



## REVIEW ARTICLE

# Predicting the outcomes of preterm neonates beyond the neonatal intensive care unit: What are we missing?

Colin J. Crilly<sup>1</sup>, Sebastien Haneuse<sup>2</sup> and Jonathan S. Litt<sup>3,4</sup>

Preterm infants are a population at high risk for mortality and adverse health outcomes. With recent improvements in survival to childhood, increasing attention is being paid to risk of long-term morbidity, specifically during childhood and young-adulthood. Although numerous tools for predicting the functional outcomes of preterm neonates have been developed in the past three decades, no studies have provided a comprehensive overview of these tools, along with their strengths and weaknesses. The purpose of this article is to provide an in-depth, narrative review of the current risk models available for predicting the functional outcomes of preterm neonates. A total of 32 studies describing 43 separate models were considered. We found that most studies used similar physiologic variables and standard regression techniques to develop models that primarily predict the risk of poor neurodevelopmental outcomes. With a recently expanded knowledge regarding the many factors that affect neurodevelopment and other important outcomes, as well as a better understanding of the limitations of traditional analytic methods, we argue that there is great room for improvement in creating risk prediction tools for preterm neonates. We also consider the ethical implications of utilizing these tools for clinical decision-making.

*Pediatric Research* (2021) 89:426–445; <https://doi.org/10.1038/s41390-020-0968-5>

**IMPACT:**

- Based on a literature review of risk prediction models for preterm neonates predicting functional outcomes, future models should aim for more consistent outcomes definitions, standardized assessment schedules and measurement tools, and consideration of risk beyond physiologic antecedents.
- Our review provides a comprehensive analysis and critique of risk prediction models developed for preterm neonates, specifically predicting functional outcomes instead of mortality, to reveal areas of improvement for future studies aiming to develop risk prediction tools for this population.
- To our knowledge, this is the first literature review and narrative analysis of risk prediction models for preterm neonates regarding their functional outcomes.

**INTRODUCTION**

Preterm infants have long been recognized as a population at high risk for mortality and adverse functional outcomes, including cerebral palsy and intellectual impairment.<sup>1</sup> As mortality rates for preterm neonates decline and more survive to childhood,<sup>2,3</sup> attention has increasingly turned towards measuring longer-term morbidities and related functional impairments during childhood and young-adulthood, as well as identifying risk factors related to these complications.<sup>4,5</sup> While child-specific characteristics, such as gestational age, birth weight, and sex, are well established as predictors of adverse neurodevelopmental outcomes,<sup>6–8</sup> recent work has identified additional factors, including bronchopulmonary dysplasia and family socioeconomic status, that are correlated with relevant outcomes, such as poor neuromotor performance and low intelligence quotient at school age.<sup>9</sup>

In clinical settings, the assessment of prognosis can vary widely across neonatologists,<sup>10</sup> making a valid and reliable predictive model for long-term outcomes a highly sought-after clinical tool. Moreover, predicting outcomes is vital when making decisions

regarding which therapeutic interventions to apply, when providing critical data to parents for informed decision-making, and when matching infants with outpatient services to best meet their needs. In addition, prediction models are useful in evaluating Neonatal Intensive Care Unit (NICU) performance and allowing for between-center comparisons with proper adjustment for the severity of cases being treated.<sup>11</sup>

Numerous prediction tools have been developed to quantify the risk of death for preterm neonates in the NICU setting, including the Score for Neonatal Acute Physiology (SNAP) and the Clinical Risk Index for Babies (CRIB).<sup>12</sup> The National Institute of Child Health and Human Development (NICHD) risk calculator, predicting survival with and without neurosensory impairment, is widely used to counsel families in the setting of threatened delivery at the edges of viability.<sup>13</sup> Furthermore, there are numerous other models that use clinical data from the NICU stay to predict risk for poor functional outcomes in infancy and school age.<sup>14,15</sup> While several studies have categorized and evaluated the risk prediction models developed and validated in recent decades

<sup>1</sup>Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, USA; <sup>2</sup>Department of Biostatistics, Harvard TH Chan School of Public Health, Boston, MA, USA; <sup>3</sup>Department of Pediatrics, Harvard Medical School, Boston, MA, USA and <sup>4</sup>Department of Neonatology, Beth Israel Deaconess Medical Center, Boston, MA, USA  
Correspondence: Jonathan S. Litt (jlitt@bidmc.harvard.edu)

Received: 16 September 2019 Revised: 13 April 2020 Accepted: 27 April 2020

Published online: 19 May 2020

for mortality,<sup>12,16</sup> no studies have compared and contrasted risk prediction models for non-mortality outcomes. Recently, Linsell et al.<sup>17</sup> published a systematic review of risk factor models for neurodevelopmental outcomes in children born very preterm or very low birth weight (VLBW). However, this review focused primarily on overall trends in model development and validation rather than a detailed consideration of individual models.

In this article, we conduct an in-depth, narrative review of the current risk models available for predicting the functional outcomes of preterm neonates, evaluating their relative strengths and weaknesses in variable and outcome selection, and considering how risk model development and validation can be improved in the future. Towards this, we first provide an overview of the different risk models developed since 1990. We then frame our review of these models in terms of the outcomes predicted, the range of predictors considered, and the statistical methods used to select the variables included in the final model, as well as to assess the predictive performance of the model. Finally, the ethical implications of integrating risk stratification into standard clinical care for preterm neonates are considered.

## METHODS

We conducted a manual search for relevant literature via PubMed, entering combinations of key terms synonymous with “prediction tool,” “preterm,” and “functional outcome” and reading the abstracts of resulting studies (Table 1). Studies with abstracts that appeared related to our review were then read in full to identify prediction models that were eligible for inclusion. Reference lists of included studies were also reviewed, as were articles that later cited these original studies. Prediction tools were defined as multivariable risk factor analyses (>2 variables) aiming to predict the probability of developing functional outcomes beyond 6 months corrected age. Models that solely investigated associations between individual risk factors and outcomes were excluded, as were models that were not evaluated for predictive ability in terms of either a validation study or an assessment for performance, discrimination, or calibration. Tests used to evaluate a model’s overall performance were  $R^2$ , adjusted  $R^2$ , and the Brier score. The use of a receiver operating characteristic (ROC) curve or a C-index evaluated a model’s discrimination, and the Hosmer–Lemeshow test was considered to evaluate a model’s calibration.<sup>18</sup> Preterm neonates were defined as <37 weeks of completed gestational age. Models with VLBW neonates <1500 g were also included, since in the past birth weight served as a substitute for measuring prematurity when gestational age could not be accurately determined. Models were excluded if they used a cohort entirely composed of infants born prior to 1 January 1990; those born after 1990 were likely to have had surfactant therapy available in the event of respiratory distress syndrome, which significantly reduced the morbidity and mortality rates among preterm neonates nationwide.<sup>19,20</sup> Models were also excluded if they limited their prediction to the outcome of survival, if they incorporated variables measured after initial NICU discharge, or if they included subjects who were not necessarily transferred to a NICU for further care following delivery. Finally, we excluded tools that only predicted outcomes to an age of <6 months corrected age, as well as case reports, narrative reviews, and tools reported in languages other than English.

### Overview of risk prediction models

Table 2 lists all 32 studies with risk prediction models that meet the inclusion and exclusion criteria.<sup>13–15,21–49</sup> From these, a total of 43 distinct models were reported.

### From mortality to neurodevelopmental impairment

Since 1990, several mortality prediction tools have been evaluated in regards to their ability to predict the likelihood of

**Table 1.** Search terms used in literature review of studies with risk prediction models for functional outcomes of preterm neonates.

Terms relating to		
Prediction tool	Logistic Model	
	Linear Model	
	Theoretical Model	
	Multivariate Analysis	
	Regression Analysis	
	Neonatal Screen	
	Outcome Assessment	
	Prognosis	
	Risk Assessment	
	Risk Factor	
	Risk Analysis	
	Preterm	Premature
		Extremely Low Birth Weight
Very Low Birth Weight		
Low Birth Weight		
Small for Gestational Age		
Neonatal Intensive Care		
Nurseries		
Functional outcome	Morbidity	
	Infant Health	
	Infant Welfare	
	Illness Severity	
	Neurodevelopmental Impairment	
	Neurologic Outcome	
	Clinical Outcome	
Child Health		

neurodevelopmental impairment (NDI) among neonates surviving to NICU discharge. One such model is the CRIB, which incorporates six physiologic variables collected within the first 12 h of the preterm infant’s life: birth weight, gestational age, presence of congenital malformations, maximum base excess, and minimum and maximum  $FiO_2$  requirement.<sup>50</sup> Fowlie et al.<sup>24</sup> evaluated how CRIB models obtained at differing time periods over the first 7 days of life predicted severe disability among a group of infants born >31 weeks gestational age or VLBW. In another study, Fowlie et al.<sup>25</sup> incorporated cranial ultrasound findings on day of life 3 along with CRIB scores between 48 and 72 h of life into their prediction model. Subsequent studies analyzed the CRIB in its original 12-h form and, with only one exception,<sup>23</sup> determined that it was not a useful tool for predicting long-term NDI or other morbidities.<sup>26–29</sup> A second example is the SNAP score.<sup>51</sup> SNAP uses 28 physiologic parameters collected over the first 24 h of life to predict survival to NICU discharge, and was modified to predict NDI at 1 year and 2–3 years of age. A subsequent assessment of both the SNAP and the SNAP with Perinatal Extension<sup>42</sup> showed a poor predictive value for morbidity at 4 years of age for children born VLBW and/or with gestational age  $\leq 31$  weeks.<sup>28</sup> Finally, the Neonatal Therapeutic Intervention Scoring System, a comprehensive exam-based prediction tool for mortality,<sup>52</sup> was found to have a poor predictive value for adverse outcomes at 4 years of age in children born very preterm or VLBW.<sup>28</sup>

Shortened forms of the early physiology-based scoring systems were developed and assessed for their ability to predict outcomes in childhood. Application of the CRIB-II on a small cohort ( $n = 107$ )

**Table 2.** Summary of studies with risk prediction models predicting functional outcomes of preterm neonates.

Risk Model (intended clinical use)	Study title, author, ref.	Cohort description, location	Variables (weighting)	Outcomes, age at outcome	Statistical analyses
NBRS (Post-discharge Care Planning)	Nursery neurobiologic risk score and outcome at 18 months, Lefebvre et al. <sup>21</sup>	Cohort born between 1987 and 1992 121 infants with $\leq 28$ weeks GA in single NICU in Canada between 1987	Upon NICU discharge Assisted ventilation (cumulative days) Blood pH (hours below certain thresholds) Seizures (severity) PVL (severity) IVH (severity) Infection (positive blood cx to meningitis) Hypoglycemia (hours of duration) Score range of 0, 1, 2, and 4 for each variable Categorized as low ( $\leq 4$ ), moderate (5–7), and high ( $\geq 8$ ) risk	NDI defined as CP or DQ $< 90$ from the Griffiths' Mental Development Scales, testing locomotor, personal-social, hearing and speech, eye and hand coordination, and performance Disability (blindness, hearing loss) 18 months corrected age	Correlation coefficients, simple and multiple regressions, ROC, sensitivity/specificity, PPV/NPV $\geq 5$ for any disability: Sensitivity: 81% Specificity: 54% PPV: 49% NPV: 84% $\geq 8$ for any disability: Sensitivity: 56% Specificity: 87% PPV: 71% NPV: 78% Predicting any disability at 18 months: AUC: 0.79
NBRS (Post-discharge Care Planning)	Evaluation of the ability of neurobiological, neurodevelopmental, and socioeconomic variables to predict cognitive outcome in premature infants; Wickremasinghe et al. <sup>22</sup>	67–129 infants born $< 32$ weeks GA in single NICU in Minnesota between November 2001 and December 2006; sample size for each outcome measurement ranged from 78 to 48	Upon NICU discharge Assisted ventilation (cumulative days) Blood pH (hours below certain thresholds) Seizures (severity) IVH (severity) PVL (severity) Infection (positive blood cx to meningitis) Hypoglycemia (hours with it) Score range of 0, 1, 2, and 4 for each variable Categorized as low ( $\leq 4$ ), moderate (5–7), and high ( $\geq 8$ ) risk	Capute Scales, also known as CAT/CLAMS Calculate DQ for language, visual-motor problem solving, and overall cognitive development 6, 12, and 24 months corrected age	Multivariable linear regression models $\rho = -0.5$
CRIB (NICU Care Planning, General)	Increased survival and deteriorating developmental outcome in 23- to 25-week-old gestation infants, 1990–1994 compared with 1984–1989, Emsley et al. <sup>23</sup>	40 infants born between 23 and 25 weeks GA in single England NICU between 1990 and 1994	First 12 h of NICU stay Birth weight (0–7) GA, weeks (0–1) Congenital malformations (0–3) Maximum base excess in first 12 h (0–3) Minimum appropriate $FiO_2$ (0–4) Maximum appropriate $FiO_2$ (0 to 5)	Disability defined as CP, blindness due to ROP or other visual impairments, deafness, DQ $< 70$ based on Griffiths Scales Range of 19 months–10 years 7 months	$P < 0.05$ for difference in mean CRIB score between those with disability (8.6) and those without (5.8)
CRIB (NICU Care Planning, General)	Measurement properties of the CRIB-12 h, and responsiveness over 7 days, Fowle et al. <sup>24,25</sup>	300 infants born $\leq 1500$ g or $< 31$ weeks GA in six Scottish NICUs surviving to 7 days of life between 1988 and 1990	First 12 h of life Birth weight (0–7) GA, weeks (0–1) Congenital malformations (0–3) Maximum base excess in first 12 h (0–3) Minimum appropriate $FiO_2$ in first 12 h (0–4) Maximum appropriate $FiO_2$ in first 12 h	Severe disability, defined as cannot sit unsupported, blind, deaf, and/or more than 12 months behind in any other field of development 2 years corrected age	OR, HL, AUC OR: 1.21 (1.02, 1.44) HL: 10.22 (0.25) AUC: 0.71 (0.13)
CRIB24 (NICU Care Planning, General)	Measurement properties of the CRIB-12 h, and responsiveness over 7 days, Fowle et al. <sup>24,25</sup>	300 infants born $\leq 1500$ g or $< 31$ weeks GA in six Scottish NICUs surviving to 7 days of life between 1988 and 1990	First 24 h of life Birth weight (0–7) GA, weeks (0–1) Congenital malformations (0–3) Maximum base excess between 13 and 24 h post delivery (0–3) Minimum appropriate $FiO_2$ between 13 and 24 h post delivery (0–4) Maximum appropriate $FiO_2$ between 13 and 24 h post delivery (0–5)	Severe disability, defined as cannot sit unsupported, blind, deaf, and/or more than 12 months behind in any other field of development 2 years corrected age	OR, HL, AUC OR: 1.18 (0.99, 1.42) HL: 5.39 (0.72) AUC: 0.71 (0.13)
CRIB48 (NICU Care Planning, General)	Measurement properties of the CRIB-12 h, and responsiveness over 7 days, Fowle et al. <sup>24,25</sup>	300 infants born $\leq 1500$ g or $< 31$ weeks GA in six Scottish NICUs surviving to 7 days of life between 1988 and 1990	First 48 h of life Birth weight (0–7) GA, weeks (0–1) Congenital malformations (0–3) Maximum base excess in first 12 h (0–3) Minimum appropriate $FiO_2$ between 25 and 48 h post delivery (0–4)	Severe disability, defined as cannot sit unsupported, blind, deaf, and/or more than 12 months behind in any other field of development 2 years corrected age	OR, HL, AUC OR: 1.28 (1.09, 1.51) HL: 13.99 (0.01) AUC: 0.76 (0.12)

