An Introduction to Artificial Intelligence in Developmental and Behavioral Pediatrics

Brandon S. Aylward, PhD,* Halim Abbas, MsC,* Sharief Taraman, MD,*†‡§||
Carmela Salomon, PhD,* Diana Gal-Szabo, PhD,* Colleen Kraft, MD, FAAP,*¶**
Louis Ehwerhemuepha, PhD,†§|| Anthony Chang, MD, MBA, MPH, MS,†‡||
Dennis P. Wall, PhD*††

ABSTRACT: Technological breakthroughs, together with the rapid growth of medical information and improved data connectivity, are creating dramatic shifts in the health care landscape, including the field of developmental and behavioral pediatrics. While medical information took an estimated 50 years to double in 1950, by 2020, it was projected to double every 73 days. Artificial intelligence (AI)–powered health technologies, once considered theoretical or research-exclusive concepts, are increasingly being granted regulatory approval and integrated into clinical care. In the United States, the Food and Drug Administration has cleared or approved over 160 health-related AI-based devices to date. These trends are only likely to accelerate as economic investment in AI health care outstrips investment in other sectors. The exponential increase in peer-reviewed AI-focused health care publications year over year highlights the speed of growth in this sector. As health care moves toward an era of intelligent technology powered by rich medical information, pediatricians will increasingly be asked to engage with tools and systems underpinned by AI. However, medical students and practicing clinicians receive insufficient training and lack preparedness for transitioning into a more AI-informed future. This article provides a brief primer on AI in health care. Underlying AI principles and key performance metrics are described, and the clinical potential of AI-driven technology together with potential pitfalls is explored within the developmental and behavioral pediatric health context.

(J Dev Behav Pediatr 44:e126–e134, 2023) **Index terms:** artificial intelligence, machine learning, autism spectrum disorder.

OVERVIEW OF ARTIFICIAL INTELLIGENCE

First coined[1–15] in 1956 by John McCarthy, artificial intelligence (AI) is an interdisciplinary field of computer science that involves the use of computers to develop systems able to perform tasks that are generally associated with intelligence in the intuitive sense. The rise of "big data," alongside the development of increasingly complex algorithms and enhanced computational power and storage capabilities, has contributed to the recent surge in AI-based technologies.[1] AI operates on a continuum, variously assisting, augmenting, or autonomizing task performance. Automating repetitive tasks, for example, may only require an "assisted" form of intelligence.[1] On the other end of the human-machine continuum, however, an "autonomous" form of intelligence is required for machines to independently make decisions in adaptive intelligent systems.[1] Within the health care context, AI is not envisioned as a technology that would supersede the need for skilled human clinicians. Rather, AI-based technologies will likely play pivotal roles in augmenting existing diagnostic and therapeutic toolkits to improve outcomes. While pediatricians may have limited familiarity with AI in a health care context, they are likely already making use of AI-powered technologies in their daily lives. Email spam filters, e-commerce platforms, and entertainment recommendation systems, for example, all rely on AI.

Machine Learning: Underlying Principles

Machine learning (ML) refers specifically to AI methodologies that incorporate an adaptive element wherein systems have the ability to "learn" using data to improve overall accuracy. At a basic level, all ML involves an

input, a function (or some mathematical calculations), and an output. In ML models, the independent variables are termed *inputs* or *features* (e.g., age, gender, medical history, clinical symptom) and the dependent variable is referred to as the *output label* or *target variable* (e.g., diagnostic label, disease level, survival time). Both structured and unstructured data can be used to train ML models (see Fig. 1). While statistical and ML techniques overlap, they are distinguishable by their underlying goals. Statistical learning is usually hypothesis-driven with the goal of inferring relationships between variables. ML is method-driven with the goal of building a model that makes actionable predictions.

Machine Learning Workflow

Clinicians familiar with the development and validation of existing behavioral screening and diagnostic tools will already have a general sense of product development workflow. Specific to the ML workflow, however, is an ability to learn from exposure to data, without the level of prespecified instructions or prior assumptions required in traditional product development. The likelihood of discovering new features and associations is therefore higher because workflow does not require the product designer to determine in advance which variables may be important. When building predictive ML models, the full data set must first be split into parts. Typically, the largest of these parts is the "training set" used for initial model training. A smaller portion of the data, the "validation set," is then used to support hyperparameter tuning and model selection. Ultimately, the data sets help "train" the system to learn from similar patients, clinical features, or outcomes, which helps the algorithm become more accurate over

time as data increase and more tests are conducted. At the end of this iterative process, a "test" set may be used to test the model's generalizability to data that was not previously seen during any prior aspect of the model's development. Prospective clinical validation studies may then be conducted to test the real-world performance of the model on data that were not previously available to model developers. Figure 2 illustrates a typical ML workflow.

Machine Learning Approaches

Data type, structure, and number of features, along with the nature of the clinical questions being explored, all inform the type of ML approach that may be taken. Classical ML, which includes *supervised learning* and *unsupervised learning*, is generally applied to less complex data sets and clinical scenarios with a small number of features.[1] Table 1 provides a brief descriptive summary of these approaches.

