Check for updates

REVIEW ARTICLE   OPEN

Big data, machine learning, and population health: predicting cognitive outcomes in childhood

Andrea K. Bowe[1✉], Gordon Lightbody[1,2], Anthony Staines[3] and Deirdre M. Murray[1]

© The Author(s) 2022

The application of machine learning (ML) to address population health challenges has received much less attention than its application in the clinical setting. One such challenge is addressing disparities in early childhood cognitive development—a complex public health issue rooted in the social determinants of health, exacerbated by inequity, characterised by intergenerational transmission, and which will continue unabated without novel approaches to address it. Early life, the period of optimal neuroplasticity, presents a window of opportunity for early intervention to improve cognitive development. Unfortunately for many, this window will be missed, and intervention may never occur or occur only when overt signs of cognitive delay manifest. In this review, we explore the potential value of ML and big data analysis in the early identification of children at risk for poor cognitive outcome, an area where there is an apparent dearth of research. We compare and contrast traditional statistical methods with ML approaches, provide examples of how ML has been used to date in the field of neurodevelopmental disorders, and present a discussion of the opportunities and risks associated with its use at a population level. The review concludes by highlighting potential directions for future research in this area.

IMPACT:

- To date, the application of machine learning to address population health challenges in paediatrics lags behind other clinical applications. This review provides an overview of the public health challenge we face in addressing disparities in childhood cognitive development and focuses on the cornerstone of early intervention.
- Recent advances in our ability to collect large volumes of data, and in analytic capabilities, provide a potential opportunity to improve current practices in this field.
- This review explores the potential role of machine learning and big data analysis in the early identification of children at risk for poor cognitive outcomes.

EARLY LIFE AND NEURODEVELOPMENTAL PLASTICITY

The phrase 'the first 1000 days', a period spanning conception to 3 years of age, is one increasingly seen in policy documents relating to child health.[1] The phrase has its origins in the work of Professor David Barker, a physician and epidemiologist, whose trio of seminal articles published in the Lancet between 1986 and 1993 prompted the development of the 'foetal origins hypothesis'.[2–5] A fundamental concept in Barker's work was developmental plasticity, formally defined as 'the ability of a single genotype to produce more than one alternative form of structure, physiological state, or behaviour in response to environmental conditions'.[6]

It is widely accepted that early life is a unique period, differing from adulthood, during which the brain shows greater potential towards plasticity.[7] This is explained in part by the highly dynamic dendritic spines, which are predominant during early development, and display huge potential for adaptability but, which over time and in adulthood are much more stable.[8] Early life is the most crucial period for shaping the developing brain and represents a window of both vulnerability and opportunity.[9] A failure to achieve early foundational cognitive skills may result in a permanent loss of opportunity to achieve full cognitive potential. Conversely, with appropriate intervention, it is a period during which there is an opportunity to disrupt the intergenerational transmission of inequity, particularly with regard to the developing brain.[9]

COGNITIVE FUNCTION AND THE IMPORTANCE OF SOCIAL DETERMINANTS

Cognitive function is a broad construct consisting of multiple domains, which can differ across literature and disciplines. Most definitions consistently refer to domains of learning, understanding, reasoning, problem solving, memory, language, attention, and decision making.[10–12] There has been considerable debate in the literature regarding the measurement of cognitive functioning, but also regarding its representativeness of overall human functioning.[13] While a thorough discussion of this debate is beyond the scope of this article, it is important to clearly acknowledge that childhood cognitive function is only one

[1]INFANT Research Centre, University College Cork, Cork, Ireland. [2]Department of Electrical and Electronic Engineering, University College Cork, Cork, Ireland. [3]School of Nursing, Psychotherapy, and Community Health, Dublin City University, Dublin, Ireland. ✉email: andrea.bowe@ucdconnect.ie

domain of overall child development. Other domains, including motor development, the development of positive social skills and emotional intelligence, and the establishment of appropriate behaviours are both intimately intertwined and of commensurate importance.[14]

For many children, particularly those exposed to adverse environmental conditions in early life, the true potential of their cognitive ability will never be reached. This gap between potential and actual ability brings with it significant implications for outcomes across multiple domains throughout the life course, including educational attainment,[15] social mobility,[16] financial well-being,[17] and survival.[18] Birth and pregnancy cohort studies have to date provided an evidence-base illustrating the important early life determinants of cognitive development. Consistently, socioenvironmental factors have been shown to be among the strongest determinants of cognitive function in childhood.[19–25] The effect of socioenvironmental disadvantage on cognitive function can be seen as early as 2 years of age, but if unabated, the adverse effect of disadvantage on cognitive function appears to be amplified over time.[22]

EVIDENCE FOR EARLY INTERVENTION
Across epidemiological, animal, and human research there is consensus agreement that early life is a period of fundamental importance in cognitive development, and strategies for intervention should ideally be targeting this unique period.[26–33]

One of the most widely cited examples supporting early cognitive intervention is the Abecedarian Study, a small randomised controlled trial that began in 1972, and involved infants at risk of poor cognitive outcomes based on an individual risk index that combined multiple factors including parental education, family income, welfare payments and the school history of family members.[34] The trial involved 111 infants, of whom 57 were randomly assigned to an intensive preschool programme, which commenced at a mean age of 4.4 months. The intervention consisted of high-quality early education delivered with high teacher to pupil ratios, 5 days per week for 50 weeks of the year, with free transportation and a focus on family involvement.[34,35] At age 3, significantly higher IQ scores were observed in the intervention group, and these persisted into adulthood. At age 21, the intervention group still had significantly higher IQ scores and academic achievement scores, as well as significantly more years of education completed compared to the control group.[36]

This pre-emptive approach to early cognitive intervention may also benefit other high-risk cohorts such as pre-term infants, who on average have childhood IQ scores approximately 10 points lower than their full-term peers.[37,38] A Cochrane review examining the effects of early intervention provided in the first 12 months of life, for infants identified as at risk based on a gestational age <37 weeks or a birth weight <2500 g, concluded that early interventions can improve cognitive outcomes in infancy, with benefits persisting into preschool age.[39]

Early pre-emptive approaches targeting high-risk individuals have also been shown to alter neurodevelopmental trajectories for other outcomes such as autism spectrum disorders (ASD). Whitehouse et al. demonstrated, in a randomised clinical trial of infants with very early signs of potential ASD, that pre-emptive intervention commenced around the first year of life reduced the odds of ASD classification at age 3.[40]

These studies illustrate the effectiveness of early pre-emptive intervention, commenced prior to overt signs of cognitive difficulty and at the time of optimal neuroplasticity. Outside of the controlled trial setting, real-world programmes, such as the federally sponsored Early Head Start programme in the United States, demonstrate that at a population level such interventions are both feasible, achievable, and effective with appropriate resourcing.[41,42] However, achieving optimal individual outcomes from early intervention is dependent on the accurate identification and involvement of infants and families who are likely to benefit most.

