

Review

Not peer-reviewed version

Artificial Intelligence in Pediatric Emergency Medicine: Applications, Challenges, and Future Perspectives

[Lorenzo Di Sarno](#) ^{*}, Anya Caroselli , Giovanna Tonin , Benedetta Graglia , Valeria Pansini , Francesco Andrea Causio , [Antonio Gatto](#) , [Antonio Chiaretti](#)

Posted Date: 25 April 2024

doi: [10.20944/preprints202404.1642.v1](https://doi.org/10.20944/preprints202404.1642.v1)

Keywords: artificial intelligence; machine learning; pediatrics; deep learning; pediatric emergency medicine



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

Artificial Intelligence in Pediatric Emergency Medicine: Applications, Challenges, and Future Perspectives

Lorenzo Di Sarno ^{1,2*}, Anya Caroselli ¹, Giovanna Tonin ³, Benedetta Graglia ¹, Valeria Pansini ³, Francesco Andrea Causio ^{2,4}, Antonio Gatto ^{2,3} and Antonio Chiaretti ^{1,2}

¹ Department of Pediatrics. Fondazione Policlinico Universitario "A. Gemelli", IRCCS, 00168, Rome, Italy. Università Cattolica del Sacro Cuore, 00168, Rome, Italy; lorenzodisarno1993@gmail.com (L.D.S.); anya0196@gmail.com (A. Caroselli); benedetta.graglia@gmail.com (B.G.); antonio.chiaretti@policlinicogemelli.it (A. Chiaretti)

² Italian Society for Artificial Intelligence in Medicine (Società Italiana Intelligenza Artificiale in Medicina, SIIAM), Rome, Italy

³ Department of Pediatrics. Fondazione Policlinico Universitario "A. Gemelli", IRCCS, 00168, Rome, Italy; giovannatoninp@gmail.com (G.T.); pansini.valeria@gmail.com (P.V.); antonio.gatto@policlinicogemelli.it (A.G.)

⁴ Section of Hygiene and Public Health, Department of Life Sciences and Public Health, Università Cattolica del Sacro Cuore, 00168, Rome, Italy; francescoandrea.causio@unicatt.it (F.A.C.)

* Correspondence: lorenzodisarno1993@gmail.com; Tel.: +39 0630154475

Abstract: The dawn of Intelligence (AI) in healthcare stands as a milestone in medical innovation. A plethora of different medical fields are heavily involved and pediatric emergency medicine is no exception. These new tools do not merely provide more advanced and efficient systems for patients' diagnosis, management and treatment. They rather concern a strict shift from traditional methods based upon broad categories towards a more personalized healthcare. AI offers many promises in pediatric emergency medicine with a wide range of applications involving clinical decision making, patients' flows management and prioritization. Main barriers to a widespread diffusion involve technological challenges but also ethical issues and the paucity of extensive datasets in pediatric contexts. We conducted a narrative review structured in two parts. The first part explores the theoretical principles of AI, providing all the necessary background to feel confident with these new state-of-the-art tools. The second part presents an informative analysis of AI models in pediatric emergencies, pointing out the actual applications and challenges until future feasible research perspectives.

Keywords: artificial intelligence; machine learning; pediatrics; deep learning; pediatric emergency medicine

1. Introduction

Artificial intelligence (AI), once considered a distant, futuristic project, has gradually become a real-world approach in a wide range of medical fields, including pediatric emergency medicine[1–3].

The innovative concept of machines capable of autonomously elaborating and processing information dates back to the 50s[4]. Over time, medicine has progressively been recognized as a breeding ground for AI[2]. Nowadays, the latest generation of AI can use large datasets to analyze interactions between the included variables and develop predictions with several use cases in healthcare[2].

Since standard clinical decision-making often relies on strict flowcharts and classifications with a well-structured approach, these new AI tools could induce a significant shift from conventional methods[3,5]. Even though effective in many cases, traditional models tend to oversimplify, occasionally missing the complexity of medical conditions. AI, through big data processing, is moving from broad categories to detailed and "point-like" classes in healthcare, ultimately tailoring the diagnosis and care pathway to the patients' needs and ultimately fostering a more personalized

medicine[6,7]. However, the development and deployment of AI solutions requires a drastic shift. AI training requires a significant amount of data, readily available and machine-readable. This can slow down the adoption of AI solutions in facilities where the digitalization process is not as advanced as necessary[3]. Additionally, the introduction of new tools requires the end users, including healthcare professionals and physicians, to be equipped with knowledge and skills to embrace them. The upskilling and reskilling are a pressing issue, hindering future advances in patient care and improvements in health outcomes. Therefore, a new approach is needed to combine AI and high-quality input data[8]. This narrative review aims to provide a conceptual introduction to AI and raise awareness of its emerging clinical tools and potential applications in pediatric emergency medicine.

2. Materials and Methods

We designed a narrative review of the literature. The first part explores the theoretical principles of AI, providing all the necessary background to feel confident with these new state-of-the-art tools. The second part presents an informative analysis of AI models in pediatric emergencies, pointing out the actual applications and challenges until future feasible research perspectives. We examined the following bibliographic electronic databases: PubMed and the Cochrane Library, from inception date until April 2024. The search was limited to English-language papers that focused on AI in pediatric emergency medicine.

Each selected paper was reviewed and analyzed in full text by two authors (L.D.S. and A. Caroselli) and any discrepancies among them were solved by debate. Due to the heterogeneity of the articles examined, we focused on a qualitative analysis.

3. Artificial Intelligence and subfields

Artificial Intelligence (AI) is the field of study that focuses on how computers learn from data and on the development of algorithms that enable this learning[3].

AI involves numerous applications capable of processing information in non-conventional ways[9]. In order to achieve the best performance, AI requires the management of large datasets known as "Big data"[10]. "Big data" is a term that was introduced in the 1990s to include datasets too large to be managed by common software[11]. The vast amount of information about patients' health in massive digital archives is the source of big data in healthcare. As a matter of fact in recent years there has been a progressive trend from paper-based to digitized data[12,13].

Big data in healthcare can be characterized by up to six main features, the so called "6 Vs", according to different authors:

- Volume: the continuous and exponentially incremental flow of data spanning from personal medical records up to 3D imaging, genomics, and biometric sensor readings ought to be carefully managed[12]. Innovations in data management, such as virtualization and cloud computing, are enabling healthcare organizations to store and manipulate large amounts of data more efficiently and cost-effectively[14];
- Velocity: the prompt and rapid transmission of data is a pivotal item nowadays, especially in scenarios like trauma monitoring, anesthesia in operating rooms, and bedside heart monitoring, where timely data analysis can be life-saving[12]. Besides, future applications, such as early infection detection and targeted treatments based upon real-time data, have the potential to notably decrease morbidity, mortality, and ultimately impact on outcome[14,15];
- Variety: the ability to analyze large datasets, including multimedia and unstructured formats, represents an innovation in healthcare[12]. The wide range of structured, unstructured, and semi-structured data analyzed, stands as a revolutionary change that adds complexity to healthcare data management[16]. Structured data can be easily stored, recalled, elaborated and manipulated by machinery. They come from a variety of sources, including diagnoses, medications, instrument readings, and lab values, and can be sorted into numeric or categorical

fields for easy analysis[12,17]. Unstructured data is commonly generated at the point of care, including free-form text such as medical notes or discharge summaries and multimedia content such as imaging[12,17]. The main challenge is to transform this data to make it suitable for AI analysis, but this process faces some obstacles. First, adding structure to unstructured data entails healthcare providers to manually review charts or images, sort the information out and enter it into the system[18]. This makes the process slow, inefficient, and prone to bias. New powerful tools such as Natural Language Processing can speed up and streamline the information extraction process[17]. Secondly, healthcare professionals' preference for the natural language simplicity of handwritten notes remains a major barrier to a widespread adoption of electronic health records, which require field coding at the point of care to provide structured inputs[12].

- Variability refers to the consistency of data over time[16]. Data flows are unpredictable, they change often and vary widely. It's essential to know how to manage daily, seasonal, and event-driven data spikes[19].
- Veracity: ensuring that big data is accurate and trustworthy is critical in healthcare, where accurate information can mean the difference between life and death[12]. Nevertheless, achieving veracity faces challenges, including variable quality and difficulties in ensuring accuracy, especially with handwritten prescriptions.
- Value consists of the worth of information to various stakeholders or decision makers[20].

Big data includes clinical data sourced from Computerized Physician Order Entry (CPOE) and Clinical Decision Support (CDS) systems, as well as patient information stored in electronic patient records (EPRs), and machine-generated/sensor data, including vital sign monitoring[12]. Big data analytics may improve care and reduce costs by identifying connections and understanding patterns and trends among different items[21]. In fact, it could potentially enhance healthcare outcomes through information elaboration, healthcare provider guidance, preventative care candidates identification, and disease profiling[12].

In regular healthcare analytics, project analysis is typically performed using easy-to-use business intelligence tools on stand-alone systems; however, in big data analytics, the processing of large datasets is distributed across multiple nodes, requiring a shift in user interfaces[22]. While traditional analytics tools are easy to use and transparent, new tools are complex, programming intensive, and require different skill sets to be most effective.

