Seminars in Fetal and Neonatal Medicine 27 (2022) 101346

Contents lists available at ScienceDirect

Seminars in Fetal and Neonatal Medicine

journal homepage: www.elsevier.com/locate/siny

Application and potential of artificial intelligence in neonatal medicine

T'ng Chang Kwok [a], Caroline Henry [a], Sina Saffaran [b], Marisse Meeus [c], Declan Bates [d], David Van Laere [c], Geraldine Boylan [e], James P. Boardman [f], Don Sharkey [a,*]

[a] Centre for Perinatal Research, School of Medicine, University of Nottingham, E Floor, East Block, Queen's Medical Centre, Nottingham, NG7 2UH, UK
[b] Faculty of Engineering Science, University College London, London, WC1E 6BT, UK
[c] Department of Neonatal Intensive Care, University Hospital Antwerp, Edegem, Belgium
[d] School of Engineering, University of Warwick, Coventry, CV4 7AL, UK
[e] INFANT Research Centre, University College Cork, Cork, Ireland
[f] MRC Centre for Reproductive Health, University of Edinburgh, Edinburgh, UK

ARTICLE INFO

Keywords:
Artificial intelligence
Neonatal medicine
Machine learning

ABSTRACT

Neonatal care is becoming increasingly complex with large amounts of rich, routinely recorded physiological, diagnostic and outcome data. Artificial intelligence (AI) has the potential to harness this vast quantity and range of information and become a powerful tool to support clinical decision making, personalised care, precise prognostics, and enhance patient safety. Current AI approaches in neonatal medicine include tools for disease prediction and risk stratification, neurological diagnostic support and novel image recognition technologies.

Key to the integration of AI in neonatal medicine is the understanding of its limitations and a standardised critical appraisal of AI tools. Barriers and challenges to this include the quality of datasets used, performance assessment, and appropriate external validation and clinical impact studies. Improving digital literacy amongst healthcare professionals and cross-disciplinary collaborations are needed to harness the full potential of AI to help take the next significant steps in improving neonatal outcomes for high-risk infants.

1. Introduction

Artificial intelligence (AI) is an integral part of our daily lives and has begun penetrating adult healthcare settings with the advent of deep learning techniques and the "big data" environment. A PubMed search of AI papers found 23,442 citations in adult healthcare as compared to just 848 citations for neonatal care, with a significant increase in publications across both groups in the last 3–4 years (Fig. 1). Despite the tremendous potential, AI application in neonatology is still in its infancy. In this review, we explore examples of AI applications, the challenges of AI integration and its potential expansion in neonatal medicine.

2. Neonatal AI application

The use of AI in neonatal care has huge potential, especially with the increasingly complex intensive care provided for high-risk infants. AI should be viewed as a tool within the healthcare professionals' (HCPs') armoury, alongside blood investigations and imaging, to support shared clinical decision making, provide efficient personalised neonatal care and reduce avoidable errors.

2.1. AI application in large neonatal clinical datasets

2.1.1. Prediction of neonatal mortality and morbidity

The establishment of high quality, validated, multi-dimensional neonatal datasets have led to the development of prediction models for neonatal mortality or morbidity using deep learning approaches [1]. A good example is preterm survival without bronchopulmonary dysplasia (BPD) [2], an important preterm birth research priority with significant respiratory and neurological problems into adulthood [3].

There were 26 BPD prediction models identified in 2012 [4] with a further 27 identified since then (TCK/DS unpublished data). None are in routine clinical use despite the need for an objective measure to identify high-risk infants for timely targeted preventative treatments such as postnatal corticosteroids. Currently, subjective approaches are used to inform preventative strategies, often based on a few birth characteristics and previous experiences. Uncertainty around who and when to treat

* Corresponding author. Neonatal Medicine and Technologies, Centre for Perinatal Research, School of Medicine, University of Nottingham, E floor, East Block, Queen's Medical Centre, Nottingham, NG7 2UH, UK.
E-mail address: don.sharkey@nottingham.ac.uk (D. Sharkey).

https://doi.org/10.1016/j.siny.2022.101346

Available online 18 April 2022

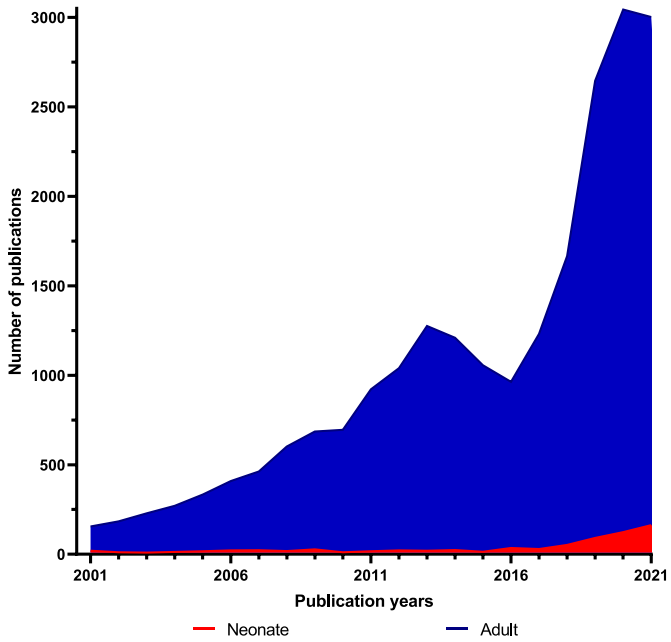

Fig. 1. Increase in the number of PubMed citations over the last 20 years comparing the use of artificial intelligence in neonates and adults using the combined Medical Subject Headings (MESH) terms of "Artificial Intelligence" OR "Machine Learning" AND either "Infant, Newborn" OR "Intensive Care, Neonatal" for neonates or "Adult" OR "Intensive Care" for adult. The search was performed on 15/03/2022.

may delay treatments and miss narrow therapeutic windows, leading to additional ventilator-induced lung injury (VILI). A dynamic approach, using deep learning to map each infant's trajectory, to provide personalised BPD risk would be invaluable in supporting clinical decision making. This approach has already been used in predicting respiratory failure in cystic fibrosis patients and could be adopted within neonatology if proven to add value and validated [5].