CURRENT POPULATION-BASED DEVELOPMENTAL SCREENING
Internationally, many countries rely on universal developmental screening programmes to identify children who may benefit from early intervention programmes, with the majority of developmental screening assessments being based on the presence of a delay in reaching developmental milestones.[43] Population-based screening programmes based on the principle of universalism apply structured developmental screening, independent of pre-existing risk factors, to all children. While this approach has many merits, there are significant limitations. Firstly, the opportunity for pre-emptive intervention is lost as these infants are already behind their typically developing counterparts at the time of detection, with valuable opportunities in the period of optimal neuroplasticity being lost. Secondly, population-based screening programmes depend on the voluntary presentation of parents and children and tend to exemplify the inverse care law in practice, whereby the most vulnerable and most at risk of developmental delay, are least likely to attend.[44–46] Thirdly, such screening programmes do not provide a systematic, robust assessment of the environmental, social, and demographic factors, to which the infant will be exposed in early life which are the strongest and most consistent predictors of cognitive outcomes in childhood.[47,48]

The birth of a child is unique in many ways, not least because there is an almost universal interaction between a mother and a healthcare professional. Therefore, the identification of high-risk mother–infant dyads in the perinatal period could provide a unique opportunity to engage families with early intervention programmes. Machine learning and artificial intelligence (AI), which have shown great promise in prediction across other fields of medicine, may have the potential to improve population-based strategies for the early identification of children at risk of poor cognitive outcomes, and this deserves further exploration.

MACHINE LEARNING AND PREDICTING RISK
Adverse cognitive outcomes in childhood are complex, heterogeneous, and result from highly interactive relationships between biological, environmental, and social factors, the mechanisms of which are often poorly understood. This real-world complexity, which may be difficult to model, particularly with more traditional statistical methods, may have to date acted as a deterrent for researchers investigating risk prediction tools for this purpose. However, there is now increasing potential to statistically model such complex interactions with machine learning. With the huge recent advances in both the quantity and quality of data available through large birth cohorts, electronic health records, imaging data, genomics, video, medical imaging, and electronic devices, the application of AI and machine learning may assist in finding the optimal predictive patterns to enable early interventions.[49]

The 'exposome' is a term used for the totality of life-course environmental exposures from the prenatal period onwards.[50] Adopting this lens in evaluating the role of environmental factors may enable a more complete understanding of the exposures involved in childhood cognitive development.[51,52] Prospective birth cohorts and linked electronic medical records provide a unique opportunity for evaluating such exposures as they can capture a wealth of valuable information on prenatal, perinatal, and early childhood exposures. Following an individual from the intrauterine period or from birth enables a longitudinal assessment of early life exposures and later life outcomes.[53] There has been an increase in both the number and the size of birth cohort studies in recent years, and in Europe alone, more than 110 birth cohorts representing 27 different countries have been formed.[54]

In addition, the capacity to collect large volumes of data, biobank serum samples, perform genomics, and use technology such as wearable devices has resulted in an explosion in the quantity, complexity, and depth of data now collected in such cohorts, providing a huge opportunity to utilise such data to advance both our understanding of the aetiology of disease, and our ability to predict outcomes. To truly harness the value of these cohorts and data sources, methods of data linkage and harmonisation require further research. Through amalgamating data sources, large data repositories could be created that would be particularly valuable for all fields of research, but especially for AI. Further work is needed in this area to address the issues around data protection, consent, and data sharing.

Generally, two classes of the statistical model are recognised—those where the aim is to understand the causal pathways leading to a certain outcome, and those where the aim is to predict as accurately as possible a certain outcome.[55] These two classes have much in common, and a tangled, but often constructive relationship.[56] Statistical models such as those intending to investigate the causal pathways between early life determinants and cognitive function in childhood estimate the effect size between a preceding factor and the outcome, adjusted for the confounding factors in this relationship, based on a formal modelling process.[57] The inclusion of additional variables in the model is based on previous evidence, changes in the size of the effect between the causative factor and the outcome, and goodness of fit.[56,57]

Prediction models however combine multiple predictors based on their ability to predict the occurrence of a particular event in the future (prognostic model) or their ability to predict the presence of a particular health state (diagnostic model).[58] Prediction model studies aim to develop and validate multivariable prediction models that combine individual predictors to predict an individual's probability of either having or developing a particular outcome.[59]

Published prediction models examining the prediction of cognitive outcomes among general paediatric populations are scarce compared to the published literature on prediction models for high-risk groups such as pre-term infants.[57] A summary of three recently published multivariable models predicting cognitive outcomes in general paediatric populations is shown in Table 1.[20–22]

A major limitation in the traditional statistical regression methods used in these models is their reliance on strong assumptions, many of which are unrealistic in complex real-world data. Traditional regression methods are often used in ways that assume that the effect of a predictor on the outcome increases or decreases uniformly throughout the range of the prediction (assumption of linearity); however, in truth, many relationships are non-linear. It is often assumed that the residuals (the difference between the observed value of the dependent variable and the model predicted value of the dependent variable) have a constant variance at each level of the predictor (assumption of homoscedasticity) and that the residuals have a normal distribution.[60] Methods exist to deal with all of these issues, but may not be properly or commonly used. Further limitations are the constraints around interaction terms in traditional statistical models. Traditional models assume predictors affect the outcome in an additive manner unless otherwise specified. Interaction terms must be determined a priori and entered into the model, a process not commonly performed despite a wealth of evidence demonstrating strong interactions between many of the predictors in the models, such as that between prematurity and maternal education, or between pre-term brain injury and socioeconomic status.[61–63]