In order to guarantee an adequate output, this huge amount of data has to be verified by valid tools. **Blockchain** is a technology characterized by the decentralization of entries, meaning that inputs are agreed upon by a peer-to-peer network through various consensus protocols, rather than a central authority controlling the content[23]. Furthermore, many blockchains offer anonymity or pseudo-anonymity[23,24]. Specifically for healthcare data management, these features ensure data security and privacy through a network of secure blocks linked by cryptographic protocols[23,25]. Another key feature of blockchain is persistency: once data is inserted into a block and added to the chain, it cannot be deleted[23]. This implies that if an inaccurate data is added to the blockchain, it becomes a permanent part of the ledger. Thus, it is important to ensure the accuracy of data before adding it to the blockchain[26].

The integration of AI and blockchain is promising: AI tools could leverage information acquired from a secure, unchangeable, and decentralized system for storing sensitive data required by AI-driven techniques[25]. An integration of AI and blockchain in the metaverse has been proposed in order to provide digital healthcare through realistic interactions[27]. By using blockchain for data security and privacy, healthcare providers and patients engage in consultations in a virtual environment: participants are represented by avatars, and consultation data is securely recorded and stored on the blockchain[27]. This data is then used by explainable AI models to predict and diagnose

diseases, ensuring logical reasoning, trust, transparency, and interpretability in the diagnostic process[27].

Machine learning (ML), a subset of computer science and artificial intelligence, seeks to identify patterns in data to boost the effectiveness of various tasks[28]. In healthcare, ML uses automated, adaptive and computationally advanced techniques to recognize patterns within complex data structures [12]. ML models improve their performance by means of a continuous auto-training process [13,14]. This approach differs from "traditional" methods and explicit human programming, which rely on certain statistical assumptions and require a predefined set of dimensions, functional relationships, and interactions [12,14] - an issue often avoided in ML.

To develop a reliable ML model, accurate training datasets are required; therefore, a preprocessing phase is usually needed[3]. Most of the data is used to train the model with preliminary analyses performed to identify the strongest relationships between variables and study outcomes. The remaining data can be used for internal validation. At this stage, the model can be tested on different datasets[3].

ML-aided tasks have already been incorporated into clinical practice, especially in imaging interpretation[29–31]. Although they are still imperfect and require a skilled supervisor, they are considered acceptable when rapid image feedback is needed and local expertise is lacking[9]. A growing number of applications have been developed. Some of them, combining clinical, genetic, and laboratory items are able to detect rare or common conditions that would otherwise be missed[9]. ML is divided into three branches, which are selected according to one of the three required research tasks[32,33]: supervised ML for prediction, unsupervised ML for description, and reinforcement learning for causal inference.

Supervised ML (SML) is a predictive model, designed to estimate the likelihood of an event occurring[3]. The predictive analytics applied span from basic computations such as correlation coefficients or risk differences to advanced pattern recognition techniques and supervised learning algorithms like random forests and neural networks, which serve as classifiers or predict the joint distribution of multiple variables[33]. The supervised ML development process involves three subsets of data. First, a training set of labeled data (e.g. histological specimens that have already been labeled as normal or diseased by a human expert) is provided for the algorithm to learn by adjusting weights to minimize the loss of function which calculates the distance between the predicted and true outcome for a given data point[32,34]. Next, the model parameters are optimized using a second validation set[32]. The validation set can also detect overfitting, which is observed when model performance is significantly better on the training set. Finally, a third set is used to evaluate the model's ability to generalize to new datasets[32]. Once the training session upon labeled data is completed, then the system is applicable to unlabeled data. In this way the trained models predict outcomes through either classification or regression, respectively in categorical or continuous types of data[34].

Decision trees (DTs) are non-parametric supervised learning algorithms used for classification[35,36]. They map attribute values to classes and have a tree-like structure, including a root node, branches, internal nodes (or decision nodes), and leaf nodes (or terminal nodes)[36,37]. Nodes conduct evaluations, based on the available information, and each node is linked to two or more subtrees or leaf nodes labeled with a class, representing a possible outcome[36,37]. Various types of DTs can be used both in classification and regression tasks[36]. DTs are often preferred over other methods in fields such as healthcare due to their interpretability, despite being less accurate. This is because they are easier to understand and explain compared to other, more complex methods that might be more accurate but relatively uninterpretable[35].

Unsupervised ML (UML) is used for descriptive tasks, with the goal of data clustering and revealing relationships within a data structure[33]. Descriptive tasks provide quantitative summaries of specific features in a certain scenario and require analytics ranging from simple calculations to complex techniques[33]. The main goal of unsupervised learning is to identify inherent groupings or clusters within a data structure, in order to find out data differences, similarities and distributions in feature space[3,28,32]. In unsupervised ML systems, training is data-driven, rather than human-

driven, and uses unlabeled data (compared to Supervised ML, whose training features labeled data and is driven by human experts)[33]. This category lacks a guiding response variable during analysis[28].

Reinforcement Learning (RL) is a computational approach where an agent learns to achieve a goal through a trial-and-error cycle in an interactive environment[32]. The agent's decision-making strategy is improved through its interaction with the environment[38]. The goal of RL is the selection of actions that will maximize future rewards[38]. This is achieved through iterative learning cycles resulting in a reward or penalty in relation to a pre-defined target[34]. For instance, since there is a need for blood glucose concentration monitoring and for an ideal determination of time and amount for insulin delivery in diabetic patients, RL algorithms are potentially capable of learning the individual glucose pattern of a diabetic patient in order to provide adaptive drug supply after a learning process[39]. Changes in the glucose level lead to an action of the agent, in terms of insulin injection or no treatment. Subsequently the agent receives a numerical reward, which along with the next glucose level will impact upon the next action[39].

AI's subsets are represented in Figure 1. ML's subtypes and algorithms are summarized in Figure 2.

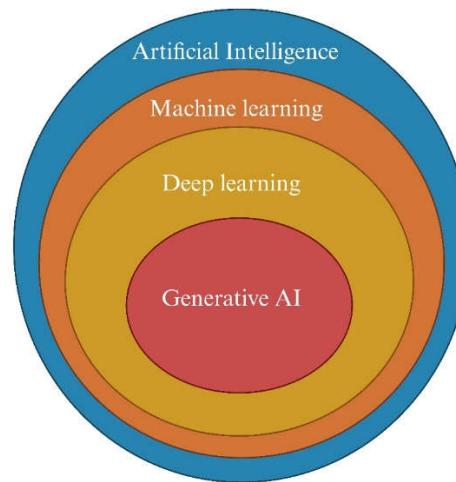


Figure 1. A comparative view of AI, machine learning, deep learning, and generative AI. Created with biorender.com.

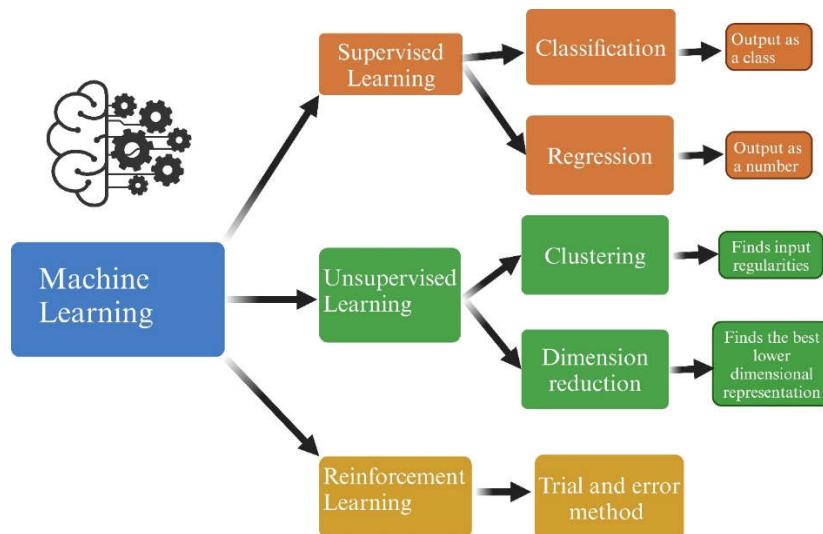


Figure 2. Machine learning algorithms. Created with biorender.com.

Deep Learning (DL) refers to an even more complex subgroup of ML based on numerous processing layers, which may use supervised, unsupervised and reinforcement ML applications[34,40]. In a certain way, it mimics how the human brain builds its own model of the world by learning from large amounts of sensory-motor data acquired through interactions with the environment[38]. DL differs from ML in a number of characteristics. ML requires "manual" feature extraction and processing[41]; it reaches a "plateau" where the quality of performance no longer increases with the volume of data; its training time is somewhat "limited"[32]. On the other hand DL is capable of automatically learning feature hierarchies; it requires a significant amount of data to make predictions, and because it is more computationally intensive than ML, it may require longer training times and state-of-the-art machines to run[32,34]. The complex architecture of DL, consisting of several processing layers that are mostly inaccessible to human users (the so-called "black box of AI"), may pose an issue for the model's accountability in healthcare[34].

The application of these models is potentially limitless, though not without risk.

Generative AI (GAI) is a type of DL technique that generates realistic facsimiles by evaluating training examples and learning their patterns and distribution[42]. GAI can produce various types of content by using existing sources such as text, audios, images, and videos[42]. One well-known example of its application is ChatGPT, an AI-driven chatbot. Its potential in supporting medical research and clinical practice is currently being assessed[43]. ChatGPT is based on Generative Pre-trained Transformers (GPT - a type of DL model that enables natural language processing) that generate human-like text based on large amounts of data[9,44]. A 2023 systematic review has evaluated its utility in several fields such as data analysis, literature review, scientific writing, medical record storage and management, up to generating diagnoses and management plans[45]. Frequently raised issues concerning generative AI in academic writing include bias, plagiarism, privacy, and legal concerns up to scientific fraud (e.g. fake image synthesis or convincing fraudulent articles resembling genuine scientific papers)[46–48]. Therefore the World Association of Medical Editors advises authors and editors to disclose chatbot use in their work and to provide themselves with tools for detecting AI-generated content[34,42]. Furthermore, GAI can also generate non-textual items (images, videos and audios).

