available, the model needs to be retrained to ensure its accuracy and effectiveness. Additionally, experiential learning is essential in the healthcare industry, where new medicines and treatments are continually being introduced. As such, ML models need to be continuously trained and adapted to incorporate the latest medical knowledge. Finally, the model is deployed for inference/predictions, and its efficiency is continually monitored and improved by adding more data to mitigate bias. In summary, the ML life cycle is an ongoing process that requires continuous training and adaptation to accommodate changes in data and context and to incorporate the latest medical knowledge.

CONCLUSION

Many healthcare providers depend on clinical decision support tools integrated within the EHRs to enhance patient safety and achieve improved outcomes. With the advancements in AI and ML, these tools have become even more crucial in the practice of medicine. In particular, AI has been integrated into many areas of pediatrics, such as predicting PTSD among children's survivors, early diagnosis of leukemia, and detecting PKU cases with high accuracy. However, AI/ML deployment in pediatrics is still facing several challenges, such as data quality issues, unnoticed, hidden clinical variables, complex clinical data, and lack of clinical labels. To ensure the successful integration of AI/ML into pediatric healthcare, specific or generalized improvements in AI design and validation are necessary.

It is important to acknowledge that AI/ML models are context-specific and need testing in diverse populations to ensure their generalizability and identify any biases within the data. While AI tools could provide valuable virtual support to clinicians and parents by answering questions and providing information about symptoms, treatments, and medications, they are limited by the quality of the data they are trained on and may be biased based on the population they are tested on [1]. Additionally, ML models lack the ability to adapt to new situations, which limits their potential for experiential learning. It is important to recognize these limitations and to highlight the importance of human judgement and expertise in the use of AI/ML in pediatric healthcare.

In conclusion, it is evident that AI/ML has the potential to significantly enhance patient outcomes in pediatric medicine. Although, there are challenges to their deployment, advancements in AI design and validation, and

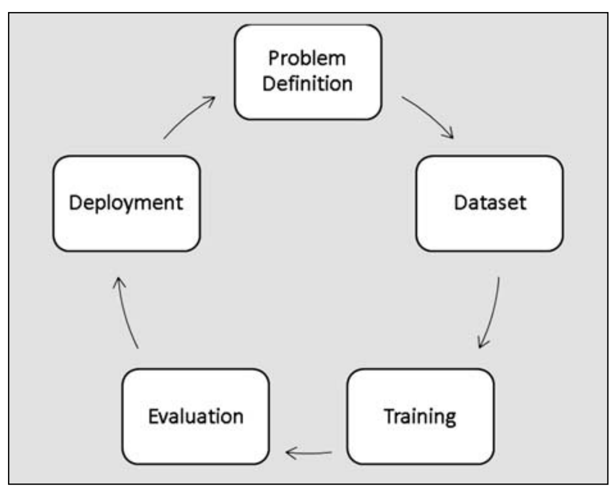

Fig. 3 Machine learning life cycle is a continuous process that requires ongoing training and adaptation of models to accommodate changes in data and context (*see text for description*).

testing in diverse populations can help to overcome these issues. Further research and practical support are recommended to explore areas not yet covered in the current literature. Ultimately, AI should be viewed as a tool to enhance and support clinical decision-making rather than a substitute for human judgement and expertise.

Contributors: YB: conceptualizing the review, providing clinical expertise, and contributing to the selection of review articles and discussion of findings. ST: important contribution to data analysis, structuring the themes, and ensuring a cohesive presentation of the findings. DW: contributed expertise in explainable AI (XAI) and data bias. All authors have contributed, designed and approved the manuscript.

Funding: None; *Competing interests:* None stated.