There is considerable debate regarding the true differences between traditional statistical approaches and machine learning, often compounded by the fact that many machine learning techniques, such as linear and logistic regression, are drawn from

traditional statistical methods.[64] While traditional statistical models and machine learning models may use similar means to reach their goal, the key difference tends to lie between the objectives of the models. Traditional statistical methods, while they can be used for prediction, tend to focus on inferring relationships between variables, while following a defined set of assumptions and rules based on the data to which the model is fitted.[65–67] The purpose of ML is to make a repeatable prediction by learning from patterns within the data, without prior assumptions or rules governing the process.[66,67] ML does not make inferences regarding the causal relationship between features and outcomes. Using designated training data, with a defined set of potential predictors and a defined outcome, an ML algorithm can learn a set of rules from the data which can then be applied to the novel, unseen data to make a prediction on the outcome.[68] ML approaches can handle enormous quantities of data, and also can combine multiple different forms of data from multiple different sources including epidemiological data, imaging data, environmental and genomic data. ML techniques can efficiently deal with missing data, and driven by the data and generalisation performance, are adept at handling feature redundancy (collinearity) and feature selection.[68]

A major advantage of machine learning is the ability to handle large numbers of highly interactive predictors and to combine them in non-linear relationships. With both the quantity and diversity of data growing, non-linear and interactive relationships within data are increasingly important and may be the key to improving predictive accuracy. ML techniques can, to a point, automatically uncover such interactive relationships in the data without the need for a priori specification of interactions.[69,70] For the purposes of predicting cognitive development, this is likely to be important. It is often the case that the effect of one predictor will change given the value of another predictor. For example, the combined effect of adverse socioenvironmental factors and pre-term brain injury results in significantly lower childhood IQ than either alone.[63]

It should not be assumed, however, that ML techniques will always result in improved predictive performance compared to traditional methods.[71] This is likely to depend on the outcome being predicted, the number of features involved in training the prediction, the type and variability of the features, and the presence of interactions between features. The optimal predictive algorithm will depend on the predictive problem, and prior to moving to more complex predictive models, traditional methods with logistic and linear regression modelling should always be trialled and reported for comparison.[71]

APPLICATIONS OF MACHINE LEARNING IN NEURODEVELOPMENTAL PAEDIATRICS

The past decade has seen a rapid expansion in the application of ML and AI, for the purposes of diagnosis and prognosis, in clinical medicine.[68] Significant progress has been made specifically in the areas of radiology and imaging, pathology, and cancer.[72] There is growing interest in the application of machine learning in the field of paediatric neurodevelopmental disorders (NDDs), a heterogeneous group of disorders impacting different domains of function including cognition, language, social, motor and behavioural, but progress has been slower.[73] The term NDD includes conditions such as intellectual disability, ASD, and attention deficit hyperactivity disorder (ADHD).[73]

Diverse data, ranging from MRI data to motor skills data collected via drag-and-drop gaming devices, have been combined with machine learning techniques for both diagnostic and prognostic purposes in the area of NDDs. Examples are provided in Table 2. Accurate, objective, diagnostic tools could be a particularly valuable aid to clinicians in this field, where many of the conditions currently rely on subjective, labour intensive, expert

Table 1. Study characteristics of multivariable prediction models for cognitive outcomes in childhood.

	Camargo-Figuera et al.[20]	Camacho et al.[21]	Eriksen et al.[22]
Authors	Camargo-Figuera et al.[20]	Camacho et al.[21]	Eriksen et al.[22]
Journal	*BMC Pediatrics*	*Paediatric Perinatal Epidemiology*	*PLoS One*
Continent	South America	Europe	Europe
Sample	Pelotas Birth Cohort	Millennium Cohort Study	Lifestyle During Pregnancy Study (sampled from the Danish National Birth Cohort)
Design	Prospective cohort	Prospective cohort	Prospective cohort
Sample size	3312	9487	1782
Year of recruitment	2004	2000–2002	1997–2003[a]
Exclusion criteria	Conditions associated with very low IQ, e.g., severe mental retardation	Nil	Multiple pregnancies, language barrier, impaired hearing or vision, congenital disabilities implying mental retardation
Age at cognitive assessment	6	3	5
Cognitive assessment	Wechsler Primary and Preschool Scales of Intelligence -III	Bracken School Readiness Assessment	Wechsler Primary and Preschool Scales of Intelligence – Revised
Cognitive outcome variable	Binary	Binary	Continuous
	Low IQ defined by a *z*-score <-1	Not school ready defined by score <1 standard deviation below mean	
Number of risk factors at outset	32	29	27
Rationale given for candidate variables	Yes—selected based on previous literature and availability	Yes—selected based on previous literature and availability	No–but broad range (>20) selected
Statistical model	Multivariable logistic regression	Multivariable logistic regression	Multivariable linear regression
Method of initial screening of candidate variables	Forward and backward stepwise selection	Forward and backward stepwise selection	Univariable association $p \geq 0.10$
Interaction terms fitted	No	No	No
Multicollinearity addressed/discussed	No	Yes	Yes
No. predictors in final model	13	13	9
Validation performed	Yes	Yes	No
	Internal and external validation performed	Internal validation only	
Predictive value measured	External validation	Internal validation	*R* squared 0.29
	Area under receiver operating curve (AUROC) 0.75	AUROC 0.80	
	Sensitivity 70.3%	Sensitivity 72%	
	Specificity 68%	Specificity 74%	
Predictors in final model	Child—gender, height-for-age deficit; head circumference-for-age deficit	Child—gender, ethnicity, developmental milestones	Child—gender, birth weight, height, head circumference
	Parental—breastfeeding, parental smoking, maternal perception of child's health, skin colour	Parental—maternal age, maternal mental health, breastfeeding	Parental—maternal BMI, breastfeeding
	Socioenvironmental—parental employment status, maternal education, income, number of siblings, number of persons per room	Socioenvironmental—socioeconomic class, maternal education, income, number of children, employment status, housing type	Socioenvironmental—maternal IQ, parental education

[a]In 2003, a prospective follow-up of 1750 mother–child pairs was initiated, sampled on the basis of maternal alcohol drinking patterns from The Danish National Birth Cohort (DNBC).

clinical assessments. In many countries, this results in long waiting lists for assessment in developmental paediatric clinics and delays in diagnosis for those affected.