**Table 2.** continued

Risk Model (intended clinical use)	Study title, author, ref.	Cohort description, location	Variables (weighting)	Outcomes, age at outcome	Statistical analyses
CRIB72 (NICU Care Planning, General)	Measurement properties of the CRIB-Reliability, validity beyond the first 12 h, and responsiveness over 7 days, Fowle et al. <sup>24,25</sup>	300 infants born $\leq 1500$ g or $< 31$ weeks GA in six Scottish NICUs surviving to 7 days of life between 1988 and 1990	Maximum appropriate $FiO_2$ between 25 and 48 h post delivery (0–5) First 72 h of life Birth weight (0–7) GA, weeks (0–1) Congenital malformations (0–3) Maximum base excess between 49 and 72 h post delivery (0–3) Minimum appropriate $FiO_2$ between 49 and 72 h post delivery (0–4) Maximum appropriate $FiO_2$ between 49 and 72 h post delivery (0–5)	Severe disability, defined as cannot sit unsupported, blind, deaf, and/or more than 12 months behind in any other field of development 2 years corrected age	OR, HL, AUC OR: 1.25 (1.04, 1.50) HL: 7.78 (0.46) AUC: 0.74 (0.12)
CRIB120 (NICU Care Planning, General)	Measurement properties of the CRIB-Reliability, validity beyond the first 12 h, and responsiveness over 7 days, Fowle et al. <sup>24,25</sup>	300 infants born $\leq 1500$ g or $< 31$ weeks GA in six Scottish NICUs surviving to 7 days of life between 1988 and 1990	First 120 h of life Birth weight (0–7) GA, weeks (0–1) Congenital malformations (0–3) Maximum base excess between 73 to 120 h post delivery (0 to 3) Minimum appropriate $FiO_2$ between 73 to 120 h post delivery (0 to 4) Maximum appropriate $FiO_2$ between 73 to 120 h post delivery (0 to 5)	Severe disability, defined as cannot sit unsupported, blind, deaf, and/or more than 12 months behind in any other field of development 2 years corrected age	OR, HL, AUC OR: 1.46 (1.17, 1.82) HL: 8.70 (0.07) AUC: 0.82 (0.10)
CRIB168 (NICU Care Planning, General)	Measurement properties of the CRIB-Reliability, validity beyond the first 12 h, and responsiveness over 7 days, Fowle et al. <sup>24,25</sup>	300 infants born $\leq 1500$ g or $< 31$ weeks GA in six Scottish NICUs surviving to 7 days of life between 1988 and 1990	First 168 h of life Birth weight (0–7) GA, weeks (0–1) Congenital malformations (0–3) Maximum base excess in first 12 h (0–3) Minimum appropriate $FiO_2$ between 121 and 168 h post delivery (0–4) Maximum appropriate $FiO_2$ between 121 and 168 h post delivery (0–5)	Severe disability, defined as cannot sit unsupported, blind, deaf, and/or more than 12 months behind in any other field of development 2 years corrected age	OR, HL, AUC OR: 1.40 (1.09, 1.79) HL: 9.11 (0.33) AUC: 0.78 (0.12)
CRIB72 + cranial ultrasound (NICU Care Planning, Medically Complex)	Predicting outcome in very low birth weight infants using an objective measure of illness severity and cranial ultrasound scanning, Fowle et al. <sup>24,25</sup>	240 infants born $\leq 1500$ g or $< 31$ weeks GA in six Scottish NICUs surviving to 3 days of life between 1988 and 1990	Day 3 of life Components of CRIB: Birth weight (0–7) GA, weeks (0–1) Congenital malformations (0–3) Maximum base excess between 48 and 72 h (0–3) Minimum appropriate $FiO_2$ between 48 and 72 h (0–4) Maximum appropriate $FiO_2$ between 48 and 72 h (0–5) Total score of 23, threshold score of 5 Cranial ultrasound (IVH grades 0–4, divided into dichotomous 0–2 vs. 3 or 4)	Severe disability, defined as inability to sit unsupported, blind, deaf, and/or more than 12 months behind in any other field of development 18 months corrected age	Regression technique, HL, AUC HL: 5.09 (0.89) AUC: 0.675
CRIB (NICU Care Planning, General)	Is the CRIB score a valid tool for predicting neurodevelopmental outcome in extremely low birth weight infants?, Lago et al. <sup>26</sup>	81 infants born GA $< 28$ weeks and/or BW $< 1000$ g in single NICU in Italy between January 1994 and December 1996	First 12 h of NICU stay Birth weight (0–7) GA, weeks (0–1) Congenital malformations (0–3) Maximum base excess in first 12 h (0–3) Minimum appropriate $FiO_2$ (0–4) Maximum appropriate $FiO_2$ (0–5) Total score of 23, threshold score of 5	Major disability (MDI $< 69$ , CP, and/or blind or deaf) 18 months corrected age	Logistic regression model, AUC AUC: 0.77
CRIB (NICU Care Planning, General)	The CRIB score and NDI at 1 year corrected age in very low birth weight infants, Bührer et al. <sup>27</sup>	352 infants born $< 1500$ g in single NICU in Germany between 1 January 1992 and 31 December 1997	First 12 h of NICU stay Birth weight (0–7) GA, weeks (0–1) Congenital malformations (0–3) Maximum base excess in first 12 h (0–3)	Death or impairment, defined as DO of 85 or below ( $< 2$ SDs) based on Griffiths developmental test 1 year corrected age	AUC, sensitivity/specificity AUC: 0.703 Sensitivity: 27.6% Specificity: 87.3%

**Table 2.** continued

Risk Model (intended clinical use)	Study title, author, ref.	Cohort description, location	Variables (weighting)	Outcomes, age at outcome	Statistical analyses
CRIB (NICU Care Planning, General)	Can severity-of-illness indices for neonatal intensive care predict outcome at 4 years of age?, Eriksson et al. <sup>28</sup>	156 infants born ≤1500 g and/or GA ≤31 weeks in 2 Swedish NICU's between 1 July 1991 and 30 June 1995	Minimum appropriate FiO <sub>2</sub> (0–4) Maximum appropriate FiO <sub>2</sub> (0–5) Total score of 23, threshold score of 5  First 12 h of NICU stay Birth weight (0–7) GA, weeks (0–1) Congenital malformations (0–3) Maximum base excess in first 12 h (0–3) Minimum appropriate FiO <sub>2</sub> (0–4) Maximum appropriate FiO <sub>2</sub> (0–5) Total score of 23, threshold score of 5	Long-term morbidity, parent-response questionnaire asking for deviations in growth and psychomotor development, any neurosensory impairment, difficulties in concentration, and any impairment in vision, hearing, or pulmonary function; validated with info from Medical Health Services records in a subgroup  4 years of age	Poor outcome (impairment or death): AUC: 0.748 Sensitivity: 34.4% Specificity: 91.3% ROC curves with AUC analysis AUC = 0.61
CRIB (NICU Care Planning, General)	Neurodevelopment of children born very preterm and free of severe disabilities: the Nord-Pas de Calais Epipage cohort study, Charkaluk et al. <sup>29</sup>	347 born before 33 weeks GA in the Nord-Pas de Calais area of France in 1997	First 12 h of NICU stay Birth weight (0–7) GA, weeks (0–1) Congenital malformations (0–3) Maximum base excess in first 12 h (0–3) Minimum appropriate FiO <sub>2</sub> (0–4) Maximum appropriate FiO <sub>2</sub> (0–5) Total score of 23, threshold score of 5	DQ based on revised Brune–Lezine Scale 2 years corrected age	Multivariate analysis p value for score >10: Global DQ: 0.26 Fine motor function: 0.57 Sociability: 0.24
Cumulative SNAP (Post-discharge Care Planning)	Chronic physiologic instability is associated with neurodevelopmental morbidity at one and two years in extremely premature infants, Mattia and deRegnier <sup>30</sup>	96 infants with GA ≤30 weeks treated in single NICU in Minnesota between 1 January 1993 and 31 December 1994	Daily throughout NICU stay to create cumulative SNAP score based on worst values recorded each day Point range from 1 to 5, weighted, categorized by <25th, 25–75th, and >75th percentile Blood pressure, high/low Heart rate, high/low Respiratory rate, high Temperature Fahrenheit, low pO <sub>2</sub> /FiO <sub>2</sub> ratio, low pCO <sub>2</sub> , high Hematocrit, high/low WBC count, low Immature: total ratio, high Absolute neutrophils, low Platelet count, low BUN, high Creatinine, high Urine output, low Indirect bilirubin, high Direct bilirubin, high Sodium, high/low Potassium, high/low Calcium, high/low Glucose, high/low Serum bicarb, high/low Serum pH, low Seizure Apnea Stool guaiac test	MDI < 70 or PDI < 70 from Bayley II, CP 2–3 years, mean 26.1 months	Regression analyses Partial R <sup>2</sup> At 1 year: MDI < 70: 0.33 PDI < 70: 0.25 At 2–3 years: MDI < 70: 0.13 PDI < 70: 0.21
SNAP (Post-discharge Care Planning)	Can severity-of-illness indices for neonatal intensive care predict outcome at 4 years of age?, Eriksson et al. <sup>28</sup>	156 infants born ≤1500 g and/or GA ≤31 weeks in 2 Swedish NICU's between 1 July 1991 and 30 June 1995	First 24 h of NICU stay Blood pressure, high/low Heart rate, high/low Respiratory rate, high Temperature Fahrenheit, low pO <sub>2</sub> , high	Long-term morbidity, parent-response questionnaire asking for deviations in growth and psychomotor development, any neurosensory impairment, difficulties in concentration, and any impairment in vision, hearing, or pulmonary function; validated	Area under ROC curve AUC: 0.59