Finally when applied to complex analysis of high-dimensional data, including clinical applications, DL can achieve remarkable outcomes[40], e.g. computer-assisted diagnosis of melanoma. In fact a Deep Convolutional Neural Network trained on images has achieved performances comparable to dermatology experts in identifying keratinocyte carcinomas and malignant melanomas[49].

Neural Networks (NNs) are the baseline architecture of DL models[50]. They are structured in multiple layers consisting of neuron-like interconnected nodes[32] (Figures 3a and 3b). Once inserted, data flows along the first layer to the structure of interconnected nodes in a "forward propagation" mechanism[32]. The signal received by each node is a result of a weighted linear combination of node outputs from the prior layer, meaning that they are multiplied by a weight assigned to each connection and summed up. A nonlinear transformation is instead applied to the node's output[32,50]. The final result from the output layer is compared to the true value and a "back propagation" algorithm optimizes results by using prediction error and adjusting weights[32]. Because NNs are highly parametrized, they might "over-fit" models to data: thus, a series of regularization strategies have been implemented to prevent it[40]. To name one, dropout is a regularization technique where random neurons are dropped, along with their connection, during training[51]. This prevents units from co-adapting too much and helps it to generalize better to unseen data.

All AI models used in clinics and research are summarized in Table 1.

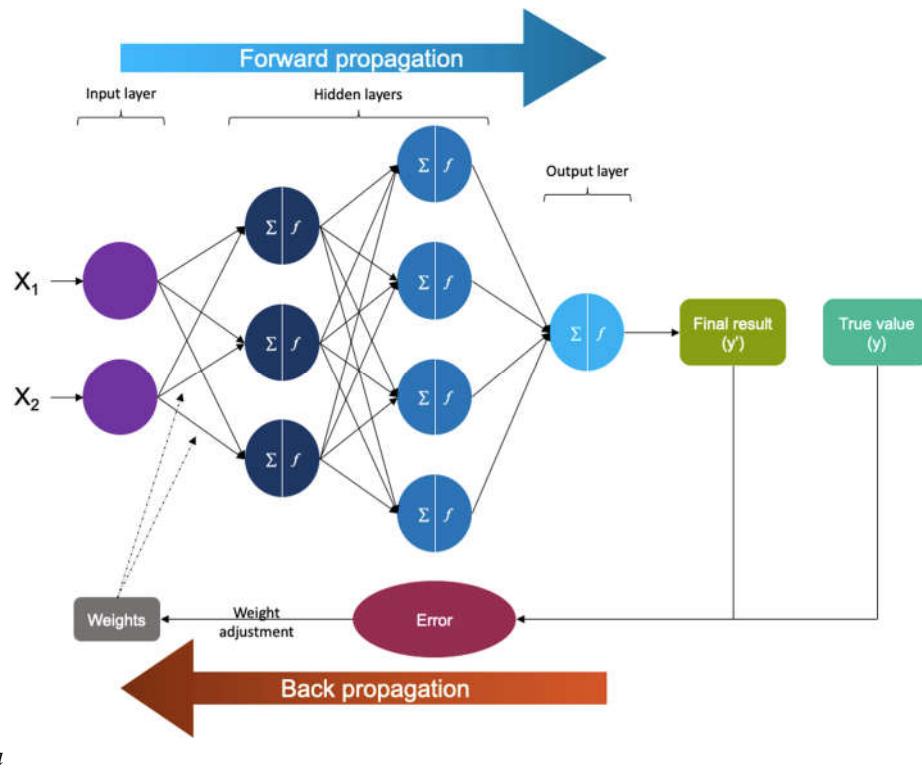
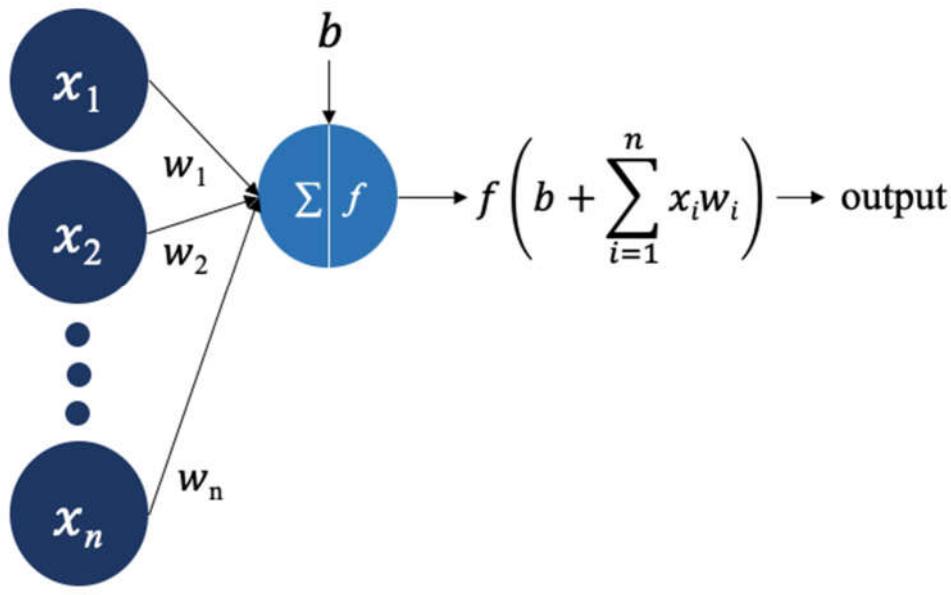
*a**b*

Figure 3. (a) Neural networks are the basic architecture of DL models. They are structured in multiple layers consisting of neuron-like interconnected nodes. Data flows through the input layer and into the structure of interconnected nodes in a "forward propagation" mechanism. The final result from the output layer is compared to the true value and a "back propagation" algorithm optimizes results by using prediction error and adjusting weights. (b) This is a close-up of a node. The signal received by each node is a result of a weighted linear combination of node outputs from the prior layer, meaning that they are multiplied by a weight assigned to each connection and summed up. A nonlinear transformation is instead applied to the node's output. $x_1 - x_n$: inputs; $w_1 - w_n$: weights; Σ : summation of weighted outputs from the previous layer; f : nonlinear transformation (activation function); b : bias.

Table 1. AI models used in clinics and research.

System	Description
Artificial Neural Network (ANN)	Nodes, akin to neurons, process information, while connections between layers, termed edges, simulate synapses with weights. Output is computed via mathematical operations on input and hidden layers, with the learning algorithm adjusting weights to minimize errors between predicted and target outputs, forming probability-weighted associations stored within the network's structure[52].
Backpropagation Neural Network	Backpropagation utilizes prediction errors to iteratively tune the weights, enabling the NN to learn patterns within the training data and enhance model accuracy over time[32].
Convolutional Neural Network (CNN)	CNNs process data that comes in the form of multiple arrays such as signals, images, audio spectrograms and videos, and is applied in the recognition of objects[50].
Deep Neural Network (DNN)	An ANN with numerous layers between the input and output layers which is capable of learning high-level features and requires high computational power[53].
Probabilistic Neural Network (PNN)	An application of DNN within probabilistic models, able to capture complex non-linear stochastic relationships between random variables[54].
Recurrent Neural Network (RNN)	A recurrent neural network (RNN) is any network whose neurons send feedback signals to each other, and are capable of modeling sequential data for sequence recognition and prediction[55,56].
Region-based Convolutional Neural Network (R-CNN)	R-CNN models use region based networks, which are capable to detect an object in an image and holds great potential especially in diagnostic imaging[57].
Multilayer Perceptron (MLP)	A feedforward type of powerful and dynamic ANN. The signals are transmitted within the network in one direction: from input to output[58].
Bayesian Inference	Bayesian statistical methods are applied to algorithms. They start with existing 'prior' beliefs, which are then updated using data to give 'posterior' beliefs, which may be used as the basis for inferential decisions[59].
Causal Associational Network (CASNET)	Three items constitute this model: patient observation, pathophysiological states, and disease classifications. Once documented, the observations are associated with the fitting states[60].
Light Gradient Boosting Machine (LightGBM)	LightGBM employs a boosting strategy to combine numerous decision trees, with each tree utilizing the negative gradient of the loss function as the residual approximation for fitting. It is designed for optimal performance, particularly in distributed systems[61].
Extreme Gradient Boosting (XGBoost)	XGBoost is a gradient boosting framework that is highly efficient and scalable. It features a proficient linear model solver and a tree learning algorithm. It enables diverse objective functions, such as regression, classification, and ranking. Its design allows for easy extension, enabling users to define custom objectives[62].
Natural Language Processing (NLP)	NLP is a subfield of AI and ML used to interpret linguistic data (e.g. clinical note analysis and decision making) [9,40].