2.1.2. Identifying hidden patterns within data

Deep learning approaches have been deployed in large clinical datasets to identify hidden data patterns. An example of such application is exploring the variation of neonatal nutritional practices and their association with clinical outcomes [6]. This could identify "optimal" nutritional practices as well as improve the understanding of the underlying pathophysiology and impact of nutritional practices on neonatal outcomes.

2.2. AI application in real-time routinely recorded neonatal intensive care vital signs

Continuous monitoring of vital signs is an essential component of cardiorespiratory care of infants admitted to the neonatal intensive care unit (NICU) [7]. The abundance of data, generated by multiple sensors, can offer insight into the infant's clinical status. Pre-set alarm thresholds in bedside monitors can alert HCPs about an acute change in the infant's condition. However, in many cases, they are drawn to the bedside by false-positive alarms leading to 'alarm-fatigue', negatively influencing HCPs' vigilance when immediate action is warranted [8]. The application of machine learning (ML) algorithms on monitoring data has been shown to improve the selection of alarms requiring immediate intervention, providing earlier recognition for pre-emptive clinical action, and directing care towards a more efficient, individualized approach [9].

Many NICU's are still dependent on infrequent, usually hourly, snapshots of vital signs, often manually registered in the patient's medical record, leaving the majority of abundantly available digital data unutilised [10]. Others have already implemented the continuous visualisation of monitoring data into their electronic medical record, enabling HCPs to look for specific patterns in the combined data. This can lead to early recognition of disease states, sometimes hours before they would become clinically apparent [10]. ML algorithms can outperform HCPs in fulfilling this task, recognising the disease and supporting clinical decision making [7,10–12].

Late onset sepsis (LOS) results in a significant burden of morbidity and mortality in preterm infants. The timely recognition of disease onset, together with prompt initiation of antibiotic therapy, is crucial to prevent adverse outcomes in these infants. Since LOS is often accompanied by changes in the infant's vital signs [10], it serves as an ideal use case for ML model development. AI algorithms can help recognise LOS before sepsis is clinically apparent [11]. The efficiency of these models is often evaluated on a small snapshot of data preceding sepsis diagnosis to identify the disease at a certain timepoint before the clinical sepsis call was made. Although timely sepsis recognition is a key component of the ML model's success, precision remains a major challenge in the model's clinical usefulness. To evaluate how the model interacts with data throughout the infant's entire NICU stay, it should be subjected to continuous performance analysis, evaluating its output at every single timepoint [7]. Moreover, ML algorithms should be compared with current clinical tools in diagnostic accuracy, be subjected to internal validation and undergo generalisability testing with external validation [12].

Once proven AI models become available, integrating them into new wireless monitoring devices [13] could potentially lead to a disruptive change in the entire NICU ecosystem and better treatment for infants in terms of comfort, family-centred care and improved outcomes. Linking these data-rich measures from fetal monitoring, care in the delivery room, through to the NICU and early home monitoring, with longer-term outcomes could help advance the understanding of optimal early life care and improve the outcomes of high-risk infants.

2.3. AI application in neuroimaging and neurophysiological investigations

2.3.1. Magnetic resonance imaging

The past five years have seen enormous strides in the use of AI for improving value and inference from brain MRI. AI methods have led to technical advances including strategies to mitigate effects of movement artefact and increase information yield [14], and improvements in tissue classification, which is an essential step in many processing pipelines [15]. These have enabled researchers to probe the developing brain with ever-increasing depth leading to new insights about the impact of perinatal adversities on structural and functional network architectures [16]. In the clinical realm, AI is enabling innovation in 3 key areas: defining neuroanatomic phenotypes, predicting outcome, and facilitating scale-up of imaging studies.

Preterm birth is closely associated with a phenotype that includes atypical brain development, and subsequent intellectual disability, cerebral palsy, autism spectrum disorder, attention deficit hyperactivity disorder, psychiatric disease and problems with language, behaviour and socioemotional functions [17]. Structural, diffusion and functional MRI have each provided fundamental insights about alterations to structural and functional networks that are common in infants born preterm [18]. However, structure-function relations are better captured by models that integrate data from two or more imaging modalities in a single framework: for instance, multimodal image analysis reveals previously unrecognised patterns of neuroanatomic variants in preterm infants that correlate with clinical exposures and predict cognitive and motor outcomes [19]. ML has taken analytic capability to the next level by enabling "fingerprinting" of an individual's brain in morphometric similarity networks (MSNs). MSNs are computed by integrating different types of MRI data in a single model including regional volumes, diffusion tensor metrics, neurite orientation dispersion and density imaging

measures (Fig. 2). This approach can classify robustly the preterm brain phenotype, and it provides an accurate estimate of chronological brain age to within 0.7 ± 0.56 weeks [20] (Fig. 2). It is anticipated this type of approach will be useful for investigating drivers of brain dysmaturation and resilience, understanding the networks that contribute most to atypical brain development, and determining neural bases of cognition and behaviour.