Neural Networks and Deep Learning

While the ML techniques described above are suitable for many clinical problems, in cases in which nonlinear and complex relationships need to be mapped, *networks and deep learning* techniques may be more appropriate. An artificial neural network is a complex form of ML model designed to mimic how the neurons in the brain work. This is achieved through multiple layers of aggregation nodes known as neurons because they simulate the function of biological neurons with mathematical functions guiding when each node neuron would fire a signal to another. Features can be multiplied or added together repeatedly. The mathematical formulation of neural networks is such that complex nonlinear

Structured Data	Unstructured Data
Structured data are commonly input into the model as a list (or vector) of numbers	
Patient's Lab Values	
Diagnosis Codes	**Images and Videos** (e.g. radiological and cardiac imaging) Images are input as grids of numbers each representing an intensity at a given pixel location.
BP Readings	
Medical Prescriptions	**Natural Language** (e.g. digital and paper clinical notes) Embeddings and bag of words are examples of unstructured natural language constructs.

Figure 1. Examples of structured and unstructured data that can be used to train machine learning models.
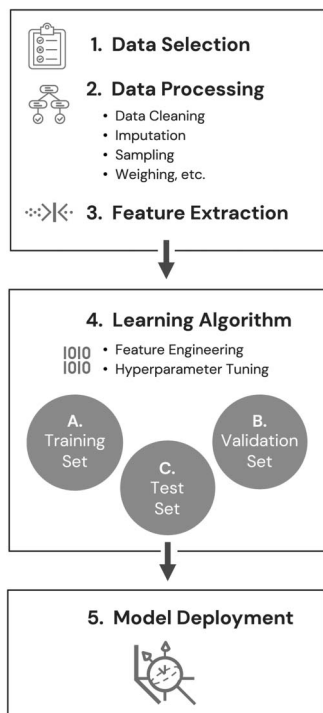
Figure 2. (1) Data are selected based on the type of prediction task to be undertaken. (2) Data are processed to prepare them for use and to address any inadequacies such as missing or biased data or incorrectly formatted data fields. (3) Features expected to contain the information most relevant to the prediction task are selected or formulated and extracted from the full feature set. (4) The data set is split into training and validation sets. Before model deployment, if sufficient data exist, a third "test set" may also be used to test the model's performance with novel data. (5) After iterative improvement, the model is deployed in real-world scenarios.

relationships can be modeled efficiently before an output layer that depends on the prediction task. Given the complexity and volume of health care data, this technique is becoming increasingly popular. Figure 3 illustrates both simple neural network and deep learning neural network.

Evaluating Model Performance

The ability to explain the output of a model and assess its performance, both in the clinical context and in relation to its intended purpose, wi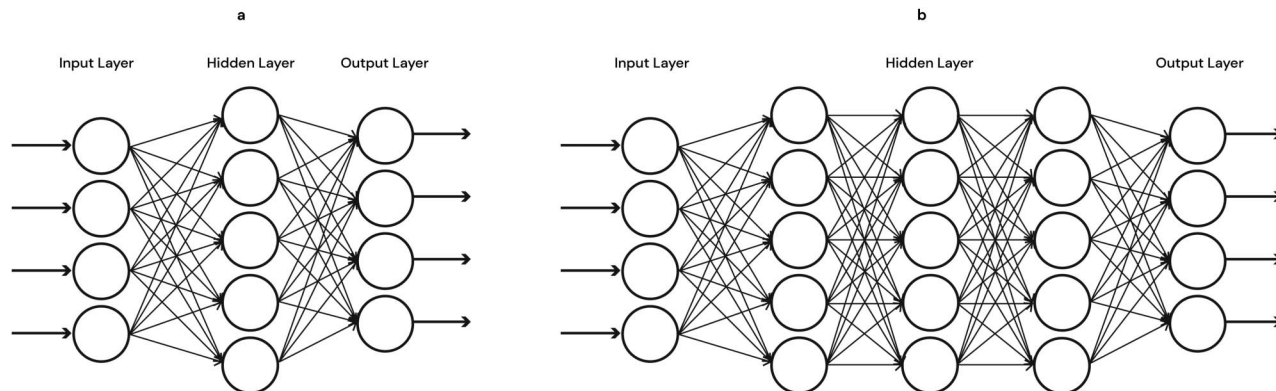ll increasingly become part of the future clinician's role. Key performance metrics relevant to ML are summarized in Table 2. We should note that these metrics do not always provide a straightforward interpretation of the performance of a model[10] because they can be biased by the nature of the data; the end result should be determined by whether clinical value can be obtained from the model. The model's accuracy and threshold, along with the disease prevalence and the model's performance compared with existing "gold-standard" non-AI–informed approaches, should all be considered.[1]

ARTIFICIAL INTELLIGENCE IN DEVELOPMENTAL AND BEHAVIORAL PEDIATRICS: OPPORTUNITIES AND CHALLENGES

Opportunities

In the field of developmental and behavioral pediatrics, artificial intelligence (AI) can assist with a broad array of tasks including diagnosis, risk prediction and stratification, treatment, administration, and regulation.[16] Core analytic tasks that may fall under the AI umbrella are depicted in Figure 4.