Random Forest Models	Random forest models use randomization to create multiple decision trees, each contributing to the final output. In classification tasks, the trees' outputs are combined through voting, while in regression tasks, they are averaged to produce a single output[63].
----------------------	--

4. Current Research and Applications

Clinical Decision Support (CDS) systems can be defined as “computer systems designed to impact clinician decision-making about individual patients at the point in time that these decisions are made”[64]. These systems can be applied in a plethora of medicine fields, including pediatric emergencies. Nevertheless, CDS are not free of limitations, and may sometimes be even perceived as intrusive or ineffective by their users[3].

Compared to traditional rule-based CDS systems, AI-implemented CDS don't rely on statistical algorithms and are occasionally defined as “non-knowledge”-based CDS[65]. As shown in Figure 4, firstly a wide range of input data is inserted into the AI system. Secondly, this data is used to make predictions. Finally, when a certain threshold is reached, a best practice alert is given to healthcare providers.

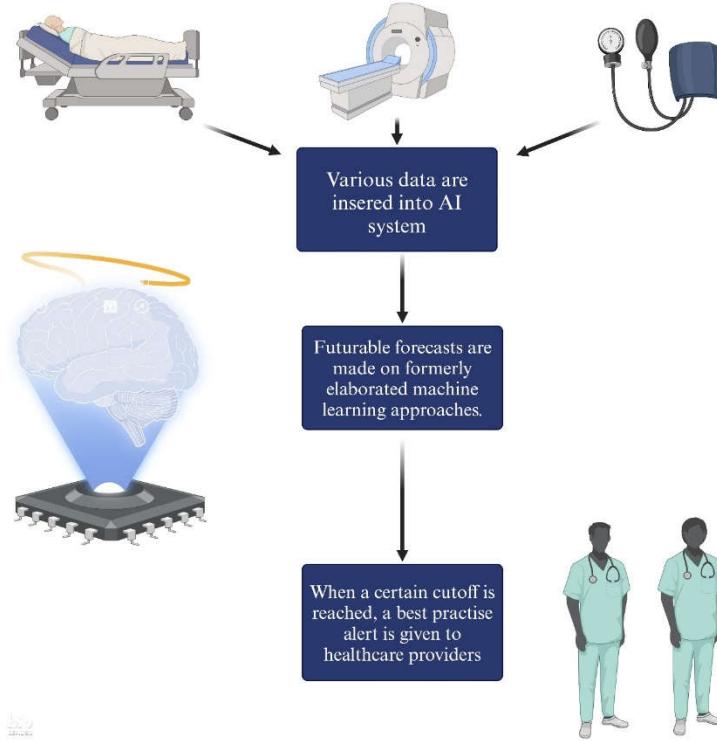


Figure 4. How AI-CDS works. Created with biorender.com.

4.1. AI for Triage Optimization

Triage is the process of quickly assessing sick children when they first present to Pediatric Emergency Departments (PED) in order to classify them into one of the following categories: those with emergency signs who require immediate treatment, those with priority signs who should be given priority in the queue so they can be treated without delay, and those who are non-urgent cases. An efficient triage, performed by healthcare providers, requires expertise but greatly relies on subjective judgement to risk-stratify patients. Several factors such as a first impression of critical care needs and a different subjective pain threshold of patients may affect triage evaluation[66–68]. By analyzing the data collected during the triage phase, AI, particularly ML, could be a useful adjunct tool for screening critical patients or those who are candidates for hospitalization[69].

Firstly in 2018 a multisite, retrospective, cross-sectional study of Emergency Department visits on adults compared an electronic triage system based on a random forest model applied to vital signs, chief complaint, and medical history to standard triage[70]. Electronic triage predictions demonstrated equivalent or improved detection of clinical patient outcomes[70].

Subsequently, a retrospective observation cohort study was conducted in Korea in 2019 on a wide range of pediatric patients[71]. They developed prognostic models of critical care and hospitalization through ML, DL, and conventional triage. The predicted variables adopted were the following: age, sex, chief complaint, symptom onset to arrival time, arrival mode, trauma, and vital signs. A DL-based algorithm was developed using a Multilayer Perceptron (MLP) [57] and derivation data, consisting of patient data from January 2014 to June 2016. Test data was then fed into the algorithm, and a risk score between 0 and 1 was obtained, corresponding to the risk of critical care, involving direct admission to the Pediatric Intensive Care Unit (PICU) from the PED or transfer to other hospitals for PICU admission. Two ML models were developed as well for performance comparison, respectively using logistic regression and random forest. The DL algorithm significantly outperformed the other methods both for intensive care and hospitalization prediction.

In 2022 a prognostic study tested the performance of ML methods in predicting clinical outcomes in children in PED and compared them to conventional approaches[72]. Four learning prediction models were developed: logistic regression with lasso regularization, random forest, gradient-boosted decision tree, and deep neural network. All of them showed better discrimination ability compared to conventional approaches for clinical outcomes, with a higher sensitivity for the critical care outcome and higher specificity for hospitalization[72]. In 2023 Sarty et al used administrative data collected through triage to train several machine learning models to predict patients who would leave the PED without being seen (LWBS) by an healthcare provider[73]. Among the models applied, XGBoost was the best-performing ML model with 95% recall and 87% sensitivity. The most influential factors in this model were PED patient load, triage hour, driving minutes from home address to PED, length of stay, and age. An earlier detection of LWBS would enable a possible development of patient-focused strategies aimed at limiting the phenomenon[73].

4.2. AI for Stress Management

The distressful experiences experimented during healthcare treatments in childhood have been linked to post-traumatic stress disorder and future avoidance of medical care as adults[74]. Stress reduction and compliance maximization are key elements in the management of pediatric patients, especially in emergency setting[75].

Innovative solutions targeting this include AI-enhanced Socially Assistive Robots (SAR), creating a patient-specific support experience aimed at reducing their discomfort during painful procedures such as peripheral intravenous line placement or surgical sutures. SAR systems generally create a relationship without physically touching patients but just through interactions that may include expressiveness, personality, dialog, empathy, and adaptation skills[76].

SARs have already been tested in the PED setting with great results in mitigating children's discomfort and pain but, to best perform, they need human inputs. They have a limited and scripted behavior and therefore, they lack a tailored and flexible feedback in this unpredictable emergency background[74,77]. Nevertheless, it is pivotal to acknowledge that the aim of the SAR is not to replace human involvement in mitigating distress and pain management, but to add a tool that healthcare providers may use.

Two clinical research articles collected concerns and expectations to implement the design of an AI-enhanced SAR[78,79]. Healthcare providers felt that the proposed SAR should be equipped with a wide range of skills to meet children's needs, involving encouraging dialogue, positive reinforcement expressions, humor, and cognitive behavioral strategies (e.g., breathing techniques and meditation) [76]. Caregivers saw the primary function of the SAR as a distractor during painful procedures. Specifically, they suggested behaviors including tricks, jokes, playing music, singing, and dancing [75].

4.3. AI for Traumatic Brain Injury Assessment

Head trauma is one of the leading causes of PED admissions worldwide[80]. Since the majority of them are constituted by minor traumas, it is essential to make a straight distinction to ensure a prompt management when needed[81]. CT (Computed Tomography) scan is considered the gold standard examination to detect a brain injury but it has several drawbacks: it is expensive, may eventually require sedation and boosts the risk of cancer in a lifetime due to radiations exposure[82]. A set of rules, known as the PECARN rules, has been validated to risk-stratify this category of patients to find out which patients ought to undergo a radiological imaging[83].

In recent years, AI approaches have been applied to the diagnosis and management of traumatic brain injury in pediatric patients[84,85].

In 2017 Dayan et al. implemented the PECARN rules with a multi-faced intervention, focused on computerized CDS, to provide the bedside pediatrician with directive indications for CT use and supporting data (e.g. risk estimates for TBI)[86]. CDS components involved: determination of whether the patient matched the age-specific PECARN very low risk criteria; a recommendation that CT was not indicated if the child met the very low risk criteria; risk estimations for clinically relevant TBI; and links to the prediction rule principles and paper. They showed this intervention was associated with modest, safe, but variable decreases in CT use without missing any relevant brain damages[86].

In 2018 Hale et al. developed an Artificial Neural Network (ANN) trained on clinical items and radiologist-interpreted imaging findings to identify patients at risk for clinically relevant TBI[87]. A total of 12,902 pediatric patients were enrolled in this study from the PECARN TBI dataset. Through the elaboration of clinical and radiological data ANN showed 99,73% sensitivity with 98.19% precision, 97.98% accuracy, 91.23% negative predictive value, 0.0027% false negative rate, and 60.47% specificity for clinically relevant TBI[87].

In 2019 Bertsimas et al. compared an approach based on optimal classification tree (OCT) with original PECARN rules in Traumatic Brain Injury (TBI) management, examining the same sample of children previously treated with PECARN rules[88]. OCTs are classification trees similar to the classification and regression trees (CARTs) that were used to derive the original PECARN rules but fitted with a novel method (Mixed Integer Optimization) that could outperform the classical CART-fitting algorithms[35]. Outcomes suggested that OCTs performed as well as or better than PECARN rules in identifying children at very low risk of clinically important traumatic brain injury[88]; thus, the potential application of OCTs may provide a valuable tool to decrease unnecessary CT scans while maintaining adequate sensitivity for identifying patients with clinically significant TBI.

In 2023 Miyagawa et al. used a decision tree method to predict the necessity of CT scans in children under 2 years of age with mild TBI. This kind of SML achieved this outcome with a rate of 95% of accuracy[89]. Focusing on the contribution of each predictor on the decision tree, days of life resulted to be the most significant. According to these findings, days of life could be used as a main factor for decision making for head trauma in children younger than 2 years of age, and could substitute age in years in clinical flowcharts[89].

