



The Opportunities and Challenges for Artificial Intelligence to Improve Sepsis Outcomes in the Paediatric Intensive Care Unit

Abdullah Tarik Aslan¹ · Budi Permana^{2,3} · Patrick N. A. Harris^{1,4} · Kuban D. Naidoo⁵ · Michael A. Pienaar⁶ · Adam D. Irwin^{1,7}

Accepted: 8 October 2023 / Published online: 31 October 2023
© The Author(s) 2023

Abstract

Purpose of Review To describe existing applications of artificial intelligence (AI) in sepsis management and the opportunities and challenges associated with its implementation in the paediatric intensive care unit.

Recent Findings Over the last decade, significant advances have occurred in the use of AI techniques, particularly in relation to medical image analysis. Increasingly, these techniques are being applied to a broad array of datasets. The availability of both structured and unstructured data from electronic health records, omics data and digital technologies (for example, portable sensors) is rapidly extending the range of applications for AI. These techniques offer the exciting potential to improve the recognition of sepsis and to help us understand the pathophysiological pathways and therapeutic targets of sepsis.

Summary Although AI has great potential to improve sepsis management in children, significant challenges need to be overcome before it can be successfully implemented to change healthcare delivery.

Keywords Artificial intelligence · Sepsis · Children · Intensive care unit · Low- and middle-income countries

Introduction

Sepsis is a life-threatening condition characterised by organ dysfunction in response to a dysregulated immune response to infection [1]. It is a global health priority estimated to cause 2.9 million deaths in children under 5 each year [2]. Although many factors contribute to sepsis-related mortality and morbidity, the failure to make a prompt and accurate diagnosis plays a key role

and is a leading cause of preventable harm. Rapid and accurate identification of sepsis is challenging since it usually presents with subtle or non-specific signs and symptoms, particularly early in the clinical course. Once sepsis is established, the risk of mortality becomes unacceptably high despite appropriate treatment interventions [3]. Early detection of sepsis and identification of patient subgroups that will benefit from specific treatment interventions are critical to improve sepsis outcomes.

✉ Adam D. Irwin
a.irwin@uq.edu.au

Abdullah Tarik Aslan
aslanabdullahtarik@gmail.com

Budi Permana
b.permana@uq.edu.au

Patrick N. A. Harris
p.harris@uq.edu.au

Kuban D. Naidoo
Kuban.Naidoo@wits.ac.za

Michael A. Pienaar
PienaarMA1@ufs.ac.za

¹ Faculty of Medicine, UQ Centre for Clinical Research, The University of Queensland, Building 71/918 Bowen Bridge Rd Herston, Brisbane, QLD, Australia

² School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD, Australia

³ Herston Infectious Diseases Institute (HeIDI), Metro North Health, Brisbane, QLD, Australia

⁴ Central Microbiology, Royal Brisbane and Women's Hospital, Pathology Queensland, Brisbane, QLD, Australia

⁵ Division of Critical Care and Chris Hani Baragwanath Academic Hospital Paediatric Intensive Care Unit, University of the Witwatersrand School of Medicine, Johannesburg, South Africa

⁶ Department of Paediatrics and Child Health, Paediatric Critical Care Unit, University of the Free State, Bloemfontein, South Africa

⁷ Infection Management and Prevention Service, Queensland Children's Hospital, Brisbane, QLD, Australia

In simple terms, the field of artificial intelligence (AI) applies computer science to data to solve problems. AI is differentiated from other statistical and computational approaches by its emphasis on learning and provides significant analytical advantages over conventional rule-based statistical approaches. The most established examples of AI in clinical medicine have been applied to image analysis, resulting in improvements in, for example, the detection of diabetic retinopathy [4], the diagnosis of colorectal cancer [5], and the prediction of lung cancer risk [6]. AI exhibits a spectrum of autonomy, meaning that as algorithms become increasingly complex, human participation and understanding declines. Deep learning approaches are however able to deal with large, multimodal and unstructured datasets and with complex data relationships. This capability lends itself to the study of a heterogeneous multi-system disorder like sepsis. Generative AI and deep learning can derive insights from diverse multi-dimensional datasets including laboratory values, demographic data, healthcare professional notes, vital signs, radiological imaging, microbiological data, histopathological images and omics-related (e.g. genomic, transcriptomic, proteomic and metabolomic) data. In other scientific domains, AI is revolutionising the approach to the development of new therapeutics, including antimicrobials, by predicting the structure of chemical compounds and identifying therapeutic targets [7].

Despite these opportunities, there are substantial challenges for AI systems to overcome if they are to improve healthcare delivery. In this narrative review, we aim to address these opportunities and challenges for AI to improve the outcomes of children with sepsis in the intensive care unit.

Opportunities for Artificial Intelligence to Improve Sepsis Outcomes in Children

Early Recognition

Existing tools to improve the early recognition of sepsis in children rely upon a constellation of physiological variables (such as heart rate), clinical risk factors, examination findings and difficult to quantify contextual information such as parental or clinician ‘concern’ [8]. Previous research efforts have sought to identify diagnostic biomarkers and to construct decision support tools able to support clinicians with the recognition of serious infections including sepsis in children [9, 10]. With few exceptions, these efforts have singularly failed to impact clinical decision making in childhood sepsis. Latterly, researchers have derived and validated gene transcript signatures to differentiate children with and without sepsis [11] or to guide treatment decisions in serious infections [12]. These more sophisticated approaches have the potential to improve diagnostic

discrimination but are themselves likely to be limited as they represent a snapshot of the immune response in an often rapidly evolving clinical presentation. The large number of potentially informative indicators of sepsis and serious infection which vary over time is a data problem which has so far proven intractable.

The capacity of AI to interrogate complex datasets presents an exciting opportunity (Table 1), placing AI in a position to broadly re-formulate the clinical evaluation of the child suspected of sepsis. Several studies have demonstrated that AI based tools derived from electronic health records (EHRs) and omics-related data can be used successfully for early prediction, phenotype characterisation, prognostication and treatment personalization of sepsis in children and neonates [13–18]. Real-time data derived from a range of bedside monitors, including electrocardiogram waveform and baby cry signals, have also been evaluated in the early prediction of sepsis in children and neonates [19, 20]. Inclusion of real-time data in a recurrent neural network model resulted in forecasts which became progressively more accurate over time [21]. The capacity of AI to incorporate a range of dynamic data inputs into accurate and continuous predictions of risk represents a step-change in diagnostic capability. Developments in large language models (LLMs) are revolutionising the integration of AI within and beyond healthcare. These LLMs, including BERT, GPT-4, and PaLM, employ advanced natural language processing (NLP) techniques, architectures and extensive data training, making them more capable in processing and providing refined outputs [22–24]. The integration of these models into chatbots, such as ChatGPT, Bard, and Claude, amplifies their functionalities through natural human-like conversational interactions [25–27]. These technologies hold significant promise. Developing domain-specific models trained with sepsis-related data and case studies is essential to generate meaningful responses. The deployment of these systems through globally available chatbots can assist individuals in low resource settings where access to specialised healthcare resources may be limited.

