Social Science & Medicine 380 (2025) 118181

Contents lists available at ScienceDirect

Social Science & Medicine

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

A heterogeneity analysis of health-related quality of life in early adults born very preterm or very low birthweight across the sociodemographic spectrum

Corneliu Bolbocean [a,*], Peter J. Anderson [b,c], Peter Bartmann [d], Jeanie L.Y. Cheong [e], Lex W. Doyle [e,f], Samantha Johnson [g], Neil Marlow [h], Dieter Wolke [i], Stavros Petrou [a], Stephen O'Neill [j]

[a] Nuffield Department of Primary Care Health Sciences, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK
[b] Turner Institute for Brain and Mental Health, School of Psychological Sciences, Monash University, Melbourne, VIC, Australia
[c] Clinical Sciences, Murdoch Children's Research Institute, Melbourne, VIC, Australia
[d] Department of Neonatology and Paediatric Intensive Care, University Hospital Bonn, Children's Hospital, Bonn, Germany
[e] Department of Obstetrics and Gynaecology, Clinical Sciences, Murdoch Children's Research Institute, Newborn Services, Royal Women's Hospital, University of Melbourne, Melbourne, VIC, Australia
[f] Department of Paediatrics, The University of Melbourne, Melbourne, VIC, Australia
[g] Department of Health Sciences, University of Leicester, Leicester, UK
[h] UCL Elizabeth Garrett Anderson Institute for Women's Health, University College London, London, UK
[i] Department of Psychology, Warwick Medical School, University of Warwick and Division of Health Sciences, Warwick, UK
[j] London School of Hygiene and Tropical Medicine, London, UK

ARTICLE INFO

Handling editor: Social Epidemiology Office

Keywords:
Heterogeneity of treatment effects
Machine learning
Preterm birth
HRQoL
Quality of life

ABSTRACT

Preterm birth and very low birthweight (VP/VLBW) are associated with poorer health-related quality of life (HRQoL) outcomes extending into adulthood, yet it remains unclear how these effects differ across sociodemographic subgroups. This study aimed to identify heterogeneity in the association of VP/VLBW on HRQoL in early adulthood, specifically examining maternal age, education, and ethnicity. Individual-level data from three longitudinal cohorts within the Research on European Children and Adults Born Preterm Consortium were analysed, including adults born VP (< 32 weeks' gestation) or VLBW (< 1500g), compared to term-born or normal birthweight controls. HRQoL was assessed using the Health Utilities Index Mark 3 (HUI3) at mean ages of 18–26 years. Bayesian Causal Forest and Shrinkage Bayesian Causal Forest methodologies were employed to estimate conditional average treatment effects. Results indicated significant heterogeneity in the effects of VP/VLBW birth on HRQoL by maternal age and education. Individuals born to mothers aged ≤25 years experienced the largest decrement in HUI3 scores (−0.08; 95 % CI −0.13, −0.02), compared to minimal or no decrements for individuals born to mothers aged ≥26 years. Similarly, lower maternal education was associated with larger decrements (−0.05; 95 % CI −0.09, −0.01), whereas high maternal education showed negligible impact (0.01; 95 % CI −0.04, 0.06). These findings highlight maternal sociodemographic characteristics as critical modifiers of VP/VLBW impacts on adult HRQoL, emphasizing the need for targeted health interventions for disadvantaged groups. Future research is warranted to examine whether modern neonatal care and changes in socioeconomic conditions can mitigate these HRQoL disparities across the life course.

1. Introduction

Very preterm birth (VP; < 32 weeks' gestation) or having very low birthweight (VLBW; < 1500 g) are associated with increased mortality (Hovi et al., 2016; Parkinson et al., 2013; Cooper et al., 2009), adverse neurodevelopmental outcomes (Doyle et al., 2021; Pascal et al., 2018; Serenius et al., 2013), and greater socio-economic disadvantage extending into early to mid-adulthood (van der Pal et al., 2020; Lund

* Corresponding author.
E-mail address: Corneliu.Bolbocean@phc.ox.ac.uk (C. Bolbocean).

https://doi.org/10.1016/j.socscimed.2025.118181
Received 3 November 2024; Received in revised form 5 May 2025; Accepted 7 May 2025
Available online 10 May 2025

et al., 2012; Allen et al., 2010; Moster et al., 2008) compared with birth at term (≥37 weeks) or with normal birthweight (≥2500 g). Prematurity is a growing public health concern as increasing preterm birth rates coupled with improvements in survival rates place increased pressures on healthcare budgets worldwide (Beam et al., 2020; Petrou et al., 2019; Horvath et al., 2017) and impose broader social costs (Treyvaud et al., 2014; Misund et al., 2014; Petrou, 2003; Petrou et al., 2001).

Preference-based health-related quality of life (HRQoL) measures are standardized, multidimensional health state classifications that include preference or utility weights derived from representative population samples. Among these measures, the Health Utilities Index Mark 3 (HUI3) is a dominant measure due to its robust psychometric properties (Bolbocean et al., 2023; Kwon et al., 2023) and it is the most widely used preference-based HRQoL measure in children (Kwon et al., 2018, 2019, 2022). Evidence from a recent meta-analysis of longitudinal prospective cohort studies indicates that VP/VLBW status is associated with a significant decrement in the HUI3 multi-attribute utility score of −0.06 (Bolbocean et al., 2022) since the minimally important difference (MID) for the HUI3 measure derived from empirical evidence is 0.03 (Samsa et al., 1999; Drummond, 2001; Grootendorst et al., 2000; Feeny et al., 2002; Norman et al., 2003; Drummond et al., 2015; Brazier et al., 2017; Noto and Uemura, 2020). However, little is known about the heterogeneity of effects of VP/VLBW status across different levels of sociodemographic factors and socioeconomic status (SES), the latter understood as a holistic construct that includes the following markers: parental education, income, and type of employment/occupation (Linsell et al., 2015).

Lower SES is a well-documented risk factor for preterm birth (Wallace et al., 2016). Research from the past two decades underscores that socioeconomic pathways are deeply intertwined with both the occurrence of preterm birth and the outcomes that follow. Maternal stress, inadequate prenatal care, and poor nutrition are key SES-related factors that increase the risk of preterm birth (Dolatian et al., 2018). These same factors (and the broader SES context they represent) contribute to differential decrements in HRQoL for those born preterm with pronounced effects on physical health, cognitive/mental development, and social functioning (Kim et al., 2023). However, not all preterm infants are exposed to the same level of adversity; those born the most preterm or with medical complications and those from more disadvantaged backgrounds tend to have greater long-term impairments in HRQoL, whereas preterm children with fewer neonatal complications and from more affluent families tend to fare better (Call et al., 2024; Brumbaugh et al., 2025). Identifying these SES-related pathways and their impact on life-course outcomes has important implications: it suggests that improving access to prenatal care, reducing maternal stress (through economic and social support), and ensuring adequate nutrition for pregnant women could not only reduce the incidence of preterm birth but also improve the health and HRQoL of preterm infants as they mature. Addressing socioeconomic disparities is seen as vital to improving physical, mental, and social outcomes for those born preterm.