One example is ADHD, the diagnosis of which relies upon a detailed clinical assessment based upon observed behaviour and reported symptoms.[74] Pupillometric data (data based on dilation of the pupil) from 50 participants aged 10–12 years used in a support vector machine (SVM) algorithm achieved a sensitivity of 0.773 and a specificity of 0.753 in the detection of ADHD, based on the results of nested 10-fold cross-validation.[75] This study however

Table 2. Examples of machine learning methods applied in neurodevelopmental paediatrics.

	Das and Khanna[75]	Sen et al.[76]	Emerson et al.[79]	Girault et al.[80]	Kim et al.[81]
Year published	2019	2018	2017	2019	2021
Purpose	Diagnostic	Diagnostic	Prognostic	Prognostic	Diagnostic
Outcome	ADHD	ADHD	ASD at 24 months of age	Cognitive ability at age 2 years	Developmental disability
Outcome measure	Diagnosed by child neurologist using DSM-IV criteria	Diagnostic criteria differed across study sites	DSM-IV criteria	Mullens Scales of Early Ability. Outcome categorised as above or below the median	Diagnosis made by paediatric psychologist
Features	Pupillometric data	Structural and functional MRI	MRI at 6 months of age	MRI at birth	Drag and drop data from computer-based game
Participants	50 (28 cases, 22 controls)	558 (279 cases, 279 controls)	59 infants with high familial risk of ASD (11 cases, 48 controls)	115 infants	370 children (147 cases, 233 controls)
Algorithms trained	LR, SVM, KNN, NB, DT, RF	SVM	SVM	DL	DL
Evaluation method	Nested 10-fold cross-validation	Independent test set (77 cases, 94 controls)	Cross-validation with nested leave one out procedure	Independent test set (22 cases, 15 controls)	10-fold cross-validation
Evaluation result (best performing models)	NB: AUROC 0.871, sensitivity 0.656, specificity 0.863	Accuracy 0.6725, sensitivity 0.4545, specificity 0.8510	Accuracy 0.966, sensitivity 0.818, specificity 1.000	Accuracy 0.838, sensitivity 0.864, specificity 0.800	AUROC 0.817, sensitivity 0.757, specificity 0.740

LR logistic regression, *SVM* support vector machine, *KNN* K-nearest neighbours, *NB* naïve Bayes, *DT* decision tree, *RF* random forest, *ADHD* attention deficit hyperactivity disorder, *ASD* autism spectrum disorder, *DSM-IV* diagnostic and statistics manual of mental disorders.

is limited by the small sample size of 50, and the absence of model evaluation on an independent test set raises questions over the generalisability of their model to wider populations. This is a limitation seen in many of the studies in this field.

MRI data has also been used in ML algorithms for the purpose of diagnosing ADHD. Sen et al. trained an SVM using structural and functional MRI data from 558 participants. The final model, which produced an accuracy of 0.673, sensitivity of 0.454, and specificity of 0.851 on an independent test set of 171 participants, is unlikely to be clinically useful in its current form but suggests there may be signals within such data that could assist in the diagnosis of ADHD.[76]

For ASD, another NDD that benefits from early intervention but for which diagnosis is often delayed, ML has been increasingly studied as a potential tool for early detection.[77] A 2019 systematic review and meta-analysis examining the accuracy of ML algorithms based on brain MRI data for detecting ASD reported a pooled area under the receiver operating curve (AUROC) of 0.90, a sensitivity of 0.83, and a specificity of 0.84 for ML models trained with structural MRI data. The results should however be interpreted with caution, as the majority of included studies performed internal validation only, and did not externally validate their models in independent test sets.[78]

While the examples outlined thus far have largely focussed on ML for diagnostic purposes, ML methods have also been deployed for the prediction of later NDDs.[79,80] A deep learning model, trained on white matter connectomes from brain MRI data obtained at the time of birth, and tested on an independent test set, correctly classified 84% of infants as having below or above the median level of cognitive ability at 2 years of age.[80] The results, which are based on a relatively small sample size of 37 in the independent test set, are promising but the clinical utility at a population level is likely limited.

Novel methods of neurodevelopmental assessment, outside of traditional medical investigations, are also being analysed using ML algorithms. For example, a deep learning model trained on data from a touch-based computer game aimed to harness this measure of motor skill to detect developmental disabilities in children. An AUROC of 0.817 based on internal cross-validation provides a basis for further research in this area.[81]

MOVING FROM CLINICAL TO POPULATION-BASED PREDICTION

Acknowledging the limitations of these studies, overall the results indicate that AI and machine learning may have an important role in the early prediction of certain NDDs. However, many of the prediction studies in this field to date are based on data, such as MRI and EEG, which is not routinely collected at birth at a population level. Therefore, in clinical practice, these studies may provide methods of prediction for higher risk populations who undergo these investigations, but are likely to be largely unsuitable for implementation at a population-based level due to resource constraints and cost. At a population level, machine learning would appear to be more advantageous when used with data that is readily available, cost effective to collect, and available early in life. This however may come at the cost of reduced predictive ability compared with models based on direct physiological and anatomical measures.

In the context of population health, the application of machine learning to address population health issues and mitigate health disparities has received much less attention than its clinical applications. Prediction models at a population level tend to use existing electronic medical records, large cohort studies, environmental databases, government databases, and internet-based data.[82] Population-based data sources have the potential to incorporate social, environmental, and other structural determinants in prediction models, potentially improving accuracy.[83]

Social determinants of health have long been studied in epidemiological research, but the findings of these studies emphasise the need for methods that are better able to capture the complex, interactive, and flexible relationships between these predictors and health outcomes.[83]

The majority of well-established predictors of cognitive function in childhood, including poverty, gestational age, drug, alcohol and tobacco use during pregnancy, birth weight, admission to the neonatal intensive care unit, parental education and socioeconomic status, are predictors that are readily available in the perinatal period. To date, the accuracy of algorithms applying traditional statistical methods to such data for the purpose of identifying children at risk of low cognitive function in childhood has shown mixed results with regard to their accuracy (Table 1). The question of how we can best harness this information for identifying high-risk children remains. Application of machine learning methods to large population-based datasets, using readily available, routinely collected information may assist in this objective but remains relatively unexplored in the literature.