While contemporary literature has focused on comparing the performance of AI-based tools to that of human decision-making, opportunities to successfully implement AI are likely to be found when healthcare practitioners actively collaborate with AI tools [28, 29]. The usefulness of human-AI collaboration (‘Humans in the loop’) will likely depend on the types of the task and the clinical context. Some studies have indicated that clinicians and AI perform better than either alone [30], whereas others have shown that human-AI collaboration provides no additional benefit over AI alone [31]. This also appears to be influenced by the clinicians’ level of expertise; trainees seem to benefit more from AI input than their more experienced colleagues [32]. The accuracy of AI

Table 1 Opportunities and challenges of artificial intelligence to improve sepsis outcomes in children

Opportunities	Explanation
Availability of multimodal datasets (e.g. EHRs, omics-related data)	While currently available AI systems can evaluate radiological images or images from pathology samples with high accuracy and speed, AI tools derived from different types of databases can also provide robust medical insights
The use of unsupervised learning in AI methods	Since the correct labelling of sepsis cases and data curation is challenging and time-consuming, these methods allow for efficient use of data even if the data is unlabelled or noisy
Prediction of antimicrobial susceptibility results	AI may reduce the time required for analysis of datasets and may support the early prediction of antimicrobial susceptibility results enabling early appropriate therapy in patients with sepsis or septic shock
Human-AI collaboration	Human-AI collaboration is necessary to enable successful clinical implementation of AI systems. It may also improve ‘explainability’ and build trust among clinicians
Large language models and natural language processing	Although there are some existing AI models that successfully use natural language processing, the use of large language models for the development of AI tools in sepsis is still in its infancy
Challenges	Explanation
Dynamic nature of sepsis	Minute-to-minute changes in clinical status of septic patients should be taken into account while constructing an AI tool
Age-related heterogeneity in the pathology of childhood sepsis	Age-related heterogeneity in a range of factors such as host factors, infection aetiology, host response and multiple organ dysfunctions, and key elements of pathogenesis of sepsis is another challenge for creating AI algorithms for children
Accurate labelling of sepsis and lack of consensus in definitions	Sepsis cases should be labelled for supervised learning AI models. However, this is a major challenge in children where sepsis definitions remain controversial
The need for large-scale, high quality and timely granular data	The size and granularity of a dataset determine what analysis can be performed on the data and whether results from that analysis lead to appropriate conclusions
The need for an external validation and model updating in independent datasets	External validation is required to determine an AI algorithm’s reproducibility and generalizability to novel and different patient populations. Alterations in the clinical practices and in patient populations cause data shift or data drift, which in turn, may adversely affect model performance. Regular updating of AI tools allows us to consistently achieve adequate performance
Explainability and AI literacy and training	Contemporary AI techniques can reveal complex relationships within multidimensional data that cannot be understood by human intelligence, creating difficulties in verifying the transparency, actionability and validity of these tools. Explainable tools may be more successfully implemented. Training for clinical end users in AI methods and their application is necessary
Compatibility of AI tools with clinical workflows	AI systems should be built with clinical workflows in mind. Establishing thresholds for alerting, the type of alert and the recipient of AI output recommendations should all be considered in the design
Functionality of the existing EHR systems	To incorporate AI tools into hospitals’ EHR systems, EHRs must have some features to support both AI algorithms and alerts sent by these algorithms to clinicians
Ethical responsibilities of developers of AI systems	AI tool developers will have some ethical responsibilities such as creating secure systems, ensuring data security and responsibly influencing public views on health
Accountability	It is currently unclear whether developers, regulators or healthcare providers should be held accountable if an AI model fails even after it has been clinically validated
There are no standards for assuring the safety of AI systems	Although there are many standards for assuring the safety of digital systems, these standards are not well-suited to AI-driven systems, which are contemplated without explicit programming but instead to automatically learn from big datasets
The need for multidisciplinary teamwork	For creation and implementation of AI systems in real life, we need experts in various fields such as ethics, bioinformatics or AI, computer engineering and clinicians
The need for validation of the AI systems in an RCT	The design and conduct of RCTs of AI interventions should follow the guidance of recent reporting extensions SPIRIT-AI and CONSORT-AI

AI artificial intelligence, *ICUs* intensive care units, *LMICs* low- and middle-income countries, *EHRs* electronic health records, *RCTs* randomised controlled trials

systems significantly affects clinicians’ willingness to collaborate with AI systems, while AI system acceptability increases when its use is supported by clinical leaders and peers [33•]. Human-AI collaboration approach might

also improve ‘explainability’ and build trust among users. Understanding how clinicians and AI systems interact to improve the care of children with sepsis is an important component of future studies.

Unsupervised Learning and Sepsis Phenotypes

Sepsis is a heterogeneous syndrome, defined crudely. The definition of paediatric sepsis is still based on the 2005 International Paediatric Sepsis Consensus Conference [34] but the definition, combining systemic inflammation with suspected infection, is too non-specific to be meaningful to either clinicians or researchers. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) advanced the development of sepsis and septic shock definitions in adults through the incorporation of organ dysfunction criteria [1] and work is underway to adapt these definitions to the specific characteristics and presentations of sepsis in children [35]. This work is necessary to understand the true burden of disease and to guide improvements in clinical care and study design. Even so, as conceived, sepsis definitions still fail to reflect the range of clinical phenotypes which are likely to have different outcomes and which may respond differently to intervention. AI using unsupervised learning can derive new clinical phenotypes in adults with sepsis with clear differences in treatment outcomes. Simulations of randomised controlled trial results in which the proportion of patients with different clinical phenotypes were varied support the hypothesis that new clinical phenotypes derived from large multi-dimensional datasets may result in effective personalised sepsis interventions [36].