REFERENCES

1. Clarke SL, Parmesar K, Saleem MA, Ramanan AV. Future of machine learning in paediatrics. Arch Dis Child. 2022; 107: 223-8.
2. Biswas SS. Role of Chat GPT in Public Health. Ann Biomed Eng. 2023;51:868-9.
3. 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;47:1-5.
4. Abdullah IS, Loganathan A, Lee RW. ChatGPT & doctors: the medical dream team. [Feb 15, 2023]. Himmelfarb Health Sciences Library, The George Washington University on University Health Sciences Research Commons. Accessed on May 10, 2023. Available from: *https://hsrc.himmelfarb. gwu.edu/smhs_URGENT_Matters?utm_source=hsrc.himmelfarb. gwu.edu%2Fsmhs_URGENT_Matters%2F1 &utm_medium =PDF&utm_campaign=PDFCoverPages*
5. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2017;2: 230-43.
6. Knake LA. Artificial intelligence in pediatrics: the future is now. Pediatr Res. 2023;93:445-6.
7. Ferguson C, et al. Europe PMC in 2020. Nucleic Acids Res. 2021;49:D1507-14.
8. Selçuk AA. A guide for systematic reviews: PRISMA. Turk Arch Otorhinolaryngol. 2019;57:57-8.
9. Ge F, Li Y, Yuan M, Zhang J, Zhang W. Identifying predictors of probable posttraumatic stress disorder in children and adolescents with earthquake exposure: A longitudinal study using a machine learning approach. J Affect Disord. 2020;264:483-93.
10. Saxe GN, Ma S, Ren J, Aliferis C. Machine learning methods to predict child posttraumatic stress: a proof of concept study. BMC Psychiatry. 2017;17:1-13.
11. Fathi E, Rezaee MJ, Tavakkoli-Moghaddam R, Alizadeh A, Montazer A. Design of an integrated model for diagnosis and classification of pediatric acute leukemia using machine learning. Proc Inst Mech Eng H. 2020;234:1051-69.
12. Zhu Z, Gu J, Genchev GZ, 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.
13. Jia J, Wang R, An Z, Guo Y, Ni X, Shi T. RDAD: a machine learning system to support phenotype-based rare disease diagnosis. Front Genet. 2018;9:587.
14. Lin H, et al. Prediction of myopia development among Chinese school-aged children using refraction data from electronic medical records: A retrospective, multicentre machine learning study. PLoS Med. 2018;15:1-17.
15. Zhang M, Gazzard G, Fu Z, Li L, Chen B, Saw SM, Congdon N. Validating the accuracy of a model to predict the onset of myopia in children. Invest Ophthalmol Vis Sci. 2011;52: 5836-41.
16. Wu TE, Chen HA, Jhou MJ, Chen YN, Chang TJ, Lu CJ. Evaluating the effect of topical atropine use for myopia control on intraocular pressure by using machine learning. J Clin Med. 2020;10:111.
17. Ogundele MO. Behavioural and emotional disorders in childhood: A brief overview for paediatricians. World J Clin Pediatr. 2018;7:9-26.
18. Aylward BS, Abbas H, Taraman S, et al. An introduction to artificial intelligence in developmental and behavioral pediatrics. J Dev Behav Pediatr. 2023;44:e126-34.
19. Anagnostopoulou P, Alexandropoulou V, Lorentzou G, Lykothanasi A, Ntaountaki P, Drigas A. Artificial intelligence in autism assessment. International Journal of Emerging Technologies in Learning. 2020;15:95-107.
20. Elankeerthana R, Hanshiya D, Pranavajothi A, Shanmugapriya N, Sujitha S. Child abuse mental symptom prediction model using machine learning techniques. *In*: The 6th International Conference on Trends in Electronics and Informatics (ICOEI). Tirunelveli, India, 2022:1459-65. Accessed on May 10, 2023. Available from: *https://ieeexplore. ieee.org/document/9777138*
21. Kissos L, Goldner L, Butman M, Eliyahu N, Lev-Wiesel R. Can artificial intelligence achieve human-level performance? A pilot study of childhood sexual abuse detection in self-figure drawings. Child Abuse Negl. 2020;109:104755.
22. Amrit C, Paauw T, Aly R, Lavric M. Identifying child abuse through text mining and machine learning. Expert Systems With Applications. 2017;88:402-18.
23. Gupta P, Gossett J, Rao Rettiganti M. Trends in mortality rates in pediatric intensive care units in the United States from 2004 to 2015. Critical Care Medicine. 2018;46:30.
24. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019;17:195.
25. Tirunagari S, Bull SC, Poh N. Automatic classification of irregularly sampled time series with unequal lengths: A case study on estimated glomerular filtration rate. *In*: The 26th IEEE International Workshop on Machine Learning for Signal Processing (MLSP). Vietri sul Mare, Italy, 2016. Accessed on May 10, 2023. Available from: *https://ieeexplore. ieee.org/*