Significant educational and racial ethnic disparities in preterm and low birthweight births have been documented in high-income countries (Fuchs et al., 2018; Tamura et al., 2018; Echevarria and Lorch, 2022). Parental education is a component of parental SES and is the single best predictor of academic and cognitive outcomes (Linsell et al., 2015; Madzwamuse et al., 2015; Benavente-Fernandez et al., 2019) in preterm-born children. SES is itself a transgenerational transmission mechanism to children (Wolke, 2019). For example, poorer maternal education is associated with poorer health behaviour before and during pregnancy, and beyond (Wolke, 2019; D'Souza et al., 2020); low parental education is not only associated with more disadvantage in housing, income, occupation, and neighborhood quality, but also with poorer access to high-quality education (Linsell et al., 2015; Wolke, 2019; D'Souza et al., 2020). While education is only a crude proxy for economic well-being (Braveman et al., 2001; Adler and Rehkopf, 2008), ethnicity (Williams et al., 2010; Bhatti et al., 2013), and age (McMaughan et al., 2020) are frequently used as indirect proxies for socioeconomic factors in health services research. To the extent that our included variables proxy for these broader socioeconomic factors, our estimation approach will allow us to capture heterogeneity in the effects of pre-term/VLBW birth on HRQoL attributable to these factors. Data limitations prevent us from considering these factors directly.

It remains unclear whether the impact of preterm birth on health-related quality of life (HRQoL) in early adulthood is more adverse for certain subgroups or mitigated in others—that is, whether the effect varies across proxies for socioeconomic status such as maternal education, age, or race/ethnicity. The value of analysing heterogeneous effects to support clinicians and decision makers has long been acknowledged, yet studies still mainly focus on associations or average treatment effects (Beam et al., 2020; Petrou et al., 2001, 2019; Horvath et al., 2017; Treyvaud et al., 2014; Misund et al., 2014; Petrou, 2003; Bolbocean et al., 2022). To address this when analysing the impact of VP/VLBW status on preference-based HRQoL outcomes, we report effect heterogeneity for subgroups categorised by markers of parental SES in addition to previously published results of effects (Bolbocean et al., 2022).

For the evaluation of heterogeneous effects, several theoretical frameworks have been suggested (Imai and Ratkovic, 2013; Grimmer et al., 2017; Hainmueller et al., 2019; Hu, 2023). However, each framework has its limitations. Traditional parametric methods that employ interaction terms provide a direct way to estimate heterogeneous treatment effects. However, these methods are limited because of the interdependence of variables, especially when several interaction terms are used. This issue can reduce the usefulness of the analysis (Elek and Bíró, 2021). The robustness of results obtained from interaction analysis can be compromised by model mis-specification (Hainmueller et al., 2019; Hu, 2023; Baranger et al., 2023). Subgroup analysis is prone to producing spurious findings (Assmann et al., 2000; Cook et al., 2004; Steyerberg, 2009) due to its tendency to be underpowered (Hainmueller et al., 2019; Petticrew et al., 2012) and its susceptibility to misinterpretation of random variation as significant treatment effects (Rothwell, 2005; Xie et al., 2012; Beckr et al., 2013; Davis and Heller, 2017).

Random Forest methods, that combine multiple decision trees are a popular machine learning method for predicting outcomes and have been shown to perform well in a range of settings. Decision trees recursively split the data such that individuals within groups defined by the splits are as similar as possible in terms of outcomes, while being as different as possible from those in other groups (Breiman, 2001).

Causal forests (CF) extend this approach to the task of causal inference, by considering splits that maximise heterogeneity of effect estimates, rather than outcomes (Wager and Athey, 2018). An attractive feature of forest-based methods is that recursive splitting allows for complex non-linearities without requiring the researcher to predetermine any functional form (Wager and Athey, 2018). To avoid overfitting and ensure valid confidence intervals, causal forests employ a technique known as "honesty" or cross-fitting, which keeps the data used to construct the tree (i.e. the splits) separate from the data used to estimate treatment effects (Wager and Athey, 2018; Athey and Imbens, 2016). The CF approach has been successfully applied in a range of fields—including healthcare decision making—to identify which patients stand to benefit most (or least) from particular interventions (Xie et al., 2012; Beckr et al., 2013; Jin et al., 2019; Bonander and Svensson, 2021; Sadique et al., 2022). In this analysis, we utilize the Bayesian Causal Forest (BCF)(Hahn et al., 2020) with its separate priors for regularization, alongside the Shrinkage Bayesian Causal Forest (SBCF) (Caron et al., 2022), which addresses the BCF's use of potentially suboptimal uniform splitting probabilities.

The primary objective of this study was to assess heterogeneity in the association between VP/VLBW birth and preference-based HRQoL outcomes in early adulthood, by pooling harmonized data from three prospective longitudinal birth cohort studies using Bayesian Causal Forest approaches. Specifically, we aimed to examine how this association varies with respect to key sociodemographic covariates, including maternal education, maternal age, and ethnicity.

2. Methods

2.1. Data

For inclusion in this analysis, prospective cohort studies were required to (1) have measured self-reported HRQoL in adulthood (defined as age ≥18 years (Arnett, 2015) of individuals born VP/VLBW using the HUI3; (2) included a comparison control group of term-born or normal birthweight individuals, and (3) contributed data to the Research on European Children and Adults Born Preterm (RECAP) Consortium (www.recap-preterm.eu), a database of cohorts of individuals born VP/VLBW. Eligible cohorts were identified by a systematic review studies of preference-based HRQoL outcomes following preterm birth or low birthweight (Petrou et al., 2020). The review identified three cohorts that measured HRQoL using the HUI3 in adulthood of individuals born VP/VLBW and a comparison group: The Bavarian Longitudinal Study (BLS) with birth years 1985-1986[41], The Victorian Infant Collaborative Study (VICS) with birth years 1991–1992 (Doyle et al., 2015) and the EPICure Study with birth year 1995 (Linsell et al., 2018). Included studies were designed to investigate the impact of VP/VLBW status on health and developmental outcomes (Darlow et al., 2020) and had received country-specific ethical reviews along with parental consent at birth and participants' written informed consent in adulthood.