Early prediction of neurodevelopmental disorders is one of the main goals of precision medicine because it would allow for early and more effective intervention, which is expected to lead to improved prognosis. To date, most studies report correlation between a perinatal exposure and a behavioural or cognitive phenotype, with many models unable to take account of other known determinants of brain development such as clinical co-morbidities, social gradients, the stress environment, genetic variants, and maternal health [21]. ML has begun to bridge the gap between biological, genetic and environmental variables [22], and it is proving useful for prediction frameworks. For example, a parsimonious ML model identified a robust set of eight out of 24 imaging and clinical features that predicted language impairment in preterm infants. The most important features were white matter microstructure, twin status, and incomplete or no exposure to antenatal corticosteroids. Female sex and breast milk exposure during NICU care reduced the risk of language delay [23]. Importantly, although these imaging and outcome prediction models have been developed in the context of preterm birth, they could be applied to many neurodevelopmental disorders.

Finally, some experimental designs require substantial scale-up using multi-site MRI data. These include defining 'normal' population variation in brain structure and function, investigating risk and resilience for brain development conferred by the (epi)genome, biomarker validation, and investigating early life origins of common neurologic and psychiatric diseases. This is challenging because of scanner and scan protocol differences, but ML can be used to enhance data interoperability, for example, by training image processing methods that are more robust to variations in input data.

2.3.2. *Continuous electroencephalography*

Electroencephalography (EEG) has become an invaluable component of neurocritical care in the NICU. EEG can be challenging to implement largely due to the lack of available experts to review the EEG in real-time and hence efforts have been ongoing for many years to automate this process.

EEG provides real-time information about brain activity and is now considered essential for the diagnosis and effective treatment of seizures in neonates [24]. A more simplified but less accurate form of EEG monitoring, the amplitude integrated EEG (aEEG), has been implemented in neonatal units for many years because of the difficulty in obtaining conventional EEG. Given that infants are born at any time, day or night, providing a 24-h neonatal EEG service is challenging for most centres. However, EEG is often needed very soon after birth particularly for infants with hypoxic-ischaemia encephalopathy (HIE) as seizures can emerge within the first 24 h [25].

As neonatal seizures are a neurological emergency, they require prompt treatment. Up to 85% of neonatal seizures may have no obvious clinical signs, particularly in infants with HIE, making recognition very difficult. The only way to recognise and promptly treat all seizures is to use continuous EEG monitoring [26]. An additional challenge for

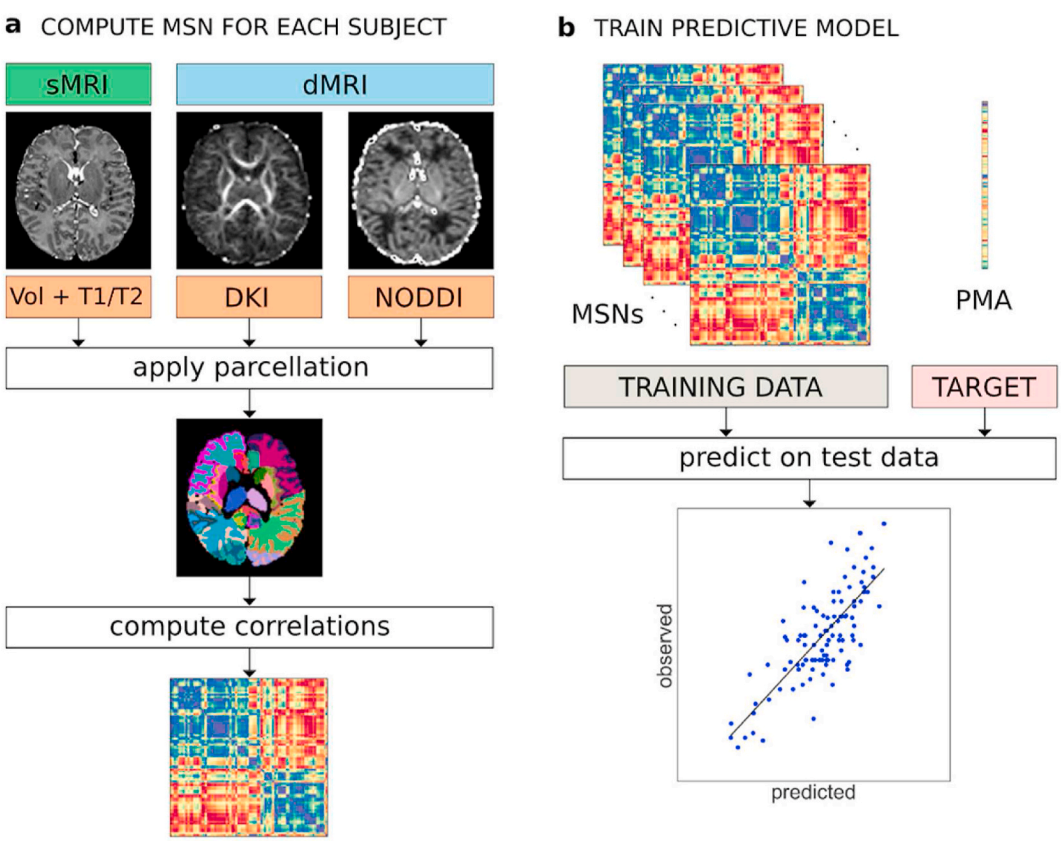

Fig. 2. a) Individual morphometric similarity networks (MSN) construction: Different metrics are extracted from structural and diffusion MRI data (sMRI, dMRI (such as diffusion kurtosis imaging (DKI) and Neurite Orientation Dispersion and Density Imaging (NODDI) models)). The same labelled atlas is applied to all image types and the average metric values are computed for 81 regions of interest. An MSN (represented here as a connectivity matrix) is built by computing the Pearson correlation between the vectors of metrics of each pair of ROIs. b) Training of a predictive model from individual MSNs: the inter-regional correlations are used as predictor variables in a machine learning model for chronological brain age. The performance of the model is evaluated on an independent test set. From Galdi P et al. Neuroimage Clin [20].

healthcare providers in the NICU is the uncoupling of clinical and EEG seizures following treatment with antiseizure medication [27]. This means that any clinical expression of the neonatal seizure that may have been present before treatment becomes invisible after treatment, yet the electrical seizure discharge continues. EEG monitoring is therefore essential to assess the efficacy of treatments.