Early Antimicrobial Optimisation

Treatment outcomes in sepsis are improved by early, effective antimicrobial use. Antimicrobial susceptibility testing (AST) is required to inform antibiotic use, but traditional culture-based AST takes days to provide meaningful results. Considering the large datasets from mass spectrometry and whole genome sequencing (WGS) analyses, AI can be used to reduce the time to pathogen identification, and to infer antimicrobial susceptibility [37, 38]. In clinical settings with a substantial burden of multidrug-resistant organisms, AI can shorten the time to effective therapy. It may also improve antimicrobial stewardship efforts to ensure that the narrowest spectrum appropriate antibiotics replaces broad spectrum empirical agents as quickly as possible, reducing collateral damage associated with broad spectrum antibiotics [39].

Challenges to Overcome for Artificial Intelligence to Improve Sepsis Outcomes in Children

Data Quality, Governance, Transparency

One of the major challenges in this context is the issue of accurately labelling children and neonates with sepsis.

The use of international classification of diseases (ICD) codes to classify sepsis fails to provide accurate labelling [40, 41]. A lack of consensus in neonatal sepsis definitions is a particular problem, leading to the development of some neonatal AI sepsis tools based on culture-positivity, and others determined by clinical assessment, independent of culture results [42, 43]. Large, well-curated and expert-labelled datasets are very important but assume the accuracy of classification and labelling, though in reality, sepsis definitions remain non-specific and controversial. What is the classifier trying to predict? This is a major hurdle for supervised learning AI systems, fostering interest in weakly supervised and unsupervised systems that require less labelling effort, meanwhile presenting opportunities for new sepsis phenotype derivation [44].

Supervised machine learning models perpetuate errors in training datasets which hamper the construction of clinically useful AI tools. Algorithms are derived on data from majority groups and may have inadequate performance in groups that are poorly represented in the dataset, a problem termed algorithmic bias. Algorithms should be developed and validated in samples representative of the population in which they will be used. Performance should be evaluated in subgroups such as age, sex, ethnicity, economic group and location [45]. Imbalanced datasets with a low proportion of sepsis cases may yield misleading classification and erroneously treat infrequent events (including sepsis) as noise [46].

The development of AI systems for paediatric sepsis is complicated by the relatively lower event rate in children and age-related heterogeneity in the pathobiology of childhood sepsis [47]. For example, risk factors such as prematurity and low-birth weight significantly increase the risk of sepsis-associated mortality in neonates, but not in older children and adolescents. Additionally, sepsis symptoms in neonates significantly overlap with normal preterm physiology creating a problem of false alarms in neonatal units already prone to alarm fatigue [48, 49]. To overcome this particular problem, technologies that record motion, sound and video may be incorporated into emerging AI systems, which will likely provide valuable insights on how babies respond to care and treatments.

Multi-centre studies are necessary to derive reliable and generalizable AI tools but present significant data governance challenges. Studies of AI in sepsis are most commonly single-centre retrospective study designs [35, 50], limiting their generalisability. In a recent meta-analysis, only three prospective studies and one RCT of AI systems for sepsis prediction were identified, and none in children [51]. A notorious pitfall in developing AI models is the risk of overfitting, which stems from application of the trained model solely to the database in which the model itself was developed or trained, and

failing to apply to an external cohort [52, 53]. Multi-centre studies require access to substantial data and computer science infrastructure and capability; however, the pre-processing of multiple datasets remains a challenging and time-consuming task. Restrictions on data sharing limit the availability of data from multiple centres and prevent the development of widely generalisable systems. Federated learning is a system in which an AI algorithm can be trained through multiple independent sessions, each using its own dataset, and may be considered as a potential solution to the challenge of data sharing between different hospitals or jurisdictions. In this system, it is possible to create a common AI model without data sharing, thereby tackling critical challenges such as data security, data privacy and access to heterogeneous data [54]. Addressing data privacy and security is vital to ensure the safe implementation of these technologies. Models that can operate locally at the device level with robust encryption and access controls will further enhance data security and mitigate privacy concerns [55].

Intriguingly, alterations in the characteristics of day-to-day practices, healthcare systems and patient populations over time can cause data shift or data drift, which in turn, can adversely affect model performance [56]. Therefore, continuous updating of AI tools, in addition to external validation, allows us to consistently achieve adequate performance when used in daily practice. The US Federal Drug Administration (FDA) has proposed a blueprint for adaptive AI systems, in which they will approve not only an initial model but also a process to update models over time [57].

The reporting of AI system performance is an additional challenge. The diagnostic discrimination of prediction models is typically reported using the area under the receiver operating characteristic curve (AUROC or AUC). AUC provides a global measure of discrimination and is equivalent to the concordance or *c*-statistic, the probability that a randomly selected subject with the outcome of interest has a higher predicted probability of having than a randomly selected subject without the outcome. AUC can be misleading in quantifying model performance because of limitations in taking prior probability into account, giving knowledge on the spatial distribution of model errors, and equal weighting of omission and commission errors [58, 59]. As a result, we may see unreliable risk estimates even in models with good discrimination accuracy [60]. Put simply, translating this global measure of discrimination into meaningful metrics to assure clinicians is challenging. Evaluations should quantify other measures of accuracy including model sensitivity and specificity at clinically meaningful thresholds and likelihood ratios which quantify post-test from pre-test odds in the presence of a positive or negative test.

Reporting Guidelines and Implementation

Most AI studies assess the performance of algorithms on retrospectively collected static datasets, whereas physicians seek to implement these models in dynamic, real-world hospital settings using real-time data [61]. This contrast limits the applicability of these algorithms in daily practice. For AI systems to be effectively implemented, the processes of algorithm development and validation need to be transparently reported. Reporting guidelines that reflect the specific requirements of AI system evaluations are now established, extending from the reporting of diagnostic or prognostic accuracy (Transparent Reporting of a multivariable prediction model of Individual Prognosis Or Diagnosis-AI [62]), through early clinical validation (Developmental and Exploratory Clinical Investigations of DEcision support systems driven by Artificial Intelligence (DECIDE-AI [63•]) to clinical trials (CONSORT-AI [64]). These guidelines are essential to improve transparency and ensure the replicability of study findings, thereby supporting successful implementation.

Challenges Specific to LLMs

In order to establish reliability and effectiveness of the LLMs in clinical practice, several challenges need to be addressed. Firstly, existing models are predominantly trained using general datasets and prone to make mistakes, termed ‘hallucinations’, whereby a model invents information. These limitations can significantly reduce the accuracy of a model’s response, particularly in specific domains like sepsis. Recent progress, such as integrating constitutional AI to produce harmless AI systems, improving training techniques, and utilising medical resources for training data, has shown promise [25, 55, 65–67]. Nevertheless, developing domain-specific models trained with sepsis-related data and case studies is essential to generate reliable model outputs.