**Table 2.** continued

Risk Model (intended clinical use)	Study title, author, ref.	Cohort description, location	Variables (weighting)	Outcomes, age at outcome	Statistical analyses
SNAP with Perinatal Extension (Post-discharge Care Planning)	Can severity-of-illness indices for neonatal intensive care predict outcome at 4 years of age?, Eriksson et al. <sup>28</sup>	156 infants born $\leq 1500$ g and/or GA $\leq 31$ weeks in 2 Swedish NICU's between 1 July 1991 and 30 June 1995	<p><math>pO_2/FiO_2</math> ratio, low  <math>pCO_2</math>, high                      Hematocrit, high/low                      WBC count, low                      Immature: total ratio, high                      Absolute neutrophils, low                      Platelet count, low                      BUN, high                      Creatinine, high                      Urine output, low                      Indirect bilirubin, high                      Direct bilirubin, high                      Sodium, high/low                      Potassium, high/low                      Calcium, high/low                      Glucose, high/low                      Serum bicarb, high/low                      Serum pH, low                      Seizure                      Apnea                      Stool guaiac test                      First 24 h of NICU stay                      Scored from 1 to 5                      Blood pressure, high/low                      Heart rate, high/low                      Respiratory rate, high                      Temperature Fahrenheit, low  <math>pO_2</math>, high  <math>pO_2/FiO_2</math> ratio, low  <math>pCO_2</math>, high                      Hematocrit, high/low                      WBC count, low                      Immature: total ratio, high                      Absolute neutrophils, low                      Platelet count, low                      BUN, high                      Creatinine, high                      Urine output, low                      Indirect bilirubin, high                      Direct bilirubin, high                      Sodium, high/low                      Potassium, high/low                      Calcium, high/low                      Glucose, high/low                      Serum bicarb, high/low                      Serum pH, low                      Seizure                      Apnea                      Stool guaiac test                      Birth weight                      Apgar at 5 min                      Small for gestational age</p>	<p>with information from Medical Health Services records in a subgroup 4 years of age</p> <p>Long-term morbidity, parent-response questionnaire asking for deviations in growth and psychomotor development, any neurosensory impairment, difficulties in concentration, and any impairment in vision, hearing, or pulmonary function; validated with info from Medical Health Services records in a subgroup 4 years of age</p> <p>Long-term morbidity, parent-response questionnaire asking for deviations in growth AUC = 0.63</p>	<p>Area under ROC curve                      AUC = 0.63</p>
Neonatal Therapeutic Intervention Scoring System (Post-discharge Care Planning)	Can severity-of-illness indices for neonatal intensive care predict outcome at 4 years of age?, Eriksson et al. <sup>28</sup>	156 infants born $\leq 1500$ g and/or GA $\leq 31$ weeks in 2 Swedish NICU's between 1 July 1991 and 30 June 1995	<p>First 24 h of NICU stay                      Total possible score of 93                      Supplemental oxygen                      Surfactant administration                      Tracheostomy care                      Tracheostomy placement                      CPAP administration                      Endotracheal intubation                      Mechanical ventilation</p>	<p>Long-term morbidity, parent-response questionnaire asking for deviations in growth and psychomotor development, any neurosensory impairment, difficulties in concentration, and any impairment in vision, hearing, or pulmonary function; validated with info from Medical Health Services records in a subgroup 4 years of age</p>	<p>Area under ROC curve                      AUC: 0.59</p>

**Table 2.** continued

Risk Model (intended clinical use)	Study title, author, ref.	Cohort description, location	Variables (weighting)	Outcomes, age at outcome	Statistical analyses
			Mechanical ventilation with muscle relaxation HFV Extracorporeal membrane oxygenation Indomethacin administration Volume expansion (15 mL/kg) Vasopressor administration (1 agent) Volume expansion (>15 mL/kg) Vasopressor administration (>1 agent) Pacemaker on standby Pacemaker used Cardiopulmonary resuscitation Antibiotic administration (2 agents) Diuretic administration (enteral) Steroid administration (postnatal) Anticonvulsant administration Aminophylline administration Other unscheduled medication Antibiotic administration (>2 agents) Diuretic administration (parenteral) Treatment of metabolic acidosis Potassium binding resin administration Frequent vital signs Cardiorespiratory monitoring Phlebotomy (5–10 blood draws) Thermoregulated environment Noninvasive oxygen monitoring Arterial pressure monitoring Central venous pressure monitoring Urinary catheter Quantitative intake and output Extensive phlebotomy (>10 blood draws) Gavage feeding Intravenous fat emulsion Intravenous amino acid solution Phototherapy Insulin administration Potassium infusion Intravenous $\gamma$ -globulin Red blood cell transfusion (15 mL/kg) Partial volume exchange transfusion Red blood cell transfusion (>15 mL/kg) Platelet transfusion White blood cell transfusion Double volume exchange transfusion Transport of patient Single chest tube in place Minor operation Multiple chest tubes in place Thoracentesis Major operation Pericardiocentesis Pericardial tube in place Dialysis Peripheral intravenous line Arterial line Central venous line Maternal education (grade completed by first follow-up visit) Maternal age Neonatal gender	Major handicap, defined as CP, blindness, deafness, mental retardation, and/or hydrocephalus MDI < 68	Stepwise regression, followed by applying to a test set ( $n = 74$ ) AUC; specificity, PPV and NPV
Pre-NNR Regression Model (NICU Care)	Prediction of neurologic morbidity in extremely low birth weight infants, Ambalavanan et al. <sup>31</sup>	144 infants born $\leq 1000$ g in single NICU in Alabama between January 1990 and December 1994			

Table 2. continued

Risk Model (intended clinical use)	Study title, author, ref.	Cohort description, location	Variables (weighting)	Outcomes, age at outcome	Statistical analyses
Planning, Medically Complex)			Birth weight Gestational age Inborn/outborn/fetal referral Race Plurality Apgar at 5 min IVH RDS BPD PVL NEC Intestinal perforation PROM PIH C-section delivery Chorioamnionitis Antenatal steroids Antenatal MgSO <sub>4</sub>	PDI < 68 12–18 months of age	at 90% and 70% sensitivity Major handicap: AUC = 0.68 Specificity (90%): 35% PPV (90%): 34% NPV (90%): 90% Specificity (70%): 50% PPV (70%): 34% NPV (70%): 82% MDI < 68: AUC = 0.66 Specificity (90%): 25% PPV (90%): 16% NPV (90%): 94% Specificity (70%): 53% PPV (70%): 19% NPV (70%): 92% PDI < 68: AUC = 0.75 Specificity (90%): 52% PPV (90%): 33% NPV (90%): 94% Specificity (70%): 69% PPV (70%): 38% NPV (70%): 89%
Neural Network Model (NICU Care Planning, Medically Complex)	Prediction of neurologic morbidity in extremely low birth weight infants, Ambalavanan et al. <sup>31</sup>	144 infants born ≤1000 g in single NICU in Alabama between January 1990 and December 1994	Maternal education (grade completed by first follow-up visit) Maternal age Neonatal gender Birth weight Gestational age Inborn/outborn/fetal referral Race Plurality Apgar at 5 min IVH RDS BPD PVL NEC Intestinal perforation PROM PIH C-section delivery Chorioamnionitis Antenatal steroids Antenatal MgSO <sub>4</sub>	Presence or absence of major handicap, defined as CP, blindness, deafness, mental retardation, and/or hydrocephalus MDI < 68 PDI < 68 12–18 months of age	Four-layer back-propagation network was trained on the training set and used to predict outcome AUC, specificity, PPV, and NPV at 90% and 70% sensitivity Major handicap: AUC = 0.62 Specificity (90%): 17% PPV (90%): 29% NPV (90%): 82% Specificity (70%): 30% PPV (70%): 27% NPV (70%): 73% MDI < 68 AUC = 0.75 Specificity (90%): 31% PPV (90%): 17% NPV (90%): 95% Specificity (70%): 62% PPV (70%): 23% NPV (70%): 95% PDI < 68 AUC = 0.69 Specificity (90%): 35% PPV (90%): 27% NPV (90%): 91% Specificity (70%): 52% PPV (70%): 28% NPV (70%): 86%
3-Morbidity Model (NICU Care Planning, Medically Complex)	Impact of BPD, brain injury, and severe retinopathy on the outcome of extremely low birth weight infants at	910 infants born 500 to 999 g internationally (TIPP trial) between 1996 and 1998	36 weeks PMA BPD Severe brain injury Severe ROP	Death, poor outcome defined as CP, MDI < 70, hearing loss requiring amplification, bilateral blindness 18 months corrected age	Logistic regression, $\chi^2$ analysis, OR Rates of death/disability: None: 18% (14% - 22%)



**Table 2.** continued

Risk Model (intended clinical use)	Study title, author, ref.	Cohort description, location	Variables (weighting)	Outcomes, age at outcome	Statistical analyses
3-Morbidity Model (NICU Care Planning, Medically Complex)	Impact at age 11 years of major neonatal morbidities in children born extremely preterm, Farooqi et al. <sup>33</sup>	97 infants born <26 weeks GA nationwide in Sweden between March 1990 and April 1992	36 weeks PMA BPD Serious brain injury Severe ROP	Poor outcome (death, survival with moderate/severe CP, severe visual impairment, moderate to severe hearing loss in both ears, severe mental retardation) 11 years	Any 1: 42% (37% - 47%) Any 2: 62% (53% - 70%) All 3: 88% (64% - 99%) OR for morbidity count: 2.9 (2.4–3.5) $\chi^2$ , Fisher's exact, multivariate logistic regression Likelihood of poor outcome: ORs None: 10% 1: 19%; OR: 3.2 to 18.8 Any 2: 58%; OR: 5.0 to 19.1 All 3: 80% (50 to 100); OR: 13.4
3-Morbidity Model (NICU Care Planning, Medically Complex)	Prediction of late death or disability at age 5 years using a count of three neonatal morbidities in very low birth weights, Schmidt et al. <sup>15</sup>	1514 infants born 500 to 1250 g internationally (caffeine trial) between 1999 and 2004	36 weeks PMA BPD Serious brain injury Severe ROP	Death, disability (motor impairment, cognitive impairment, behavior problems, poor general health, deafness, and/or blindness) "poor general health" defined as supplemental oxygen, need for positive airway pressure, feeding tube, seizure more than once per month, or recent ICU admission for NICU-related complication 5 years corrected age	Dichotomous values compared with Fisher's exact test, various logistic regression models Likelihood of poor general health: ORs: None: 11.2%; OR: 1.0 Any 1: 22.9%; OR: 2.4 Any 2: 34.9%; OR: 6.2 All 3: 61.5%; OR: 12.7
4-Morbidity Model (NICU Care Planning, Medically Complex)	Effect of severe neonatal morbidities on long-term outcome in extremely low birth weight infants, Koo et al. <sup>34</sup>	80 infants born <1000 g in single NICU in Korea between 1 January 1997 and 31 December 2007	36 weeks PMA BPD Serious brain injury Severe ROP Parenteral nutrition-associated cholestasis	Poor outcome (death or survival with neurosensory impairment defined as CP, delayed development, hearing loss, or blindness) 18–24 months corrected age	Multiple logistic regression analysis Rates of death/disability: None: 9% Any 1: 46% Any 2: 69% Any 3: 100%
Classification Tree Model (NICU Care Planning, General)	Early prediction of poor outcome in extremely low birth weight infants by classification tree analysis, Ambalavanan et al. <sup>35</sup>	1046 infants 500–999 g in birth weight surviving to 8 days GA enrolled internationally (TIPP) between January 1996 and March 1998	Upon birth, 4 days GA, and 8 days GA Classification trees where model branches as it progresses Antenatal model: GA $\leq$ 25.5 weeks Non-white race Three-day model: BW $\leq$ 787 g Total fluid intake <101 ml/kg/day Seven-day model: BW $\leq$ 787 g RBC transfusion >3 ml/kg/day	Death or NDI, defined as CP, cognitive delay (MDI < 70), hearing loss, and bilateral blindness 18–22 months corrected age	Tree models using development set of data (75%), choosing set of binary splits Validation set (25%) Antenatal model: Sensitivity: 49% (40–59%) Specificity: 71% (63–78%) PPV: 58% (47–67%) NPV: 64% (56–71%) Accuracy: 61% (55–67%) Three-day model: Sensitivity: 31% (23–40%) Specificity: 86% (80–91%) PPV: 64% (50–77%) NPV: 61% (54–68%) Accuracy: 62% (56–68%) Seven-day model: Sensitivity: 34% (25–43%) Specificity: 83% (76–89%) PPV: 61% (48–73%) NPV: 61% (54–68%) Accuracy: 61% (55–67%)
NRI (Post-discharge Care Planning)	Consequences and risks of <1000 g birth weight for neuropsychological skills, achievement, and adaptive functioning, Taylor et al. <sup>36</sup>	204 infants born <1000 g in single Ohio NICU between January 1992 and December 1995	Upon NICU discharge Ultrasound evidence for grade I/II IVH (1) Grade III/IV IVH, PVL, or ventricular dilatation (2) Oxygen requirement at 28 days (1)	Significantly decreased NEPSY, which in turn is a mediator for decreased academic performance measured by Woodcock Johnson Tests of Achievement 3rd Edition (WJ-III; letter word identification, spelling, resources	Linear regression analysis, fitted separately and adjusted sex, race, parental SES, family stressors, and family resources