Nowadays a significant proportion of pediatric neuroimaging performed attempts to generate outcome prediction of brain injury, either hypoxic and traumatic[90,91]. Subtle mild TBI anomalies not visualized on CT or on conventional MRI (Magnetic Resonance Imaging) can be detected through state-of-the-art neuroimaging. One of the most significant among them is diffusion MRI, which enables for qualitative and quantitative assessment of specific white matter tracts in the nervous system[92]. The study of the network of white nerve fibers, known as the "connectome," has recently received increasing attention[93]. Connectome mapping using post-processing methods through diffusion MRI-based fiber tracking, such as track density imaging and edge density imaging, is a new frontier in research because it can reveal abnormalities even in mild TBI, such as white matter damage not seen on CT and MRI[94]. Raji et al. used Support Vector Machines (SVMs) to analyze patients with TBI based on edge density imaging[95]. SVMs examine and group labelled data into classes, split by the widest plane (support vector). They are often employed when there is a non-linear correlation among data, and, as such, a separation line is not easily recognizable[96]. In their study,

Raji et al. identified three white matter regions distinguishing mild TBI from controls using edge density imaging maps. Bilateral tapetum, sagittal stratum and callosal splenium identified mild TBI subjects with sensitivity of 79% and specificity of 100%; accuracy from the area under the ROC curve (AUC) was 94%. In this study edge density imaging could provide better diagnostic delineation of pediatric mild TBI than neurocognitive assessment of memory or attention[95].

4.4. AI for Pediatric Sepsis Prediction

Sepsis is the main cause of death worldwide in pediatric patients resulting in an estimated 7.5 million deaths annually[97]. Sepsis is a life-threatening organ dysfunction associated with infection. Prompt and accurate identification of sepsis requires data-driven screening tools with affordable precision and high sensitivity[98]. Several scores for sepsis identification have been designed over years, even though they often lack in sensibility and specificity[99,100].

Systemic Inflammatory Response Syndrome (SIRS) criteria were previously included in pediatric sepsis definition by Goldstein et al. but they had poor predictive properties[101].

In 2024 the Society of Critical Care Medicine task force suggested that sepsis in children is defined by a Phoenix Sepsis Score of at least 2 points in children with suspected infection, meaning potentially life-threatening dysfunction of respiratory, cardiovascular, coagulation, or neurological systems[98]. Even though the Phoenix sepsis criteria performed well, future independent validation is needed, especially in low-resources settings. In recent years computerized sepsis prediction systems have been developed to overcome intrinsic limitations of pediatric sepsis scores. By accessing electronic health record (EHR) data for clinical decision support, these systems can early detect septic patients whose treatment would otherwise be delayed[102]. In 2018 Kamaleswaran et al. conducted an observational cohort study. They analyzed continuous minute-by-minute physiological data of 493 PICU patients over a timeline of 24 hours to assess the onset of severe sepsis[103]. 20 of this cluster of patients developed severe sepsis. The authors demonstrated that AI could identify patients with severe sepsis before they clinically show relevant findings by just assessing quantifiable physiomarkers, such as heart rate (HR), mean blood pressure (MBP), systolic blood pressure (SBP), diastolic blood pressure (DBP), and oxygen saturation (Spo2). Furthermore, these algorithms were able to detect severe sepsis 8 hours earlier than a currently implemented real-time electronic screening tool in critically ill children. These findings pointed out how is pivotal for bedside monitors to be combined with artificial intelligence to improve their predictivity of severe sepsis.

In 2019 Le et al. tested a ML-based prediction algorithm using EHR[104]. The ML system adopted was based upon boosted ensembles of decision trees. Ensemble classifiers sum up the output from weak apprentices, each of which would be inadequate to solve the problem autonomously, generating a more efficient learner. Every single baseline learner in this paper was a decision tree. Each tree was constructed by repeatedly splitting the feature space, acquiring thresholds within the features which most reduce entropy, and therefore enhance information[104]. Their algorithm outperformed in terms of sensibility, specificity and accuracy the Pediatric Logistic Organ Dysfunction score (PELOD-2) ($p < 0.05$) and the pediatric Systemic Inflammatory Response Syndrome (SIRS) score ($p < 0.05$) in the prediction of severe sepsis[104].

In 2022 Stella et al. used a different approach which focused on predicting the need of resuscitation within 6 hours of triage, rather than diagnosing sepsis[105]. In this way the model could provide actionable decisional support. Data was extracted from EHC and involved demographics, triage vitals, triage nurse comments, chief complaint information, as well as orders placed and medications administered within 6 hours of arrival. Several models were employed including: standard and regularized regression, random forests, gradient boosted trees and generalized additive[105]. Moving the aim from the diagnosis of severe sepsis to the provision of resuscitative care allowed avoiding the development of severe sepsis, rather than treating the full blown condition[105].

In 2023 Mercurio et al. conducted a retrospective observational study of children presenting to a PED at a tertiary care children's hospital with fever, hypotension, or an infectious disease International Classification of Diseases (ICD)-10 diagnosis[106]. They proved that combining clinical

and sociodemographic variables, sensibility and specificity performance of ML methods were as high as 93% and 84% respectively in identifying patients with diagnosis of sepsis. The random forest classifier performed the best, followed by a classification and regression tree[106]. The maximum recorded heart rate and the maximum mean arterial pressure resulted the two most significant factors in determining the model. Other unexpected variables such as age, immunization status and demographics data have proved relevant in early detection of sepsis as well[106].

5. Discussion

Although further validation is required, AI might represent an useful supporting tool for PED decision-making and has the potential to improve the timely allocation of resources and interventions[72]. Recognizing and addressing a series of barriers is crucial for a safe development of efficient AI tools[107,108].

The first key challenge is ensuring AI an accurate input. Expert clinicians ought to select features to be accurately included in datasets for ML training because inaccurate training datasets lead to suboptimal diagnostic accuracy[109]. ML subsequent iterations may amplify these errors, reinforcing biases introduced in the early phases[32]. In most cases monocentric datasets fail to correctly address the heterogeneity of pediatric conditions, particularly in the Emergency Department. Pediatric populations accessing different PEDs may differ heavily, particularly when comparing rural and urban PEDs, leading to significant biases in the training phase. Failure to account for this can lead to misdiagnosis[110]. Additionally, the peculiarities of the emergency-urgency network of a certain region may lead to some centers being the reference hubs for a subset of pathologies (e.g., neurological disorders), leading to an overrepresentation of certain conditions in the dataset.

These datasets should be the result of a multidisciplinary team of professionals working towards a predefined goal. As a notorious quote says: “garbage in – garbage out”, meaning input datasets are crucial in determining the final outcome[111]. Ensuring data quality requires a dedicated infrastructure, i.e. patient health records, diagnostic images, and real-time data monitoring. Relying excessively on paper-based systems and having operator-built databases introduces several potential pitfalls for error, potentially reducing the final accuracy and reproducibility of the algorithms[112]. Confirmation bias is another item to be considered. AI recognizes patterns on which it has trained upon but unlikely identifies what it is not taught[113].

To this purpose, ensuring external validation is vital for quality control. When applied to an external dataset a certain clinical support tool could result neither sensible or specific[91]. Besides, the algorithm's performance can quickly reduce as clinical practices dynamically evolve. Hence, a continuous influx of data is pivotal to refine the model and keep it current[32].

Some of these limitations are particularly true for pediatrics. Children stand only for a marginal portion of healthcare resources and their datasets are frequently small[3]. Granular data (i.e., displaying a high level of detail in the way data is structured), ideal for ML applications, is rarely available in pediatrics. Targeting a right balance between granularity and simplicity is a key factor in optimizing AI performance and ensuring significant outcomes from complex datasets[114]. Furthermore, in children data is not homogenously distributed resulting in inequality with some data sharply prevailing on others due to significative variations in features according to the patient's age. AI models ought to account for variations and changes in disease risk that occur according to age[115]. Vital and auxological parameters have to be interpreted according to age as well. Even radiologic images need to be interpreted taking into account the age and the resulting changes in anatomy, physiopathology and possible differential diagnosis. All the previously mentioned issues could be challenging in validating an automated analysis[116].

The absence of evidence-based and variability in care are other factors limiting the application of AI in pediatrics. In fact, a wide range of pediatric conditions has no gold standard care universally shared and treatment significantly varies among different institutions[117,118].

The integration of AI and ML into pediatric emergency wards demands a workforce that is proficient in these technologies. Healthcare professionals must be upskilled or reskilled to understand the capabilities and limitations of AI and ML, interpret the outputs of these systems, and

integrate this information into clinical decision-making processes[119]. Training programs and continuous education initiatives are essential to equip healthcare professionals with the knowledge and skills needed to work alongside AI and ML technologies effectively[120]. This not only enhances the quality of patient care but also ensures that healthcare professionals can remain competitive and adapt to the evolving landscape of healthcare technology.

Most of the studies retrieved by our review were conducted outside of the European Union. This can be partly linked to the attention that AI-assisted software and medical devices receive in the EU, e.g. compared to the United States. The EU's recently approved AI Act[121] categorizes AI as "high risk" when it is implemented in health and care; this wary approach is not far from what was previously observed in the General Data Protection Regulation and similar norms, which impose stricter regulations on the use of health data for research and training purposes[122]. Most of the studies we reviewed were conducted in the U.S.A., whose Food and Drug Administration has issued several documents defining AI and ML software as "medical devices"[123], providing additional guidance on good practice to develop them. A recent scoping-review suggested similar results on the broader field of clinical trials[124].