Challenges of Implementing Artificial Intelligence Tools in Daily Practice

Systems

Significant challenges remain if we are to harness AI in healthcare systems. The translation of AI innovations into clinical care is limited by healthcare systems in which data are rarely available for real-time analysis. Data may be siloed in a range of systems which separately document pathology results, physiological variables and clinical notes. Unified data formats such as Fast Healthcare Interoperability

Resources may support the aggregation of data but adequate investment in local analytics infrastructure is inevitable to allow data collection, curation, transformation and analytics [68]. Learning healthcare systems are an initiative that aim to incorporate research and innovation to enable rapid improvements in the quality of care driven by continuous analysis of data derived from routine care and may have a role in successfully implementing AI systems [69].

Demonstrating the Clinical Validity of AI Systems

Following the development and validation of novel AI algorithms, deployment may be undertaken first in ‘silent mode’, that is, running the algorithm but not incorporating its results into routine practice. In the process of deployment of an AI system in daily practice, the necessity of a multi-faceted approach including all relevant stakeholders should be kept in mind [70]. Large-scale, well-conducted RCTs are arguably the most important step toward illuminating the impact of AI tools on clinically meaningful outcomes. In order to interpret the results of RCTs and rigorously investigate the usefulness of the AI systems, methods for conducting and reporting of RCTs should be standardised [71, 72]. There are presently no published RCTs of an AI-driven approach to sepsis in children. Once an AI system has been integrated into the hospital system, it is vitally important to give recurring training to healthcare professionals to ensure its proper implementation. Since the usefulness of AI tools can be greatly influenced by the way people provide input and evaluate output, regulatory authorities should mandate that human factors be tested and adequate training provided for end users of AI systems [73].

Explainability and Collaboration Between Clinicians and AI

Contemporary AI techniques can reveal complex relationships within multidimensional data that are challenging to understand and explain to clinicians. Clinical implementation is likely to succeed best where models derived from AI are not only accurate, but transparent, interpretable and actionable [74]. The consequences associated with incorrect model predictions, in this context, for example, the recognised consequences of a missed diagnosis of sepsis, argue strongly in favour of explainability. An explainable system allows for the recognition of errors and the identification of bias or confounding. The best performing models, such as deep neural networks, are the least explainable, however leading to a trade-off between performance and explainability. Work to enhance the explainability of complex AI systems by reporting the relative importance of constituents of model is developing rapidly. Measures such as the shapley additive explanation, for example, have

been developed to estimate and visualise the contribution of individual model features in complex and highly accurate ensemble decision trees [75].

This issue can be partially resolved by involving clinical end users (and consumers) in every step of the development process of these tools. Clinicians need to understand how the application of AI tools will impact clinical workflows and improve patient care, something rarely examined and reported. In the pre-deployment phase of an AI system, it should be determined whether the management of the system outputs will be undertaken by a designated team or by attending physicians. Although the first option seems to reduce alarm fatigue in attending physicians, some physicians may find this method disruptive. System alerts are required to balance sensitivity with alert fatigue taking into consideration the clinical importance of sepsis, its prevalence and acuity of the patient population [76]. The alerts generated may be accompanied by additional recommendations or information and sent to healthcare providers as messages via emails, phones or personal devices or through existing hospital EHR systems. This affirms the importance of EHR systems with the functionality to support complex AI systems [77]. Additionally, surveillance of an AI tool’s diagnostic performance should be ongoing after implementation, permitting making alterations of alerting thresholds if necessary. As another solution for alert fatigue, alert types can be classified as hard alerts that should elicit an immediate response, or soft alerts that can be managed more flexibly [33•, 78••, 79].

There is limited evidence to guide the process of AI implementation in sepsis or other clinical domains. Frameworks such as the recently published SALIENT framework offer insight into the barriers and enablers and a roadmap to successful AI implementations [61].

Ethical and Regulatory Challenges

Accountability for the clinical impact of AI systems is presently unclear. If a clinically validated AI system fails or produces predictions which result in harm, will healthcare providers, developers, or regulators, be held accountable [80]? AI developers may begin to exert substantial influence on healthcare service provision. These developers have a responsibility to create safe systems, ensure data security, and to responsibly influence public views on health [81]. The implementation and maintenance of AI tools are affected by the different ownership of input data and AI system. Significant investments must be made in data transfer and processing to ensure optimal data security, to bring stakeholders together, develop trust-based relationships and define roles and responsibilities. There are many standards for assuring the safety of digital systems [82, 83]. However, these standards are not well-suited to AI-driven systems,

which evolve through learning and without explicit intervention [84]. Although, the FDA took some steps to address these issues, many healthcare systems lack a regulatory framework to assure the safe application of AI systems in healthcare [67]. A recent report from the UK Care Quality Commission concluded that further studies are needed to provide clarity on how healthcare facilities should implement AI systems into clinical workflows [85].

An aspiration for the use of AI should be to make medicine more human, not less, prioritising more time between clinicians and patients and limiting excessive automation. The potential transformation of healthcare delivery through AI should prompt serious reflection collectively from clinicians, healthcare providers and AI developers regarding the ethics of healthcare provision.

Barriers to Implementing Artificial Intelligence for Sepsis in Low- and Middle-Income Countries

The global burden of sepsis-associated deaths in children and neonates falls disproportionately upon low- and middle-income countries (LMICs) [2]. The vast inequity in resources between high- and low-income settings means that despite the enormous public health impact, communities most affected by sepsis are least likely to benefit from technological progress. An ethical approach to AI implementation would ensure that its benefits are shared. Not only is sepsis more prevalent in children in LMICs, it is also more difficult to treat, being more often associated with antimicrobial resistant infections [86]. There is an urgent need for better diagnostics both to improve early sepsis recognition and to guide more judicious use of antibiotics. Many novel diagnostic platforms have limited suitability in LMICs because of cost and the requirements for technical expertise and laboratory facilities [87, 88]. Paradoxically, AI tools can be a solution to this unmet need. While AI systems derived from extensive EHR datasets are unlikely to be widespread, alternative approaches to data infrastructure may allow high performance computing situated in organisations such as universities to interact with near ubiquitous smart or mobile devices, even in remote locations. As discussed earlier, AI systems implemented in populations of children and neonates in LMICs should be developed in representative groups. AI developers and healthcare providers should resist the temptation to import solutions derived in entirely different patient groups.