This study uses records from the start of data collection (birth/antenatal) up to the most recent assessments in early adulthood (BLS at 26 years, VICS at 18 years, and EPICure at 19 years) for the main analyses. Table 1 details the background characteristics of the samples in each cohort, including eligibility criteria for VP/VLBW, at term or normal birthweight controls, age(s) of assessment in adulthood and the composition of control groups. Additional details for each study can be found in published research as follows: BLS (Madzwamuse et al., 2015), VICS (Doyle et al., 2015), and EPICure (Linsell et al., 2018). Harmonization dictionaries were developed to guide harmonization of all variables analysed and available at *https://recap − preterm.eu/* or could be found in the Supplementary Material.

Table 1
Summary statistics of background characteristics of cohorts.

	BLS (Germany)		VICS (Australia)		EPICure (UK&Ireland)	
	VPT/ VLBW	Controls	EP/ ELBW	Controls	EP	Controls
Number completing MAUI	231	224	186	137	110	62
Age at assessment (years) Mean (SD)	26.3 (0.68)	26.3 (0.69)	17.9 (0.78)	18.1 (0.88)	19.3 (0.56)	19.2 (0.54)
GA at birth (weeks), Mean (SD)	30.6 (2.2)	39.7 (1.2)	26.7 (2.1)	39.2 (1.4)	24.5 (0.7)	N/A
Birth weight (gr), Mean (SD)	1330 (320)	3360 (448)	887 (155)	3419 (468)	746 (119)	N/A
Sex, n (%) male	125 (54.1)	105 (46.9)	84 (45.2)	60 (43.8)	51 (46.4)	23 (37.1)
Study Name	Bavarian Longitudinal Study		Victorian Infant Collaborative Study		EPICure	
Birth Year	1985–1986		1991–1992		1995	
Eligibility Criteria VP/VPBW	VPT\VLBW (GA < 32wk or BW < 1500g)		EPT\ELBW (GA < 28wk or BW < 1000g)		EPT (GA<26wk)	
Controls	Recruited in the same obstetric hospitals.		Normal birth weight, contemporaneously recruited.		Recruited at school age.	

Notes: VP = Very Preterm (< 32 weeks), VLBW = Very Low Birth Weight (< 1500 g birth weight), EP = Extremely Preterm (< 26 weeks for EPICure and < 28 weeks for VICS), ELBW = Extremely Low Birth Weight (< 1000 g birth weight). NA = Not applicable/measurement not administered, MAUI – multi-attribute utility index. GA - gestational age.

2.2. Outcome variables

This study examined the HUI3 multi-attribute utility score as the outcome of interest. Study participants in the BLS, VICS and EPICure cohorts completed the Health Utilities Index 15-item questionnaire for usual health status assessment, which was obtained from the Health Utilities Index developers and covers the HUI3 health status classification system. The HUI3 was developed to describe HRQoL in general population and clinical contexts, and comprises eight attributes: ambulation, dexterity, cognition, vision, hearing, speech, emotion, and pain (Furlong et al., 1998, 2001; Feeny et al., 1995). The level of function within each attribute is scored on a 5- or 6-point scale ranging from normal/optimal function to severe impairment. Responses are mapped onto an eight-attribute health status vector. Algorithms reflecting the preferences of the general public for the HUI3 health states can be used to convert responses to the measure's eight attributes into multi-attribute utility scores. We applied the Canadian algorithms (Furlong et al., 1998, 2001; Feeny et al., 1995; Torrance et al., 1996) that have been most widely used in the health economics literature (Kwon et al., 2022; Noto and Uemura, 2020; Horsman et al., 2003) and research has shown that using different national weights may not drastically change the relative ordering of health states or treatment effects (Noto and Uemura, 2020). For example, UK and Canadian populations have valued health states similarly, meaning that while absolute utility values shifted slightly, the consistency in the direction and rank of results remained intact (Fang et al., 2016; Achana et al., 2022; Audhya et al., 2023).

Many utility measures have utility scales that include negative ranges (Kwon et al., 2022). The Canadian value set generates HUI3 multi-attribute utility (HUI3 MAU) scores that can span from −0.36 (theoretical minimum value on the utility scale range) to 1.0 representing optimal health (theoretical maximum) (Feeny et al., 1995, 2002; Torrance et al., 1996). Death is represented by 0.0, and a negative HUI3 score does not necessarily mean an individual personally regards their state as worse than death, but rather that in population-level valuation exercises, the average societal preference weight for some combination of severe functional losses are considered worse than death. For example, this includes states like total dependency or permanent tubes (Rubin et al., 2016), advanced dementia, inability to communicate, social isolation (Auriemma et al., 2022), constantly depressed, severely cognitively impaired, extreme or excruciating pain have been valued as worse than death (Feeny et al., 1995, 2002; Torrance et al., 1996).

The MID for HUI3 MAU scores derived from empirical studies is 0.03[24–31], consequently a difference of ≥0.03 is considered clinically significant from a population based perspective. The mean age at assessment varied across the BLS, VICS, EPICure cohorts (18–26 years). Additional details can be found in Table 1.

2.3. Main exposure

The main independent variable in this study was an indicator for VP or VLBW birth, i.e. whether an individual was born < 32 weeks' gestation or < 1500 g.

2.4. Covariates

Independent variables incorporated into the analysis were: individual's sex (male (referent)), age at assessment (measured in years), and mother's education (low level, medium level and high level) as previous literature has revealed that these factors affect HRQoL among preterm-born individuals (van Lunenburg et al., 2013; Saigal et al., 2006). Maternal educational level was harmonized according to the International Standard Classification of Education (ISCED) into low (ISCED levels 0–2), medium (ISCED levels 3–5), and high (ISCED levels 68) (UNESCO Institute for Statistics, 2012). Low maternal education was a referent category in all models. The subgroups analysed included maternal age at birth (categorised into ≤25 years, 26−30 years, and ≥31 years) and maternal ethnicity (White (referent), and non-White). This age categorization was chosen to balance subgroup sizes and ensure an adequate number of participants in each group.

We additionally accounted for possible cohort effects using indicator variables for each cohort. In addition to the covariates mentioned above, the eligibility criteria for each cohort, as described in Table 1, indicate that cases and controls in the BLS and VICS cohorts were contemporaneously recruited from the same hospitals. Controls for the EPICure study were recruited from the same schools as the cases. Thus, our models implicitly control for the contextual and environmental factors associated with hospital and school settings within each cohort, reducing potential confounding effects related to differences in healthcare access as well as educational environments. This approach ensures a more accurate comparison of outcomes between cases and controls within each setting.