Finally, AI poses ethical questions, especially when it comes to liability. Will pediatricians be responsible for eventual consequences that do not fit predictions? At present time clinicians may be liable for harm to patients if they observe AI indications to use nonstandard care methods. Current law only protects doctors from liability when they follow the standard care. Nevertheless, as AI gets integrated in gold standards, we could speculate physicians would probably avoid liability when following AI indications[125].

6. Conclusions

AI offers many promises in pediatric healthcare with a wide range of applications involving clinical decision making, patients' flows management and prioritization. In a bunch of years AI could integrally reshape how we approach pediatric emergencies. On the other side it raises technical, professional, and moral queries as well. Main barriers to a widespread diffusion involve technological challenges but also ethical issues, age-dependent variations in data interpretation, and the paucity of extensive datasets in pediatric contexts.

Overall, the next steps in pediatric AI are aimed at matching the actual gap between research and clinical practice, striking a delicate balance between its transformative potential and current limitations.

Author Contributions: Conceptualization, L.D.S. and A. Chiaretti; writing – original draft preparation, L.D.S. and A. Caroselli; writing – review and editing, G.T., B.G., L.D.S., A. Caroselli, and F.A.C.; Supervision, A. Chiaretti, A.G., and V.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Manickam P, Mariappan SA, Murugesan SM, *et al.* Artificial Intelligence (AI) and Internet of Medical Things (IoMT) Assisted Biomedical Systems for Intelligent Healthcare. *Biosensors* 2022;12:562. doi:10.3390/bios12080562
2. Yu K-H, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng* 2018;2:719–31. doi:10.1038/s41551-018-0305-z
3. Ramgopal S, Sanchez-Pinto LN, Horvat CM, *et al.* Artificial intelligence-based clinical decision support in pediatrics. *Pediatr Res* 2023;93:334–41. doi:10.1038/s41390-022-02226-1
4. Turing AM. Computing Machinery and Intelligence. *Mind* 1950;59:433–60. doi:10.1093/mind/LIX.236.433
5. Rajula HSR, Verlato G, Manchia M, *et al.* Comparison of Conventional Statistical Methods with Machine Learning in Medicine: Diagnosis, Drug Development, and Treatment. *Med* 2020;56:455. doi:10.3390/medicina56090455
6. Cirillo D, Valencia A. Big data analytics for personalized medicine. *Curr Opin Biotechnol* 2019;58:161–7. doi:10.1016/j.copbio.2019.03.004

7. Mesko B. The role of artificial intelligence in precision medicine. *Expert Rev Precis Med Drug Dev* 2017;2:239–41. doi:10.1080/23808993.2017.1380516
8. Nijman J, Zoodsma RS, Koomen E. A Strategy for Artificial Intelligence With Clinical Impact—Eyes on the Prize. *JAMA Pediatr* 2024;178:219. doi:10.1001/jamapediatrics.2023.6259
9. Hossain E, Rana R, Higgins N, et al. Natural Language Processing in Electronic Health Records in relation to healthcare decision-making: A systematic review. *Comput Biol Med* 2023;155:106649. doi:10.1016/j.combiomed.2023.106649
10. Benke K, Benke G. Artificial Intelligence and Big Data in Public Health. *Int J Environ Res Public Health* 2018;15:2796. doi:10.3390/ijerph15122796
11. Mallappallil M, Sabu J, Guessner A, et al. A review of big data and medical research. *SAGE Open Med* 2020;8:205031212093483. doi:10.1177/2050312120934839
12. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Heal Inf Sci Syst* 2014;2:3. doi:10.1186/2047-2501-2-3
13. Frost, Sullivan. Drowning in Big Data? Reducing Information Technology Complexities and Costs for Healthcare Organizations. *WSEAS Trans. Comput. Res.* 2016;4. 123-131. https://www.researchgate.net/publication/310416741_Healthcare_Big_Data_and_Cloud_Computing (accessed 12 Feb 2024).
14. Feldman B, Martin E, Skotnes T. Big Data in Healthcare: Hype and Hope. 2012. <https://www.yumpu.com/en/document/view/29226285/big-data-in-healthcare-hype-and-hope> (accessed 13 Feb 2024).
15. Hoover W. Transforming Health Care Through Big Data: Strategies for leveraging big data in the health care industry. 2013.
16. Ristevski B, Chen M. Big Data Analytics in Medicine and Healthcare. *J Integr Bioinform* 2018;15. doi:10.1515/jib-2017-0030
17. Li I, Pan J, Goldwasser J, et al. Neural Natural Language Processing for unstructured data in electronic health records: A review. *Comput Sci Rev* 2022;46:100511. doi:10.1016/j.cosrev.2022.100511
18. Kamran S. Natural Language Processing in Healthcare Explained. 2023. <https://www.consensus.com/blog/natural-language-processing-in-healthcare/> (accessed 15 Feb 2024).
19. SAS Big Data - What It Is and Why It Matters. https://www.sas.com/en_us/insights/big-data/what-is-big-data.html (accessed 21 Feb 2024).
20. Hermon R, Williams PAH. Big data in healthcare: What is it used for? In: *Australian Ehealth Informatics and Security Conference*. Perth, Western Australia: : SRI Security Research Institute, Edith Cowan University 2014. 40–9. doi:10.4225/75/57982b9431b48
21. Wang Y, Kung L, Byrd TA. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technol Forecast Soc Change* 2018;126:3–13. doi:10.1016/j.techfore.2015.12.019
22. Elgendi N, Elragal A. Big Data Analytics: A Literature Review Paper. In: *Lecture Notes in Computer Science*. 2014. 214–27. doi:10.1007/978-3-319-08976-8_16
23. Hasselgren A, Kraljevska K, Gligoroski D, et al. Blockchain in healthcare and health sciences—A scoping review. *Int J Med Inform* 2020;134:104040. doi:10.1016/j.ijmedinf.2019.104040
24. Lax G, Russo A. Blockchain-Based Access Control Supporting Anonymity and Accountability. *J Adv Inf Technol* 2020;11:186–91. doi:10.12720/jait.11.4.186-191
25. Tagde P, Tagde S, Bhattacharya T, et al. Blockchain and artificial intelligence technology in e-Health. *Environ Sci Pollut Res* 2021;28:52810–31. doi:10.1007/s11356-021-16223-0
26. Zheng Z, Xie S, Dai H, et al. An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends. In: *2017 IEEE International Congress on Big Data (BigData Congress)*. IEEE 2017. 557–64. doi:10.1109/BigDataCongress.2017.85
27. Ali S, Abdullah, Armand TPT, et al. Metaverse in Healthcare Integrated with Explainable AI and Blockchain: Enabling Immersiveness, Ensuring Trust, and Providing Patient Data Security. *Sensors* 2023;23:565. doi:10.3390/s23020565
28. Jiang T, Gradus JL, Rosellini AJ. Supervised Machine Learning: A Brief Primer. *Behav Ther* 2020;51:675–87. doi:10.1016/j.beth.2020.05.002
29. Maghami M, Sattari SA, Tahmasbi M, et al. Diagnostic test accuracy of machine learning algorithms for the detection intracranial hemorrhage: a systematic review and meta-analysis study. *Biomed Eng Online* 2023;22:114. doi:10.1186/s12938-023-01172-1
30. Schaffter T, Buist DSM, Lee CI, et al. Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms. *JAMA Netw Open* 2020;3:e200265. doi:10.1001/jamanetworkopen.2020.0265
31. Zhang F. Application of machine learning in CT images and X-rays of COVID-19 pneumonia. *Medicine (Baltimore)* 2021;100:e26855. doi:10.1097/MD.0000000000026855
32. Mueller B, Kinoshita T, Peebles A, et al. Artificial intelligence and machine learning in emergency medicine: a narrative review. *Acute Med Surg* 2022;9:1–10. doi:10.1002/ams2.740