While EEG is hugely beneficial for neonatal seizure monitoring and treatment, it is also very useful in helping support the diagnosis of neonatal encephalopathy, particularly HIE. The majority of AI research using neonatal EEG has emerged in the area of automated seizure detection and more recent algorithms have used deep learning. The seizure detection accuracy of these algorithms is impressive and comparable to that of human experts [28]. The report of the first multicentre trial of a neonatal seizure detection algorithm [29] demonstrates that an AI algorithm can be implemented in the NICU, is acceptable to staff and most importantly, given algorithms are 'always on', more seizures were detected in real-time. Most of the seizure detection algorithms developed to date have been for full terms neonates, but some for preterm infants are now emerging [30]. The bottleneck in the development of these algorithms is always the availability of sufficient data for testing, training and validation. In addition, all seizures in the datasets must be labelled by several experts before they can be used for testing and training.

Studies are also underway to develop algorithms, many using deep learning methods, that can assess brain maturation, estimate sleep states [31] and grade the background EEG patterns in conditions such as HIE [32] (Fig. 3). Appropriate and sufficient data that has been expertly labelled is key to this effort. Physiological data like EEG is noisy and subject to both movement and biological artefact, which must also be considered during testing and training. As a result, a multidisciplinary approach is critical to its success. The recently funded European Cooperation in Science and Technology (EU COST) AI4NICU Action (https://www.cost.eu/actions/CA20124/#tabs|Name:Working%20Groups) aims to speed up the development of AI technologies that detect brain injuries in neonates through such a multidisciplinary collaboration.

2.4. AI application in image recognition

Deep learning techniques such as convolutional neural networks (CNNs) and computer vision approaches can address the challenges of the potentially small and skewed datasets that often characterise neonatal image recognition tasks. These rely on a two-stage technique: automated segmentation of image or scene of interest followed by identification of the outcome of interest. Examples include the estimation of the gestational age of infants at birth using images of newborn infants [33]; analysing videos of clinical procedures, such as newborn resuscitation [34], and assessment of pain [35].

Gestational age at birth often guides treatment delivered by HCPs. Early dating prenatal ultrasound scan is the gold standard for assessing the gestational age of an infant [33]. However, this may not always be available particularly in low-resource settings. A novel automated postnatal gestational age estimation tool [36] used images of a newborn infants' face, ear and foot. The images were initially passed through a segmentation stage using fully convolutional networks (FCNs) (Fig. 4). The system was taught how to identify the body part of interest from the image and disregard the background. FCNs create binary masks, labelling the pixels of the body part of interest, and then comparing these to masks created when the images were manually annotated. Following segmentation to identify the region of interest, the images were passed through CNNs to learn features and classify the images into gestational classes for each body part. Following training, test images are passed through the CNNs and assigned a probability vector for each body part. The probability vectors are then combined with the newborn's weight and regression performed, outputting an estimate of the gestational age in days rather than a broad gestational class. The tool estimated gestational age to within approximately six days, surpassing that of postnatal clinical examination, including the Ballard score and last menstrual period [37]. Integration of the algorithms into smartphones could allow rapid estimation of gestational age, based on photos, in settings without other antenatal gestation dating resources. Similar approaches could be adapted to support the diagnosis of neonatal conditions including syndromes especially when linked with other data [38].

During newborn resuscitation, HCPs adhere to a defined protocol consisting of a sequence of actions by several people. Due to the high-pressure environment during resuscitation, errors can be made. Smith et al. [34] developed an automated scene segmentation and action recognition tool using deep learning to analyse videos of newborn resuscitation with good performance from as little as 20 training images. The tool would not only be useful in the training of HCPs by providing real-time feedback, but also has the potential to develop early warning systems used in resuscitation settings when HCPs deviate from standard protocols.

2.5. AI application in predicting response to neonatal treatment

Premature infants are commonly diagnosed with respiratory distress syndrome, requiring intubation and mechanical ventilation (MV). MV of preterm infants presents several challenges, including specific oxygenation targets and minimising ventilator-induced lung injury (VILI). Personalised treatment requires rapid and frequent interventions based