**Table 2.** continued

Risk Model (intended clinical use)	Study title, author, ref.	Cohort description, location	Variables (weighting)	Outcomes, age at outcome	Statistical analyses
"Early" Clinical Model (NICU Care Planning, Medically Complex)	Clinical data predict neurodevelopmental outcome better than head ultrasound in extremely low birth weight infants, Broitman et al. <sup>37</sup>	2103 extremely low birth weight (401–1000 g) infants admitted to any NRN NICU from 1 January 1998 to 30 June 2001	Oxygen requirement at 36 weeks (2) CA Septicemia (1) Jaundice of prematurity (1) Apnea of prematurity (1) NEC (1) Increased risk with NRI > 3	calculation, and academic skills) 8 years	Model 1 (NRI adjusted for socioeconomic factors) $R^2$ : WJ-III Letter Word ID: 0.26 WJ-III Spelling: 0.21 WJ-III Calculation: 0.16 WJ-III Academic Skills: 0.26 Model 2 (addition of NEPSY) $R^2$ : WJ-III Letter Word ID: 0.56 Beta weights: NRI > 3 to NEPSY: -0.24 NEPSY to WJ-III: 0.68 Sobel test = -4.39, $p < 0.001$ Stepwise variable selection followed by multiple logistic regression analysis, AUC Divided cohort into development set (70%) and validation set (30%) AUC: 0.68
"All" Clinical Model (NICU Care Planning, Medically Complex)	Clinical Data predict neurodevelopmental outcome better than head ultrasound in extremely low birth weight infants, Broitman et al. <sup>37</sup>	2103 extremely low birth weight (401–1000 g) infants admitted to any NRN NICU from 1 January 1998 to 30 June 2001	Early = postnatal day 28 "Early" model (variables that could be assessed by postnatal day 28) • Male • Race (black and non-Hispanic vs. others) • Birth weight (100 g increments) • Surfactant • Late-onset sepsis • Seizure • HFV Up to 120 days of age • Male • Race • Surfactant • Late-onset sepsis • Supplemental O <sub>2</sub> at 36 weeks PMA (CLD) • Threshold for ROP	NDI defined as MDI < 70, PDI < 70, CP, deafness, and/or blindness; used Bayley Scales II 18–22 months corrected age	Stepwise variable selection followed by multiple logistic regression analysis Divided cohort into development set (70%) and validation set (30%) AUC: 0.68
5-Factor Model (NICU Care Planning, General)	Intensive care for extreme prematurity—moving beyond gestational age, Tyson et al. <sup>13</sup>	4446 infants born 22–25 weeks GA and between 401 and 1000 g in US nationwide (NICHD) born between 1 January 1998 and 31 December 2003	Upon delivery Gestational age Sex Exposure to antenatal corticosteroids within 7 days prior to delivery Plurality Birth weight	Survival without NDI, survival with NDI, survival with profound NDI NDI defined as PDI/MDI < 70, moderate to severe CP, bilateral blindness, or bilateral hearing loss Profound ≤ 50 Bayley or adult assistance required to move 18–22 months corrected age	Logistic mixed model, bootstrap estimates AUC, HL: AUC (death): 0.753 AUC (death or profound NDI): 0.751 HL: not significant
5-Factor Model (NICU Care Planning, General)	Infant outcomes after periviable birth: external validation of the NRN estimator with the BEAM trial, Marrs et al. <sup>38</sup>	289 infants born at 23 0/7–25 6/7 weeks GA who were mechanically ventilated between December 1997 and May 2004	Upon delivery Gestational age Sex Exposure to antenatal corticosteroids within 7 days prior to delivery Plurality Birth weight	Survival without NDI, survival with NDI, survival with profound NDI NDI defined as PDI/MDI < 70, moderate to severe CP, bilateral blindness, or bilateral hearing loss Profound ≤ 50 Bayley or adult assistance required to move 18–22 months corrected age	AUC, HL, Brier score: AUC (death): 0.70 HL (death): not significant Brier (death): 0.17 AUC (death or profound NDI): 0.64 HL (death or profound NDI): significant Brier (death or profound NDI): 0.22 AUC (death or NDI): 0.71 Brier (death or NDI): 0.21 HL (death or NDI): not significant
CRIB-II (NICU Care Planning, General)	CRIB score for the prediction of neurodevelopmental outcomes at 3 years of age in infants of very low birth weight, Lodha et al. <sup>39</sup>	107 infants born < 1250 g in single NICU in Canada between January 2000 and December 2001	Within first hour of birth Gender Gestational age (weeks) Birth weight (g) Admission temperature	Major neurodevelopmental disability (CP, neurosensory hearing loss requiring amplification, legal blindness, severe seizure disorder, and/or cognitive score more than 2 SD below mean for adjusted age determined by Wechsler Preschool and	Logistic regression and bootstrap estimates, AUC, sensitivity/Specificity AUC (major NDI): 0.84 Sensitivity: 83.3% Specificity: 84.2%

**Table 2.** continued

Risk Model (intended clinical use)	Study title, author, ref.	Cohort description, location	Variables (weighting)	Outcomes, age at outcome	Statistical analyses
CRIB-II (NICU Care Planning, General)	Can the early condition at admission of a high-risk infant aid in the prediction of mortality and poor neurodevelopmental outcome? A population study in Australia, Greenwood et al. <sup>40</sup>	1328 infants born <29 weeks gestation admitted to 10 NICUs in Australia between 1 January 1998 and 31 December 2003	Base deficit (total score of 13 or above out of 27) Within first hour of birth Gender Gestational age (weeks) Birth weight (g) Admission temperature Base deficit (total score of 13 or above out of 27)	Primary Scale of Intelligence, Bayley Scales II, or revised Leiter International Performance Scale) 3 years corrected age Moderate to severe functional disability, defined as $\leq 2$ SDs on GMDS or BSID-II, non-ambulatory CP, bilateral blindness, or bilateral deafness 2–3 years corrected age	AUC (death or major NDI): 0.82 Area under ROC curve AUC: 0.68
Autism Prediction Model (Post-discharge Care Planning)	Autism spectrum disorders in extremely preterm children, Johnson et al. <sup>41</sup>	189 infants born <26 weeks GA in United Kingdom and Ireland between March and December 1995	Prior to NICU discharge Male sex Gestational age $\leq 24$ weeks Any breast milk Abnormal cranial US scanning results Vaginal breech delivery	Autism spectrum disorder, based on parental screening questionnaire called Social Communication Questionnaire 11 years of age	Univariate and multivariate linear regression $R^2$ : 19.7%
SNAP-II (NICU Care Planning, General)	SNAP-II and SNAPPE-II and the risk of structural and functional brain disorders in extremely low gestational age newborns: the ELGAN study, Dammann et al. <sup>42</sup>	1149 infants <28 weeks GA in US nationwide (ELGAN study) born between 2002 and 2004	12 h post delivery Lowest mean BP Lowest temp. $pO_2/FIO_2$ Lowest serum pH Presence of multiple seizures Low urine output	Cerebral US lesions in the NICU, low MDI, low PDI, autism spectrum disorder, small head circumference (Z-score < -2) 24 months corrected age	Logistic regression ORs Bayley Scales: SNAP-II in highest decile for gestational age predicting PDI < 55: 1.8 (1.1–3.2) MDI < 55: 2.0 (1.1–3.5) Autism positive screening: SNAP-II in highest quartile for GA: 1.7 (1.2–2.4) Small head circumference at 24 months (Z-score < -2): SNAP-II in highest quartile for GA: 1.8 (1.1–2.9)
SNAPPE-II (NICU Care Planning, General)	SNAP-II and SNAPPE-II and the risk of structural and functional brain disorders in extremely low gestational age newborns: the ELGAN study, Dammann et al. <sup>42</sup>	1149 infants <28 weeks GA in US nationwide (ELGAN study) born between 2002 and 2004	12 h post delivery Lowest mean BP Lowest temp. $pO_2/FIO_2$ Lowest serum pH Presence of multiple seizures Low urine output Birth weight Apgar at 5 min Small for gestational age	Cerebral US lesions in the NICU, low MDI, low PDI, autism spectrum disorder, small head circumference (Z-score < -2) 24 months corrected age	Logistic regression ORs Bayley Scales: SNAPPE-II in highest quartile for gestational age predicting PDI < 55: 1.8 (1.2–2.7) MDI < 55: 1.8 (1.2–2.8) Autism positive screening: SNAPPE-II in highest quartile for GA: 1.8 (1.3–2.5) Small head circumference at 24 months (Z-score < -2): SNAPPE-II in highest quartile for GA 2.3 (1.4–3.7)
SNAP-II (NICU Care Planning, General)	Early postnatal illness severity scores predict NDIs at 10 years of age in children born extremely preterm, Logan et al. <sup>43</sup>	874 infants born <28 weeks GA in nationwide NICUs (ELGAN study) between 2002 and 2004	Within 12 h of life Lowest mean BP Lowest temp. $pO_2/FIO_2$ Lowest serum pH Presence of multiple seizures Low urine output Higher risk with score $\geq 30$ Intermediate category with score 20–29	Cognitive impairment (IQ, executive function, language ability) Adverse neurological outcomes (epilepsy, impaired gross motor function) Behavioral abnormalities (attention deficit disorder and hyperactivity), social dysfunction (autism spectrum disorder) and education-related adversities (school achievement and need for educational supports) 10 years	ORs comparing categories of SNAP-II score and association with Z-score $\leq 1$ for each outcome 11 of 18 cognitive outcomes associated with SNAP-II $\geq 30$ and 6 of 18 associated with SNAP-II of 20 to 29; ORs ranged from 1.4 to 2.1 2 of 8 social dysfunctions associated with SNAP-II $\geq 30$ and 3 of 8 social dysfunctions associated with SNAP-II 20 to 29; OR ranged from 1.6 to 2.3

Table 2. continued

Risk Model (intended clinical use)	Study title, author, ref.	Cohort description, location	Variables (weighting)	Outcomes, age at outcome	Statistical analyses
Social Risk with Cranial MRI Abnormalities (Post-discharge Care Planning)	High prevalence/low severity language delay in preschool children born very preterm, Foster-Cohen et al. <sup>44</sup>	110 very preterm children born ≤33 weeks gestation or ≤1500 g birth weight in one NICU in New Zealand between November 1998 and December 2000	<p>Term-equivalent age</p> <p>Minority ethnicity</p> <p>Early motherhood (&lt;25 at childbirth)</p> <p>Maternal education (left secondary school early or only completed secondary school)</p> <p>Single parent family at birth</p> <p>SES assessed with Eley-Irving Socioeconomic Index (higher vs. lower)</p> <ul style="list-style-type: none"> <li>• Professional</li> <li>• Managerial</li> <li>• Clerical/technical</li> <li>• Skilled</li> <li>• Semi-skilled (lower)</li> <li>• Unskilled (lower)</li> <li>• Unemployed (lower)</li> <li>• White matter abnormalities on MRI at term-equivalent age, graded according to five 3-point scales (white matter signal abnormalities, periventricular white matter volume loss, presence of periventricular white matter cysts, ventricular dilation, and thinning of corpus callosum); divided into three groups: <ul style="list-style-type: none"> <li>• No abnormalities</li> <li>• Mild abnormalities</li> <li>• Moderate to severe abnormalities</li> </ul> </li> </ul>	Language development using CELF-P; delay defined as score greater than 1 SD below mean from full-term group 4 years corrected age	Multiple regression analysis $R^2$ : 0.18
Rehospitalization Risk Regression Model (Post-discharge Care Planning)	Identification of extremely premature infants at high risk of rehospitalization, Ambalavanan et al. <sup>45</sup>	3787 infants born between 400 and 1000 g in nationwide NICUs (NIHCD) between 2002 and 2005	<p>Upon NICU discharge</p> <p>Yes/No point-based model, where every "No" is 1.0 point</p> <p>Shunt for hydrocephalus (4.5)</p> <p>Infant in hospital &gt;120 days due to pulmonary reasons (1.9)</p> <p>Proven NEC or spontaneous gastrointestinal perforation (1.6)</p> <p>FiO<sub>2</sub> at 36 weeks: 0.21–0.28 (1.3), &gt;0.28 (1.6)</p> <p>Male (1.3)</p> <p>Score range of 5–10.9</p>	Rehospitalization, defined as at least one overnight stay in hospital since initial discharge from hospital or to a chronic care facility 18–22 months corrected age	<p>Logistic regression, AUC, HL, sensitivity/specificity/PPV/NPV</p> <p>AUC: 0.63</p> <p><math>R^2</math> = 0.059</p> <p>Max rescaled <math>R^2</math> = 0.079</p> <p>HL: 0.16</p> <p>Sensitivity:</p> <ul style="list-style-type: none"> <li>&gt;5.0 = 100</li> <li>&gt;5.3 = 75</li> <li>&gt;5.6 = 42</li> <li>&gt;8.5 = 5</li> </ul> <p>Specificity:</p> <ul style="list-style-type: none"> <li>&gt;5.0 = 0</li> <li>&gt;5.3 = 40</li> <li>&gt;5.6 = 79</li> <li>&gt;8.5 = 99</li> </ul> <p>PPV: <ul style="list-style-type: none"> <li>&gt;5.0 = 46</li> <li>&gt;5.3 = 52</li> <li>&gt;5.6 = 63</li> <li>&gt;8.5 = 80</li> </ul> </p> <p>NPV: <ul style="list-style-type: none"> <li>&gt;5.0 = 0</li> <li>&gt;5.3 = 65</li> <li>&gt;5.6 = 61</li> <li>&gt;8.5 = 55</li> </ul> </p>
Rehospitalization Risk CART Model (Post-discharge Care Planning)	Identification of extremely premature infants at high risk of rehospitalization, Ambalavanan et al. <sup>45</sup>	3787 infants born between 400 and 1000 g in nationwide NICUs (NIHCD) between 2002 and 2005	<p>Upon NICU discharge</p> <p>Step-based yes/no classification tree</p> <p>Hospital for pulmonary reasons &gt;120 days</p> <p>Shunt for hydrocephalus</p> <p>≥43 days on mechanical ventilation</p> <p>Male</p>	Rehospitalization, defined as at least one overnight stay in hospital since initial discharge from hospital or to a chronic care facility 18–22 months corrected age	Misclassification rate: 0.40