Enhancing Clinical Decision-Making, Risk Prediction, and Diagnosis

Massive and constantly expanding quantities of medical data including electronic medical records (EMRs), high-resolution medical images, public health data sets, genomics, and wearables have exceeded the limits of human analysis.[10] AI offers opportunities to harness and derive clinically meaningful insights from this ever-growing volume of health care data in ways that traditional analytic techniques cannot. Neural networks, for example, trained on much larger quantities of data than any single clinician could possibly be exposed to in the course of their career, can support the identification of

Figure 3. Illustration of a simple neural network and a deep learning neural network. Compared with the simple neural network (A) that contains only 1 hidden layer, deep learning networks (B) contain multiple hidden layers. The more layers that are added, the "deeper" the model becomes. Inputs can be multiplied and added together many times. The outputs of this process then become the inputs that are fed into the next hierarchical layer. The process can be repeated many times before a prediction is made.

Table 1. Examples of Machine Learning Approaches

Machine Learning Type	Key Features	Clinical Example
Supervised learning useful when the outcome variable of interest in the training data set is known (e.g., presence or absence of disorder)	**Task-driven classification or regression:** We 'supervise' the program's learning by giving it labeled input (or 'features') and corresponding output (or 'target variables'). The program then 'learns' the relationship between them. Learning depends on a mathematical function (the loss function) that estimates the goodness of fit that is used to determine adjustments required to improve model performance. The goal is to use the training data to minimize errors such as misclassification. If successful, the model can eventually be used to predict an outcome such as the presence or absence of a disease from new data.	Distinguish between benign and malignant lesions
Unsupervised learning useful when the outcome of interest is unknown, unlabeled, or undefined or when we want to explore and identify patterns within the data	**Data-driven clustering:** The model examines a collection of unlabeled examples, then groups them in an 'unsupervised' manner based on some shared commonalities it detects. Unsupervised learning can include clustering and/or dimensionality reduction. In dimensionality reduction, the goal is to reduce the number of variables of the data set while keeping the principal ones that explain the most variation in the data. Ultimately, the output of unsupervised learning can serve as inference and inputs for supervised learning.	A researcher who is interested in how symptoms or features of ADHD in a data set tend to cluster together may use unsupervised learning techniques to identify groups of patients.

ADHD = attention-deficit/hyperactivity disorder.

subtle nonlinear data patterns. Such approaches promise to significantly enhance interpretation of medical scans, pathology slides, and other imaging data that rely on pattern recognition.[10] By observing subtle nonlinear correlations in the data,[17,18] machine learning (ML) approaches have potential to augment risk prediction and diagnostic processes and ultimately provide an enhanced quality of care.

A number of studies within the field of developmental and behavioral pediatrics demonstrate the potential for AI to enhance risk prediction practices. ML techniques were used, for example, to analyze the EMRs of over a million individuals to identify risk for Fragile X based on associations with comorbid medical conditions. The resulting predictive model was able to flag Fragile X cases 5 years earlier than current practice without relying on genetic or familial data.[19] A similar approach was taken to predict autism spectrum disorder (ASD) risk based on high-prevalence comorbidity clusters detected in EMRs.[20] Digital biomarkers inferred from deep comorbidity patterns have also been leveraged to develop an ASD comorbid risk score with a superior predictive performance than some questionnaire-based screening tools.[21] Automated speech analysis was combined with ML in another study to accurately predict risk for psychosis onset in clinically vulnerable youths.[22] In this proof-of-principle study, speech features from transcripts of interviews with at-risk youth were fed into a classification algorithm to assess their predictive value for psychosis.

Research has also highlighted the potential for AI to streamline diagnostic pathways for conditions such as ASD and attention-deficit/hyperactivity disorder (ADHD). Such research is promising given that streamlining diagnosis could allow for earlier treatment initiation during the critical neurodevelopmental window. For example, researchers have used ML to examine behavioral phenotypes of children with ASD with a high rate of accuracy and to shorten the time for observation-based screening and diagnosis.[23-27] In addition, ML has been used to differentiate between ASD and ADHD with high accuracy using a small number of measured behaviors.[28] Research has also explored facial expression analysis based on dynamic deep learning and 3D behavior analysis to detect and distinguish between ADHD, ASD, and comorbid ADHD/ASD presentations.[29]

Expanding Treatment Options

Along with risk prediction and diagnostics, AI holds potential to enhance treatment delivery in the field of developmental and behavioral pediatrics. AI robots have been used in a number of intervention studies to enhance social skill acquisition and spontaneous language development in children with ASD.[30,31] The potential utility of AI-enabled technologies to treat disruptive behaviors and mood and anxiety disorders in children[32] is also being explored. In the field of ADHD, a number of emerging technologies show promise to infer behaviors that can then be used to tailor feedback to enhance self-regulation.[33] For example, in 1 study, a neural network was used to learn behavioral intervention delivery techniques based on human demonstrations. The trained network then enabled a robot to autonomously deliver a similar intervention to children with ADHD.[34]