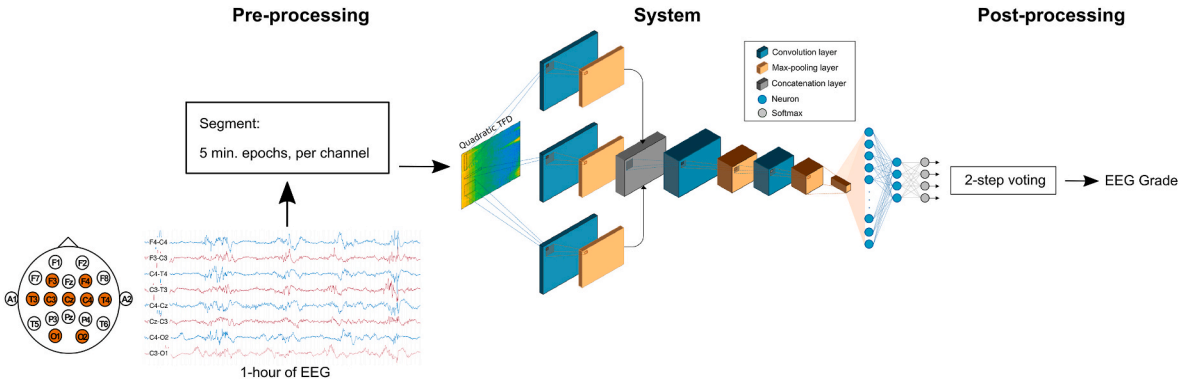

Fig. 3. Overview of an AI system to grade the neonatal EEG. The system assigns a grade (1–4) based on 1 h of multi-channel EEG. Pre-processing: the EEG is filtered and downsampled to 64 Hz, then divided into 5-min segments. System model: each 5-min segment, per channel, is converted to a quadratic time-frequency distribution (TFD). This 2-dimensional image is passed through a pre-trained convolutional neural network (CNN). The CNN has parallel layers to extract information across time, frequency, and in the joint time-frequency direction of the TFD. The CNN produces a separate probability for each of the 4 grades. Post-processing: the probabilities from the CNN are combined over all 5-min segments and across channels to give a final grade prediction. Adapted from Raurale et al. J Neural Eng [32].

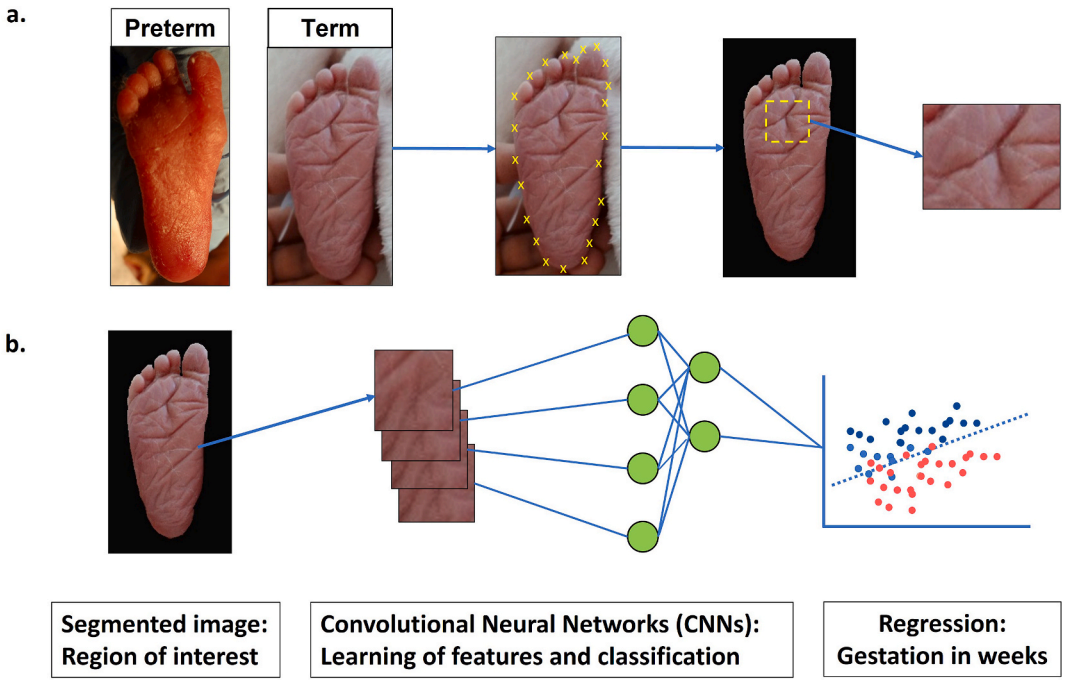

Fig. 4. a) Plantar surface photograph of preterm and post-term infant. Manually labelled region of interest (ROI) used to train system to identify foot and remove background with subsequent plantar crease pattern recognition. This uses fully Convolutional Neural Networks (CNNs) to segment images and provide per-pixel classification. b) Images undergo deep machine learning, with CNNs passing through different layers, to be classified into one of five classes (extremely preterm, very preterm, moderate preterm, term and post-term). This probability vector is then combined with the weight of the infant to improve the decision-making process during regression and the output is an estimate of gestational age in weeks. Adapted from Torres et al. Image and Vision Computing [36].

on changes in the patient's state that are often not achievable within current NICU constraints. For example, oxygen saturation targets in mechanically ventilated neonates were achieved only 40% of the time [39], something AI techniques could potentially address.

Minimising VILI is one of the key challenges in neonatal care to reduce the risk of associated adverse outcomes contributing to morbidity, mortality, and poor long term quality of life. Weaning from MV remains a complex clinical problem in the NICU, with 15%–40% of infants failing extubation [40]. AI techniques have been employed to help design decision support tools to predict extubation readiness and neonatal outcomes [41,42]. Precup et al. [41] suggested a predictor model, based on support vector machines, could reduce the extubation failure rate by more than 80% as compared to current clinical measures. The performance of predictive models of extubation outcome based on different ML algorithms have been studied [42]. Although some models showed satisfactory performance, HCPs' predictions still outperformed all developed models, indicating the need for further refinement of detailed mechanistic models that can be fully validated against individual patient data.