RISKS OF INDIVIDUAL RISK PREDICTIONS AND MACHINE LEARNING ALGORITHMS

There are undoubtedly benefits to the early identification of children at risk of poor cognitive outcomes, namely providing early interventions during the period of optimal neuroplasticity and potentially altering the cognitive trajectory of the individual child. However, this approach is not without risk.

Early identification of those at risk is not synonymous with labelling, yet even the use of terms such as 'at risk' or 'higher risk' could have detrimental effects on both the child and their family. Labelling a child at a very early stage of life, as being at risk of a poor cognitive outcome, can perpetuate a self-fulfilling prophecy due to well-documented effects on parent and teacher expectations, as well as the potentially detrimental effect on a child's own self-concept.[84–86] Furthermore, using a label based largely on social factors, which are transmitted through generations and rooted in inequities generated by social and political policy, may shift the blame of societal failures onto the individual child or family. These risks are especially pertinent for children and families who may be incorrectly identified as at risk on a screening or risk prediction tool and who are unlikely to benefit from early intervention. For those correctly identified, the question is whether these risks are outweighed by the risk of overlooking a child until their cognitive difficulties manifest in overt signs such as academic failures, a point at which intervention may be too late.[87]

There are also broader ethical considerations with the use of machine learning algorithms in healthcare in general.[88–90] There is a growing body of literature examining algorithmic fairness, a topic warranting urgent attention following examples of ML algorithms exacerbating health inequities, raising serious ethical concerns.[91] Each step of development of an ML algorithm is fraught with potential for bias—from the initial decision to use ML for the specific problem, to how it is deployed in practice.[92] A major threat to algorithmic fairness is the inherent bias in the data on which many of these models are built. For example, healthcare utilisation data represents those who attend for healthcare, but not necessarily those who need it; cohort data tend not to represent marginalised populations such as the homeless who may be unable to participate in longitudinal follow-up. A further threat is the use of 'black box' models that lack transparency, cannot be interpreted by the user, and therefore have inherent risks for producing bias. There are opportunities to harness the power of ML in addressing population health challenges, but stringent assessment of potential for bias, quality, and applicability are required prior to dissemination in practice.[93]

CONCLUSIONS AND FUTURE RESEARCH

There is an optimal window in early life during which the trajectory of cognitive development is more amenable to change.[94] Infants at risk of poor cognitive outcomes due to prenatal, birth-related, or socioenvironmental risk factors, whom we fail to identify and provide with appropriate interventions, may never meet their true cognitive potential. For children who begin their educational journey lacking basic cognitive skills, a downward spiral of repeated failure, grade retention, and behavioural difficulties may culminate in school dropout, and its associated future implications that include unemployment and substance abuse.[95] Equally, it is the case that these children at higher risk are likely to live in poorer households, in more deprived areas, and be members of marginalised groups, so population-level interventions including educational policy, planning policy, and economic policy, also have a central role to play.

It is estimated that less than half of children with developmental delay are detected prior to the school entrance, with the vast majority of those detected receiving no intervention in the very early years.[96,97] It is abundantly clear that the current practice of waiting for delays to present prior to the intervention is inadequate and goes against the large body of evidence highlighting the optimal period of neuroplasticity in early life.[98] The risk factors for poor cognitive outcomes in childhood are well established in the literature, are amenable to identification in the perinatal period, and are often routinely collected. Yet, to date, we have been largely unable to harness this information in a personalised and meaningful way to identify the high-risk mother–infant dyads and provide appropriate interventions.

Our improved capacity to collect large volumes of rich data and to apply novel statistical methods particularly suited to prediction provides an opportunity for researchers and clinicians to investigate alternative methods of identifying, at an earlier stage, children at risk of developmental and cognitive delay. Utilising, to the best of their predictive ability, currently available data in the form of birth cohorts, electronic health records, and population registries should be the first step.

Addressing health and social inequity is an age-old problem crying out for novel solutions. In this review, we have focussed primarily on the domain of cognitive development. This is not to put undue emphasis on cognitive ability as a measure of success, nor to endorse a meritocratic society where a child's emotional, social, or creative skills are viewed as less important. The purpose is to strive towards creating a society whereby every child is given the opportunity to meet their cognitive potential, irrespective of the circumstances into which they are born. To achieve this, high-risk infants and their families should be supported with appropriate, evidence-based interventions at the earliest possible stage, prior to the onset of cognitive delay. This requires early, accurate identification of those who will benefit. Combining the strengths of machine learning, big data analysis, and paediatric public health for a potential solution deserves thorough exploration for the individual, community, economic, and societal benefits that could result.

REFERENCES
1. Moore, T. G., Arefadib, N., Deery, A. & West, S. *The First Thousand Days: An Evidence Paper* (Parkville, Victoria, 2017).
2. Barker, D. et al. Fetal nutrition and cardiovascular disease in adult life. *Lancet* **341**, 938–941 (1993).
3. Barker, D. & Osmond, C. Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. *Lancet* **1**, 1077–1081 (1986).
4. Barker, D. et al. Weight in infancy and death from ischaemic heart disease. *Lancet* **9**, 577–580 (1989).
5. Wadhwa, P. D., Buss, C., Entringer, S. & Swanson, J. M. Developmental origins of health and disease: brief history of the approach and current focus on epigenetic mechanisms. *Semin. Reprod. Med.* **27**, 358–368 (2009).

