TINY PATIENTS, BIG DATA: HOW ARTIFICIAL INTELLIGENCE CAN TRANSFORM NEONATOLOGY

Zdraveska N

[1]*University Clinic for Pediatric Diseases, Faculty of Medicine, Ss Cyril and Methodius University in Skopje, North Macedonia*

Artificial intelligence (AI) has emerged as a transformative technology, becoming an integral part of our daily lives, particularly in healthcare. It has the potential to harness vast amounts of data, serving as a powerful tool to support clinical decisions, provide personalized care, deliver accurate prognostics, and enhance patient safety.

The term "artificial intelligence" refers to the ability of computing algorithms to replicate human decision-making processes. "Machine learning" (ML) is a subset of AI that encompasses techniques that enable machines to learn from large datasets without explicit programming, primarily aimed at developing predictive models. Both AI and ML can analyze significantly more data than humans can handle, exploring nonlinear relationships and complex interactions among biological, genetic, and environmental factors. Their main focus is on prediction and detection rather than establishing causal relationships. (1)

Neonatology, as a pediatric subspecialty, along with neonatal intensive care, generates significant volumes of multidimensional data. This presents opportunities for both AI and ML to transform neonatology, enhance early diagnostics, improve monitoring processes, and develop targeted treatment strategies.

There has been a significant increase in the number of PubMed citations over the last 15 years on the use of AI in both adults and children. A PubMed search using the combined Medical Subject Headings (MeSH) terms "Artificial Intelligence" or "Machine Learning" and either "Infant, Newborn" or "Intensive Care, Neonatal" for neonates retrieved over 1,000 publications, with a quarter of these published in the last two years.

A recently published systematic review aimed at mapping the current evidence on ML in pediatrics and adolescent medicine found that neonatology was the dominant specialty, accounting for 24% of the 363 articles reviewed. Furthermore, the majority of these studies were conducted in high-income and upper-middle-income countries, whereas low- and lower-middle-income countries (LMIC) were represented in only 3% of the studies. Despite the high academic interest, the implementation of AI in clinical settings remains limited. In addition, the most impactful applications of AI in neonatal medicine will be addressed. (2)

AI Application in Routinely Recorded Neonatal Vital Signs

Continuous real-time monitoring of vital signs—such as heart rate, respiratory rate, and pulse oximetry—is essential in specialized neonatal care. Data collected from multiple sensors provides valuable insights into the infant's clinical status. Vital signs are influenced by complex physiological mechanisms, and any deviations from the normal patterns can be early indicators of potential medical complications. Clinical care is typically influenced by interval assessments of vital signs, and with periodic assessments, the majority of data on bedside monitors remains unanalyzed. Pre-set alarm thresholds can alert healthcare providers to acute changes in an infant's condition. However, the prevalence of false-positive alarms may lead to alarm fatigue, which can adversely impact clinical decision-making during critical moments when prompt intervention is required. (3)

Recent studies indicate that cumulative analysis of continuously recorded vital sign trends can serve as a valuable tool for predicting imminent clinical deterioration and the onset of pathological conditions such as sepsis. (4)

Late-onset sepsis (LOS) poses a significant burden of morbidity and mortality among preterm infants. Timely recognition, along with prompt initiation of antibiotic therapy, is essential for preventing adverse outcomes in these infants. (4,5) However, due to the subtle and nonspecific signs and symptoms, clinicians often find themselves at a challenging crossroad between early detection and the risks of over-treatment and antimicrobial resistance.