AI techniques can be exploited for optimising the administration of drug doses. Prediction models have used optimised support vector machines, decision tree ensembles, and deep learning to predict the effectiveness of therapeutic caffeine regimens in preventing apnoeas and reducing the need for prolonged MV [43].

Artificial neural networks have been designed to suggest the most protective ventilator settings and minimise VILI, maintaining blood gases within acceptable ranges demonstrating advantages over rule-based systems [44]. Recurrent neural networks for predicting future ventilation parameters have demonstrated good accuracy when predicting 1.5 s ahead [45], but performed poorly when attempting to predict further into the future, again indicating the need for more refined models with higher predictive capabilities [46].

3. Challenges of AI in neonatology

Alongside recognising the far-reaching AI potential in neonatal medicine, it is also crucial to understand its limitations and pitfalls if we are to integrate it within routine care pathways. Currently, there is a lack of uniformity in critically appraising the development and performance assessment of AI healthcare tools in the following areas [47,48].

3.1. *Dataset quality*

High-quality data is needed to train AI tools. The datasets used to develop AI tools should be reported fully by researchers and appraised carefully by reviewers [49]. Some of the common pitfalls are small sample sizes, inappropriate handling of missing values and heterogeneity assessment in different population subsets or healthcare settings [47]. Caution is also needed to look for biases against underserved groups that may have been unintentionally embedded within the developed AI tool.

3.2. *Model performance assessment based on the dataset type*

Cross-validation and bootstrapping are commonly used in assessing model performance to utilise the full range of data collected. Depending on the aim of the AI tool, performance should also be continuously assessed on the full dataset throughout the neonatal admission, rather than on a snapshot of the data.

The most reported performance measure of AI tools is the area under the receiver operating characteristics curve (AUROC), a discrimination measure. However, this can be misleading in certain settings. In some healthcare timeseries datasets with class imbalance, AUROC may provide false assurance of the model performance as AUROC depends on the true negative rate which is high in imbalanced datasets. Hence, other performance measures must be assessed including precision-recall curve, specificity/sensitivity and calibration measures [48].

3.3. External validation and clinical impact

Overfitting is a common issue plaguing AI prediction tools. Complex ML algorithms may be very sensitive in detecting nuances within the dataset, producing excellent performance in the initial dataset used to develop the algorithm. However, because the tool models the training data too well, its performance may suffer when tested on a different dataset or in clinical practice, making external validation essential [47].

Clinical impact studies are also required to provide robust evidence to inform AI application in healthcare settings. These assess the performance of AI tools in terms of their discrimination and calibration characteristics, and their impact on the clinical workflow (e.g. changes in the behaviour of HCPs or parents/carers) and patient outcomes. How the AI tool integrates seamlessly within the current clinical workflow, providing the right information to HCPs and parents/carers, is crucial to bridge the development-implementation gap to clinical practice [49].

3.4. Interpretability

Interpretability and transparency are crucial to achieving augmented intelligence. There have been increasing innovative efforts in exploring approaches to improve AI's interpretability [50] and explaining how the prediction is derived. Interpretability should be inherently considered in model construction for bedside clinical use, guiding the choice of algorithm to train a model.

3.5. Critical appraisal, regulatory and monitoring guidance

There has been an enormous effort in updating current critical appraisal, regulatory and monitoring guidance for AI healthcare devices [48,49]. When developed, these will help address some of the methodological, critical appraisal and medicolegal challenges as well as the monitoring needed to ensure its safe efficient use for the intended purpose.

4. Future

AI is likely to become an indispensable part of the neonatal care toolkit to support HCPs and parents/carers in providing improved, efficient and safer neonatal care. For this to become a reality, two crucial steps need to be taken. Firstly, the digital literacy among HCPs in understanding AI's principles and limitations needs to be improved. This enables HCPs to appraise newly developed AI tools and monitor their safety and appropriate use in clinical practice. Secondly, there is a need for cross-disciplinary, international collaborations that includes data and computer scientists, HCPs, lawyers and policymakers to design and apply AI tools that will overcome the challenges highlighted.

5. Conclusion

AI will be an integral part of the data-rich environment of neonatal care and this review highlights important areas of its application under investigation. These include mortality and disease prediction, image analysis and clinical decision support tools. However, current AI application is lagging behind adult specialities and a concerted effort is needed to accelerate neonatal AI research and translation into meaningful clinical application.

6. Practice points

- AI will be an integral part of future healthcare, especially in complex settings such as neonatal intensive care, to support shared clinical decision making, providing efficient personalised neonatal care while reducing avoidable errors.
- Novel application of AI in neonatal medicine is already improving our understanding of neuroanatomic phenotypes, prediction and disease modelling, seizure analysis and optimal ventilatory strategies.
- Addressing the challenges, improving digital literacy amongst HCPs and multi-disciplinary collaborations, are needed to harness the full AI potential in neonatal care.

Declaration of competing interest

TCK and DS are supported by the National Institute of Health Research (NIHR) Children and Young People MedTech Co-operative (CYP MedTech). DS has received funding for technology development from the Medical Research Council, NIHR and Action Medical Research, and is a non-executive director of SurePulse Medical who are developing monitoring solutions for neonatal care.

DVL received funding for technology development through a research grant provided by the University Hospital of Antwerp and a private fund. DVL is the co-owner and an executive board member of Innocens BV who are developing an AI based Clinical decision support system for Neonatal Intensive Care.

CH, MM, SS, DB, GB and JPB have no conflict of interest to disclose.

The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or of the Department of Health.