33. Hernán MA, Hsu J, Healy B. A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks. *CHANCE* 2019;32:42–9. doi:10.1080/09332480.2019.1579578
34. Theodosiou AA, Read RC. Artificial intelligence, machine learning and deep learning: Potential resources for the infection clinician. *J Infect* 2023;87:287–94. doi:10.1016/j.jinf.2023.07.006
35. Bertsimas D, Dunn J. Optimal classification trees. *Mach Learn* 2017;106:1039–82. doi:10.1007/s10994-017-5633-9
36. Decision Trees. <https://www.ibm.com/topics/decision-trees> (accessed 23 Feb 2024).
37. Podgorelec V, Kokol P, Stiglic B, et al. Decision trees: an overview and their use in medicine. *J Med Syst* 2002;26(5):445–63. doi:10.1023/a:1016409317640
38. Matsuo Y, LeCun Y, Sahani M, et al. Deep learning, reinforcement learning, and world models. *Neural Networks* 2022;152:267–75. doi:10.1016/j.neunet.2022.03.037
39. Bothe MK, Dickens L, Reichel K, et al. The use of reinforcement learning algorithms to meet the challenges of an artificial pancreas. *Expert Rev Med Devices* 2013;10:661–73. doi:10.1586/17434440.2013.827515
40. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction to natural language processing. *BMC Med Res Methodol* 2021;21:158. doi:10.1186/s12874-021-01347-1
41. Stafie CS, Sufaru I-G, Ghicicu CM, et al. Exploring the Intersection of Artificial Intelligence and Clinical Healthcare: A Multidisciplinary Review. *Diagnostics* 2023;13:1995. doi:10.3390/diagnostics13121995
42. Zielinski C, Winker M, Aggarwal R, et al. Chatbots, generative AI, and scholarly manuscripts. WAME recommendations on chatbots and generative artificial intelligence in relation to scholarly publications. 2023.<https://wame.org/page3.php?id=106>. (accessed 16 Feb 2024).
43. Stokel-Walker C, Van Noorden R. What ChatGPT and generative AI mean for science. *Nature* 2023;614:214–6. doi:10.1038/d41586-023-00340-6
44. Birhane A, Kasirzadeh A, Leslie D, et al. Science in the age of large language models. *Nat Rev Phys* 2023;5:277–80. doi:10.1038/s42254-023-00581-4
45. Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. *Healthcare* 2023;11:887. doi:10.3390/healthcare11060887
46. Liverpool L. AI intensifies fight against ‘paper mills’ that churn out fake research. *Nature* 2023;618:222–3. doi:10.1038/d41586-023-01780-w
47. Gu J, Wang X, Li C, et al. AI-enabled image fraud in scientific publications. *Patterns* 2022;3:100511. doi:10.1016/j.patter.2022.100511
48. Májovský M, Černý M, Kasal M, et al. Artificial Intelligence Can Generate Fraudulent but Authentic-Looking Scientific Medical Articles: Pandora’s Box Has Been Opened. *J Med Internet Res* 2023;25:e46924. doi:10.2196/46924
49. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542:115–8. doi:10.1038/nature21056
50. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436–44. doi:10.1038/nature14539
51. Nitish S, Hinton G, Krizhevsky A, et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J Mach Learn Res* 2014;1929–58.
52. Si T, Bagchi J, Miranda PBC. Artificial Neural Network training using metaheuristics for medical data classification: An experimental study. *Expert Syst Appl* 2022;193:116423. doi:10.1016/j.eswa.2021.116423
53. Sze V, Chen Y-H, Yang T-J, et al. Efficient Processing of Deep Neural Networks: A Tutorial and Survey. *Proc IEEE* 2017;105:2295–329. doi:10.1109/JPROC.2017.2761740
54. Masegosa AR, Cabañas R, Langseth H, et al. Probabilistic Models with Deep Neural Networks. *Entropy* 2021;23:117. doi:10.3390/e23010117
55. Grossberg S. Recurrent neural networks. *Scholarpedia* 2013;8:1888. doi:10.4249/scholarpedia.1888
56. Salehinejad H, Sankar S, Barfett J, et al. Recent Advances in Recurrent Neural Networks. *arXiv Prepr* Published Online First: 28 December 2017. doi:10.48550/arXiv.1801.01078
57. Choi BW, Kang S, Kim HW, et al. Faster Region-Based Convolutional Neural Network in the Classification of Different Parkinsonism Patterns of the Striatum on Maximum Intensity Projection Images of [18F]FP-CIT Positron Emission Tomography. *Diagnostics* 2021;11:1557. doi:10.3390/diagnostics11091557
58. Popescu MC, Balas VE, L. P-P, et al. Multilayer perceptron and neural networks. *WSEAS Trans Circuits Syst* 2009;:579–558.
59. Spiegelhalter D, Rice K. Bayesian statistics. *Scholarpedia* 2009;4:5230. doi:10.4249/scholarpedia.5230
60. Raita Y, Camargo CA, Liang L, et al. Big Data, Data Science, and Causal Inference: A Primer for Clinicians. *Front Med* 2021;8. doi:10.3389/fmed.2021.678047
61. Ji X, Chang W, Zhang Y, et al. Prediction Model of Hypertension Complications Based on GBDT and LightGBM. *J Phys Conf Ser* 2021;1813:012008. doi:10.1088/1742-6596/1813/1/012008
62. Chen T, He T, Benesty M, et al. Xgboost: extreme gradient boosting. *R Packag version* 2015;:0.4-2, 1(4), 1–4.<https://cran.ms.unimelb.edu.au/web/packages/xgboost/vignettes/xgboost.pdf>
63. Rigatti SJ. Random Forest. *J Insur Med* 2017;47:31–9. doi:10.17849/insm-47-01-31-39.1

64. Berner ES, La Lande TJ. Overview of Clinical Decision Support Systems. In: *Clinical Decision Support Systems*. 2007. 3–22. doi:10.1007/978-0-387-38319-4_1

65. Sutton RT, Pincock D, Baumgart DC, *et al.* An overview of clinical decision support systems: benefits, risks, and strategies for success. *npj Digit Med* 2020;3:17. doi:10.1038/s41746-020-0221-y

66. Green NA, Durani Y, Brecher D, *et al.* Emergency Severity Index Version 4. *Pediatr Emerg Care* 2012;28:753–7. doi:10.1097/PEC.0b013e3182621813

67. Thomas D, Kircher J, Plint AC, *et al.* Pediatric Pain Management in the Emergency Department: The Triage Nurses' Perspective. *J Emerg Nurs* 2015;41:407–13. doi:10.1016/j.jen.2015.02.012

68. Di Sarno L, Gatto A, Korn D, *et al.* Pain management in pediatric age. An update. *Acta Biomed* 2023;94:e2023174. doi:10.23750/abm.v94i4.14289

69. Hwang S, Lee B. Machine learning-based prediction of critical illness in children visiting the emergency department. *PLoS One* 2022;17:1–11. doi:10.1371/journal.pone.0264184

70. Levin S, Toerper M, Hamrock E, *et al.* Machine-Learning-Based Electronic Triage More Accurately Differentiates Patients With Respect to Clinical Outcomes Compared With the Emergency Severity Index. *Ann Emerg Med* 2018;71:565–574.e2. doi:10.1016/j.annemergmed.2017.08.005

71. Kwon J, Jeon K-H, Lee M, *et al.* Deep Learning Algorithm to Predict Need for Critical Care in Pediatric Emergency Departments. *Pediatr Emerg Care* 2021;37:e988–94. doi:10.1097/PEC.00000000000001858

72. Goto T, Camargo CA, Faridi MK, *et al.* Machine Learning-Based Prediction of Clinical Outcomes for Children During Emergency Department Triage. *JAMA Netw Open* 2019;2:e186937. doi:10.1001/jamanetworkopen.2018.6937

73. Sarty J, Fitzpatrick EA, Taghavi M, *et al.* Machine learning to identify attributes that predict patients who leave without being seen in a pediatric emergency department. *CJEM* 2023;25:689–94. doi:10.1007/s43678-023-00545-8

74. Trost MJ, Ford AR, Kysh L, *et al.* Socially Assistive Robots for Helping Pediatric Distress and Pain. *Clin J Pain* 2019;35:451–8. doi:10.1097/AJP.0000000000000688

75. Sanchez Cristal N, Staab J, Chatham R, *et al.* Child Life Reduces Distress and Pain and Improves Family Satisfaction in the Pediatric Emergency Department. *Clin Pediatr* 2018;57:1567–75. doi:10.1177/0009922818798386

76. Trost MJ, Chrysilla G, Gold JI, *et al.* Socially-Assistive Robots Using Empathy to Reduce Pain and Distress during Peripheral IV Placement in Children. *Pain Res Manag* 2020;2020:1–7. doi:10.1155/2020/7935215

77. Chita-Tegmark M, Scheutz M. Assistive Robots for the Social Management of Health: A Framework for Robot Design and Human–Robot Interaction Research. *Int J Soc Robot* 2021;13:197–217. doi:10.1007/s12369-020-00634-z

78. Nishat F, Hudson S, Panesar P, *et al.* Exploring the needs of children and caregivers to inform design of an artificial intelligence-enhanced social robot in the pediatric emergency department. *J Clin Transl Sci* 2023;7:e191. doi:10.1017/cts.2023.608

79. Hudson S, Nishat F, Stinson J, *et al.* Perspectives of Healthcare Providers to Inform the Design of an AI-Enhanced Social Robot in the Pediatric Emergency Department. *Children* 2023;10:1–13. doi:10.3390/children10091511

80. Mastrangelo M, Midulla F. Minor Head Trauma in the Pediatric Emergency Department: Decision Making Nodes. *Curr Pediatr Rev* 2018;13. doi:10.2174/1573396313666170404113214

81. Schutzman SA, Greenes DS. Pediatric minor head trauma. *Ann Emerg Med* 2001;37:65–74. doi:10.1067/mem.2001.109440

82. Da Dalt L, Parri N, Amigoni A, *et al.* Italian guidelines on the assessment and management of pediatric head injury in the emergency department. *Ital J Pediatr* 2018;44:7. doi:10.1186/s13052-017-0442-0

83. Kuppermann N, Holmes JF, Dayan PS, *et al.* Identification of children at very low risk of clinically-important brain injuries after head trauma: a prospective cohort study. *Lancet* 2009;374:1160–70. doi:10.1016/S0140-6736(09)61558-0

84. Tunthanathip T, Oearsakul T. Application of machine learning to predict the outcome of pediatric traumatic brain injury. *Chinese J Traumatol = Zhonghua chuang shang za zhi* 2021;24:350–5. doi:10.1016/j.cjtee.2021.06.003