In addition to the covariates used in the adjusted analysis, the eligibility criteria applied within BLS, VICS, and EPICure inherently controlled for contextual and environmental factors associated with neighborhood, hospital, or school-level effects. This implicit control further mitigates unmeasured confounding linked to the different settings from which participants were recruited.

3. Statistical analysis

We used frequency distributions, and measures of central tendency and dispersion, such as means and standard deviations to describe the characteristics of study participants. We compared the balance of covariates across the VP/VLBW and controls using Student *t*-test for continuous variables, and Pearson's chi-squared test for categorical variables across the following variables: age group, sex, maternal education, maternal ethnicity, maternal age at birth, and outcome variables. The HUI3 multi-attribute utility score was the main outcome of interest in all analyses performed.

CF methods are a generalization of the random forest (Breiman, 2001) tailored to the estimation of treatment effects. Where not all confounders are observed, the method provides robust estimates of association that are less dependent on model specification. Below we refer to "effects" rather than robust "associations" as this terminology is more common in the literature.

3.1. Random forests and causal forests

A random forest comprises an ensemble of decision trees that recursively split the dataset based on the response variable such that individuals within groups defined by the splits are as similar as possible in terms of outcomes, while being as different as possible from those in other groups (Breiman, 2001). Recursive splitting continues until a set stopping criterion (e.g. a minimum number of observations within a group) is met. This procedure is repeated multiple times over random data subsets, which mitigates the risk of overfitting that plagues single decision trees.

The utilization of random forests has been popular in economics, health, and environmental science due to their robust predictive capabilities and their robustness to potential confounding effects (Nie and Wager, 2021). Comparative studies have demonstrated that random forests can yield comparable or superior predictions relative to traditional methods such as ordinary least squares (Dandl et al., 2022). This advantage stems from the model's flexibility in handling both linear and non-linear relationships and intricate inter-variable interactions, all without the need for predefined model structures.

Analogously to decision trees and random forests, causal trees and causal forests have been proposed that estimate expected differences in outcomes between groups (e.g. treatment effect estimates) rather than predicted outcomes. Causal trees are built recursively by splitting observations where sample splits are formed such that there are 'treated' and 'control' observations in each 'leaf' of the tree and such that estimated treatment effect is as homogenous as possible within a leaf, and as different as possible between leaves. A causal forest is an ensemble of causal trees fitted to random subsets of the data to increase the robustness of estimates, analogously to random forests being an ensemble of decision trees.

3.2. Bayesian Causal Forests

Here, we instead apply the BCF (Caron et al., 2022) which estimates effects building on Bayesian Additive Regression Trees (BART) in place of decision trees, which has been shown to perform well in estimating conditional expectations, in place of the random forests underlying Causal Forest. The outcome (Y_i) can be expressed as:

$$Y_i = \mu(X_i, \pi(\widetilde{X}_i)) + \tau(W_i)T_i + \epsilon_i,$$

where μ is the prognostic (baseline outcome) model depending on covariates Xi and the propensity score π for the treatment T_i is represented by $\pi(\widetilde{x}_i) = P(T_i = 1|\widetilde{\boldsymbol{X}}_i = \widetilde{x}_i)$, and $\tau(\boldsymbol{W}_i)$ is the moderating (treatment effect) component that depends on modifiers $\boldsymbol{W}_i$ and $\widetilde{\boldsymbol{X}}_i$ are the covariates influencing treatment assignment (T_i). Separate priors can be placed on the prognostic score μ(·) and on CATE τ(·) directly, which determine the probability of each variable being split by a tree. In BCF, a default prior is used for the prognostic score, while a stronger regularization is used for the CATE, under the assumption that the effect structure is less complex than the equation determining outcomes. This is achieved by assigning higher probability mass to simpler trees than deeper trees (Caron et al., 2022). The default version of BART (and hence BCF) places a uniform distribution on the splitting variable, meaning that each predictor has an equal chance of being used as a splitting variable.

3.2.1. Shrinkage Bayesian Causal Forests

The SBCF extends this approach by incorporating Dirichlet priors over the splitting probabilities, in addition to the priors described above. This induces sparsity in the estimation of prognostic and moderating effects, with the level of sparsity controlled by hyper-parameters α_μ and α_τ, with lower values of α implying fewer variables will tend to be included in the corresponding model ($\mu_j(\cdot)$ or $\tau_j(\cdot)$). Posterior splitting probabilities can be intuitively viewed as a measure of the variables' importance (Breiman, 2001; Caron et al., 2022). While the estimates obtained using BCF and SBCF methods have a causal interpretation, provided the assumption of selection on observables holds (i.e. all relevant confounders are observed), so we view this approach as providing more robust measures of association rather than causal effects.

After estimating $\tau(W_i)$ i.e. the 'effects' of VP/VLBW for each individual, we estimated conditional average treatment effects (CATEs) for our pre-specified subgroups by aggregating the estimated individual-level treatment effects of group average VP/VLBW (Kreif et al., 2022). The factors considered to influence heterogeneity in $\tau(W_i)$ were informed by relevant literature (Bolbocean et al., 2022; van Lunenburg et al., 2013; Saigal et al., 2006) and clinical judgement. We considered the following pre-specified subgroup variables: sex (male, female), maternal education (low, medium, high), maternal ethnicity (White, non-White), and maternal age group (≤25), 26–30, ≥31. We also considered maternal education as a binary variable, distinguishing between high or medium levels of education versus low maternal education. Analyses were conducted in the statistical package R (Version 4.3.2).

3.2.2. Implementation of causal forest approaches

We estimate the BCF and SBCF Forest using the SparseBCF package (Caron et al., 2022) in R as follows:

1. We restrict the sample to complete cases.
2. We next estimate a propensity score for each individual as a function of their covariates. While we could estimate this using a parametric model (e.g., probit), here we follow Caron et al. (2022) (Caron et al., 2022) in using a 1-hidden-layer neural net (the *nnet* package) in R. We include 10 neurons in the hidden layer. We set the parameter for weight decay (decay) to 0.01 to avoid over-fitting, increase the maximum number of iterations (*maxit*) from 100 to 2000, and reduce the stopping criterion (*abstol*) from 1.0e-4 to 1.0e-8.
3. We estimate the BCF using the SparseBCF package (Version 1.2), but set the option sparse = FALSE. This models both the outcome as a function of the covariates, treatment, and the propensity score. We use the default hyperparameters, except for setting *Nburn*, which controls the number of burn-in Markov chain Monte Carlo (MCMC) iterations, to 2000, and *Npost*, which controls the number of MCMC iterations, to 10,000 and *nsim* which controls the number of MCMC iterations to save after burn-in, to 5000.
4. We then estimate the SBCF using the same function but with sparse = TRUE.
5. The posterior draws of $\hat{\tau}(x)$ are averaged to obtain an effect estimate for each individual or subgroup. A $(1-\alpha)$% credible interval is obtained by calculating the upper $\alpha/2$ and lower $\alpha/2$ quintiles of the posterior draws.