6. Barker, D. Developmental origins of adult health and disease. *J. Epidemiol. Community Health* **58**, LP–115 (2004).

7. Fu, M. & Zuo, Y. Experience-dependent structural plasticity in the cortex. *Trends Neurosci.* **34**, 177–187 (2011).

8. Cioni, G., Inguaggiato, E. & Sgandurra, G. Early intervention in neurodevelopmental disorders: underlying neural mechanisms. *Dev. Med. Child Neurol.* **58**, 61–66 (2016).

9. Spencer, N., Raman, S., O'Hare, B. & Tamburlini, G. Addressing inequities in child health and development: towards social justice. *BMJ Pediatr. Open* **3**, e000503–e000503 (2019).

10. American Psychological Association. *APA Dictionary of Psychology* (2020).

11. Nouchi, R. & Kawashima, R. Improving cognitive function from children to old age: a systematic review of recent smart ageing intervention studies. *Adv. Neurosci.* **2014**, 235479 (2014).

12. National Research Council (US) Panel to Review the Status of Basic Research on School-Age Children. *Development During Middle Childhood: The Years From Six to Twelve* (National Academies Press (US), Washington (DC), 1984).

13. Ganuthula, V. R. R. & Sinha, S. The looking glass for intelligence quotient tests: the interplay of motivation, cognitive functioning, and affect. *Front. Psychol.* **10**, 2857 (2019).

14. Drigas, A. S. & Papoutsi, C. A new layered model on emotional intelligence. *Behav. Sci. (Basel).* **8**, 45, https://doi.org/10.3390/bs8050045. (2018). PMID: 29724021; PMCID: PMC5981239.

15. Lager, A., Bremberg, S. & Vågerö, D. The association of early IQ and education with mortality: 65 year longitudinal study in Malmö, Sweden. *BMJ (Clin. Res. Ed.)* **339**, b5282 (2009).

16. Forrest, L. F., Hodgson, S., Parker, L. & Pearce, M. S. The influence of childhood IQ and education on social mobility in the Newcastle Thousand Families birth cohort. *BMC Public Health* **11**, 895 (2011).

17. Furnham, A. & Cheng H. Childhood cognitive ability predicts adult financial well-being. *J. Intell..* **5**, 3 https://doi.org/10.3390/jintelligence5010003 (2017).

18. Whalley, L. J. & Deary, I. J. Longitudinal cohort study of childhood IQ and survival up to age 76. *BMJ (Clin. Res. Ed.)* **322**, 819 (2001).

19. Tong, S., Baghurst, P., Vimpani, G. & McMichael, A. Socioeconomic position, maternal IQ, home environment, and cognitive development. *J. Pediatr.* **151**, 288.e1 (2007).

20. Camargo-Figuera, F. A., Barros, A. J. D., Santos, I. S., Matijasevich, A. & Barros, F. C. Early life determinants of low IQ at age 6 in children from the 2004 Pelotas Birth Cohort: a predictive approach. *BMC Pediatr.* **14**, 308 (2014).

21. Camacho, C., Straatmann, V. S., Day, J. C. & Taylor-Robinson, D. Development of a predictive risk model for school readiness at age 3 years using the UK Millennium Cohort Study. *BMJ Open* **9**, e024851 (2019).

22. Eriksen, H. L. F. et al. Predictors of intelligence at the age of 5: family, pregnancy and birth characteristics, postnatal influences, and postnatal growth. *PLoS One* **8**, e79200 (2013).

23. Lawlor, D. A. et al. Early life predictors of childhood intelligence: findings from the Mater-University study of pregnancy and its outcomes. *Pediatr. Perinat. Epidemiol.* **20**, 148–162 (2006).

24. von Stumm, S. & Plomin, R. Socioeconomic status and the growth of intelligence from infancy through adolescence. *Intelligence* **48**, 30–36 (2015).

25. Schoon, I., Jones, E., Cheng, H. & Maughan, B. Family hardship, family instability, and cognitive development. *J. Epidemiol. Community Health* **66**, 716–722 (2012).

26. Bugental, D. B., Corpuz, R. & Schwartz, A. Preventing children's aggression: outcomes of an early intervention. *Dev. Psychol.* **48**, 1443–1449 (2012).

27. Vinen, Z., Clark, M., Paynter, J. & Dissanayake, C. School age outcomes of children with autism spectrum disorder who received community-based early interventions. *J. Autism Dev. Disord.* **48**, 1673–1683 (2018).

28. Yoshinaga-Itano, C., Sedey, A. L., Mason, C. A., Wiggin, M. & Chung, W. Early intervention, parent talk, and pragmatic language in children with hearing loss. *Pediatrics* **146**, S270–S277 (2020).

29. Freitag, H. & Tuxhorn, I. Cognitive function in preschool children after epilepsy surgery: rationale for early intervention. *Epilepsia* **46**, 561–567 (2005).

30. Ramey, C. T. & Ramey, S. L. Prevention of intellectual disabilities: early interventions to improve cognitive development. *Preventive Med.* **27**, 224–232 (1998).

31. Gillette, Y. Family-centered early intervention: an opportunity for creative practice in speech-language pathology. *Clin. Commun. Disord.* **2**, 48–60 (1992).

32. Pungello, E. P. et al. Early educational intervention, early cumulative risk, and the early home environment as predictors of young adult outcomes within a high-risk sample. *Child Dev.* **81**, 410–426 (2010).

33. Watanabe, K., Flores, R., Fujiwara, J. & Tran, L. T. H. Early childhood development interventions and cognitive development of young children in rural Vietnam. *J. Nutr.* **135**, 1918–1925 (2005).

34. Masse, L. & Barnett, S. A benefit cost analysis of the abecedarian early childhood intervention (2002). New Brunswick NJ. National Institute for Early Education Research., Rutgers University.[cited Feb 3rd 2022]. Available at https://nieer.org/wp-content/uploads/2002/11/AbecedarianStudy.pdf

35. Campbell, F. A. & Ramey, C. T. Effects of early intervention on intellectual and academic achievement: a follow-up study of children from low-income families. *Child Dev.* **65**, 684–698 (1994).

36. Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J. & Miller-Johnson, S. Early childhood education: young adult outcomes from the Abecedarian Project. *Appl. Dev. Sci.* **6**, 42–57 (2002).