**Table 2.** continued

Risk Model (intended clinical use)	Study title, author, ref.	Cohort description, location	Variables (weighting)	Outcomes, age at outcome	Statistical analyses
Rehospitalization for Pulmonary Reasons Risk Regression Model (Post-discharge Care Planning)	Identification of extremely premature infants at high risk of rehospitalization, Ambalavanan et al. <sup>45</sup>	3438 infants born between 400 and 1000 g in nationwide NICUs (NIHCD) between 2002 and 2005	<p>Upon NICU discharge Yes/No point-based model, where every "No" is 1.0 point</p> <p>Discharge on bronchodilators (2.6)</p> <p>Infant in hospital &gt;120 days due to pulmonary reasons (1.9)</p> <p>No private insurance (1.5)</p> <p>Episodes of late-onset culture-negative infection treated with antibiotics for &gt;5 days: 1 (1.3), &gt;1 (1.4)</p> <p>Male (1.3)</p> <p>Score range of 5–7.8</p>	Rehospitalization, defined as at least one overnight stay in hospital since initial discharge from hospital or to a chronic care facility, due a respiratory cause 18–22 months corrected age	<p>Logistic regression, AUC, HL, sensitivity/specificity/PPV/NPV</p> <p>AUC = 0.63</p> <p>R<sup>2</sup> = 0.026</p> <p>Max-rescaled R<sup>2</sup> = 0.045</p> <p>HL: 0.35</p> <p>Sensitivity:</p> <p>&gt;5.0 = 100</p> <p>&gt;5.5 = 86</p> <p>&gt;6.0 = 45</p> <p>&gt;6.5 = 25</p> <p>Specificity:</p> <p>&gt;5.0 = 0</p> <p>&gt;5.5 = 23</p> <p>&gt;6.0 = 74</p> <p>&gt;6.5 = 88</p> <p>PPV:</p> <p>&gt;5.0 = 15</p> <p>&gt;5.5 = 16</p> <p>&gt;6.0 = 23</p> <p>&gt;6.5 = 26</p> <p>NPV:</p> <p>&gt;5.0 = 0</p> <p>&gt;5.5 = 91</p> <p>&gt;6.0 = 89</p> <p>&gt;6.5 = 87</p>
Rehospitalization for Pulmonary Reasons Risk CART Model (Post-discharge Care Planning)	Identification of extremely premature infants at high risk of rehospitalization, Ambalavanan et al. <sup>45</sup>	3787 infants born between 400 and 1000 g in nationwide NICUs (NIHCD) between 2002 and 2005	<p>Upon NICU discharge</p> <p>Step-based yes/no classification tree</p> <p>If ≥22 days on mechanical ventilation: FIO<sub>2</sub> at 36 weeks PMA &gt; 0.25</p> <p>Maternal age ≤23 years</p> <p>If &lt;22 days on mechanical ventilation: Discharged on oxygen at &gt;120 days</p> <p>Late-onset culture-negative sepsis requiring antibiotics for ≥5 days</p>	Rehospitalization, defined as at least one overnight stay in hospital since initial discharge from hospital or to a chronic care facility, due a respiratory cause 18–22 months corrected age	Misclassification rate: 0.42
Autism Spectrum Disorder Model with NRN-Derived Cohort (Post-discharge Care Planning)	Screening for autism spectrum disorders in extremely preterm infants, Stephens et al. <sup>46</sup>	554 infants born <27 weeks at 15 NICUs that were part of NRN, enrolled at follow-up between 1 November 2008 and 8 April 2010	<p>Upon NICU discharge</p> <p>Birth weight (kg)</p> <p>Male</p> <p>Multiple birth</p> <p>BPD</p> <p>NEC</p> <p>Sepsis</p> <p>IVH/PVL</p> <p>Days in hospital</p> <p>Maternal age &gt;35</p> <p>Race (non-White)</p> <p>Maternal education (less than high school grad, high school grad, or more than high school grad)</p> <p>Medicaid</p>	Autism spectrum disorder using three tests: Pervasive Developmental Disorders Screening Test, Second edition Stage 2, Response to Joint Attention, and Response to Name 18 months corrected age	<p>Logistic regression, AUC</p> <p>AUC: 0.77</p>
Outcome Trajectory (delivery, death/NDI) (NICU Care Planning, General)	Outcome trajectories in extremely preterm infants, Ambalavanan et al. <sup>45</sup>	8713 infants born between 401 and 1000 g BW nationwide (NICHD) between 1 January 1998 and 31 December 2005	<p>Delivery</p> <p>Birth weight (g)</p> <p>Apgar at 5 min</p> <p>Male gender</p> <p>GA (weeks)</p> <p>Antenatal steroids</p>	Death/NDI (MDI < 70, PDI < 70, CP, bilateral blindness, or bilateral need for hearing aids) 18–22 months corrected age	<p>Multivariable forward stepwise logistic regression models, AUC</p> <p>AUC: 0.783</p>
Outcome Trajectory (delivery, death/NDI) (NICU Care Planning, General)	Outcome trajectories in extremely preterm infants, Ambalavanan et al. <sup>45</sup>	8713 infants born between 401 and 1000 g BW nationwide (NICHD) between 1 January 1998 and 31 December 2005	<p>Delivery</p> <p>Birth weight (g)</p> <p>Apgar at 5 min</p> <p>Male gender</p> <p>GA (weeks)</p> <p>Antenatal steroids</p>	Death/NDI (MDI < 70, PDI < 70, CP, bilateral blindness, or bilateral need for hearing aids) 18–22 months corrected age	<p>Multivariable forward stepwise logistic regression models, AUC</p> <p>AUC: 0.783</p>

**Table 2.** continued

Risk Model (intended clinical use)	Study title, author, ref.	Cohort description, location	Variables (weighting)	Outcomes, age at outcome	Statistical analyses
NDI) (NICU Care Planning, General)		between 1 January 1998 and 31 December 2005	Male gender Intubation Inborn	NDI (MDI < 70, PDI < 70, CP, bilateral blindness, bilateral need for hearing aids)	models, AUC AUC: 0.659
Outcome Trajectory (7 days, death/NDI) (NICU Care Planning, General)	Outcome trajectories in extremely preterm infants, Ambalavanan et al. <sup>45</sup>	6996 infants born between 401 and 1000 g BW nationwide (NICHD) between 1 January 1998 and 31 December 2005	7 days of life Birth weight (g) Highest FiO <sub>2</sub> on day 7 Male gender IVH grade (diagnosed by day 7) Days on CPAP	Death/NDI (MDI < 70, PDI < 70, CP, bilateral blindness, bilateral need for hearing aids) 18–22 months corrected age	Multivariable forward stepwise logistic regression models, AUC AUC: 0.724
Outcome Trajectory (7 days, NDI) (NICU Care Planning, General)	Outcome trajectories in extremely preterm infants, Ambalavanan et al. <sup>45</sup>	6996 infants born between 401 and 1000 g BW nationwide (NICHD) between 1 January 1998 and 31 December 2005	7 days of life Birth weight (g) Male gender Days on CPAP (to day 7) Apgar at 5 min Days on HFV (to day 7) Days on CV (to day 7)	NDI (MDI < 70, PDI < 70, CP, bilateral blindness, bilateral need for hearing aids) 18–22 months corrected age	Multivariable forward stepwise logistic regression models, AUC AUC: 0.676
Outcome Trajectory (28 days, death/NDI) (NICU Care Planning, Medically Complex)	Outcome trajectories in extremely preterm infants, Ambalavanan et al. <sup>45</sup>	6241 infants born between 401 and 1000 g BW nationwide (NICHD) between 1 January 1998 and 31 December 2005	28 days of life Highest FiO <sub>2</sub> on day 28 Episodes of late-onset culture-negative clinical infection Days parenteral feeding to day 28 Days on CPAP to day 28 Male gender Birth weight (g)	Death/NDI (MDI < 70, PDI < 70, CP, bilateral blindness, bilateral need for hearing aids) 18–22 months corrected age	Multivariable forward stepwise logistic regression models, AUC AUC: 0.721
Outcome Trajectory (28 days, NDI) (NICU Care Planning, Medically Complex)	Outcome trajectories in extremely preterm infants, Ambalavanan et al. <sup>45</sup>	6241 infants born between 401 and 1000 g BW nationwide (NICHD) between 1 January 1998 and 31 December 2005	28 days of life Episodes of late-onset culture-negative clinical infection Days on CV to 28 days Days on HFV to 28 days Male gender	NDI (MDI < 70, PDI < 70, CP, bilateral blindness, bilateral need for hearing aids) 18–22 months corrected age	Multivariable forward stepwise logistic regression models, AUC AUC: 0.686
Outcome Trajectory (36 weeks PMA, death/NDI) (NICU Care Planning, Medically Complex)	Outcome trajectories in extremely preterm infants, Ambalavanan et al. <sup>45</sup>	5118 infants born between 401 and 1000 g BW nationwide (NICHD) between 1 January 1998 and 31 December 2005	36 weeks PMA Days on CV to 36 weeks Days on HFV to 36 weeks Ventricular size enlarged on ultrasound (day 28 to 36 weeks) Male gender PVL or porencephalic cyst on ultrasound (day 28 to 36 weeks) BPD (on ventilator or CPAP at 36 weeks)	Death/NDI (MDI < 70, PDI < 70, CP, bilateral blindness, bilateral need for hearing aids) 18–22 months corrected age	Multivariable forward stepwise logistic regression models, AUC AUC: 0.741
Outcome Trajectory (36 weeks PMA, NDI) (NICU Care Planning, Medically Complex)	Outcome trajectories in extremely preterm infants, Ambalavanan et al. <sup>45</sup>	5118 infants born between 401 and 1000 g BW nationwide (NICHD) between 1 January 1998 and 31 December 2005	36 weeks PMA Days on CV (to 36 weeks) PVL or porencephalic cyst on ultrasound (day 28 to 36 weeks) Male gender Ventricular size enlarged on ultrasound (day 28 to 36 weeks)	NDI (MDI < 70, PDI < 70, CP, bilateral blindness, bilateral need for hearing aids) 18–22 months corrected age	Multivariable forward stepwise logistic regression models, AUC AUC: 0.716
Perinatal Risk Factor Model (Post-discharge Care Planning)	Perinatal risk factors for neurocognitive impairments in preschool children born very preterm, Potharst et al. <sup>47</sup>	102 infants born <30 weeks GA or <1000 g in single Netherlands NICU between 2007 and 2009	Upon NICU discharge SES (level of parental education, low, medium, high) Parental foreign country of birth Gestational age Very small for gestational age Sex BPD Indomethacin for PDA Abnormal ultrasound Sepsis and/or meningitis	Relative to control group, lower: Intelligence based on Wechsler Processing speed Executive functioning Attention Visual–motor coordination Face recognition Emotion recognition 5 years corrected age	Regression analyses R <sup>2</sup> (Visual–motor): 0.19 R <sup>2</sup> (reaction time/attention): 0.23 R <sup>2</sup> (emotion/face recognition): 0.18 R <sup>2</sup> (accuracy/attention): 0.10 R <sup>2</sup> (working memory): 0.21