As health data infrastructure expands in size and sophistication, future AI technologies could potentially support increasingly personalized treatment options.[6] Once comprehensive and integrated biologic, anatomic, physiologic, environmental, socioeconomic, behavioral, and pharmacogenomic patient data become routinely available, AI-based nearest-neighbor analysis could be used to identify "digital twins."[10] Patients with similar genomic and clinical features could be identified and clustered, for example, to allow for highly targeted treatments.[35] Such approaches, while currently largely

Table 2. Common Metrics Used in Machine Learning Model Evaluation

Metric	Clinical Utility
Precision: (correctly predicted positive)/(total predicted positive) = TP/Tp + FP	Used together to help provide an indicator of the effectiveness of an algorithm's performance. These metrics usually work in opposition. In other words, increasing recall will decrease precision and vice versa.
Precision is also referred to as positive predictive value (PPV)	Precision, also known as PPV, tells us the percentage of the cases our model identified as positive that were actually truly positive. For example, if a model testing for presence or absence of a disease had a precision of 91%, this would indicate that 91% of people whom the model determined to have the disease actually had it.
Recall: (correctly predicted positive)/(total correct positive observation) = TP/Tp + FN. In a binary classification, recall is also referred to as "sensitivity"	Recall tells us the proportion of cases that were truly positive that our model managed to correctly identify as such.
Specificity: (correctly predicted negative)/(total correct negative observation) = TN/TN + FP	This figure tells us the proportion of all the negative cases we were able to correctly identify as negative.
Accuracy: The number of correctly classified examples (TP + TN)/the total number of examples that were classified (TP + TN + FP + FN)	The fraction of predictions the model made correctly. Accuracy on its own should not be used to evaluate model performance, especially in cases in which there are significant differences in the number of positive and negative cases in a population.
F1 score: The harmonic mean of the *Precision* and *Recall* scores	This score combines *Precision* and *Recall* to help measure the model's accuracy. In cases in which there is a class imbalance between negative and positive cases, the F1 score can provide a better measure of the incorrectly classified cases than accuracy. F1 scores range from 0 to 1, with higher scores indicating better performance of the model.
The receiver operating characteristic curve (ROC curve): A graph that displays the *sensitivity* of the model on the y-axis and the *false positive ratio* on the x-axis. The graph plots these metrics at different classification thresholds.	Allows us to visually assess how the model performs across its entire operating range (i.e., with thresholds from 0.0 to 1.0). Can help us to understand the tradeoff between sensitivity (true-positive rate) and specificity (1 false positive rate) at different thresholds.
The area under the ROC curve (ROC-AUC or AUROC): A single number that summarizes the efficacy of the algorithm as measured by the ROC curve. Ranges in value from 0 to 1.	AUROC can help us to understand the probability that the model will rank a random positive example higher than a random negative sample. An AUC of 0.0 means the model's predictions are wrong 100% of the time. An AUC of 1.0 indicates the model's predictions are correct 100% of the time. An AUC of 0.5 indicates the model is no better than a completely random classifier. An AUC of 0.9 or higher generally indicates good model classification performance.[a]
Precision recall curve (AUPRC): A graph that displays *precision* on the y-axis and *recall* on the x-axis at various thresholds. Unlike the ROC curve, the PR curve does not use the number of true-negative results.	More robust to imbalanced data than ROC curves. Typically, the higher of 2 curves appearing on a PR plot would likely represent the better performing one. Like for the ROC-AUC, an area under the PC curve can be calculated (AUPRC).

[a]This number will vary depending on context and condition.

theoretical, could help predict therapeutic and adverse medication responses more accurately and also form an evidence base for personalized treatment pathways.[36]

Streamlining Care and Enhancing Workplace Efficiency

Artificial intelligence algorithms have potential to automate many arduous administrative tasks, thereby streamlining care pathways and freeing clinicians to spend more time with patients.[37] Natural language processing solutions, for example, are being developed to decrease reliance on human scribes in clinical encounters.[38] Such approaches may be used to process and transform clinical-free text into meaningful structured outcomes,[39] automate some documentation practices through text summarization,[40] and scan text-based reports to support accurate and rapid diagnostic recommendations.[41] Natural language processing solutions may prove particularly valuable in fields such as child and adolescent psychiatry and developmental and behavioral pediatrics that are extremely text-heavy.[36] Developmental behavioral assessments often involve text-

heavy tasks such as documentation of complex patient histories, results of extensive testing, and collateral history from multiple informants. Given the current national US shortage of child and adolescent psychiatrists with a median of 11 psychiatrists per 100,000 children,[42] AI-based approaches that streamline and automate administrative tasks seem particularly promising. AI-assisted image interpretation has also shown potential to increase workplace productivity and provide considerable cost savings over current practice.[10,43] Techniques to streamline medical research and drug discovery by using natural language processing to rapidly scan biomedical literature and data mine molecular structures are also being developed.[10]