document/7738901
26. Xiao C, Choi E, Sun J. Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. J Am Med Inform Assoc. 2018;25:1419-28.
27. Vogl TM. Artificial intelligence and organizational memory in government: the experience of record duplication in the child welfare sector in Canada. *In*: The 21st Annual International Conference on Digital Government Research, June 2020. Accessed on May 10, 2023. Available from: *https://dl.acm.org/doi/abs/10.1145/3396956.3396971*
28. McCartney M. Margaret McCartney: AI in medicine must be rigorously tested. BMJ. 2018;361:k1752.
29. Davendralingam N, Sebire NJ, Arthurs OJ, Shelmerdine SC. Artificial intelligence in paediatric radiology: Future opportunities. Br J Radiol. 2021;94:20200975.
30. Chen KC, Yu HR, Chen WS, et al. Diagnosis of common pulmonary diseases in children by X-ray images and deep learning. Sci Rep. 2020;10:17374.
31. Reid JE, Eaton E. Artificial intelligence for pediatric ophthalmology. Curr Opin Ophthalmol. 2019;30:337-46.
32. Deliu M, Fontanella S, Haider S, et al. Longitudinal trajectories of severe wheeze exacerbations from infancy to school age and their association with early-life risk factors and late asthma outcomes. Clin Exp Allergy. 2020;50:315-24.
33. Bravo À, Piñero J, Queralt-Rosinach N, Rautschka M, Furlong LI. Extraction of relations between genes and diseases from text and large-scale data analysis: implications for translational research. BMC Bioinformatics. 2015;16:55.
34. Ziegler A, Wilichowski E, Schara U, et al. Handlungsempfehlungen zur Gentherapie der spinalen Muskelatrophie mit Onasemnogene Abeparvovec – AVXS-101 : Konsensuspapier der deutschen Vertretung der Gesellschaft für Neuropädiatrie (GNP) und der deutschen Behandlungszentren unter Mitwirkung des Medizinisch-Wissenschaftlichen Beirates der Deutschen Gesellschaft für Muskelkranke (DGM) e./ V [Recommendations for gene therapy of spinal muscular atrophy with onasemnogene abeparvovec-AVXS-101: Consensus paper of the German representatives of the Society for Pediatric Neurology (GNP) and the German treatment centers with collaboration of the medical scientific advisory board of the German Society for Muscular Diseases (DGM)]. Nervenarzt. 2020;91:518-29. German.
35. Hunt X, Tomlinson M, Sikander S, et al. Artificial intelligence, big data, and mhealth: the frontiers of the prevention of violence against children. Front Artif Intell. 2020;3:543305.
36. Tornese G, Pellegrin MC, Barbi E, Ventura A. Pediatric endocrinology through syndromes. Eur J Med Genet. 2020;63:103614.
37. Filipow N, Main E, Sebire NJ, et al. Implementation of prognostic machine learning algorithms in paediatric chronic respiratory conditions: a scoping review. BMJ Open Respir Res. 2022;9:e001165.
38. Shu LQ, Sun YK, Tan LH, Shu Q, Chang AC. Application of artificial intelligence in pediatrics: past, present and future. World J Pediatr. 2019;15:105-8.
39. Adadi A, Berrada M. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE Access. 2018;6:52138-60.
40. van der Velden BHM, Kuijf HJ, Gilhuijs KGA, Viergever MA. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. Med Image Anal. 2022;79:102470.
41. 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.
42. Vokinger KN, Feuerriegel S, Kesselheim AS. Mitigating bias in machine learning for medicine. Commun Med (Lond). 2021;1:25.