4. Results

4.1. Baseline characteristics

Table 1 details the background characteristics of each cohort, including eligibility criteria, and age(s) of assessment in adulthood. Table 2 shows the baseline characteristics by VP/VLBW birth status for the combined meta-cohort. Differences between groups in baseline characteristics were observed for the following variables: maternal education level at birth or during childhood and age at assessment. The mean HUI3 multi-attribute utility score was 0.90 for the controls and 0.84 for the VP/VLBW adults and medians in HUI3 multi-attribute utility scores were 0.93 for controls vs 0.91 for VP/VLBW adults. Results show that there were no statistically significant differences between groups in child sex (p-value = 0.13), and maternal ethnicity (p-value = 0.06), at the 5 % level of significance.

4.2. Heterogeneity of VP/VLBW effects in HUI3 multi-attribute utility scores by predefined subgroups

Table 3 shows the CATE for each subgroup estimated using BCF and SBCF outputs combined. The results from both BCF and SBCF methods revealed highly similar patterns of association. As expected, the SBCF generally provided estimates with greater precision, though this improvement was modest in our analysis. The evidence shows that the impact of VP/VLBW status on HUI3 MAU scores differs across the subgroups considered. The effect of VP/VLBW status varied by maternal age and maternal education. In particular, the effect ranges from −0.08 (95 % CI −0.13,−0.02) for individuals born to mothers ≤25 years to 0.01 (95 % CI −0.04,0.06) for the individuals born to mothers with high education. Across all subgroups considered, except for individuals born to mothers with high education, the effect of VP/VLBW individuals on HUI3 multi-attribute utility scores is negative. VP/VLBW individuals born to the youngest mothers (≤25 years) were found to have a decrement of −0.08 in HUI3 multi-attribute utility scores, while those born to mothers aged 26–30 years and ≥31 years had decrements of −0.02 (95 % CI −0.05,0.01) and −0.03 (95 % CI −0.06,0.00), respectively. We found a decrement of −0.05 (95 % CI −0.09,−0.01) for VP/VLBW individuals born to mothers with the lowest educational level (ISCED levels 0–2), a decrement of −0.04 (95 % CI −0.08,0.00) for those born to mothers with a medium level of education (ISCED levels 3–5), and an increment of 0.01 (95 % CI −0.04,0.06) for those born to mothers with high education (ISCED levels 6–8) (UNESCO Institute for Statistics, 2012). There was little difference in the decrements in preterm individuals for those born to white mothers and those born to non-white mothers.

Table 2
Summary statistics of demographic characteristics of study participants.

	VP/VLBW	Controls	*P*-value	Missings/*N* (Pct)
N(%)	527 (55.5)	423 (44.5)		0/950 (0.00)
Age at QoL assessment (years), mean (SD)	21.9 (4.0)	22.6 (4.0)	0.01	7/950 (0.7)
Child sex, *N*(%)				
Male	260 (49.3)	188 (44.4)		
Female	267 (50.7)	235 (55.6)	0.13	0/950 (0.00)
Gestational age (weeks), mean (SD)	27.9 (3.14)	39.5 (1.30)	<0.001	62/950 (6.53)
Birth weight (gr), mean (SD)	1052 (345)	3382 (456)	< 0.001	62/950 (6.53)
Maternal age at birth (years), mean (SD)	29.1 (5.2)	29.5 (4.6)	0.29	64/950 (6.74)
Maternal education level at birth or during childhood, *N*(%)				
Low level (equivalent to ISCED 0 to 2)	129 (29.7)	108 (31.1)		
Medium level (equivalent to ISCED 3 to 5)	254 (58.5)	165 (47.6)		
High level (equivalent to ISCED 6 to 8)	51 (11.8)	74 (21.3)	< 0.001	169/950 (17.79)
Maternal ethnicity, *N*(%)				
White	475 (92.8)	343 (95.8)		
Non-White	37 (7.2)	15 (4.2)	0.06	80/950 (8.42)
HUI3 MAU score, mean (SD)	0.84 (0.21)	0.90 (0.14)	< 0.001	0/950 (0.00)
HUI3 MAU score, median (min; max)	0.91 (−0.13; 1.00)	0.93 (0.10; 1.00)	< 0.001	0/950 (0.00)

Notes: Age of assessment was measured in years, BLS - Bavarian Longitudinal Study, VICS - Victorian Infant Collaborative Study, EPICure - EPICure Study. VP/VLBW – very preterm or very low birth weight. ISCED - International Standard Classification of Education. QoL – quality of life. HUI3 MAU – Health Utility Index Mark 3, multi-attribute utility score.

Consistent with selection criteria across participating cohorts, the effect of VP/VLBW status in the EPICure cohort was the largest compared with the BLS or VICS cohorts. This is consistent with selection criteria given that the EPICure Study had selected infants with ≤26 weeks' gestation. These findings highlight the importance of considering subgroup-specific characteristics when assessing the impact of preterm birth or low birthweight. However, the effect estimates provide limited evidence of variation in effects, with the confidence intervals of the subgroup effect estimates including the previously reported overall effect (−0.06) (Bolbocean et al., 2022) in most cases.

Fig. 1 shows that the SBCF had high and moderately uniform probability of splitting on each variable, with no variable having a splitting probability close to zero, which implies that all of the variables incorporated in our analyses are relevant.

5. Discussion

This is the first study to use novel machine learning methods to study

Table 3
Bayesian causal forest and sparse Bayesian causal forest.