**Table 2.** continued

Risk Model (intended clinical use)	Study title, author, ref.	Cohort description, location	Variables (weighting)	Outcomes, age at outcome	Statistical analyses
5-Morbidity Model (Post-discharge Care Planning)	Correlations between initial neonatal and early childhood outcomes following preterm birth, Manuck et al. <sup>48</sup>	459 infants born <34 weeks gestation in multiple NICU's nationwide between December 1997 and May 2004 (NICHD Network from Magnesium trial)	Upon NICU discharge IVH and/or PVL BPD ROP NEC Sepsis Model controlled for gestational age, chorioamnionitis, maternal education, maternal race, fetal sex, treatment group assignment (magnesium vs. control), and maternal use of tobacco, alcohol, and/or drugs during pregnancy	Death or NDI as determined by moderate or severe CP and/or PDI and/or MDI scores >2SD below the mean 2 years corrected age	Multiple logistic regression models, AUC For entire preterm group AUCs: None: 0.66 (0.63–0.69) Any 1: 0.62 (0.58–0.66) Any 2: 0.65 (0.61–0.69) Any 3: 0.64 (0.60–0.69) Any 4: 0.65 (0.60–0.69) All 5: 0.61 (0.56–0.65) Individual combinations ranged from 0.61 to 0.68 <28 weeks GA AUCs: None: 0.64 (0.59–0.69) Any 1: 0.66 (0.58–0.74) Any 2: 0.66 (0.58–0.73) Any 3: 0.69 (0.60–0.77) Any 4: 0.74 (0.64–0.83) All 5: 0.67 (0.53–0.82) Individual combinations ranged from 0.64 to 0.77
Perinatal Post-Prematurity Respiratory Disease Model (Post-discharge Care Planning)	BPD and perinatal characteristics predict one-year respiratory outcomes in extremely low gestational age newborns, Keller et al. <sup>49</sup>	697 infants born at 23 0/7–28 6/7 weeks GA at six academic centers	At delivery GA (per week) IUGR Male sex Smoking during pregnancy Intubation at birth Infant race Public insurance Parent with asthma	Respiratory morbidity on at least two caregiver questionnaires given 3, 6, 9, and 12 months corrected age (hospitalization for respiratory indication, home respiratory support, respiratory medication administration, or respiratory symptoms)	Multivariate analyses, ROC, and c-statistic c-Statistic = 0.858
30 Days PMA Post-Prematurity Respiratory Disease (Post-discharge Care Planning)	BPD and perinatal characteristics predict one-year respiratory outcomes in extremely low gestational age newborns, Keller et al. <sup>49</sup>	679 infants born at 23 0/7–28 6/7 weeks GA at six academic centers	36 weeks PMA or at NICU discharge IUGR Male sex Smoking during pregnancy Intubation at birth Infant race Public insurance Low breast milk exposure BPD (none/mild, moderate, severe)	Respiratory morbidity on at least two caregiver questionnaires given 3, 6, 9, and 12 months corrected age (hospitalization for respiratory indication, home respiratory support, respiratory medication administration, or respiratory symptoms)	Multivariate analyses, ROC and c-statistic c-Statistic = 0.856

*NBRS* Nursery Neurobiologic Risk Score, *IVH* intraventricular hemorrhage, *PVL* periventricular leukomalacia, *MDI* neurodevelopmental impairment, *CP* cerebral palsy, *DO* developmental quotient, *ROC* receiver operating characteristic, *PPV* positive predictive value, *NPV* negative predictive value, *CRIB* Clinical Risk Index for Babies, *ROP* retinopathy of prematurity, *MDI* mental development index, *SNAP* Score for Neonatal Acute Physiology, *PDI* psychomotor development index, *NRN* Neonatal Research Network, *RDS* respiratory distress syndrome, *BPD* bronchopulmonary dysplasia, *NEC* necrotizing enterocolitis, *PROM* premature rupture of membranes, *PIH* pregnancy-induced hypertension, *PMA* post-menstrual age, *SES* socioeconomic status, *NICHD* National Institute of Child Health and Human Development, *CPAP* continuous positive airway pressure, *HFV* high-frequency ventilator, *CV* conventional ventilator, *PDA* patent ductus arteriosus, *IUGR* intrauterine growth restriction, *CAT/CLAMS* Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale, *OR* odds ratio, *HL* Hosmer–Lemeshow, *AUC* area under ROC curve, *ROC* receiver operating characteristic curve, *TIPP* Trial of Indomethacin Prophylaxis in Preterms, *IQ* intelligence quotient, *NI/Neonatal Risk Index*, *NEPSY* Developmental Neuropsychological Assessment, *GMDS* Griffiths Mental Developmental Scales, *CELF-P* Clinical Evaluation of Language Fundamentals, *SD* standard deviation.

of infants born <1250 g predicted significant NDI at 3 years of age.<sup>39</sup> However, a subsequent evaluation in a much larger cohort ( $n = 1328$ ) of preterm infants <29 weeks gestational age concluded that the CRIB-II did no better than gestational age or birth weight alone in predicting moderate to severe functional disability at 2–3 years of age.<sup>40</sup> Studies have supported an association between the SNAP-II and SNAPPE-II scores and neurodevelopmental outcomes and small head circumference at 24 months corrected age. High SNAP-II scores were shown to correlate with adverse neurological, cognitive, and behavioral outcomes up to 10 years of age within a large cohort ( $n = 874$ ) of children born very preterm.<sup>43</sup>

#### Antenatal risk factors

Several groups have used data from the NICHD's Neonatal Research Network (NRN) to design and test various risk prediction models for extremely low birth weight (ELBW) newborns. One of the most widely used risk prediction tools developed from this cohort was by Tyson et al.,<sup>13</sup> using data from ELBW infants in NRN centers between 1998 and 2003. The model includes five variables available prior to delivery—gestational age, estimated birth weight, sex, plurality, and antenatal corticosteroid exposure—to predict risk for death or profound NDI for infants born between 22 and 25 completed weeks gestation. The model has been incorporated into an online calculator that facilitates the counseling of families facing delivery at the margins of viability. It has also been validated by two separate studies—Lee et al.<sup>53</sup> only evaluated the model for predicting death before discharge, but Marrs et al.<sup>38</sup> evaluated the model for risk of death or NDI as well.

#### Postnatal morbidity

A large cohort study ( $n = 910$ ) from Schmidt et al.<sup>15,32</sup> used data from ELBW neonates 500–999 g enrolled in the international Trial of Indomethacin Prophylaxis in Preterms (TIPP). They found that the presence of three morbidities at 36 weeks post-menstrual age—bronchopulmonary dysplasia, serious brain injury, and severe retinopathy of prematurity—had a significant and additive effect on the risk for death or poor neurologic outcome at 18 months corrected age. They developed a model from this relationship that has been corroborated in two studies with smaller samples and by Schmidt et al.<sup>15</sup> in a separate, large cohort in which the definition of poor outcome was expanded from solely NDI to “poor general health.”<sup>33,34</sup>

#### Letting the machines decide

Some innovative work has been recently performed by Ambalavanan et al.<sup>14,35</sup> in creating several risk prediction models.<sup>45</sup> Along with studies developing risk prediction tools with data from the NRN and the TIPP to predict the outcomes of death and NDI or solely NDI, the group made the only risk prediction tool for the outcome of rehospitalization, both general and specifically for respiratory complications, using a combination of physiologic and socioeconomic variables incorporated into a decision tree approach. They have also been the only group to create neural network-trained models, using the same small cohort to predict major handicap, low mental development index (MDI), or low psychomotor development index (PDI). The advantage of using neural networks—algorithms that can “learn” mathematical relationships between a series of independent variables and a set of outcomes—is the ability to model complex or nonlinear relationships that can be elucidated by the model without having to consider these relationships a priori (as is typically required when using multiple regression models). Despite the use of innovative approaches, however, none of these models differed from other studies in predictive strength or even had high predictive efficacy.<sup>31</sup>

#### Limitations of prior approaches

The above literature review highlights the substantial interest in developing a clinically useful risk prediction model and the limits of efforts to date. Notwithstanding their differing inclusion and exclusion criteria, existing risk prediction models are relatively similar in terms of variables selected, outcomes analyzed, and statistical strategies employed. With few exceptions, the limitations of existing risk prediction models are especially apparent in their reliance on solely biologic variables and traditional analytic methods ill-equipped to handle the statistical complexity necessary for risk modeling.

#### Conceptual considerations

*Identifying important outcomes.* The majority of risk prediction models defined NDI as their primary outcome of interest. Making a determination of impairment often relies on standardized measures of cognition in concert with neurosensory deficits. Yet, researchers often define NDI in different ways, making between-study comparisons difficult. NDI is a construct relating to global abilities encompassing cognition, language, motor function, and vision and hearing. While the tools used to identify NDI are often also used to make diagnoses of developmental delay, NDI is not a clinical term or diagnosis in and of itself. Many of the remaining studies also predicted functional outcomes, such as academic performance, executive function, language ability, and autism spectrum disorder (ASD). These outcomes may be more meaningful to parents and providers than NDI.<sup>54</sup>

To date, only four studies have considered outcomes unrelated to neurodevelopment, such as impaired pulmonary function, “poor general health,” and rehospitalization rates.<sup>15,28,45,49</sup> While the emphasis on NDI is unsurprising given the high-risk population, moderate to severe NDI only affects a minority of the preterm population.<sup>55,56</sup> Studies have revealed numerous additional adverse outcomes that preterm individuals are more likely to experience compared to their full-term counterparts, such as impaired respiratory, cardiovascular, and metabolic function.<sup>57–66</sup> Neurodevelopment has been linked to chronic health problems in later childhood.<sup>67</sup> Limiting risk prediction to moderate to severe NDI therefore ignores other, more common complications that preterm infants are likely to face that have an impact on neurodevelopment. This represents a missed opportunity for researchers to better understand what variables influence the likelihood that these problems occur.

The impact of developmental disability on the child and family is completely absent from current risk models. Health-related quality of life (HRQL), which distinguishes itself as a personal rather than third-party valuation of a patient's physical and emotional well-being, is being increasingly appreciated as an important metric necessary to fully understand the impact of prematurity.<sup>68</sup> In a French national survey, the majority of neonatologists, obstetricians, and pediatric neurologists stated that predicting HRQL in the long term for preterm infants would be beneficial for consulting parents about what additional responsibilities they can anticipate in caring for their child.<sup>69</sup> The trajectory of HRQL from childhood to young-adulthood appears to improve in both VLBW and extremely low gestational age populations.<sup>70</sup> Prediction modeling might aid in determining which factors could positively or negatively impact HRQL in this vulnerable population.

Finally, we must consider the age at which outcomes are being predicted. It is evident that lower gestational age is inversely proportional to rates of NDI and academic achievement in adolescence.<sup>71,72</sup> However, the vast majority of risk prediction models assessed outcomes at the age of 3 years or less, with only three studies doing so at 10 years of age or above. Although early childhood outcomes may give clues about later development, many problems do not manifest until later in childhood, such as learning disabilities and certain psychiatric disorders.



Developmental disability severity can fluctuate throughout childhood, with catch-up occurring in early preterm children and worsening delay in some moderate and late preterm children.<sup>73,74</sup> Although cohorts of preterm infants are not usually followed for more than several years, likely due to lack of resources and expense, recent studies have used data from national registries to link neonatal clinical data to sampled adults, providing evidence of increased rates of adverse neurodevelopmental, behavioral, and educational outcomes among adults born preterm.<sup>75,76</sup> Opportunities are therefore available to use long-term data to extend risk prediction models beyond the first few years of life.

**Variable selection.** Most of the risk models reviewed relied primarily on physiologic and clinical measures obtained during the NICU stay. While an emphasis on biologic risk factors is clearly reasonable given the known associations between perinatal morbidities and long-term outcomes, there is strong evidence in the literature suggesting associations between sociodemographic factors like parental race, education, and age, and outcomes such as cognitive impairment, cerebral palsy, and mental health disorders in children born preterm. More specific socioeconomic variables such as lower parental education, maternal income, insurance status, foreign country of birth by a parent, and socioeconomic status as defined by the Elly-Irving Socioeconomic Index have been repeatedly correlated with reduced mental development index, psychomotor development index, intelligence quotient, and social competence throughout childhood.<sup>71,72,77–82</sup> The geographic area in which preterm neonates are raised could also have a profound influence on their development. Neighborhood poverty rate, high school dropout rate, and place of residence (metropolitan vs. non-metropolitan) have all been correlated with academic skills and rate of mental health disorders among low birth weight children.<sup>83,84</sup>

Only 12 of the 43 models reviewed included socioeconomic variables. This may be due, at least in part, to the difficulty in obtaining social, economic, and demographic data; these variables are often not collected upon hospital admission. Additionally, socioeconomic information is often poorly, inaccurately, and variably recorded or is largely missing.<sup>85</sup> Some risk prediction models collected socioeconomic variables at the follow-up visit when outcomes were assessed. This is an imperfect method given that factors such as household setting and family income may change substantially in the years following NICU discharge and affect children's health.<sup>86,87</sup>

In some models, socioeconomic variables were not included because they did not significantly improve the model's predictive ability.<sup>45</sup> Testing the effects of social factors on infant and child outcomes requires samples that are socially and economically diverse. Even large, diverse study populations may become more homogeneous over time, as subjects of lower socioeconomic status and non-white race are more likely to drop out of studies dependent on long-term follow-up.<sup>41</sup> And treating socioeconomic variables as statistically independent factors rather than interrelated might minimize the impact of contextual information on neurodevelopmental outcomes.