The challenges associated with this digital transformation should not be underestimated. Much of the burden of sepsis mortality is attributable to a lack of access to healthcare services. Civil registration systems in LMICs

may be limited which presents difficulties in identifying the burden of sepsis in children [89, 90]. Even where healthcare is accessible, many healthcare facilities in LMICs do not have EHR systems to support a sophisticated use of clinical data. Accessing and harmonising data from highly heterogeneous clinical systems is a significant technical and governance challenge. Significant investment must be made to develop the infrastructures necessary for high-quality and large-scale data storage in LMICs [91]. A widespread use of big data to create AI tools is made even more difficult by the different regulatory approaches within and between LMICs [92]. Factors such as linguistic and cultural differences contribute to the challenge of data sharing between LMICs [93].

Sepsis in children in LMICs is associated with unacceptably high mortality and long-term disability. Data derived from EHRs or mobile and smart devices may enable AI tools to relieve the burden on healthcare practitioners and inform evidence-based personalised management of children and neonates with sepsis. While transforming the data ecosystem in LMICs AI-based systems is appealing, the priorities of existing health systems must be considered seriously.

Conclusion

Artificial intelligence offers a rapidly evolving approach to the analysis of routine clinical and biological data with the potential to transform the care of children with suspected sepsis. Rapid developments in the application of LLMs offer exciting and unpredictable opportunities for innovation. However, there are many challenges impeding the development and implementation of safe, reliable and acceptable AI systems in sepsis care. The unique challenges of LMICs should be urgently addressed to reflect the disproportionate impact of sepsis in these settings.

Author Contributions ADI and KN conceived the idea. ATA and BP performed the literature review. ATA and ADI drafted the manuscript text. All authors reviewed and contributed to manuscript editing. All authors approved the final submitted version of the manuscript.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

Compliance with Ethical Standards

Conflict of Interest Dr Harris has received grants from Gilead, and reports personal fees from OpGen, Biomerieux, Pfizer, Sandoz, and MSD, outside the submitted work. Dr Naidoo reports personal fees from Nutricia, personal fees from Fresenius Kabi, outside the submitted work. Other authors have nothing to disclose.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Papers of particular interest, published recently, have been highlighted as:

- Of importance
- Of major importance

1. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA*. 2016;315(8):801–10. <https://doi.org/10.1001/jama.2016.0287>.
2. Rudd KE, Johnson SC, Agesa KM, Shackelford KA, Tsoi D, Kievlan DR, et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. *Lancet*. 2020;395(10219):200–11. [https://doi.org/10.1016/S0140-6736\(19\)32989-7](https://doi.org/10.1016/S0140-6736(19)32989-7).
3. Angus DC, Bindman AB. Achieving diagnostic excellence for sepsis. *JAMA*. 2022;327(2):117–8. <https://doi.org/10.1001/jama.2021.23916>.
4. Wolf RM, Channa R, Abramoff MD, Lehmann HP. Cost-effectiveness of autonomous point-of-care diabetic retinopathy screening for pediatric patients with diabetes. *JAMA Ophthalmol*. 2020;138(10):1063–9. <https://doi.org/10.1001/jamaophthalmol.2020.3190>.
5. Zhou D, Tian F, Tian X, Sun L, Huang X, Zhao F, et al. Diagnostic evaluation of a deep learning model for optical diagnosis of colorectal cancer. *Nat Commun*. 2020;11(1):2961. <https://doi.org/10.1038/s41467-020-16777-6>.
6. Huang P, Lin CT, Li Y, Tammemagi MC, Brock MV, Atkar-Khattra S, et al. Prediction of lung cancer risk at follow-up screening with low-dose CT: a training and validation study of a deep learning method. *Lancet Digit Health*. 2019;1(7):e353–62. [https://doi.org/10.1016/S2589-7500\(19\)30159-1](https://doi.org/10.1016/S2589-7500(19)30159-1).
7. Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, et al. A deep learning approach to antibiotic discovery. *Cell*. 2020;181(2):475–83. <https://doi.org/10.1016/j.cell.2020.04.001>.
8. Van den Bruel A, Haj-Hassan T, Thompson M, Buntinx F, Mant D. European Research Network on Recognising Serious Infection i. Diagnostic value of clinical features at presentation to identify serious infection in children in developed countries: a systematic review. *Lancet*. 2010;375(9717):834–45. [https://doi.org/10.1016/S0140-6736\(09\)62000-6](https://doi.org/10.1016/S0140-6736(09)62000-6).
9. Van den Bruel A, Thompson MJ, Haj-Hassan T, Stevens R, Moll H, Lakhanpaul M, et al. Diagnostic value of laboratory tests in identifying serious infections in febrile children: systematic review. *BMJ*. 2011;342: d3082. <https://doi.org/10.1136/bmj.d3082>.
10. Nijman RG, Vergouwe Y, Thompson M, van Veen M, van Meurs AH, van der Lei J, et al. Clinical prediction model to aid emergency doctors managing febrile children at risk of serious bacterial infections: diagnostic study. *BMJ*. 2013;346: f1706. <https://doi.org/10.1136/bmj.f1706>.
11. Wong HR, Cvijanovich N, Allen GL, Lin R, Anas N, Meyer K, et al. Genomic expression profiling across the pediatric systemic inflammatory response syndrome, sepsis, and septic shock spectrum. *Crit Care Med*. 2009;37(5):1558–66. <https://doi.org/10.1097/CCM.0b013e31819fcc08>.
12. Herberg JA, Kaforou M, Wright VJ, Shailes H, Eleftherohorinou H, Hoggart CJ, et al. Diagnostic test accuracy of a 2-transcript host RNA signature for discriminating bacterial vs viral infection in febrile children. *JAMA*. 2016;316(8):835–45. <https://doi.org/10.1001/jama.2016.11236>.
13. Abbas M, El-Manzalawy Y. Machine learning based refined differential gene expression analysis of pediatric sepsis. *BMC Med Genomics*. 2020;13(1):122. <https://doi.org/10.1186/s12920-020-00771-4>.
14. Banerjee S, Mohammed A, Wong HR, Palaniyar N, Kamaleswaran R. Machine learning identifies complicated sepsis course and subsequent mortality based on 20 genes in peripheral blood immune cells at 24 H post-ICU admission. *Front Immunol*. 2021;12: 592303. <https://doi.org/10.3389/fimmu.2021.592303>.
15. But S, Celar B, Fister P. Tackling neonatal sepsis-can it be predicted? *Int J Environ Res Public Health*. 2023;20(4):3644. <https://doi.org/10.3390/ijerph20043644>.
16. Lamping F, Jack T, Rubsamen N, Sasse M, Beerbaum P, Mikolajczyk RT, et al. Development and validation of a diagnostic model for early differentiation of sepsis and non-infectious SIRS in critically ill children - a data-driven approach using machine-learning algorithms. *BMC Pediatr*. 2018;18(1):112. <https://doi.org/10.1186/s12887-018-1082-2>.
17. Stocker M, Daunhawer I, van Herk W, El Helou S, Dutta S, Schuerman F, et al. Machine learning used to compare the diagnostic accuracy of risk factors, clinical signs and biomarkers and to develop a new prediction model for neonatal early-onset sepsis. *Pediatr Infect Dis J*. 2022;41(3):248–54. <https://doi.org/10.1097/INF.0000000000003344>.
18. Wong HR, Cvijanovich N, Lin R, Allen GL, Thomas NJ, Willson DF, et al. Identification of pediatric septic shock subclasses based on genome-wide expression profiling. *BMC Med*. 2009;7:34. <https://doi.org/10.1186/1741-7015-7-34>.
19. Khalilzad Z, Hasasneh A, Tadj C. Newborn cry-based diagnostic system to distinguish between sepsis and respiratory distress syndrome using combined acoustic features. *Diagnosics (Basel)*. 2022;12(11):2802. <https://doi.org/10.3390/diagnostics12112802>.
20. Tarricone F, Brunetti A, Buongiorno D, Altini N, Bevilacqua V, Del Vecchio A, et al. Intelligent neonatal sepsis early diagnosis system for very low birth weight infants. *Appl Sci*. 2021;11(1):404. <https://doi.org/10.3390/app11010404>.
21. Aczon MD, Ledbetter DR, Laksana E, Ho LV, Wetzel RC. Continuous prediction of mortality in the PICU: a recurrent neural network model in a single-center dataset. *Pediatric critical care medicine : a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies*. 2021;22(6):519–29. <https://doi.org/10.1097/PCC.0000000000002682>.
22. Chowdhery A, Narang S, Devlin J, Bosma M, Mishra G, Roberts A, et al. Palm: Scaling language modeling with pathways. *arXiv preprint arXiv:220402311*. 2022.

23. Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, et al. Language models are few-shot learners. *Adv Neural Inf Process Syst*. 2020;33:1877–901.
24. Devlin J, Chang M-W, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*. 2018.
25. OpenAI. Introducing ChatGPT. 2022. <https://openai.com/blog/chatgpt>. Accessed 12 July 2023.
26. Manyika J. An overview of Bard: an early experiment with generative AI. 2023. <https://ai.google/static/documents/google-about-bard.pdf>. Accessed 12 July 2023.
27. Anthropic. Introducing Claude. 2023. <https://www.anthropic.com/index/introducing-claude>. Accessed 12 July 2023.
28. Chen PC, Gadepalli K, MacDonald R, Liu Y, Kadowaki S, Nagpal K, et al. An augmented reality microscope with real-time artificial intelligence integration for cancer diagnosis. *Nat Med*. 2019;25(9):1453–7. <https://doi.org/10.1038/s41591-019-0539-7>.
29. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit Health*. 2019;1(6):e271–97. [https://doi.org/10.1016/S2589-7500\(19\)30123-2](https://doi.org/10.1016/S2589-7500(19)30123-2).
30. Park A, Chute C, Rajpurkar P, Lou J, Ball RL, Shpanskaya K, et al. Deep learning-assisted diagnosis of cerebral aneurysms using the HeadXNet Model. *JAMA Netw Open*. 2019;2(6):e195600. <https://doi.org/10.1001/jamanetworkopen.2019.5600>.
31. Kim HE, Kim HH, Han BK, Kim KH, Han K, Nam H, et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multi-reader study. *Lancet Digit Health*. 2020;2(3):e138–48. [https://doi.org/10.1016/S2589-7500\(20\)30003-0](https://doi.org/10.1016/S2589-7500(20)30003-0).
32. Tschandl P, Rinner C, Apalla Z, Argenziano G, Codella N, Halpern A, et al. Human-computer collaboration for skin cancer recognition. *Nat Med*. 2020;26(8):1229–34. <https://doi.org/10.1038/s41591-020-0942-0>.
33. Henry KE, Adams R, Parent C, Soleimani H, Sridharan A, Johnson L, et al. Factors driving provider adoption of the TREWS machine learning-based early warning system and its effects on sepsis treatment timing. *Nat Med*. 2022;28(7):1447–54. <https://doi.org/10.1038/s41591-022-01895-z>. **Accompanied a large multi-centre evaluation of a machine-learning early warning system and reported the factors which supported its adoption.**
34. Goldstein B, Giroir B, Randolph A. International Consensus Conference on Pediatric Sepsis. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatric critical care medicine : A Journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies*. 2005;6(1):2–8. <https://doi.org/10.1097/01.PCC.0000149131.72248.E6>.
35. Menon K, Schlapbach LJ, Akech S, Argent A, Biban P, Carrol ED, et al. Criteria for pediatric sepsis—a systematic review and meta-analysis by the pediatric sepsis definition taskforce. *Crit Care Med*. 2022;50(1):21–36. <https://doi.org/10.1097/CCM.0000000000005294>.
36. Seymour CW, Kennedy JN, Wang S, Chang CH, Elliott CF, Xu Z, et al. Derivation, validation, and potential treatment implications of novel clinical phenotypes for sepsis. *JAMA*. 2019;321(20):2003–17. <https://doi.org/10.1001/jama.2019.5791>.
37. Ferreira I, Beisken S, Lueftinger L, Weinmaier T, Klein M, Bacher J, et al. Species identification and antibiotic resistance prediction by analysis of whole-genome sequence data by use of ARESdb: an Analysis of isolates from the unyvero lower respiratory tract infection trial. *J Clin Microbiol*. 2020;58(7):e00273–e320. <https://doi.org/10.1128/JCM.00273-20>. **Illustrated the potential to use apply machine learning to whole genome sequencing data to infer antibiotic susceptibility.**
38. Wijesinghe HGS, Hare DJ, Mohamed A, Shah AK, Harris PNA, Hill MM. Detecting antimicrobial resistance in *Escherichia coli* using benchtop attenuated total reflectance-Fourier transform infrared spectroscopy and machine learning. *Analyst*. 2021;146(20):6211–9. <https://doi.org/10.1039/d1an00546d>.
39. Kanjilal S, Oberst M, Boominathan S, Zhou H, Hooper DC, Sontag D. A decision algorithm to promote outpatient antimicrobial stewardship for uncomplicated urinary tract infection. *Sci Transl Med*. 2020;12(568):eaay5067. <https://doi.org/10.1126/scitranslmed.aay5067>.
40. Iwashyna TJ, Odden A, Rohde J, Bonham C, Kuhn L, Malani P, et al. Identifying patients with severe sepsis using administrative claims: patient-level validation of the angus implementation of the international consensus conference definition of severe sepsis. *Med Care*. 2014;52(6):e39–43. <https://doi.org/10.1097/MLR.0b013e318268ac86>.
41. Ramanathan R, Leavell P, Stockslager G, Mays C, Harvey D, Duane TM. Validity of International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) screening for sepsis in surgical mortalities. *Surg Infect (Larchmt)*. 2014;15(5):513–6. <https://doi.org/10.1089/sur.2013.089>.
42. Hayes R, Hartnett J, Semova G, Murray C, Murphy K, Carroll L, et al. Neonatal sepsis definitions from randomised clinical trials. *Pediatr Res*. 2023;93(5):1141–8. <https://doi.org/10.1038/s41390-021-01749-3>.
43. Henry CJ, Semova G, Barnes E, Cotter I, Devers T, Rafae A, et al. Neonatal sepsis: a systematic review of core outcomes from randomised clinical trials. *Pediatr Res*. 2022;91(4):735–42. <https://doi.org/10.1038/s41390-021-01883-y>.
44. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med*. 2022;28(1):31–8. <https://doi.org/10.1038/s41591-021-01614-0>.
45. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med*. 2019;17(1):195. <https://doi.org/10.1186/s12916-019-1426-2>.
46. Haixiang G, Yijing L, Shang J, Mingyun G, Yuanyue H, Bing G. Learning from class-imbalanced data: review of methods and applications. *Expert Syst Appl*. 2017;73:220–39.
47. Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, et al. Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock, 2012. *Intensive Care Med*. 2013;39(2):165–228. <https://doi.org/10.1007/s00134-012-2769-8>.
48. Joshi R, van Pul C, Atallah L, Feijs L, Van Huffel S, Andriessen P. Pattern discovery in critical alarms originating from neonates under intensive care. *Physiol Meas*. 2016;37(4):564–79. <https://doi.org/10.1088/0967-3334/37/4/564>.
49. Winters BD, Cvach MM, Bonafide CP, Hu X, Konkani A, O'Connor MF, et al. Technological Distractions (Part 2): a summary of approaches to manage clinical alarms with intent to reduce alarm fatigue. *Crit Care Med*. 2018;46(1):130–7. <https://doi.org/10.1097/CCM.0000000000002803>.
50. Weiss SL, Peters MJ, Alhazzani W, Agus MS, Flori HR, Inwald DP, et al. Surviving sepsis campaign international guidelines for the management of septic shock and sepsis-associated organ dysfunction in children. *Intensive Care Med*. 2020;46:10–67.
51. Fleuren LM, Klausch TLT, Zwager CL, Schoonmade LJ, Guo T, Roggeveen LF, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive Care Med*. 2020;46(3):383–400. <https://doi.org/10.1007/s00134-019-05872-y>.