Alterations in heart rate characteristics (HRC), including frequent small accelerations and decelerations, occur as part of the pathophysiologic response to systemic infection. In particular, sepsis is the most significant cause of abnormal heart rate in preterm infants. Since LOS is often associated with variations in vital signs, it offers an ideal opportunity for developing ML models. AI algorithms can help detect LOS before sepsis becomes clinically apparent. (5)

The first commercially available device, called the HeRO (Heart Rate Observation) monitor, was developed to analyze abnormal heart rate characteristics and provide clinicians with a score indicating the risk of an infant developing sepsis within the next 24 hours. (6)

Results from the largest randomized controlled trial to date involving very low-birth-weight (VLBW) infants, including 3,003 participants, showed statistically significant and clinically meaningful outcomes. There was a 22% relative reduction in mortality for infants whose HeRO scores were displayed, translating to one additional survivor for every 48 VLBW infants monitored or for every 23 extremely low birth weight (ELBW) infants monitored, and a 40% reduction of sepsis mortality within 30 days. (7,8)

However, the abnormal HRC is not exclusively indicative of neonatal sepsis; other potential conditions must be considered, as some studies have shown a 10% increase in blood cultures and a 5% rise in antibiotic use. (5) Therefore, by combining physical examination findings, interviewing the attending caregiver, and analyzing trends in the HeRo score, we can significantly enhance the diagnostic accuracy. Ultimately, it is up to the physician to determine the next steps, whether that involves a "wait and watch closely" approach or initiating investigations and empirical treatment.

A recent meta-analysis, including nine articles and twelve prediction models, evaluated the diagnostic accuracy of existing prediction models for sepsis in neonates. Among these, three models demonstrated a sensitivity of at least 95%, with significantly lower specificity and positive predictive value.(9) Consequently, the prediction models should be used as guidance rather than absolute indicators, as most of them have limited diagnostic accuracy. It's essential to always consider the clinical context. Therefore, future research will aim to enhance predictive models by integrating targeted physio and laboratory biomarkers with cardiorespiratory algorithms, as this is likely to have a much greater impact than either strategy alone. Finally, the integration of validated AI models into wireless monitoring devices has the potential to transform newborn care in terms of comfort, family-centered care, and better outcomes. (10)

AI and Neuromonitoring

Neonatal seizures are a neurological emergency that requires prompt treatment. However, diagnosing seizures is particularly difficult in preterm infants, as over 85% of cases may not exhibit obvious clinical signs, and the events are usually electrographic. Additionally, differentiating neonatal seizures from other physiological or abnormal, but non-epileptic movements in neonates presents a significant challenge. As a result, infants may be undertreated if their condition goes undiagnosed or overtreated if misdiagnosed.

Continuous conventional electroencephalography (cEEG) is the gold standard for diagnosing neonatal seizures. However, its implementation can be expensive and time-consuming. Furthermore, real-time interpretation requires the presence of highly skilled medical personnel 24/7, which may not be feasible, even in high-income countries. (11) A simplified, though less accurate form of EEG monitoring is the amplitude-integrated EEG (aEEG). This method has been used in neonatal units for many years, primarily due to the difficulties in obtaining a conventional EEG.

The majority of AI research using neonatal EEG has focused on developing algorithms for automated seizure detection, achieving impressive detection accuracy comparable to that of human experts.

One of the most impactful studies was a multicenter randomized controlled trial conducted across eight NICUs in Ireland, the Netherlands, Sweden, and the UK. All neonates included in the study underwent EEG monitoring due to clinically suspected seizures or a high risk of seizures. An automated EEG-based seizure detection software, known as the Algorithm for Neonatal Seizure Recognition (ANSeR), was used in real time to assess the diagnostic accuracy of detecting neonatal electrographic seizures, serving as a support tool for clinicians at the bedside. The neonates were randomly assigned (1:1) into two groups: one group received continuous EEG (cEEG) monitoring with ANSeR (the algorithm group), while the other group received routine cEEG monitoring without the algorithm (the non-algorithm group), which is considered standard care at the participating centers. All cEEG recordings were reviewed by two independent expert neurophysiologists. The study's findings revealed a higher percentage of seizure hours correctly identified in the algorithm group. The use of the algorithm was safe and did not lead to an increase in antiseizure medication prescriptions. (12) An additional noteworthy finding was a significant difference in seizure recognition between weekdays and weekends, with seizures being less likely to be recognized during weekends without the support of the algorithm. (12) This implies that in limited resource settings and in centers with less experience in interpreting neonatal EEG there is a potentially greater benefit of using an algorithm.