Subgroup	N	$CATE_{BCF}$	95 % CI_{BCF}	$CATESH-BCF$	95 % $CISH-BCF$
Age of Assessment ≤ 20	467	−0.04	−0.08, 0.00	−0.04	−0.08, 0.01
Age of Assessment > 20 and ≤ 25	21	−0.09	−0.19, 0.00	−0.01	−0.21, −0.01
Age of Assessment > 25	455	−0.03	−0.07, 0.00	−0.03	−0.07, 0.00
VICS	323	−0.01	−0.06, 0.03	−0.01	−0.06, 0.03
BLS	455	−0.03	−0.07, 0.00	−0.03	−0.07, 0.00
EPICure	172	−0.08	−0.15, 0.00	−0.08	−0.15, 0.00
Male	448	−0.03	−0.07, 0.00	−0.03	−0.07, 0.00
Female	502	−0.04	−0.08, 0.00	−0.04	−0.08, 0.00
Maternal Education Low	237	−0.05	−0.09, 0.00	−0.05	−0.09, −0.01
Maternal Education Medium	419	−0.04	−0.08, 0.00	−0.04	−0.08, 0.00
Maternal Education High	125	0.01	−0.04, 0.06	0.01	−0.04, 0.06
Maternal Education Medium or High	544	−0.03	−0.06, 0.00	−0.03	−0.06, 0.00
White	818	−0.04	−0.07, 0.00	−0.04	−0.07, 0.00
Non-White	52	−0.03	−0.10, 0.04	−0.03	−0.10, 0.03
Maternal Age ≤ 25	200	−0.07	−0.12, −0.01	−0.08	−0.13, −0.02
Maternal Age 26–30	324	−0.02	−0.06, 0.01	−0.02	−0.05, 0.01
Maternal Age ≥ 31	362	−0.03	−0.06, 0.00	−0.03	−0.06, 0.00

Notes: CI- credible intervals. Age of assessment was measured in years, BLS - Bavarian Longitudinal Study, VICS - Victorian Infant Collaborative Study, EPICure - EPICure Study. 95%CI_{BCF} - 95 % credible intervals for BCF. $CATE_{BCF}$ - conditional average treatment effect estimated using Bayesian Causal Forest; 95%CI_{SH-BCF} - 95 % credible intervals for SH-BCF. $CATE_{HS-BCF}$ - conditional average treatment effect estimated using Shrinkage Bayesian Causal Forest.

heterogeneity in the impact of VP/VLBW status on preference-based HRQoL outcomes in early adulthood, leveraging the power of causal forest models to examine how different patient demographics respond to VP birth or VLBW. This study introduces a comparative analysis between BCF and SBCF methods, highlighting the latter's advantage in providing estimates with greater precision around VP/VLBW effects, although the gain was small overall here given that we included a small number of explanatory variables and these were all deemed to be relevant *a priori*. The novelty lies in the granular examination of the effect of VP/VLBW status across various maternal age subgroups and educational levels, as well as ethnic subgroups. The findings reveal variations in preference-based HRQoL outcomes, which have not been thoroughly explored in previous research. Additionally, the use of SBCF represents a methodological advancement that can enhance the precision of future studies in this area. Our study contributes to the growing literature which highlights the value of CF models in health research by examining heterogeneity across groups of individuals (Duong et al., 2024). Furthermore, our study shows how BCF and SBCF enable researchers to explore how maternal socio-demographic factors or clinical factors modify long-term outcomes and the utility of the BART (Caron et al., 2022; Kokandakar et al., 2023).

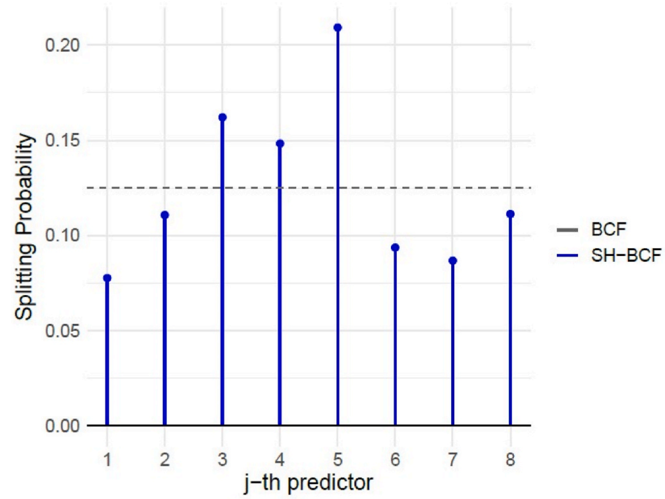

Fig. 1. Probability Splitting Graph.
Notes: $CATE_{BCF}$ - conditional average treatment effect estimated using Bayesian Causal Forest; $CATE_{HS-BCF}$ - conditional average treatment effect estimated using Shrinkage Bayesian Causal Forest. 1st predictor - BLS cohort, 2nd - EPICure cohort, 3rd - age of assessment, 4th - indicator for sex, 5th - maternal age in years, 6th - indicator for white mother, 7th - indicator for medium maternal education, 8th - indicator for high maternal education.

Our analysis revealed differences across maternal age and maternal education subgroups. The effect size of −0.08 for the youngest mothers (≤25 years) shows a large negative effect compared with older mothers as it is greater than the MID for HUI3 MAU scores of 0.03. Furthermore, we found a decrement of −0.05 for VP/VLBW individuals born to mothers with the lowest educational level (ISCED levels 0–2). Furthermore, mothers with low education and young mothers are less likely to be married or to be of high socio-economic status. By contrast, we found that individuals born to mothers with higher education or older mothers had only a slight decrement in HUI3 multi-attribute utility scores compared with term-born or normal birthweight controls. This suggests that maternal socio-economic factors are likely to play a crucial role in determining HRQoL outcomes among VP/VLBW individuals that extend into adulthood.

Our findings are broadly consistent with previous systematic reviews, which identified that higher levels of maternal education are associated with higher multi-attribute utility scores, including in adulthood (van der Pal et al., 2020; Petrou et al., 2019). In fact, the differences in HUI3 multi-attribute utility scores between with mothers with low versus high education are of a similar quantum to being born VP/VLBW versus being born at term. This has previously been shown for differences in functional outcomes such as intelligence (Wolke, 2019; Eves et al., 2021).

Our study extends the previous research on methodological grounds by using novel machine learning methods designed to identify and estimate the effect of VP/VLBW by sociodemographic covariates. As far as we are aware this has not been done previously in this context. Overall, our findings suggest that high socioeconomic status might be a protective factor because we found that amongst individuals born to mothers with high education the effect of VP/VLBW status was almost zero, while amongst individuals born to mothers with low education the effect was more negative. This is consistent with an earlier study that found that high maternal education and higher SES serve as protective factors for VP/VLBW individuals (Bolbocean et al., 2022). Previous research that examined cognitive outcomes and compared those born at high risk, i.e. VP/VLBW with those born at term found that high SES was a protective factor amongst both VP/VLBW individuals and term-born children and adults (Madzwamuse et al., 2015; Benavente-Fernandez et al., 2019). The protective effect is also evident in socioeconomic outcomes in a Canadian study which noted that preterm-born individuals from higher-income families had smaller income deficits in adulthood than those from low-income families (Ahmed et al., 2024).