#### Statistical considerations

**Model development.** Of the 32 papers included in the review, 12 reported on de novo risk prediction tools. The other 20 studies either evaluated a previous model or adjusted a prior model by changing the times at which data were collected or by adding additional variables. The approach to prediction tool development was almost uniform among the studies, with nine of the models solely using regression techniques to select variables. Ambalavanan et al. deviated from this method in three separate studies: two using classification tree analysis,<sup>35,45</sup> and one using a four-layer back-propagation neural network.<sup>31</sup>

Each new model—with the exception of the neural network-based model by Ambalavanan et al.<sup>35,45</sup>—depended on an approach in which individual variables were selected and treated as independent of one another as they were analyzed in their ability to predict the outcome of interest. Yet, variables may, in fact, not act independently. While parsing the roles of potential interrelationships may be computationally onerous and treating them independently may lead to a more parsimonious model, this may be at the expense of accuracy. Alternative computational approaches are needed to account for the differential likelihoods of certain outcomes on the causal pathway from preterm birth to later childhood outcome. Nonlinear statistical tools should be further utilized in risk prediction model development to examine the relationships between variables and outcomes of interest. Machine learning, for instance, is a method of inputting a group of variables and generating a predictive model without making assumptions of independence between the factors or that specific factors would contribute the most to the model.<sup>88</sup> Different forms of machine learning have already been employed in NICU's to extract the most important variables for predicting outcomes such as days to discharge.<sup>89</sup>

The non-independence of risk factors is also complicated by the role of time in models of human health and development. The life-course framework describes how an accumulation or "chains" of risk experienced over time and at certain critical periods impact later health outcomes.<sup>90</sup> In the context of preterm birth, the risk of being born early is not uniform across populations and dependent on a given set of maternal risks. In turn, the degree of prematurity imparts differential risk for developing complications such as bronchopulmonary dysplasia, necrotizing enterocolitis, or retinopathy of prematurity. These morbidities then, in turn, increase risks for further medical and developmental impairment. These time-varying probabilities can be modeled and incorporated into prediction tools to more accurately capture the longitudinal and varying relationships between exposures and outcomes and improve thereby estimations of risk.<sup>91–93</sup>

A final methodological concern regarding model development is whether and how the competing risk of death is considered when the outcome being predicted is non-terminal. Consider, for example, the task of developing a model for the risk of NDI at 10 years of age. How one handles death can have a dramatic effect on the model, especially since mortality is relatively high among preterm infants. Moreover, if death is treated simply as a censoring mechanism, as it is often done in time-to-event analyses such as those based on the Cox model, then the overall risk of NDI will be artificially reduced; those children who die before being diagnosed with NDI will be viewed as remaining at risk even though they cannot possibly be subsequently diagnosed with NDI. While an alternative to this would be to use a composite outcome of time the first of NDI or death, doing so may result in a model that is unable to predict either event well. Instead, one promising avenue is to frame the development of a prediction model for NDI within the semi-competing risks paradigm.<sup>94,95</sup> Briefly, semi-competing risks refer to settings where one event is a competing risk for the other, but not vice versa. This is distinct from standard competing risks, where each event is competing for the other (e.g., death due to one cause or another). To the best of our knowledge, however, semi-competing risks have not been applied to the study of long-term outcomes among preterm infants.

**Model evaluation.** Waljee et al.<sup>18</sup> provide a summary of methods for assessing the performance of a predictive model, categorizing them into three types: overall model performance, which focuses on the extent of variation in risk explained by the model; calibration, which assesses differences between observed and predicted event rates; and discrimination, which assesses the ability to distinguish between patients who do and do not experience the outcome of interest. The majority of studies in our

review assessed their models with ROC curve analysis, a method of assessing discrimination. While widely used, there is some debate with regard to ROC-based assessments, specifically in regard to its lack of sensitivity in assessing differences between good predictive models.<sup>96</sup> Although several novel performance measures for comparing discrimination among models have been proposed, none have been employed in the context of comparing risk prediction tools for preterm neonates.<sup>97,98</sup>

Few studies employed analyses other than ROC. Only six in our review assessed overall performance with  $R^2$  or partial  $R^2$ , and five evaluated calibration using the Hosmer–Lemeshow test. Another four studies assessed internal validation with either an internal validation set or bootstrapping techniques.<sup>99</sup> There were nine studies meeting inclusion criteria solely because they had models that were externally validated via other studies. Schmidt et al.<sup>32</sup> reported odds ratio associations for their 3-morbidity model, which are not a reliable method of determining the strength of risk prediction tools.<sup>100</sup> Future risk model assessments for preterm neonates should at minimum include an ROC curve analysis, although assessments of overall performance and calibration would also be helpful. Validation with a different sample from the development set is also advised, ideally with a population outside the original cohort.<sup>18</sup>

## CONCLUSION

Risk assessment and outcomes prediction are valuable tools in medical decision-making. Fortunately, infants born prematurely enjoy ever-increasing likelihood of survival. Research over the past several decades has highlighted the many influences, physiologic and psychosocial, affecting neurodevelopment, HRQL, and health services utilization. Yet, the wealth of knowledge gained from longitudinal studies of growth and development is not reflected in current risk prediction models. Moreover, some of the most well-known and widely used tools today, such as Tyson et al.'s<sup>13</sup> five-factor model, were developed nearly two decades ago. As advances in neonatal intensive care progressively reduce the risk of certain outcomes, it is clear that these older models require updating if they are to be of continued clinical use. It should be recognized that there are potential ethical ramifications to incorporating more psychosocial factors and outcomes into risk prediction models, such as crossing the line from risk stratification to “profiling” patients and offering different treatment decisions based on race or class.<sup>101</sup> However, physician predictions without the aid of prediction tools are highly inconsistent during counseling at the margins of viability, and further research is needed regarding the level of influence that physicians actually have on caregiver decision-making during counseling, as well as the extent to which risk prediction tools would change their approach to counseling.<sup>10</sup> In addition, despite recent innovation in statistical approaches to risk modeling, such as machine learning, most prediction tools rely on standard regression techniques. Insofar that risk prediction models will continue to be developed for preterm neonatal care, making use of the clinical data available in most modern electronic health records and taking into consideration the analytic challenges related to unequal prior probabilities of exposures, non-independence of variables, and semi-competing risk can only strengthen our approach to predicting outcomes. We therefore recommend taking a broader view of risk, incorporating these concepts in creating stronger risk prediction tools that can ultimately serve to benefit the long-term care of preterm neonates.

## AUTHOR CONTRIBUTIONS

C.J.C. and J.S.L. designed and carried out this literature review. C.J.C., J.S.L., and S.H. worked jointly in the analysis and interpretation of the literature review results, as

well as the drafting and revision of this article. All three authors gave final approval of the version to be published.

## ADDITIONAL INFORMATION

**Competing interests:** The authors declare no competing interests.

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

## REFERENCES

- Little, W. J. On the influence of abnormal parturition, difficult labours, premature birth, and asphyxia neonatorum, on the mental and physical condition of the child, especially in relation to deformities. *Clin. Orthop. Relat. Res.* **46**, 7–22 (1966).
- Stoll, B. J. et al. Trends in care practices, morbidity, and mortality of extremely preterm neonates, 1993–2012. *JAMA* **314**, 1039–1051 (2015).
- Patel, R. M., Rysavy, M. A., Bell, E. F. & Tyson, J. E. Survival of infants born at periviable gestational ages. *Clin. Perinatol.* **44**, 287–303 (2017).
- Allen, M. C., Cristofalo, E. A. & Kim, C. Outcomes of preterm infants: morbidity replaces mortality. *Clin. Perinatol.* **38**, 441–454 (2011).
- Institute of Medicine Committee on Understanding Premature Birth and Assuring Healthy Outcomes. *Preterm Birth: Causes, Consequences, and Prevention* (National Academies Press, Washington, 2007).
- Shapiro, S. Influence of birth weight, sex, and plurality on neonatal loss in United States. *Am. J. Public Health Nations Health* **44**, 1142–1153 (1954).
- Manuck, T. A. et al. Preterm neonatal morbidity and mortality by gestational age: a contemporary cohort. *Am. J. Obstet. Gynecol.* **215**, 103.e1–103.e14 (2016).
- Salas, A. A. et al. Gestational age and birthweight for risk assessment of neurodevelopmental impairment or death in extremely preterm infants. *Arch. Dis. Child Fetal Neonatal Ed.* **101**, F494–F501 (2016).
- Mikkola, K. et al. Neurodevelopmental outcome at 5 years of age of a national cohort of extremely low birth weight infants who were born in 1996–1997. *Pediatrics* **116**, 1391–1400 (2005).
- Tucker Edmonds, B., McKenzie, F. M., Panoch, J. E., Barnato, A. E. & Frankel, R. M. Comparing neonatal morbidity and mortality estimates across specialty in periviable counseling. *J. Matern. Fetal Neonatal Med.* **28**, 2145–2149 (2015).
- Moons, K. G., Royston, P., Vergouwe, Y., Grobbee, D. E. & Altman, D. G. Prognosis and prognostic research: what, why, and how? *BMJ* **338**, b375 (2009).
- Dorling, J. S., Field, D. J. & Manktelow, B. Neonatal disease severity scoring systems. *Arch. Dis. Child Fetal Neonatal Ed.* **90**, F11–F16 (2005).
- Tyson, J. E., Parikh, N. A., Langer, J., Green, C. & Higgins, R. D. Intensive care for extreme prematurity—moving beyond gestational age. *N. Engl. J. Med.* **358**, 1672–1681 (2008).
- Ambalavanan, N. et al. Outcome trajectories in extremely preterm infants. *Pediatrics* **130**, e115–e125 (2012).
- Schmidt, B. et al. Prediction of late death or disability at age 5 years using a count of 3 neonatal morbidities in very low birth weight infants. *J. Pediatr.* **167**, 982–986.e2 (2015).
- Medlock, S., Ravelli, A. C. J., Taminga, P., Mol, B. W. M. & Abu-Hanna, A. Prediction of mortality in very premature infants: a systematic review of prediction models. *PLoS ONE* **6**, e23441 (2011).
- 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).
- Waljee, A. K., Higgins, P. D. & Singal, A. G. A primer on predictive models. *Clin. Transl. Gastroenterol.* **5**, e44 (2014).
- Jobe, A. H. Pulmonary surfactant therapy. *N. Engl. J. Med.* **328**, 861–868 (1993).
- Willson, D. F. & Notter, R. H. The future of exogenous surfactant therapy. *Respir. Care* **56**, 1369–1388 (2011).
- Lefebvre, F., Grégoire, M. C., Dubois, J. & Glorieux, J. Nursery Neurobiologic Risk Score and outcome at 18 months. *Acta Paediatr.* **87**, 751–757 (1998).
- Wickremasinghe, A. C. et al. Evaluation of the ability of neurobiological, neurodevelopmental and socio-economic variables to predict cognitive outcome in premature infants. *Child Care Health Dev.* **38**, 683–689 (2012).
- Emsley, H. C. A., Wardle, S. P., Sims, D. G., Chiswick, M. L. & D'Souza, S. W. Increased survival and deteriorating developmental outcome in 23 to 25 week old gestation infants, 1990–4 compared with 1984–9. *Arch. Dis. Child Fetal Neonatal Ed.* **78**, F99–F104 (1998).
- Fowle, P. W., Gould, C. R., Tamow-Mordi, W. O. & Strang, D. Measurement properties of the Clinical Risk Index for Babies—reliability, validity beyond the first 12 hours, and responsiveness over 7 days. *Crit. Care Med.* **26**, 163–168 (1998).