37. Sejer, E. P. F., Bruun, F. J., Slavensky, J. A., Mortensen, E. L. & Schiøler Kesmodel, U. Impact of gestational age on child intelligence, attention and executive function at age 5: a cohort study. *BMJ Open* **9**, e028982 https://doi.org/10.1136/bmjopen-2019-028982 (2019).

38. Turpin, H. et al. The interplay between prematurity, maternal stress and children's intelligence quotient at age 11: a longitudinal study. *Sci. Rep.* **9**, 450 (2019).

39. Spittle, A. J., Orton, J., Doyle, L. W. & Boyd, R. Early developmental intervention programs post hospital discharge to prevent motor and cognitive impairments in preterm infants. *Cochrane Database Syst. Rev.* CD005495 (2007).

40. Whitehouse, A. J. O. et al. Effect of preemptive intervention on developmental outcomes among infants showing early signs of autism: a randomized clinical trial of outcomes to diagnosis. *JAMA Pediatr.* **175**, e213298–e213298 (2021).

41. Love, J. M., Chazan-Cohen, R., Raikes, H. & Brooks-Gunn, J. What makes a difference: Early Head Start evaluation findings in a developmental context. *Monogr. Soc. Res. Child Dev.* **78**, 1–173 (2013).

42. Morris, P. A. et al. New findings on impact variation from the head start impact study: informing the scale-up of early childhood programs. *AERA Open* **4**, 2332858418769287 (2018).

43. Kerstjens, J. M. et al. Support for the global feasibility of the Ages and Stages Questionnaire as developmental screener. *Early Hum. Dev.* **85**, 443–447 (2009).

44. Hirai, A. H., Kogan, M. D., Kandasamy, V., Reuland, C. & Bethell, C. Prevalence and variation of developmental screening and surveillance in early childhood. *JAMA Pediatr.* **172**, 857–866 (2018).

45. Wolf, E. R. et al. Gaps in well-child care attendance among primary care clinics serving low-income families. *Pediatrics* **142**, e20174019 (2018).

46. Edwards, K. et al. Improving access to early childhood developmental surveillance for Children from Culturally and Linguistically Diverse (CALD) Background. *Int. J. Integr. Care* **20**, 3 (2020).

47. Flensborg-Madsen, T., Falgreen Eriksen, H.-L. & Mortensen, E. L. Early life predictors of intelligence in young adulthood and middle age. *PLoS One* **15**, e0228144 (2020).

48. Bradley, R. H. & Corwyn, R. F. Socioeconomic status and child development. *Annu. Rev. Psychol.* **53**, 371–399 (2002).

49. Dwyer, D. & Koutsouleris, N. Annual research review: translational machine learning for child and adolescent psychiatry. *J. Child Psychol. Psychiatry* **63**, 421–433 https://doi.org/10.1111/jcpp.13545 (2022).

50. Santos, S. et al. Applying the exposome concept in birth cohort research: a review of statistical approaches. *Eur. J. Epidemiol.* **35**, 193–204 (2020).

51. Vrijheid, M. et al. Early-life environmental exposures and childhood obesity: an exposome-wide approach. *Environ. Health Perspect.* **128**, 1–14. (2020).

52. Steer, C. D., Bolton, P. & Golding, J. Preconception and prenatal environmental factors associated with communication impairments in 9 Year old children using an exposome-wide approach. *PLoS One* **10**, e0118701 https://doi.org/10.1371/journal.pone.0118701 (2015).

53. Canova, C. & Cantarutti, A. Population-based birth cohort studies in epidemiology. *Int. J. Environ. Res. Public Health* **17**, 5276 (2020).

54. Pansieri, C., Pandolfini, C., Clavenna, A., Choonara, I. & Bonati, M. An inventory of European birth cohorts. *Int. J. Environ. Res. Public Health* **17**, 3071 (2020).

55. Ramspek, C. L. et al. Prediction or causality? A scoping review of their conflation within current observational research. *Eur. J. Epidemiol.* **36**, 889–898 (2021).

56. Blakely, T., Lynch, J., Simons, K., Bentley, R. & Rose, S. Reflection on modern methods: when worlds collide-prediction, machine learning and causal inference. *Int. J. Epidemiol.* **49**, 2058–2064 (2021).

57. Linsell, L., Malouf, R., Morris, J., Kurinczuk, J. J. & Marlow, N. Risk factor models for neurodevelopmental outcomes in children born very preterm or with very low birth weight: a systematic review of methodology and reporting. *Am. J. Epidemiol.* **185**, 601–612 (2017).

58. Wolff, R. F. et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann. Intern. Med.* **170**, 51–58 (2019).

59. Steyerberg, E. W. et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med.* **10**, e1001381 (2013).

60. Barker, L. E. & Shaw, K. M. Best (but oft-forgotten) practices: checking assumptions concerning regression residuals. *Am. J. Clin. Nutr.* **102**, 533–539 (2015).

61. Benavente-Fernández, I., Siddiqi, A. & Miller, S. P. Socioeconomic status and brain injury in children born preterm: modifying neurodevelopmental outcome. *Pediatr. Res.* **87**, 391–398 (2020).

62. Patra, K., Greene, M. M., Patel, A. L. & Meier, P. Maternal education level predicts cognitive, language, and motor outcome in preterm infants in the second year of life. *Am. J. Perinatol.* **33**, 738–744 (2016).

63. Benavente-Fernández, I. et al. Association of socioeconomic status and brain injury with neurodevelopmental outcomes of very preterm children. *JAMA Netw. Open* **2**, e192914 (2019).

64. Bzdok, D., Altman, N. & Krzywinski, M. Statistics versus machine learning. *Nat. Methods* **15**, 233–234 (2018).

65. Azzolina, D. et al. Machine learning in clinical and epidemiological research: isn't it time for biostatisticians to work on it? *Epidemiol. Biostat. Public Health* **16** https://doi.org/10.2427/13245 (2019).

66. Bzdok, D., Altman, N. & Krzywinski, M. Statistics versus machine learning. *Nat. Methods* **15**, 233–234 (2018).

67. Rajula, H. S. R., Verlato, G., Manchia, M., Antonucci, N. & Fanos, V. Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment. *Medicine (Kaunas)* **56**, 455 (2020).