Promoting Equity and Access

Artificial intelligence has potential to address several bias and access disparities apparent in existing care models. While access to developmental and behavioral specialists is extremely limited in much of the world, it is estimated that over 50% of the global population has access to a smartphone.[36] Digital AI–based diagnostic and treatment platforms could thus potentially expand
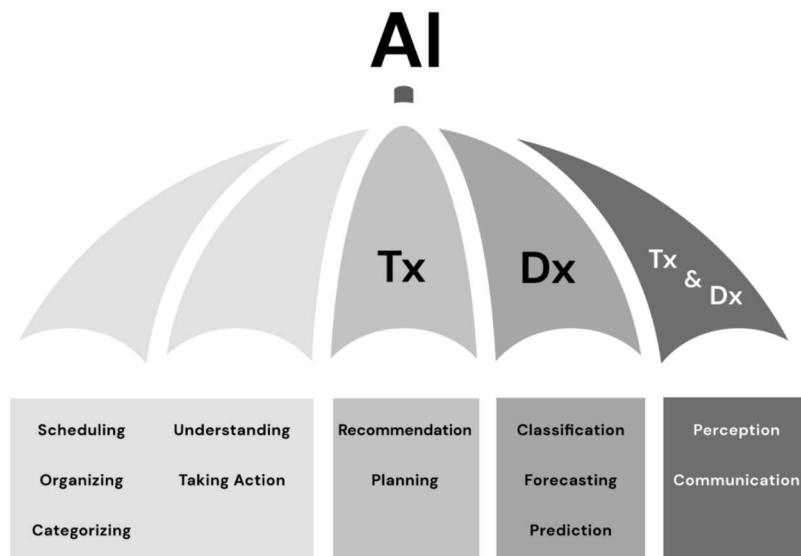
Figure 4. Core analytic tasks that may fall under the AI umbrella. Dx, diagnosis; Tx, therapeutic.

access to underserved and geographically remote populations.[6] Thoughtful use of AI may also help to address racial, socioeconomic, and gender biases. In the field of developmental and behavioral pediatrics, for example, it has been noted that despite ASD prevalence rates being roughly equal across racial/ethnic and socioeconomic groups, human clinicians are more likely to diagnose Black, Latinx, and Asian children, as well as children from low-income families, at a later date than White children and children with a higher socioeconomic status.[44] By integrating and training on large racial-conscious and gender-conscious data sets, AI algorithms can assess thousands of traits and features and build on the findings to assist clinicians in making more accurate, timely, and less biased ASD diagnoses.[44]

Challenges

While AI presents multiple opportunities to the field of developmental and behavioral pediatrics, to date, a very few AI-based technologies have been broadly integrated into clinical practice. Challenges to the widespread deployment of AI in health care settings, along with potential solutions, are outlined below.

Data Bias

Any AI algorithm is only as good as the data from which it was derived. Simply put, the performance of the algorithm and the quality of its prediction are dependent on the quality of the data supplied. If the data are biased, imbalanced, or otherwise an incomplete representation of the target group, the derivative model's generalizability will be limited.[45] A class imbalance problem can occur, for example, in cases in which the total number of 1 class of data (i.e., "girls" or "disease positive status") is far less than the total number of another class of data (i.e., "boys" or "disease negative status"). These biases can sometimes be identified and addressed through techniques such as over- or undersampling, but at other times with "black-box" learning, it is harder to detect and fix the bias in the algorithm(s). Representative populations for most conditions, including in pediatrics, are not a homogeneous group. Thus, it is essential that the data sets used to train these models include balanced data with diverse representations of the clinical symptoms across gender, age, and race in order for the model to be generalizable. Without such safeguards, models may, in fact, perpetuate or amplify stereotypes or biases.[37,45]

Data Sharing and Privacy Concerns

To robustly train an algorithm, sufficient data are required, yet pediatric data sets can be limited by small sample sizes, especially when split by age group. For the case of rare pediatric diseases, extremely low prevalence rates mean the amount and type of data available to train a model on very limited.[46] Creative use of deep learning techniques such as generative adversarial networks may be required in such cases to counteract the lack of data.[1] Generative adversarial networks pit one neural network against another for the purpose of generating synthetic (yet realistic) data to support a variety of tasks such as image and voice generation. While such techniques can be extremely useful when appropriately applied, overreliance on synthetic data also comes with its own set of risks.[1] Multiple data sets may also be combined to produce a sufficient volume of data for model training. However, combining data sets presents its own set of challenges including data privacy and ownership issues[16] and difficulties integrating data with heterogeneous features. Federated learning is an emerging ML technique with potential to address some of these data sharing and privacy concerns. Federated learning allows algorithms to train across many decentralized servers or edge devices, exchanging parameters (i.e., the models' weights and biases) without explicitly exchanging the data samples themselves. This technique obviates many of the privacy issues

engendered by uploading highly sensitive health data from different sources onto a single server.[47]