25. Fowlie, P. W., Tarnow-Mordi, W. O., Gould, C. R. & Strang, D. Predicting outcome in very low birthweight infants using an objective measure of illness severity and cranial ultrasound scanning. *Arch. Dis. Child Fetal Neonatal Ed.* **78**, F175–F178 (1998).
26. Lago, P. et al. Is the CRIB score (Clinical Risk Index for babies) a valid tool in predicting neurodevelopmental outcome in extremely low birth weight infants? *Biol. Neonate* **76**, 220–227 (1999).
27. Bühner, C., Grimmer, I., Metzke, B. & Obladen, M. The CRIB (Clinical Risk Index for Babies) score and neurodevelopmental impairment at one year corrected age in very low birth weight infants. *Intens. Care Med.* **26**, 325–329 (2000).
28. Eriksson, M., Bodin, L., Finnström & Schollin, J. Can severity-of-illness indices for neonatal intensive care predict outcome at 4 years of age? *Acta Paediatr.* **91**, 1093–1100 (2002).
29. Charkaluk, M. L. et al. Neurodevelopment of children born very preterm and free of severe disabilities: the Nord-Pas de Calais Epipage cohort study. *Acta Paediatr.* **99**, 684–689 (2010).
30. Mattia, F. R. & deRegnier, R. O. Chronic physiologic instability is associated with neurodevelopmental morbidity at one and two years in extremely premature infants. *Pediatrics* **102**, E35 (1998).
31. Ambalavanan, N. et al. Prediction of neurologic morbidity in extremely low birth weight infants. *J. Perinatol.* **20**, 496–503 (2000).
32. Schmidt, B. et al. Impact of bronchopulmonary dysplasia, brain injury, and severe retinopathy on the outcome of extremely low-birth-weight infants at 18 months: results from the trial of indomethacin prophylaxis in preterms. *JAMA* **289**, 1124–1129 (2003).
33. Farooqi, A., Hägglöf, B., Sedin, G. & Serenius, F. Impact at age 11 years of major neonatal morbidities in children born extremely preterm. *Pediatrics* **127**, e1247–e1257 (2011).
34. Koo, K. Y. et al. Effect of severe neonatal morbidities on long term outcome in extremely low birthweight infants. *Korean J. Pediatr.* **53**, 694–700 (2010).
35. Ambalavanan, N. et al. Early prediction of poor outcome in extremely low birth weight infants by classification tree analysis. *J. Pediatr.* **148**, 438–444 (2006).
36. Taylor, H. G., Klein, N., Drotar, D., Schluchter, M. & Hack, M. Consequences and risks of <1000-g birth weight for neuropsychological skills, achievement, and adaptive functioning. *J. Dev. Behav. Pediatr.* **27**, 459–469 (2006).
37. Broitman, E. et al. Clinical data predict neurodevelopmental outcome better than head ultrasound in extremely low birth weight infants. *J. Pediatr.* **151**, 500–505.e1–2 (2007).
38. Marrs, C. C., Pedroza, C., Mendez-Figueroa, H., Chauhan, S. P. & Tyson, J. E. Infant outcomes after periviable birth; external validation of the neonatal research network estimator with the BEAM Trial. *Am. J. Perinatol.* **33**, 569–576 (2016).
39. Lodha, A., Sauv e, R., Chen, S., Tang, S. & Christianson, H. Clinical Risk Index for Babies score for the prediction of neurodevelopmental outcomes at 3 years of age in infants of very low birthweight. *Dev. Med. Child Neurol.* **51**, 895–900 (2009).
40. Greenwood, S., Abdel-Latif, M. E., Bajuk, B. & Lui, K., NSW and ACT Neonatal Intensive Care Units Audit Group. Can the early condition at admission of a high-risk infant aid in the prediction of mortality and poor neurodevelopmental outcome? A population study in Australia. *J. Paediatr. Child Health* **48**, 588–595 (2012).
41. Johnson, S. et al. Autism spectrum disorders in extremely preterm children. *J. Pediatr.* **156**, 525–531.e2 (2010).
42. Dammann, O. et al. SNAP-II and SNAPPE-II and the risk of structural and functional brain disorders in extremely low gestational age newborns: the ELGAN study. *Neonatology* **97**, 71–82 (2010).
43. Logan, J. W. et al. Early postnatal illness severity scores predict neurodevelopmental impairments at 10 years of age in children born extremely preterm. *J. Perinatol.* **37**, 606–614 (2017).
44. Foster-Cohen, S. H., Friesen, M. D., Champion, P. R. & Woodward, L. J. High prevalence/low severity language delay in preschool children born very preterm. *J. Dev. Behav. Pediatr.* **31**, 658–667 (2010).
45. Ambalavanan, N. et al. Identification of extremely premature infants at high risk of rehospitalization. *Pediatrics* **128**, e1216–e1225 (2011).
46. Stephens, B. E. et al. Screening for autism spectrum disorders in extremely preterm infants. *J. Dev. Behav. Pediatr.* **33**, 535–541 (2012).
47. Potharst, E. S. et al. Perinatal risk factors for neurocognitive impairments in preschool children born very preterm. *Dev. Med. Child Neurol.* **55**, 178–184 (2013).
48. Manuck, T. A., Sheng, X., Yoder, B. A. & Varner, M. W. Correlation between initial neonatal and early childhood outcomes following preterm birth. *Am. J. Obstet. Gynecol.* **210**, 426.e1–9 (2014).
49. Keller, R. L. et al. Bronchopulmonary dysplasia and perinatal characteristics predict 1-year respiratory outcomes in newborns born at extremely low gestational age: a prospective cohort study. *J. Pediatr.* **187**, 89–97.e3 (2017).
50. The International Neonatal Network. The CRIB (Clinical Risk Index for Babies) score: a tool for assessing initial neonatal risk and comparing performance of neonatal intensive care units. *Lancet* **342**, 193–198 (1993).
51. Richardson, D. K., Gray, J. E., McCormick, M. C., Workman, K. & Goldmann, D. A. Score for neonatal acute physiology: a physiologic severity index for neonatal intensive care. *Pediatrics* **91**, 617–623 (1993).
52. Gray, J. E., Richardson, D. K., McCormick, M. C., Workman-Daniels, K. & Goldmann, D. A. Neonatal therapeutic intervention scoring system: a therapy-based severity-of-illness index. *Pediatrics* **90**, 561–567 (1992).
53. Lee, H. C. et al. Prediction of death for extremely premature infants in a population-based cohort. *Pediatrics* **126**, 644–650 (2010).
54. Jaworski, M., Janvier, A., Lefebvre, F. & Luu, T. M. Parental perspectives regarding outcomes of very preterm infants: toward a balanced approach. *J. Pediatr.* **200**, 58–63.e1 (2018).
55. Kerstjens, J. M., De Winter, A. F., Bocca-Tjeertes, I. F., Bos, A. F. & Reijneveld, S. A. Risk of developmental delay increases exponentially as gestational age of preterm infants decreases: a cohort study at age 4 years. *Dev. Med. Child Neurol.* **54**, 1096–1101 (2012).
56. Blencowe, H. et al. Preterm birth-associated neurodevelopmental impairment estimates at regional and global levels for 2010. *Pediatr. Res.* **74**, 17–34 (2013).
57. Greenough, A. Late respiratory outcomes after preterm birth. *Early Hum. Dev.* **83**, 785–788 (2007).
58. Hennessy, E. M. et al. Respiratory health in pre-school and school age children following extremely preterm birth. *Arch. Dis. Child* **93**, 1037–1043 (2008).
59. Jaakkola, J. J. et al. Preterm delivery and asthma: a systematic review and meta-analysis. *J. Allergy Clin. Immunol.* **118**, 823–830 (2006).
60. Sonnenschein-van der Voort, A. M. et al. Preterm birth, infant weight gain, and childhood asthma risk: a meta-analysis of 147,000 European children. *J. Allergy Clin. Immunol.* **133**, 1317–1329 (2014).
61. Luu, T. M., Katz, S. L., Leeson, P., Th ebaud, B. & Nuyt, A. Preterm birth: risk factor for early-onset chronic diseases. *CMAJ* **188**, 736–746 (2016).
62. Lewandowski, A. J. et al. Preterm heart in adult life: cardiovascular magnetic resonance reveals distinct differences in left ventricular mass, geometry, and function. *Circulation* **127**, 197–206 (2013).
63. Lewandowski, A. J. et al. Right ventricular systolic dysfunction in young adults born preterm. *Circulation* **128**, 713–720 (2013).
64. Lewandowski, A. J. et al. Elevated blood pressure in preterm-born offspring associates with a distinct antiangiogenic state and microvascular abnormalities in adult life. *Hypertension* **65**, 607–614 (2015).
65. Parkinson, J. R. et al. Preterm birth and the metabolic syndrome in adult life: a systematic review and meta-analysis. *Pediatrics* **131**, e1240–e1263 (2013).
66. Kopec, G., Shekawat, P. S. & Mhanna, M. J. Prevalence of diabetes and obesity in association with prematurity and growth restriction. *Diabetes Metab. Syndr. Obes.* **10**, 285–295 (2017).
67. McCormick, M. C., Litt, J. S., Smith, V. C. & Zupancic, J. A. Prematurity: an overview and public health implications. *Annu. Rev. Public Health* **32**, 367–379 (2011).
68. Saigal, S. & Tyson, J. Measurement of quality of life of survivors of neonatal intensive care: critique and implications. *Semin. Perinatol.* **32**, 59–66 (2008).
69. Einaudi, M. A., Gire, C., Loundou, A., Le Coz, P. & Auquier, P. Quality of life assessment in preterm children: physicians' knowledge, attitude, belief, practice —a KABP study. *BMC Pediatr.* **13**, 58 (2013).
70. Vederhus, B. J. et al. Health-related quality of life and emotional and behavioral difficulties after extreme preterm birth: developmental trajectories. *PeerJ* **3**, e738 (2015).
71. Linsell, L., Malouf, R., Morris, J., Kurinczuk, J. J. & Marlow, N. Prognostic factors for poor cognitive development in children born very preterm or with very low birth weight: a systematic review. *JAMA Pediatr.* **169**, 1162–1172 (2015).
72. Linsell, L., Malouf, R., Morris, J., Kurinczuk, J. J. & Marlow, N. Prognostic factors for cerebral palsy and motor impairment in children born very preterm or very low birthweight: a systematic review. *Dev. Med. Child Neurol.* **58**, 554–569 (2016).
73. Luu, T. M., Vohr, B. R., Allan, W., Schneider, K. C. & Ment, L. R. Evidence for catch-up in cognition and receptive vocabulary among adolescents born very preterm. *Pediatrics* **128**, 313–322 (2011).
74. Johnston, K. M. et al. The economic burden of prematurity in Canada. *BMC Pediatr.* **14**, 93 (2014).
75. Hintz, S. R., Newman, J. E. & Vohr, B. R. Changing definitions of long-term follow-up: should "long term" be even longer? *Semin. Perinatol.* **40**, 398–409 (2016).
76. Saigal, S. Functional outcomes of very premature infants into adulthood. *Semin. Fetal Neonatal Med.* **19**, 125–130 (2014).
77. Jones, K. M., Champion, P. R. & Woodward, L. J. Social competence of preschool children born very preterm. *Early Hum. Dev.* **89**, 795–802 (2013).
78. Potharst, E. S. et al. Prediction of cognitive abilities at the age of 5 years using developmental follow-up assessments at the age of 2 and 3 years in very preterm children. *Dev. Med. Child Neurol.* **54**, 240–246 (2012).

79. Hansen, B. M. et al. Perinatal risk factors of adverse outcome in very preterm children: a role of initial treatment of respiratory insufficiency? *Acta Paediatr.* **93**, 185–189 (2004).
80. Messinger, D. et al. The relationship between behavior ratings and concurrent mental and motor performance in toddlers born at extremely low birth weight. *J. Early Inter.* **32**, 214–233 (2010).
81. Linsell, L. et al. Prognostic factors for behavioral problems and psychiatric disorders in children born very preterm or very low birth weight: a systematic review. *J. Dev. Behav. Pediatr.* **37**, 88–102 (2016).
82. Vohr, B. R., Wright, L. L., Poole, W. K. & McDonald, S. A. Neurodevelopmental outcomes of extremely low birth weight infants <32 weeks' gestation between 1993 and 1998. *Pediatrics* **116**, 635–643 (2005).
83. Andreias, L. et al. Neighborhood influences on the academic achievement of extremely low birth weight children. *J. Pediatr. Psychol.* **35**, 275–283 (2010).
84. Singh, G. K., Kenney, M. K., Ghandour, R. M., Kogan, M. D. & Lu M. C. Mental health outcomes in US children and adolescents born prematurely or with low birthweight. *Depress Res. Treat.* **2013** 570743 (2013).
85. Shavers, V. L. Measurement of socioeconomic status in health disparities research. *J. Natl Med. Assoc.* **99**, 1013–1023 (2007).
86. Kendzor, D. E., Caughy, M. O. & Owen, M. T. Family income trajectory during childhood is associated with adiposity in adolescence: a latent class growth analysis. *BMC Public Health* **12**, 611 (2012).
87. Poonawalla, I. B., Kendzor, D. E., Owen, M. T. & Caughy, M. O. Family income trajectory during childhood is associated with adolescent cigarette smoking and alcohol use. *Addict. Behav.* **39**, 1383–1388 (2014).
88. Waljee, A. K. & Higgins, P. D. Machine learning in medicine: a primer for physicians. *Am. J. Gastroenterol.* **105**, 1224–1226 (2010).
89. Temple, M. W., Lehmann, C. U. & Fabbri, D. Predicting discharge dates from the NICU using progress note data. *Pediatrics* **136**, e395–e405 (2015).
90. Kuh, D. & Ben-Shlomo, Y. *A Life Course Approach to Chronic Diseases Epidemiology* 2nd edn. A Life Course Approach to Adult health (Oxford University Press, New York, 2004).
91. Amrhein, V., Greenland, S. & McShane, B. Scientists rise up against statistical significance. *Nature* **567**, 305–307 (2019).
92. Wasserstein, R. L. & Lazar, N. A. The ASA's Statement on *p*-values: context, process, and purpose. *Am. Statistician* **70**, 129–133 (2016).
93. Lewis, R. J. & Angus, D. C. Time for clinicians to embrace their inner Bayesian? Reanalysis of results of a clinical trial of extracorporeal membrane oxygenation. *JAMA* **320**, 2208–2210 (2018).
94. Haneuse, S. & Lee, K. H. Semi-competing risks data analysis: accounting for death as a competing risk when the outcome of interest is nonterminal. *Circ. Cardiovasc. Qual. Outcomes* **9**, 322–331 (2016).
95. Jazić, I., Schrag, D., Sargent, D. J. & Haneuse, S. Beyond composite endpoints analysis: semicompeting risks as an underutilized framework for cancer research. *J. Natl Cancer Inst.* **108**, 12 (2016).
96. Cook, N. R. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* **115**, 928–935 (2007).
97. Steyerberg, E. W. et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* **21**, 128–138 (2010).
98. Pencina, M. J., D'Agostino, R. B. & Demler, O. V. Novel metrics for evaluating improvement in discrimination: net reclassification and integrated discrimination improvement for normal variables and nested models. *Stat. Med.* **31**, 101–113 (2012).
99. Harrell, F. E., Lee, K. L. & Mark, D. B. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat. Med.* **15**, 361–387 (1996).
100. Pepe, M. S., Janes, H., Longton, G., Leisenring, W. & Newcomb, P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am. J. Epidemiol.* **159**, 882–890 (2004).
101. Bærøe, K. & Bringedal, B. Just health: on the conditions for acceptable and unacceptable priority settings with respect to patients' socioeconomic status. *J. Med. Ethics* **37**, 526–529 (2011).