85. Ellethy H, Chandra SS, Nasrallah FA. The detection of mild traumatic brain injury in paediatrics using artificial neural networks. *Comput Biol Med* 2021;135:104614. doi:10.1016/j.combiomed.2021.104614

86. Dayan PS, Ballard DW, Tham E, *et al.* Use of Traumatic Brain Injury Prediction Rules With Clinical Decision Support. *Pediatrics* 2017;139. doi:10.1542/peds.2016-2709

87. Hale AT, Stonko DP, Lim J, *et al.* Using an artificial neural network to predict traumatic brain injury. *J Neurosurg Pediatr* 2019;23:219–26. doi:10.3171/2018.8.PEDS18370

88. Bertsimas D, Dunn J, Steele DW, *et al.* Comparison of Machine Learning Optimal Classification Trees with the Pediatric Emergency Care Applied Research Network Head Trauma Decision Rules. *JAMA Pediatr* 2019;173:648–56. doi:10.1001/jamapediatrics.2019.1068

89. Miyagawa T, Saga M, Sasaki M, *et al.* Statistical and machine learning approaches to predict the necessity for computed tomography in children with mild traumatic brain injury. *PLoS One* 2023;18:e0278562. doi:10.1371/journal.pone.0278562

90. Ellethy H, Chandra SS, Nasrallah FA. Deep Neural Networks Predict the Need for CT in Pediatric Mild Traumatic Brain Injury: A Corroboration of the PECARN Rule. *J Am Coll Radiol* 2022;19:769–78. doi:10.1016/j.jacr.2022.02.024

91. Wong A, Otles E, Donnelly JP, *et al.* External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients. *JAMA Intern Med* 2021;181:1065. doi:10.1001/jamainternmed.2021.2626

92. Eierud C, Craddock RC, Fletcher S, *et al.* Neuroimaging after mild traumatic brain injury: Review and meta-analysis. *NeuroImage Clin* 2014;4:283–94. doi:10.1016/j.nicl.2013.12.009

93. Shah HA, Mehta NH, Saleem MI, *et al.* Connecting the connectome: A bibliometric investigation of the 50 most cited articles. *Clin Neurol Neurosurg* 2022;223:107481. doi:10.1016/j.clineuro.2022.107481

94. Payabvash S, Palacios EM, Owen JP, *et al.* White Matter Connectome Edge Density in Children with Autism Spectrum Disorders: Potential Imaging Biomarkers Using Machine-Learning Models. *Brain Connect* 2019;9:209–20. doi:10.1089/brain.2018.0658

95. Raji CA, Wang MB, Nguyen N, *et al.* Connectome mapping with edge density imaging differentiates pediatric mild traumatic brain injury from typically developing controls: proof of concept. *Pediatr Radiol* 2020;50:1594–601. doi:10.1007/s00247-020-04743-9

96. Ben-Hur A, Weston J. A User's Guide to Support Vector Machines. 2010. 223–39. doi:10.1007/978-1-60327-241-4_13

97. Ruth A, McCracken CE, Fortenberry JD, *et al.* Pediatric Severe Sepsis. *Pediatr Crit Care Med* 2014;15:828–38. doi:10.1097/PCC.0000000000000254

98. Schlapbach LJ, Watson RS, Sorce LR, *et al.* International Consensus Criteria for Pediatric Sepsis and Septic Shock. *JAMA* 2024;331:665. doi:10.1001/jama.2024.0179

99. Yu SC, Shivakumar N, Betthauser K, *et al.* Comparison of early warning scores for sepsis early identification and prediction in the general ward setting. *JAMIA Open* 2021;4. doi:10.1093/jamiaopen/ooab062

100. Uffen JW, Oosterheert JJ, Schweitzer VA, *et al.* Interventions for rapid recognition and treatment of sepsis in the emergency department: a narrative review. *Clin Microbiol Infect* 2021;27:192–203. doi:10.1016/j.cmi.2020.02.022

101. Goldstein B, Giroir B, Randolph A. International pediatric sepsis consensus conference: Definitions for sepsis and organ dysfunction in pediatrics*. *Pediatr Crit Care Med* 2005;6:2–8. doi:10.1097/01.PCC.0000149131.72248.E6

102. Balamuth F, Alpern ER, Grundmeier RW, *et al.* Comparison of Two Sepsis Recognition Methods in a Pediatric Emergency Department. *Acad Emerg Med* 2015;22:1298–306. doi:10.1111/acem.12814

103. Kamaleswaran R, Akbilgic O, Hallman MA, *et al.* Applying Artificial Intelligence to Identify Physiomarkers Predicting Severe Sepsis in the PICU. *Pediatr Crit Care Med* 2018;19:e495–503. doi:10.1097/PCC.0000000000001666

104. Le S, Hoffman J, Barton C, *et al.* Pediatric Severe Sepsis Prediction Using Machine Learning. *Front Pediatr* 2019;7. doi:10.3389/fped.2019.00413

105. Stella P, Haines E, Aphinyanaphongs Y. Prediction of Resuscitation for Pediatric Sepsis from Data Available at Triage. In: *AMIA Annu Symp Proc*. 2021. 1129–38. <http://www.ncbi.nlm.nih.gov/pubmed/35308977>

106. Mercurio L, Pou S, Duffy S, *et al.* Risk Factors for Pediatric Sepsis in the Emergency Department. *Pediatr Emerg Care* 2023;39:e48–56. doi:10.1097/PEC.0000000000002893

107. Rajkomar A, Hardt M, Howell MD, *et al.* Ensuring Fairness in Machine Learning to Advance Health Equity. *Ann Intern Med* 2018;169:866. doi:10.7326/M18-1990

108. Amershi S, Weld D, Vorvoreanu M, *et al.* Guidelines for Human-AI Interaction. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: : ACM 2019. 1–13. doi:10.1145/3290605.3300233

109. Wiens J, Saria S, Sendak M, *et al.* Do no harm: a roadmap for responsible machine learning for health care. *Nat Med* 2019;25:1337–40. doi:10.1038/s41591-019-0548-6

110. Ventura CAI, Denton E. Artificial Intelligence Chatbots and Emergency Medical Services: Perspectives on the Implications of Generative AI in Prehospital Care. *Open Access Emerg Med* 2023;15:289–92. doi:10.2147/OAEM.S420764

111. Kilkenny MF, Robinson KM. Data quality: “Garbage in – garbage out”. *Heal Inf Manag* 2018;47:103–5. doi:10.1177/1833358318774357

112. He J, Baxter SL, Xu J, *et al.* The practical implementation of artificial intelligence technologies in medicine. *Nat Med* 2019;25:30–6. doi:10.1038/s41591-018-0307-0

113. Challen R, Denny J, Pitt M, *et al.* Artificial intelligence, bias and clinical safety. *BMJ Qual Saf* 2019;28:231–7. doi:10.1136/bmjqqs-2018-008370

114. Ramgopal S, Adler MD, Horvat CM. Application of the Improving Pediatric Sepsis Outcomes Definition for Pediatric Sepsis to Nationally Representative Emergency Department Data. *Pediatr Qual Saf* 2021;6:e468. doi:10.1097/pq9.0000000000000468
115. Lee B, Kim K, Hwang H, *et al.* Development of a machine learning model for predicting pediatric mortality in the early stages of intensive care unit admission. *Sci Rep* 2021;11:1263. doi:10.1038/s41598-020-80474-z
116. Padash S, Mohebbian MR, Adams SJ, *et al.* Pediatric chest radiograph interpretation: how far has artificial intelligence come? A systematic literature review. *Pediatr Radiol* 2022;52:1568–80. doi:10.1007/s00247-022-05368-w
117. Marshall TL, Rinke ML, Olson APJ, *et al.* Diagnostic Error in Pediatrics: A Narrative Review. *Pediatrics* 2022;149. doi:10.1542/peds.2020-045948D
118. Di Sarno L, Cammisa I, Curatola A, *et al.* A scoping review of the management of acute mastoiditis in children: What is the best approach? *Turk J Pediatr* 2023;65:906–18. doi:10.24953/turkjped.2023.320
119. Causio FA, Beccia F, Hoxhaj I, *et al.* Integrating China in the International Consortium for Personalized Medicine: A Position Paper on Personalized Medicine in Sustainable Healthcare. *Public Health Genomics* 2024;27:1–11. doi:10.1159/000534010
120. Paranjape K, Schinkel M, Nannan Panday R, *et al.* Introducing Artificial Intelligence Training in Medical Education. *JMIR Med Educ* 2019;5:e16048. doi:10.2196/16048
121. Proposal For A Regulation Of The European Parliament And Of The Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts. 2021. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>
122. Causio FA, Hoxhaj I, Beccia F, *et al.* Big data and ICT solutions in the European Union and in China: A comparative analysis of policies in personalized medicine. *Digit Heal* 2022;8:205520762211290. doi:10.1177/20552076221129060
123. Good Machine Learning Practice for Medical Device Development: Guiding Principles. U.S. Food Drug Adm. 2021. <https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles> (accessed 15 Mar 2024).
124. Cascini F, Beccia F, Causio FA, *et al.* Scoping review of the current landscape of AI-based applications in clinical trials. *Front Public Heal* 2022;10. doi:10.3389/fpubh.2022.949377
125. Michelson KN, Klugman CM, Kho AN, *et al.* Ethical Considerations Related to Using Machine Learning-Based Prediction of Mortality in the Pediatric Intensive Care Unit. *J Pediatr* 2022;247:125–8. doi:10.1016/j.jpeds.2021.12.069

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.