References

[1] Mangold C, Zoretic S, Thallapureddy K, Moreira A, Chorath K. Machine learning models for predicting neonatal mortality: a systematic review. Neonatology 2021;118:394–405.
[2] Dai D, Chen HY, Dong XR, Chen JL, Mei M, Lu YL, et al. Bronchopulmonary dysplasia predicted by developing a machine learning model of genetic and clinical information. Front Genet 2021;12.
[3] Duley L, Uhm S, Oliver S. Preterm Birth Priority Setting P. Top 15 UK research priorities for preterm birth. Lancet 2014;383:2041–2.
[4] Onland W, Debray TP, Laughon MM, Miedema M, Cools F, Askie LM, et al. Clinical prediction models for bronchopulmonary dysplasia: a systematic review and external validation study. BMC Pediatr 2013;13:207.
[5] Lee C, Yoon J, Schaar MV. Dynamic-DeepHit: a deep learning approach for dynamic survival analysis with competing risks based on longitudinal data. IEEE Trans Biomed Eng 2020;67:122–33.
[6] Greenbury SF, Ougham K, Wu J, Battersby C, Gale C, Modi N, et al. Identification of variation in nutritional practice in neonatal units in England and association with clinical outcomes using agnostic machine learning. Sci Rep 2021;11:7178.
[7] Van Laere D, Meeus M, Beirnaert C, Sonck V, Laukens K, Mahieu L, et al. Machine learning to support hemodynamic intervention in the neonatal intensive care unit. Clin Perinatol 2020;47:435–48.
[8] Ostojic D, Guglielmini S, Moser V, Fauchère JC, Bucher HU, Bassler D, et al. Reducing false alarm rates in neonatal intensive care: a new machine learning approach. Adv Exp Med Biol 2020;1232:285–90.
[9] Joshi R, Peng Z, Long X, Feijs L, Andriessen P, Van Pul C. Predictive monitoring of critical cardiorespiratory alarms in neonates under intensive care. IEEE J Trans Eng Health Med 2019;7:2700310.
[10] Kumar N, Akangire G, Sullivan B, Fairchild K, Sampath V. Continuous vital sign analysis for predicting and preventing neonatal diseases in the twenty-first century: big data to the forefront. Pediatr Res 2020;87:210–20.
[11] Cabrera-Quiros L, Kommers D, Wolvers MK, Oosterwijk L, Arents N, van der Sluijs-Bens J, et al. Prediction of late-onset sepsis in preterm infants using monitoring signals and machine learning. Crit Care Explor 2021;3:e0302.
[12] Zimmet AM, Sullivan BA, Fairchild KD, Moorman JR, Isler JR, Wallman-Stokes AW, et al. Vital sign metrics of VLBW infants in three NICUs: implications for predictive algorithms. Pediatr Res 2021;90:125–30.
[13] Chung HU, Rwei AY, Hourlier-Fargette A, Xu S, Lee K, Dunne EC, et al. Skin-interfaced biosensors for advanced wireless physiological monitoring in neonatal and pediatric intensive-care units. Nat Med 2020;26:418–29.
[14] Liu S, Thung K-H, Qu L, Lin W, Shen D, Yap P-T. Learning MRI artefact removal with unpaired data. Nat Mach Intell 2021;3:60–7.
[15] Serag A, Blesa M, Moore EJ, Pataky R, Sparrow SA, Wilkinson AG, et al. Accurate Learning with Few Atlases (ALFA): an algorithm for MRI neonatal brain extraction and comparison with 11 publicly available methods. Sci Rep 2016;6:23470.
[16] Blesa M, Galdi P, Cox SR, Sullivan G, Stoye DQ, Lamb GJ, et al. Hierarchical complexity of the macro-scale neonatal brain. Cerebr Cortex 2021;31:2071–84.
[17] Dean B, Ginnell L, Boardman JP, Fletcher-Watson S. Social cognition following preterm birth: a systematic review. Neurosci Biobehav Rev 2021;124:151–67.
[18] Batalle D, Edwards AD, O'Muircheartaigh J. Annual Research Review: not just a small adult brain: understanding later neurodevelopment through imaging the neonatal brain. JCPP (J Child Psychol Psychiatr) 2018;59:350–71.
[19] Ball G, Aljabar P, Nongena P, Kennea N, Gonzalez-Cinca N, Falconer S, et al. Multimodal image analysis of clinical influences on preterm brain development. Ann Neurol 2017;82:233–46.