Our findings underscore the profound influence of the social gradient on long-term health outcomes following VP/VLBW birth (Marmot et al., 2010; Wong and Edwards, 2013; Bilsteen et al., 2022). The observed heterogeneity, particularly the protective association of higher maternal education, aligns with established social determinants of health frameworks, which posit that education influences health through multiple mechanistic pathways (Marmot et al., 2010; Marmot, 2005; Cutler and Lleras-Muney, 2010; Braveman and Gottlieb, 2014; Zimmerman et al., 2015). These include improved health literacy, better access to and navigation of healthcare systems, enhanced economic resources leading to better nutrition and living conditions, and potentially reduced exposure to chronic stress (Benavente-Fernandez et al., 2019; Bilsteen et al., 2022). For VP/VLBW individuals, whose development can be more vulnerable to environmental exposures, the resources and stability associated with higher maternal SES may be particularly crucial for mitigating long-term health deficits (Benavente-Fernandez et al., 2019; Braveman and Gottlieb, 2014; McHale et al., 2024).

These findings suggest the need for targeted health interventions for VP/VLBW individuals born to younger mothers and those with lower educational levels, as they are at a higher risk of poorer HRQoL outcomes in early adulthood than those born to older mothers and with higher educational levels (Saigal et al., 2000; Currie and Hyson, 1999). Our results underscore the importance of addressing the social determinants of health that influence outcomes in this vulnerable population (Benavente-Fernandez et al., 2019; Braveman and Gottlieb, 2014; McHale et al., 2024; Case et al., 2002). Improved maternal education and socioeconomic status are associated with better access to resources (including healthcare and educational support), healthier environments, enhanced health literacy, and greater capacity to support a child's development, all of which can act as protective factors mitigating the long-term impacts of VP/VLBW birth on HRQoL. Thus, policies and programs aimed at improving maternal socioeconomic status and educational attainment – such as early childhood support programs, maternal education subsidies, or targeted social safety nets for vulnerable families – are likely to be particularly important (Olds et al., 1997; Ludwig and Phillips, 2007; Currie, 2009; Heckman et al., 2010; Hoynes et al., 2016; Puthussery et al., 2018). These interventions, by addressing socioeconomic factors early on, have the potential to indirectly but substantially improve long-term HRQoL outcomes among VP/VLBW populations, helping to mitigate adverse health trajectories and reduce disparities in adult quality of life.

5.1. Strengths and limitations

The use of individual and harmonized participant data from multiple cohorts strengthens the validity of the findings, as it accounts for some of the potential variability in data sources and selection criteria across cohorts. Our research included diverse populations across different geographical regions and socio-demographic backgrounds, which enhances the generalizability of our findings. An additional strength of our study lies in the large sample of VP/VLBW individuals and the longitudinal research design, which enhance both internal and external validity. The contributing cohorts utilized reliable and valid recruitment methods and maintained moderately to high participation rates throughout the follow-up period for both study cases and controls.

In addition to the covariates used in the adjusted analysis, the eligibility criteria applied within BLS, VICS, and EPICure inherently controlled for contextual and environmental factors associated with neighborhood, hospital, or school-level effects. This implicit control further strengthens the robustness of our findings by mitigating unmeasured confounding linked to the different settings from which participants were recruited. Overall, this approach minimizes concerns about omitted variable bias and strengthens the validity of the selection on observables assumption. Finally, the use of adult self-reported HRQoL data helps to avoid biases inherent in proxy parental reporting. Literature on the measurement of childhood HRQoL shows that child-proxy agreement is generally poor, especially for subjective constructs such as emotion and pain (Khanna et al., 2022).

We acknowledge the limitations of our analysis. Data from this study come from observational cohorts and hence are prone to confounding. We conduct a complete case analysis, so caution is required in generalizing results to the full population. The methods used here are based on an assumption of selection on observables that cannot be verified with observed data (Imbens and Wooldridge, 2009; Arnold et al., 2010). Thus, our findings are best viewed as robust associations rather than causal effects, given concerns about the plausibility of the selection on observables assumption here. Thus, this study may not fully account for all potential confounding variables that could influence the relationship between VP/VLBW status and HRQoL outcomes in early adulthood. For example, our dataset does not include data on direct assessments of parenting skills. However, evidence shows that parents of VP/VLBW children are as sensitive in their parenting as parents of term-born or normal birthweight children (Bilgin and Wolke, 2015). Nonetheless other differences between the groups may remain, motivating future research using richer datasets. Our findings are best viewed as robust associations rather than causal effects, given uncertainty about the plausibility of selection on observables here.

We acknowledge that we are not able to measure direct measures of socioeconomic status such as income or employment. Specifically, we used maternal education, which has been shown to be a crude proxy for economic well-being (Braveman et al., 2001; Adler and Rehkopf, 2008). There was little variation in ethnicity in included cohort studies which does not reflect the current population trends especially in South Germany (Bavaria). It was not feasible to conduct analysis separately for each of the three cohorts given the available sample sizes. Furthermore, our sample size did not allow for further subdivisions by maternal age (e.g., 35+ years); more granular analyses of maternal age, particularly at the upper end of the maternal age spectrum, would be valuable in larger datasets in the future.

The study provides a snapshot of the effects of VP/VLBW status but does not consider HRQoL changes over time as individuals age. Future research should focus on longitudinal studies that track the HRQoL of VP/VLBW individuals over time, providing insights into how the impact evolves with age and across stages of the life-course. Due to data limitations, we considered a limited number of subgroups. Further investigation into other potential subgroups, such as those defined by economic status or parental occupation, access to healthcare, and parental health status, could provide a more comprehensive understanding of the effects of VP/VLBW birth.

Modern clinical practice may differ from practice during the range of birth years across the cohorts considered (1985–1995). However, it is plausible that the underlying biological impacts of VP/VLBW remains informative as the evidence shows that despite changes in practices there is no improvements in outcome (Cheong et al., 2017; Spittle et al., 2018; Marlow et al., 2021; Ni et al., 2022; Larsen et al., 2024). Future research is needed to assess whether improvements in care have modified long-term outcomes using more recent data.

Finally, the results of our study may not be generalizable to low- or middle-income countries, as the data were collected exclusively from high-income countries. Additionally, our findings might not fully apply to populations in the Americas, Asia, or Africa, as the study cohorts were limited to Western European countries and Australia.

6. Conclusions

Our CF approach provided heterogeneous estimates of CATE of VP/VLBW status on preference-based HRQoL outcomes in early adulthood, suggesting that VP/VLBW status has the worst impact on HRQoL in early adulthood among individuals born to young and lower-educated mothers. Evidence of heterogeneity in HRQoL based on sociodemographic factors, such as maternal education and maternal age at birth, underscores the importance of assessing the sensitivity of health economic evaluations related to preterm birth to the choice of health utility parameters. It also highlights the need to consider the distributional implications when conducting health economic evaluations related to preterm birth.