Algorithmic Transparency and Explainability

A lack of transparency in certain types of ML algorithms such as deep neural networks has raised concern about their clinical trustworthiness.[37] Many models used to analyze images and text, for example, include levels of complexity and multidimensionality that exceed intuitive understanding or interpretation.[48] In cases in which a clinician is unable to understand how the algorithm produces an output, should the algorithm be relied on as part of their clinical decision-making process? Such concerns have led to calls for algorithmic deconvolution before use in health care settings.[10] Other researchers[48] argue, however, that current approaches to explainability disregard the reality that local explanations can be unreliable or too superficial to be meaningful and that rigorous model validation before deployment may be a more important marker of trustworthiness.

Consumer and Clinician Preparedness

If patients and clinicians mistrust AI-based technologies, and/or lack sufficient training to understand, in broad terms, how they function, clinical adoption may be delayed. As with all new tools, implementation matters, and discipline is required to ensure safe deployment of AI-based devices without loss of clinician skill. Over-reliance on AI-based imaging at the expense of history and physical examination, for example, should be avoided. Patient reservations that will require consideration include safety, cost, choice, data bias, and data security concerns.[49] While the American Medical Association has called for research into how AI should be addressed in medical education,[50] current medical training lacks a consistent approach to AI education, and key licensing examinations do not test on this content.[3] Clinicians and medical students alike have identified knowledge gaps and reservations regarding the use of AI in health care.[11-13] A number of preliminary frameworks for integrating AI curriculum into medical training have been proposed[3,5,51]; however, additional research is urgently required to develop and then integrate standardized AI content into medical training pathways.

Ethical Ambiguities

From an ethical standpoint, users of this technology must consider the direct impact and unintended consequences of AI implementation in general as well as specific implications within the clinical context.[52] A number of well-publicized non–health-related cases have illustrated AI data privacy concerns, along with the ethically problematic potential for AI to amplify social, racial, and gender biases.[53] There are also ethical concerns around the magnitude of harm that could occur if an ML algorithm, deployed clinically at scale, were to malfunction; associated impacts could far exceed the harm caused by a single clinician's malpractice.[10] The use of AI for clinical decision-making also raises questions of accountability, such as who is liable if unintended consequences result from use of the technology (e.g., missed diagnosis),[52] or what course of action an autonomous therapy chatbot might take if it detects speech patterns indicative of risk for self-harm.[30] Ethical AI frameworks addressing such concerns are under development,[54,55] and researchers are calling for AI technologies to undergo robust simulation, validation, and prospective scrutiny before clinical adoption.[10]

Regulatory and Payment Barriers

Notwithstanding data, provider adoption, and ethical safeguards, AI technologies face several systematic challenges to be readily implemented into clinical practice, including regulatory, interoperability with EMRs and data exchange, and payment barriers. Given that AI devices can learn from data and alter their algorithms accordingly, traditional medical device regulatory frameworks might not be sufficient.[56] As a result, the Food and Drug Administration has developed a proposed regulatory framework that includes a potential "Predetermined Change Control Plan" for premarket submissions, including "Software Pre-Specifications" and an "Algorithm Change Protocol," to address the iterative nature of AI/ML-based Software as a Medical Device[57] technologies. Health care organizations and practices will also need to establish a data infrastructure and privacy policy for data that are stored across multiple servers and sources (e.g., medical records, health sensors, medical devices, etc). Development of new digital medical software and devices that use AI are likely to outpace the current health care payment structure. New billing codes associated with new treatments and procedures require formal approvals by national organizations with subsequent adoption by insurances, both public and private. This process can take many months to years. To facilitate provider and patient adoption of new AI technologies which may improve quality of care, streamlined development of billing codes for technologies using AI should be developed.

SUMMARY

Artificial intelligence (AI) in health care is not just a futuristic premise, and adoption has shifted from the "early adopter" fringe to a mainstream concept.[15] The convergence of enhanced computational power and cloud storage solutions, increasingly sophisticated machine learning (ML) approaches and rapidly expanding volumes of digitized health care data, has ushered in this new wave of AI-based technologies.[35] Strong economic investment in the AI health care sector, together with the growing number of AI-driven devices being granted regulatory approval, underscores the increasing role of AI in the future health care landscape.[15] In the field of developmental and behavioral pediatrics, we are at an inflection point at which AI-driven technologies show potential to augment clinical decision-making, risk prediction, diagnostics, and treatment delivery. In addition, AI may be leveraged to automate certain time-intensive and arduous clinical tasks and to streamline workflows. Future research is still needed to address

impediments to widespread clinical adoption. These include data bias, privacy, ownership and integration issues, disquietude over a perceived lack of algorithmic transparency, regulatory and payer bottlenecks, ethical ambiguities, and lack of rigorous and standardized AI-focused clinician training. AI technologies are not meant to replace the practicing physician or his/her clinical judgment, nor will they serve as a panacea to all the shortcomings of modern health care. However, we are optimistic about the future of AI in health care, including developmental and behavioral pediatrics. By enhancing the efficiency and impact of health care processes, AI approaches promise to reduce barriers to care and maximize the time clinicians are able to spend with their patients.