68. Obermeyer, Z. & Emanuel, E. J. Predicting the future—big data, machine learning, and clinical medicine. *N. Engl. J. Med.* **375**, 1216–1219 (2016).

69. Touw, W. G. et al. Data mining in the Life Sciences with Random Forest: a walk in the park or lost in the jungle? *Brief. Bioinforma.* **14**, 315–326 (2013).

70. Mooney, S. J., Westreich, D. J. & El-Sayed, A. M. Commentary: epidemiology in the era of big data. *Epidemiology* **26**, 390–394 (2015).

71. Christodoulou, E. et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J. Clin. Epidemiol.* **110**, 12–22 (2019).

72. Rajpurkar, P., Chen, E., Banerjee, O. & Topol, E. J. AI in health and medicine. *Nat. Med.* **28**, 31–38 https://doi.org/10.1038/s41591-021-01614-0 (2022).

73. Uddin, M., Wang, Y. & Woodbury-Smith, M. Artificial intelligence for precision medicine in neurodevelopmental disorders. *npj Digital Med.* **2**, 112 (2019).

74. National Institute for Health and Care Excellence (NICE). *Attention Deficit Hyperactivity Disorder: Diagnosis and Management* (NICE, London, accessed 12 Apr 2022; https://www.nice.org.uk/guidance/ng87/chapter/Recommendations#diagnosis (2018).

75. Das, W. & Khanna, S. A robust machine learning based framework for the automated detection of ADHD using pupillometric biomarkers and time series analysis. *Sci. Rep.* **11**, 16370 (2021).

76. Sen, B., Borle, N. C., Greiner, R. & Brown, M. R. G. A general prediction model for the detection of ADHD and autism using structural and functional MRI. *PLoS One* **13**, e0194856 (2018).

77. Zwaigenbaum, L. & Penner, M. Autism spectrum disorder: advances in diagnosis and evaluation. *BMJ* **361**, k1674 (2018).

78. Moon, S. J., Hwang, J., Kana, R., Torous, J. & Kim, J. W. Accuracy of machine learning algorithms for the diagnosis of autism spectrum disorder: systematic review and meta-analysis of brain magnetic resonance imaging studies. *JMIR Ment. Health* **6**, e14108 (2019).

79. Emerson, R. W. et al. Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age. *Sci. Transl. Med* **9**, eaag2882 (2017).

80. Girault, J. B. et al. White matter microstructural development and cognitive ability in the first 2 years of life. *Hum. Brain Mapp.* **40**, 1195–1210 (2019).

81. Kim, H. H., An, J. & il, Park, Y. R. A prediction model for detecting developmental disabilities in preschool-age children through digital biomarker-driven deep learning in serious games: development study. *JMIR Serious Games* **9**, e23130 (2021).

82. Morgenstern, J. D. et al. Predicting population health with machine learning: a scoping review. *BMJ Open* **10**, e037860 (2020).

83. Mhsawade, V., Zhao, Y. & Chunara, R. Machine learning and algorithmic fairness in public and population health. *Nat. Mach. Intell.* **3**, 659–666 (2021).

84. Jussim, L. & Harber, K. D. Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies. *Personal. Soc. Psychol. Rev.* **9**, 131–155 (2005).

85. Shifrer, D. Stigma of a label: educational expectations for high school students labeled with learning disabilities. *J. Health Soc. Behav.* **54**, 462–480 (2013).

86. Taylor, L. M., Hume, I. R. & Welsh, N. Labelling and self-esteem: the impact of using specific vs. generic labels. *Educ. Psychol.* **30**, 191–202 (2010).

87. Leigh, J. E. Early labelling of children: Concerns and alternatives. *Top. Early Child. Spec. Educ.* **3**, 1–6 (1983).

88. Char, D. S., Shah, N. H. & Magnus, D. Implementing machine learning in health care - addressing ethical challenges. *N. Engl. J. Med* **378**, 981–983 (2018).

89. McCradden, M. D., Joshi, S., Mazwi, M. & Anderson, J. A. Ethical limitations of algorithmic fairness solutions in health care machine learning. *Lancet Digit Health* **2**, e221–e223. (2020).

90. Mhsawade, V., Zhao, Y. & Chunara, R. Machine learning and algorithmic fairness in public and population health. *Nat. Mach. Intell.* **3**, 659–666 (2021).

91. Obermeyer, Z., Powers, B., Vogeli, C. & Mullainathan, S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* **366**, 447–453 (2019).

92. Striving for health equity with machine learning. *Nat. Mach. Intell.* 2021; **3**, 653 (2021).

93. de Hond, A. A. H. et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. *npj Digital Med.* **5**, 2 (2022).

94. Almond, D. & Currie, J. Human capital development before age five. In *Handbook of Labor Economics* (eds Card, D. & Ashenfelter OBT-H of LE) 1315–1486 (Elsevier, 2011).

95. Lansford, J. E., Dodge, K. A., Pettit, G. S. & Bates, J. E. A public health perspective on school dropout and adult outcomes: a prospective study of risk and protective factors from age 5 to 27 years. *J. Adolesc. Health.* **58**, 652–658 (2016).

96. Morelli, D. L. et al. Challenges to implementation of developmental screening in urban primary care: a mixed methods study. *BMC Pediatr.* **14**, 16 (2014).

97. Dearlove, J. & Kearney, D. How good is general practice developmental screening? *BMJ (Clin. Res. Ed.)* **300**, 1177–1180 (1990).

98. Klebanov, P. & Brooks-Gunn, J. Cumulative, human capital, and psychological risk in the context of early intervention: links with IQ at ages 3, 5, and 8. *Ann. N. Y. Acad. Sci.* **1094**, 63–82 (2006).

AUTHOR CONTRIBUTIONS

Substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data: A.K.B. and D.M.M. Drafting the article or revising it critically for important intellectual content: A.K.B., D.M.M., A.S. and G.L. Final approval of the version to be published: A.K.B., D.M.M., A.S. and G.L.

FUNDING

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Andrea K. Bowe.

Reprints and permission information is available at http://www.nature.com/reprints

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