Ongoing research focuses on developing algorithms that can assess brain maturation, estimate sleep states, and analyze background EEG patterns in conditions such as hypoxic-ischemic encephalopathy (HIE). By applying machine learning algorithms to readily available clinical data, we may be able to accurately identify infants at risk of developing HIE following perinatal asphyxia. This approach could be integrated into bedside decision-support tools, ensuring timely and precise initiation of therapeutic hypothermia for infants who would benefit the most. (13-15)

The multidisciplinary approach is critical to success; thus, a recently funded European Cooperation in Science and Technology (EU COST) AI4NICU Action (https://www.cost.eu/actions/CA20124), to which our university participated, aims to foster the development of AI technologies that detect brain injuries in neonates through such multidisciplinary collaboration.

AI and Neonatal Jaundice

Neonatal jaundice is a prevalent condition, affecting 60-80% of healthy newborns. While it is generally benign, severe hyperbilirubinemia can lead to permanent brain damage, a condition known as "kernicterus," associated with lifelong complications and disabilities. The American Academy of Pediatrics recommends universal screening for bilirubin levels in all newborns prior to discharge, along with a close follow-up for jaundice assessment after 48-72 hours. Since bilirubin levels typically peak around the third postnatal day — when many newborns are already at home — accurately detecting jaundice can be particularly challenging, as it primarily relies on visual inspection. Hyperbilirubinemia presents a prevalent cause of neonatal readmission, particularly in low and LMICs. This can cause considerable stress for newborn mothers and may result in anxiety, depression, or cessation of breastfeeding. Consequently, a need to develop an inexpensive, widely available technology to screen newborns for jaundice globally emerged. Recent studies have introduced innovative methods for assessing neonatal jaundice using digital images and smartphones. Researchers have developed a physics-based system that estimates bilirubin levels from these images using deep learning and machine learning models. In this system, images are captured with a smartphone and color-calibrated using a reference card that matches the skin tone of the area over a newborn's sternum. The calibrated image data is then transmitted via the Internet to a computer server, where it is compared to a large database containing pairs of colors and their corresponding bilirubin levels. (16-18)

Taylor and his group developed BiliCam, a smartphone application, and assessed its accuracy in a diverse sample of newborns at 7 sites across the United States, showing an impressive overall correlation of 0.91 and a sensitivity of 84.6% for identifying newborns in high-risk zones. (17)

Another smartphone application, "Picterus," was developed by researchers at the Norwegian University of Technology and Science. This innovative tool has demonstrated remarkable effectiveness for screening neonatal jaundice, particularly in detecting cases of severe jaundice (total serum bilirubin > 250 μmol/L). The image-derived bilirubin were strongly correlated with total serum bilirubin levels in both Caucasian and non-Caucasian populations in low- and LMICs, including individuals with moderate dark skin tones. (18,19)

We are currently conducting a study to evaluate the accuracy of the Picterus application during phototherapy for neonates hospitalized for jaundice in the Department of Neonatology at the University Children's Hospital. The use of mobile health technology holds great promise for advancing neonatal care, as it allows for a noninvasive alternative to bilirubin monitoring.

Similar approaches could be adapted to support the diagnosis of neonatal conditions, including syndromes, especially when linked with other data, in low and middle-income countries with limited access to genetic screening services. (20)

In conclusion, artificial intelligence will play a crucial role in the future of healthcare, particularly in complex areas such as neonatology and neonatal intensive care. It will enhance clinical decision-making, facilitate efficient and personalized care for newborns, and help minimize avoidable errors. To fully harness the AI capabilities in neonatal care, it is crucial to enhance digital literacy among healthcare professionals and encourage multi-disciplinary collaboration. AI should be seen as an invaluable tool in the healthcare professionals' toolkit, alongside standardized blood tests and imaging, to provide effective neonatal care in everyday clinical practice.