[20] Galdi P, Blesa M, Stoye DQ, Sullivan G, Lamb GJ, Quigley AJ, et al. Neonatal morphometric similarity mapping for predicting brain age and characterizing neuroanatomic variation associated with preterm birth. Neuroimag Clin 2020;25:102195.
[21] Boardman JP, Counsell SJ. Invited Review: factors associated with atypical brain development in preterm infants: insights from magnetic resonance imaging. Neuropathol Appl Neurobiol 2020;46:413–21.
[22] Krishnan ML, Wang Z, Aljabar P, Ball G, Mirza G, Saxena A, et al. Machine learning shows association between genetic variability in PPARG and cerebral connectivity in preterm infants. Proc Natl Acad Sci U S A 2017;114:13744–9.
[23] Valavani E, Blesa M, Galdi P, Sullivan G, Dean B, Cruickshank H, et al. Language function following preterm birth: prediction using machine learning. Pediatr Res 2021. https://doi.org/10.1038/s41390-021-01779-x.
[24] DeLaGarza-Pineda O, Mailo JA, Boylan G, Chau V, Glass HC, Mathur AM, et al. Management of seizures in neonates with neonatal encephalopathy treated with hypothermia. Semin Fetal Neonatal Med 2021;26:101279.
[25] Pavel AM, Rennie JM, de Vries LS, Blennow M, Foran A, Shah DK, et al. Neonatal seizure management: is the timing of treatment critical? J Pediatr 2022;243:61-68.e2. https://doi.org/10.1016/j.jpeds.2021.09.058.
[26] Shellhaas RA. Continuous long-term electroencephalography: the gold standard for neonatal seizure diagnosis. Semin Fetal Neonatal Med 2015;20:149–53.
[27] Scher MS, Alvin J, Gaus L, Minnigh B, Painter MJ. Uncoupling of EEG-clinical neonatal seizures after antiepileptic drug use. Pediatr Neurol 2003;28:277–80.
[28] O'Shea A, Lightbody G, Boylan G, Temko A. Neonatal seizure detection from raw multi-channel EEG using a fully convolutional architecture. Neural Network 2020;123:12–25.
[29] Pavel AM, Rennie JM, de Vries LS, Blennow M, Foran A, Shah DK, et al. A machine-learning algorithm for neonatal seizure recognition: a multicentre, randomised, controlled trial. Lancet Child Adolesc Health 2020;4:740–9.
[30] O'Shea A, Ahmed R, Lightbody G, Pavlidis E, Lloyd R, Pisani F, et al. Deep learning for EEG seizure detection in preterm infants. Int J Neural Syst 2021;31:2150008.
[31] Ansari AH, Pillay K, Dereymaeker A, Jansen K, Van Huffel S, Naulaers G, et al. A deep shared multi-scale inception network enables accurate neonatal quiet sleep detection with limited EEG channels. IEEE J Biomed Health Inf 2022;26(3):1023–33. https://doi.org/10.1109/JBHI.2021.3101117.
[32] Raurale SA, Boylan GB, Mathieson S, Marnane WP, Lightbody G, O'Toole JM. Grading hypoxic-ischemic encephalopathy in neonatal EEG with convolutional neural networks and quadratic time-frequency distributions. J Neural Eng 2021;18(4):046007. https://doi.org/10.1088/1741-2552/abe8ae.
[33] Campbell S, Warsof SL, Little D, Cooper DJ. Routine ultrasound screening for the prediction of gestational age. Obstet Gynecol 1985;65:613–20.
[34] Smith MV TJ, Sharkey D, Crowe J. Clinical scene segmentation with tiny datasets. In: 2019 IEEE/CVF international conference on computer vision workshop (ICCVW). Clinical Scene Segmentation with Tiny Datasets; 2019. p. 1637–45. https://doi.org/10.1109/ICCVW.2019.00203.
[35] Egede J, Valstar M, Torres MT, Sharkey D. Automatic neonatal pain estimation: an acute pain in neonates database. In: 2019 8th international conference on affective computing and intelligent interaction (ACII); 2019. p. 1–7.
[36] Torres M, Valstar Mi, Henry C, Ward C, Sharkey D. Postnatal gestational age estimation of newborns using Small Sample Deep Learning. Image Vis Comput 2019;83–84:87–99.
[37] Lynch CD, Zhang J. The research implications of the selection of a gestational age estimation method. Paediatr Perinat Epidemiol 2007;21(Suppl 2):86–96.
[38] Porras AR, Rosenbaum K, Tor-Diez C, Summar M, Linguraru MG. 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;3:e635–43.
[39] Hagadorn JI, Furey AM, Nghiem TH, Schmid CH, Phelps DL, Pillers DA, et al. Achieved versus intended pulse oximeter saturation in infants born less than 28 weeks' gestation: the AVIOx study. Pediatrics 2006;118:1574–82.
[40] Dimitriou G, Fouzas S, Vervenioti A, Tzifas S, Mantagos S. Prediction of extubation outcome in preterm infants by composite extubation indices. Pediatr Crit Care Med 2011;12:e242–9.
[41] Precup D, Robles-Rubio CA, Brown KA, Kanbar L, Kaczmarek J, Chawla S, et al. Prediction of extubation readiness in extreme preterm infants based on measures of cardiorespiratory variability. Annu Int Conf IEEE Eng Med Biol Soc 2012;2012:5630–3.
[42] Mueller M, Almeida JS, Stanislaus R, Wagner CL. Can machine learning methods predict extubation outcome in premature infants as well as clinicians? J Neonatal Biol 2013;2.
[43] Shirwaikar RD. Estimation of caffeine regimens: a machine learning approach for enhanced clinical decision making at a neonatal intensive care unit (NICU). Crit Rev Biomed Eng 2018;46:93–115.
[44] Snowden S, Brownlee KG, Smye SW, Dear PR. An advisory system for artificial ventilation of the newborn utilizing a neural network. Med Inform 1993;18:367–76.
[45] Brigham K, Gupta S, Brigham JC. Predicting responses to mechanical ventilation for preterm infants with acute respiratory illness using artificial neural networks. Int J Num Method Biomed Eng 2018:e3094.
[46] Saffaran S, Das A, Algarni S, Hardman J, Sharkey D, Bates D. Computational simulation of mechanically ventilated neonatal patients in the intensive care unit. In: 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC); 2019.
[47] Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung MC, Ursprung S, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. Nat Mach Intell 2021;3:199–217.
[48] 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:e048008.
[49] Group D-AS. DECIDE-AI: new reporting guidelines to bridge the development-to-implementation gap in clinical artificial intelligence. Nat Med 2021;27:186–7.
[50] Hanif AM, Beqiri S, Keane PA, Campbell JP. Applications of interpretability in deep learning models for ophthalmology. Curr Opin Ophthalmol 2021;32:452–8.