REFERENCES

1. Chang AC. *Intelligence-based Medicine: Artificial Intelligence and Human Cognition in Clinical Medicine and Healthcare*. Cambridge, MA: Academic Press; 2020.
2. Katznelson G, Gerke S. The need for health AI ethics in medical school education. *Adv Health Sci Edu*. 2021;26:1447–1458.
3. Paranjape K, Schinkel M, Nannan Panday R, et al. Introducing artificial intelligence training in medical education. *JMIR Med Educ*. 2019;5:e16048.
4. Densen P. Challenges and opportunities facing medical education. *Trans Am Clin Climatol Assoc*. 2011;122:48–58.
5. Kolachalama VB, Garg PS. Machine learning and medical education. *NPJ Digital Med*. 2018;1:954–957.
6. Buch VH, Ahmed I, Maruthappu M. Artificial intelligence in medicine: current trends and future possibilities. *Br J Gen Pract*. 2018;68:143–144.
7. Kokol P, Završnik J, Blažun Vošner H. Artificial intelligence and pediatrics: a synthetic mini review. *Pediatr Dimensions*. 2017; 2:1–5.
8. Meskó B, Görög M. A short guide for medical professionals in the era of artificial intelligence. *NPJ Digit Med*. 2020;3:126.
9. Wartman SA, Combs CD. Medical education must move from the information age to the age of artificial intelligence. *Acad Med*. 2018;93:1107–1109.
10. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25:44–56.
11. Banerjee M, Chiew D, Patel KT, et al. The impact of artificial intelligence on clinical education: perceptions of postgraduate trainee doctors in London (UK) and recommendations for trainers. *BMC Med Educ*. 2021;21:429.
12. Pinto dos Santos D, Giese D, Brodehl S, et al. Medical students' attitude towards artificial intelligence: a multicentre survey. *Eur Radiol*. 2019;29:1640–1646.
13. Sit C, Srinivasan R, Amlani A, et al. Attitudes and perceptions of UK medical students towards artificial intelligence and radiology: a multicentre survey. *Insights Imaging*. 2020;11:14–16.
14. Dumić-Čule I, Orešković T, Brkljačić B, et al. The importance of introducing artificial intelligence to the medical curriculum–assessing practitioners' perspectives. *Croat Med J*. 2020;61:457–464.
15. Greenhill AT, Edmunds BR. A primer of artificial intelligence in medicine. *Tech Innov Gastrointest Endosc*. 2020;22:85–89.
16. He J, Baxter SL, Xu J, et al. The practical implementation of artificial intelligence technologies in medicine. *Nat Med*. 2019;25:30–36.
17. Le S, Hoffman J, Barton C, et al. Pediatric severe sepsis prediction using machine learning. *Front Pediatr*. 2019;7:413.
18. Chahal D, Byrne MF. A primer on artificial intelligence and its application to endoscopy. *Gastrointest Endosc*. 2020;92: 813–820.
19. Movaghar A, Page D, Scholze D, et al. Artificial intelligence–assisted phenotype discovery of fragile X syndrome in a population-based sample. *Genet Med*. 2021;23:1273–1280.
20. Lingren T, Chen P, Bochenek J, et al. Electronic health record based algorithm to identify patients with autism spectrum disorder. *PLoS One*. 2016;11:e0159621.
21. Onishchenko D, Huang Y, van Horne J, et al. Reduced false positives in autism screening via digital biomarkers inferred from deep comorbidity patterns. *Sci Adv*. 2021;7:eabf0354.
22. Bedi G, Carrillo F, Cecchi GA, et al. Automated analysis of free speech predicts psychosis onset in high-risk youths. *NPJ Schizophr*. 2015;1:15030.
23. Abbas H, Garberson F, Liu-Mayo S, et al. Multi-modular AI approach to streamline autism diagnosis in young children. *Sci Rep*. 2020;10: 5014–5018.
24. Kosmicki JA, Sochat V, Duda M, et al. Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning. *Transl Psychiatry*. 2015;5:e514.
25. Levy S, Duda M, Haber N, et al. Sparsifying machine learning models identify stable subsets of predictive features for behavioral detection of autism. *Mol Autism*. 2017;8:65.
26. Rahman MM, Usman OL, Muniyandi RC, et al. A review of machine learning methods of feature selection and classification for autism spectrum disorder. *Brain Sci*. 2020;10:949.
27. Tariq Q, Daniels J, Schwartz JN, et al. Mobile detection of autism through machine learning on home video: a development and prospective validation study. *Plos Med*. 2018;15:e1002705.
28. Duda M, Haber N, Daniels J, et al. Crowdsourced validation of a machine-learning classification system for autism and ADHD. *Transl Psychiatry*. 2017;7:e1133.
29. Jaiswal S, Valstar MF, Gillott A, et al. *Automatic Detection of ADHD and ASD From Expressive Behaviour in RGBD Data*. 2017.12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017); Washington, DC, USA.
30. Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *J Med Internet Res*. 2019;21:e13216.
31. Pennisi P, Tonacci A, Tartarisco G, et al. Autism and social robotics: a systematic review. *Autism Res*. 2016;9:165–183.
32. Rabbitt SM, Kazdin AE, Scassellati B. Integrating socially assistive robotics into mental healthcare interventions: applications and recommendations for expanded use. *Clin Psychol Rev*. 2015;35: 35–46.
33. Cibrian FL, Lakes KD, Schuck SE, et al. The potential for emerging technologies to support self-regulation in children with ADHD: a literature review. *Int J Child-Computer Interaction*. 2022;31: 100421.
34. Clark-Turner M, Begum M. *Deep Recurrent Q-Learning of Behavioral Intervention Delivery by a Robot From Demonstration Data*. 26th IEEE International Symposium on Robot and Human Interactive Communication. RO-MAN); 2017: Lisbon, Portugal. 2017.
35. Shu L-Q, Sun Y-K, Tan L-H, et al. Application of artificial intelligence in pediatrics: past, present and future. *World J Pediatr*. 2019;15: 105–108.
36. Lovejoy CA. Technology and mental health: the role of artificial intelligence. *Eur Psychiatry*. 2019;55:1–3.
37. Greenfield D. *Artificial Intelligence in Medicine: Applications, Implications, and Limitations*. 2019. Available at: https://sitn.hms.harvard.edu/flash/2019/artificial-intelligence-in-medicine-applications-implications-and-limitations/
38. Coiera E, Kocaballi B, Halamka J, et al. The digital scribe. *NPJ Dig Med*. 2018;1:58.
39. Kreimeyer K, Foster M, Pandey A, et al. Natural language processing systems for capturing and standardizing unstructured clinical information: a systematic review. *J Biomed Inform*. 2017;73:14–29.