References:

1. Rowe M. An introduction to machine learning for clinicians. Acad Med. 2019;94:1433–6.
2. Hoodbhoy Z, Masroor Jeelani S, Aziz A, et al. Machine learning for child and adolescent health: a systematic review. Pediatrics. 2021 Jan;147(1):e2020011833.
3. Ostojic D, Guglielmini S, Moser V, et al. Reducing false alarm rates in neonatal intensive care: a new machine learning approach. Adv Exp Med Biol. 2020;1232:285-290.
4. Kumar N, Akangire G, Sullivan B, et al. Continuous vital sign analysis for predicting and preventing neonatal diseases in the twenty-first century: big data to the forefront. Pediatr Res. 2020 Jan;87(2):210-220.
5. Cabrera-Quiros L, Kommers D, Wolvers MK, et al. Prediction of late-onset sepsis in preterm infants using monitoring signals and machine learning. Crit Care Explor. 2021 Jan 27;3(1):e0302.
6. Hicks JF, Fairchild K. HeRO monitoring in the NICU: sepsis detection and beyond. Infant 2013; 9(6): 187-91.
7. Moorman JR, Carlo WA, Kattwinkel J, et al. Mortality reduction by heart rate characteristic monitoring in very low birth weight neonates: a randomized trial. J Pediatr. 2011 Dec;159(6):900-6.e1.
8. Fairchild KD, Schelonka RL, Kaufman DA, et al. Septicemia mortality reduction in neonates in a heart rate characteristics monitoring trial. Pediatr Res. 2013 Nov;74(5):570-5.
9. Verstraete EH, Blot K, Mahieu L, et al. Prediction models for neonatal health care-associated sepsis: a meta-analysis. Pediatrics. 2015 Apr;135(4):e1002-14.
10. Chung HU, Rwei AY, Hourlier-Fargette A, et al. Skin-interfaced biosensors for advanced wireless physiological monitoring in neonatal and pediatric intensive-care units. Nat Med. 2020 Mar;26(3):418-429.
11. Shellhaas RA. Continuous long-term electroencephalography: the gold standard for neonatal seizure diagnosis. Semin Fetal Neonatal Med. 2015 Jun;20(3):149-53.
12. Pavel AM, Rennie JM, de Vries LS, et al. A machine-learning algorithm for neonatal seizure recognition: a multicentre, randomised, controlled trial. Lancet Child Adolesc Health. 2020 Oct;4(10):740-749.
13. Raurale SA, Boylan GB, Mathieson SR, et al. Grading hypoxic-ischemic encephalopathy in neonatal EEG with convolutional neural networks and quadratic time-frequency distributions. J Neural Eng. 2021 Mar 19;18(4):046007.
14. Mooney C, O'Boyle D, Finder M, et al. Predictive modelling of hypoxic ischaemic encephalopathy risk following perinatal asphyxia. Heliyon. 2021 Jun 29;7(7):e07411.
15. O'Boyle DS, Dunn WB, O'Neill D, et al.. Improvement in the prediction of neonatal hypoxic-ischemic encephalopathy with the integration of umbilical cord metabolites and current clinical makers. J Pediatr. 2021 Feb;229:175-181.e1.

16. Vartdal G, Randeberg Lise L, Aune A, Kringstad A, Picterus AS. Image based bilirubin determination. EP patent EP 3393353 A4. 2019 2016/12/15.

17. Taylor JA, Stout JW, de Greef L, et al. Use of a smartphone app to assess neonatal jaundice. Pediatrics. 2017 Sep;140(3):e20170312.

18. Aune A, Vartdal G, Bergseng H, et al. Bilirubin estimates from smartphone images of newborn infants' skin correlated highly to serum bilirubin levels. Acta Paediatr. 2020 Dec;109(12):2532-2538.

19. Aune A, Jímenez-Díaz G, Gierman LM, et al. Smartphone-based screening of neonatal jaundice in three populations in low and middle-income countries: a cross-sectional study. BMJ Paediatr Open. 2025 May 8;9(1):e002242.

20. Porras AR, Rosenbaum K, Tor-Diez C, et al. Development and evaluation of a machine learning-based point-of-care screening tool for genetic syndromes in children: a multinational retrospective study. Lancet Digit Health. 2021 Oct;3(10):e635-e643.