52. Habib AR, Lin AL, Grant RW. The epic sepsis model falls short—the importance of external validation. *JAMA Intern Med.* 2021;181(8):1040–1. <https://doi.org/10.1001/jamainternmed.2021.3333>.
53. Kreitmann L, Bodinier M, Fleurie A, Imhoff K, Cazalis MA, Peronnet E, et al. Mortality prediction in sepsis with an immune-related transcriptomics signature: a multi-cohort analysis. *Front Med (Lausanne).* 2022;9: 930043. <https://doi.org/10.3389/fmed.2022.930043>.
54. Narmadha K, Varalakshmi P. Federated learning in healthcare: a privacy preserving approach. *Studies in Health Technology and Informatics: IOS Press.* 2022.
55. Ghahramani Z. Introducing PaLM 2. 2023. <https://blog.google/technology/ai/google-palm-2-ai-large-language-model/>. Accessed 12 July 2023.
56. Subbaswamy A AR, Saria S. Evaluating model robustness and stability to dataset shift. arXiv. 2021. abs/2010.15100. <https://doi.org/10.48550/arxiv.2010.15100>.
57. Food, Administration D. Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD). 2019.
58. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* 1996;15(4):361–87. [https://doi.org/10.1002/\(SICI\)1097-0258\(19960229\)15:4<361::AID-SIM168%3e3.0.CO;2-4](https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168%3e3.0.CO;2-4).
59. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med.* 1999;130(6):515–24. <https://doi.org/10.7326/0003-4819-130-6-199903160-00016>.
60. Diamond GA. What price perfection? Calibration and discrimination of clinical prediction models. *J Clin Epidemiol.* 1992;45(1):85–9. [https://doi.org/10.1016/0895-4356\(92\)90192-p](https://doi.org/10.1016/0895-4356(92)90192-p).
61. van der Vegt AH, Scott IA, Dermawan K, Schnettler RJ, Kalke VR, Lane PJ. Deployment of machine learning algorithms to predict sepsis: systematic review and application of the SALIENT clinical AI implementation framework. *J Am Med Inform Assoc.* 2023;30(7):1349–61. <https://doi.org/10.1093/jamia/ocad075>.
62. Collins GS, Dhiman P, Andaur Navarro CL, Ma J, Hooft L, Reitsma JB, et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. *BMJ Open.* 2021;11(7): e048008. <https://doi.org/10.1136/bmjopen-2020-048008>.
63. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat Med.* 2022;28(5):924–33. <https://doi.org/10.1038/s41591-022-01772-9>. **A framework to improve the transparency of reporting of AI-based decision support systems, essential to ensure safety and reproducibility of published systems.**
64. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, Spirit AI, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med.* 2020;26(9):1364–74. <https://doi.org/10.1038/s41591-020-1034-x>.
65. Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, et al. Publisher correction: large language models encode clinical knowledge. *Nature.* 2023. <https://doi.org/10.1038/s41586-023-06455-0>.
66. Lightman H, Kosaraju V, Burda Y, Edwards H, Baker B, Lee T, et al. Let’s verify step by step. arXiv preprint arXiv:230520050. 2023.
67. Bai Y, Kadavath S, Kundu S, Askell A, Kernion J, Jones A, et al. Constitutional ai: Harmlessness from ai feedback. arXiv preprint arXiv:221208073. 2022.
68. Macias CG, Remy KE, Barda AJ. Utilizing big data from electronic health records in pediatric clinical care. *Pediatr Res.* 2023;93(2):382–9. <https://doi.org/10.1038/s41390-022-02343-x>.
69. Menear M, Blanchette MA, Demers-Payette O, Roy D. A framework for value-creating learning health systems. *Health Res Policy Syst.* 2019;17(1):79. <https://doi.org/10.1186/s12961-019-0477-3>.
70. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med.* 2019;25(9):1337–40. <https://doi.org/10.1038/s41591-019-0548-6>.
71. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK, Spirit AI, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension. *BMJ.* 2020;370: m3164. <https://doi.org/10.1136/bmj.m3164>.
72. Rivera SC, Liu X, Chan AW, Denniston AK, Calvert MJ, Spirit AI, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI Extension. *BMJ.* 2020;370: m3210. <https://doi.org/10.1136/bmj.m3210>.
73. Gerke S, Babic B, Evgeniou T, Cohen IG. The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. *NPJ Digit Med.* 2020;3:53. <https://doi.org/10.1038/s41746-020-0262-2>.
74. Kersting K. Machine Learning and artificial intelligence: two fellow travelers on the quest for intelligent behavior in machines. *Front Big Data.* 2018;1:6. <https://doi.org/10.3389/fdata.2018.00006>.
75. Nohara Y, Matsumoto K, Soejima H, Nakashima N. Explanation of machine learning models using shapley additive explanation and application for real data in hospital. *Comput Methods Programs Biomed.* 2022;214: 106584. <https://doi.org/10.1016/j.cmpb.2021.106584>.
76. Scott HF, Colborn KL, Sevick CJ, Bajaj L, Kissoon N, Deakyne Davies SJ, et al. Development and validation of a predictive model of the risk of pediatric septic shock using data known at the time of hospital arrival. *J Pediatr.* 2020;217(145–51): e6. <https://doi.org/10.1016/j.jpeds.2019.09.079>.
77. Sendak MP, Ratliff W, Sarro D, Alderton E, Futoma J, Gao M, et al. Real-world integration of a sepsis deep learning technology into routine clinical care: implementation study. *JMIR Med Inform.* 2020;8(7): e15182. <https://doi.org/10.2196/15182>.
78. Adams R, Henry KE, Sridharan A, Soleimani H, Zhan A, Rawat N, et al. Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis. *Nat Med.* 2022;28(7):1455–60. <https://doi.org/10.1038/s41591-022-01894-0>. **A large multi-centre study of an AI-based decision support system for sepsis which demonstrated a mortality benefit associated with its implementation.**
79. Sandhu S, Lin AL, Brajer N, Sperling J, Ratliff W, Bedoya AD, et al. Integrating a machine learning system into clinical workflows: qualitative study. *J Med Internet Res.* 2020;22(11): e22421. <https://doi.org/10.2196/22421>.
80. Price WN 2nd, Gerke S, Cohen IG. Potential liability for physicians using artificial intelligence. *JAMA.* 2019;322(18):1765–6. <https://doi.org/10.1001/jama.2019.15064>.
81. Duncan CF, Youngstein T, Kirrane MD, Lonsdale DO. Diagnostic challenges in sepsis. *Curr Infect Dis Rep.* 2021;23(12):22. <https://doi.org/10.1007/s11908-021-00765-y>.