40. Goldstein A, Shahar Y. An automated knowledge-based textual summarization system for longitudinal, multivariate clinical data. *J Biomed Inform.* 2016;61:159–175.

41. Bressem KK, Adams LC, Gaudin RA, et al. Highly accurate classification of chest radiographic reports using a deep learning natural language model pre-trained on 3.8 million text reports. *Bioinformatics.* 2020;36:5255–5261.

42. American Academy of Child & Adolescent Psychiatry. *AACAP Releases Workforce Maps Illustrating Severe Shortage of Child and Adolescent Psychiatrists.* Washington, DC: American Academy of Child & Adolescent Psychiatry.

43. Beam AL, Kohane IS. Translating artificial intelligence into clinical care. *JAMA.* 2016;316:2368–2369.

44. Aylward BS, Gal-Szabo DE, Taraman S. Racial, ethnic, and sociodemographic disparities in diagnosis of children with autism spectrum disorder. *J Dev Behav Pediatr.* 2021;42:682–689.

45. Zhang H, Lu AX, Abdalla M, et al. *Hurtful words: quantifying biases in clinical contextual word embeddings.* Proceedings of the ACM Conference on Health. 2020:110–120.

46. Long E, Lin H, Liu Z, et al. An artificial intelligence platform for the multihospital collaborative management of congenital cataracts. *Nat Biomed Eng.* 2017;1:0024–0028.

47. Xu J, Glicksberg BS, Su C, et al. Federated learning for healthcare informatics. *J Healthc Inform Res.* 2021;5:1–19.

48. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digital Health.* 2021;3:e745–e750, .

49. Richardson JP, Smith C, Curtis S, et al. Patient apprehensions about the use of artificial intelligence in healthcare. *Npj Digit Med.* 2021; 4:140–146.

50. American Medical Association. *Augmented Intelligence in Health Care Policy Report*; 2018. Available at: https://www.ama-assn.org/system/files/2019-01/augmented-intelligence-policy-report.pdf.

51. Wartman SA, Combs CD. Reimagining medical education in the age of AI. *AMA J Ethics.* 2019;21:146–152.

52. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J.* 2019;6:94–98.

53. Wiens J, Saria S, Sendak M, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med.* 2019;25: 1337–1340.

54. Floridi L, Cowls J, Beltrametti M, et al. An ethical framework for a good AI society: opportunities, risks, principles, and recommendations. *Minds and Machines.* 2018;28:19–39.

55. Miller KW. Moral responsibility for computing artifacts: the rules. *IT Prof.* 2011;13:57–59.

56. Kagiyama N, Shrestha S, Farjo PD, et al. Artificial intelligence: practical primer for clinical research in cardiovascular disease. *J Am Heart Assoc.* 2019;8:e012788.

57. FDA. *Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD): Discussion Paper and Request for Feedback*; 2019. Available at: https://www.regulations.gov/document/FDA-2019-N-1185-0001;