82. Habli I, White S, Sujan M, Harrison S, Ugarte M. What is the safety case for health IT? A study of assurance practices in England. *Saf Sci*. 2018;110:324–35.
83. Digital N. DCB0160: Clinical risk management: its application in the deployment and use of health IT systems, 2018. <https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/dcb0160-clinical-risk-management-its-application-in-the-deployment-and-use-of-health-it-systems>. Accessed 14 July 2023.
84. Jia Y, Lawton T, Burden J, McDermid J, Habli I. Safety-driven design of machine learning for sepsis treatment. *J Biomed Inform*. 2021;117: 103762. <https://doi.org/10.1016/j.jbi.2021.103762>.
85. Agency CQCaMaHpR. Using machine learning in diagnostic services: A report with recommendations from CQC's regulatory sandbox, 2020. https://www.cqc.org.uk/sites/default/files/20200324%20CQC%20sandbox%20report_machine%20learning%20in%20diagnostic%20services.pdf. Accessed 14 July 2023.
86. Wen SCH, Ezure Y, Rolley L, Spurling G, Lau CL, Riaz S, et al. Gram-negative neonatal sepsis in low- and lower-middle-income countries and WHO empirical antibiotic recommendations: a systematic review and meta-analysis. *PLoS Med*. 2021;18(9): e1003787. <https://doi.org/10.1371/journal.pmed.1003787>.
87. Mancini N, Carletti S, Ghidoli N, Cichero P, Burioni R, Clementi M. The era of molecular and other non-culture-based methods in diagnosis of sepsis. *Clin Microbiol Rev*. 2010;23(1):235–51. <https://doi.org/10.1128/CMR.00043-09>.
88. de Souza DC, Goncalves Martin J, Soares Lanziotti V, de Oliveira CF, Tonial C, de Carvalho WB, et al. The epidemiology of sepsis in paediatric intensive care units in Brazil (the Sepsis PREvalence Assessment Database in Pediatric population, SPREAD PED): an observational study. *Lancet Child Adolesc Health*. 2021;5(12):873–81. [https://doi.org/10.1016/S2352-4642\(21\)00286-8](https://doi.org/10.1016/S2352-4642(21)00286-8).
89. Narayanan I, Nsungwa-Sabiti J, Lusyati S, Rohsiswatmo R, Thomas N, Kamalarathnam CN, et al. Facility readiness in low and middle-income countries to address care of high risk/ small and sick newborns. *Matern Health Neonatol Perinatol*. 2019;5:10. <https://doi.org/10.1186/s40748-019-0105-9>.
90. Wyber R, Vaillancourt S, Perry W, Mannava P, Folaranmi T, Celi LA. Big data in global health: improving health in low- and middle-income countries. *Bull World Health Organ*. 2015;93(3):203–8. <https://doi.org/10.2471/BLT.14.139022>.
91. Dornan L, Pinyopornpanish K, Jiraporncharoen W, Hashmi A, Dejkriengkraikul N, Angkurawaranon C. Utilisation of electronic health records for public health in asia: a review of success factors and potential challenges. *Biomed Res Int*. 2019;2019:7341841. <https://doi.org/10.1155/2019/7341841>.
92. Keddy KH, Saha S, Okeke IN, Kalule JB, Qamar FN, Kariuki S. Combating childhood infections in LMICs: evaluating the contribution of Big Data Big data, biomarkers and proteomics: informing childhood diarrhoeal disease management in low- and middle-income countries. *EBioMedicine*. 2021;73: 103668. <https://doi.org/10.1016/j.ebiom.2021.103668>.
93. Alter GC, Vardigan M. Addressing Global Data Sharing Challenges. *J Empir Res Hum Res Ethics*. 2015;10(3):317–23. <https://doi.org/10.1177/1556264615591561>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.