CRediT authorship contribution statement

Corneliu Bolbocean: Conceptualization, Research Design, Methodology, Investigation, Data curation, formal analysis, data analysis and software, Visualization, Validation, Writing – original draft, Writing – review & editing, Project administration, Funding. **Peter J. Anderson:** Writing – review & editing, Data curation, Validation. **Peter Bartmann:** Writing – review & editing, Data curation, Validation. **Jeanie L.Y. Cheong:** Writing – review & editing, Data curation, Validation. **Lex W. Doyle:** Writing – review & editing, Data curation, Validation. **Samantha Johnson:** Writing – review & editing, Data curation, Validation. **Neil Marlow:** Writing – review & editing, Data curation, Validation. **Dieter Wolke:** Writing – review & editing, Data curation, Validation. **Stavros Petrou:** Writing – review & editing, Data curation, Validation. **Stephen O'Neill:** Conceptualization. Research Design, Methodology, Investigation, Data curation, formal analysis, data analysis and software, Visualization, Validation, Writing – original draft, Writing – review & editing.

Ethics approval

The Ethics Committees of the participating cohorts have approved the study protocol.

Declaration of competing interest

This work was funded by the European Union's Horizon 2020 research and innovation program (Research on European Children and Adults Born Preterm), under grant agreement 733280. Corneliu Bolboceau is funded by School of Primary Care Research Award and NIHR Applied Research Collaboration for Oxford and the Thames Valley Award. Dieter Wolke is funded by a UKRI Frontier Research grant (EP/X023206/1) guarantee of an ERCAdG award.

Acknowledgements

This work was funded by the European Union's Horizon 2020 research and innovation program (Research on European Children and Adults Born Preterm), under grant agreement 733280. Corneliu Bolboceau is funded by School of Primary Care Research Award (Award Number 713) and NIHR Applied Research Collaboration for Oxford and the Thames Valley Award. Dieter Wolke is funded by a UKRI Frontier Research grant (EP/X023206/1) guarantee of an ERCAdG award.

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2025.118181.

Appendix A. Summary Statistics of Demographic Covariates for Included Cohorts

Table A.1
Summary Statistics of Demographic Covariates for BLS Cohort

	VP/VLBW	Controls	*p*-value	Missings/*N* (Pct)
N(%)	231 (50.8)	224 (49.2)		0/455 (0.0)
Age at QoL assessment (years), mean (SD)	26.32 (0.68)	26.29 (0.69)	0.62	0/455 (0.0)
Child sex, *N*(%)				
Male	125 (54.1)	105 (46.9)		
Female	106 (45.9)	119 (53.1)	0.12	0/455 (0.0)
Maternal age at birth (years), mean (SD)	29.06 (4.74)	29.12 (4.59)	0.90	0/455 (0.0)
Maternal education level at birth or during childhood, *N*(%)				
Low level (equivalent to ISCED 0 to 2)	72 (31.9)	99 (44.4)		
Medium level (equivalent to ISCED 3 to 5)	126 (55.8)	88 (39.5)		
High level (equivalent to ISCED 6 to 8)	28 (12.4)	36 (16.1)	< 0.001	6/455 (1.3)
Maternal ethnicity, *N*(%)				
White	231 (100.0)	224 (100.0)	0.54	0/455 (0.0)
HUI MAU score, mean (SD)	0.85 (0.18)	0.89 (0.14)	0.01	0/455 (0.0)
HUI MAU score, median (min; max)	0.91 (0.10; 1.00)	0.93 (0.10; 1.00)	< 0.001	0/455 (0.0)

Notes: Age of assessment was measured in years. VP/VLBW – very preterm or very low birth weight. ISCED - International Standard Classification of Education. QoL – quality of life. HUI3 MAU – Health Utility Index Mark 3, multi-attribute utility score.

Table A.2
Summary Statistics of Demographic Covariates for VICS Cohort

	VP/VLBW	Controls	*p*-value	Missings/*N* (Pct)
N(%)	186 (57.6)	137 (42.4)		0/323 (0.0)
Age at QoL assessment (years), mean (SD)	17.92 (0.78)	18.07 (0.88)	0.11	7/323 (2.2)
Child sex, *N*(%)				
Male	84 (45.2)	60 (43.8)		
Female	102 (54.8)	77 (56.2)	0.81	0/323 (0.0)
Maternal age at birth (years), mean (SD)	29.12 (5.71)	30.12 (4.70)	0.10	0/323 (0.0)
Maternal education level at birth or during childhood, *N*(%)				

(*continued on next page*)

Table A.2 (*continued*)

	VP/VLBW	Controls	*p*-value	Missings/*N* (Pct)
Low level (equivalent to ISCED 0 to 2)	37 (35.6)	6 (9.7)		
Medium level (equivalent to ISCED 3 to 5)	47 (45.2)	30 (48.4)		
High level (equivalent to ISCED 6 to 8)	20 (19.2)	26 (41.9)	< 0.001	157/323 (48.6)
Maternal ethnicity, *N*(%)				
White	151 (87.3)	119 (88.8)		
non-White	22 (12.7)	15 (11.2)	0.68	16/323 (5.0)
HUI MAU score, mean (SD)	0.86 (0.20)	0.90 (0.15)	0.08	0/323 (0.0)
HUI MAU score, median (min; max)	0.93 (−0.09; 1.00)	0.95 (0.19; 1.00)	0.18	0/323 (0.0)

Notes: Age of assessment was measured in years. VP/VLBW – very preterm or very low birth weight. ISCED - International Standard Classification of Education. QoL – quality of life. HUI3 MAU – Health Utility Index Mark 3, multi-attribute utility score.

Table A.3
Summary Statistics of Demographic Covariates for EPICure Cohort

	VP/VLBW	Controls	*p*-value	Missings/*N* (Pct)
N(%)	110 (64.0)	62 (36.0)		0/172 (0.0)
Age at QoL assessment (years), mean (SD)	19.29 (0.56)	19.19 (0.54)	0.27	0/172 (0.0)
Child sex, *N*(%)				
Male	51 (46.4)	23 (37.1)		
Female	59 (53.6)	39 (62.9)	0.24	0/172 (0.0)
Maternal age at birth (years), mean (SD)	29.31 (5.47)	0 (0.0)	N/A	64/172 (37.2)
Maternal education level at birth or during childhood, *N*(%)				
Low level (equivalent to ISCED 0 to 2)	20 (19.2)	3 (4.8)		
Medium level (equivalent to ISCED 3 to 5)	81 (77.9)	47 (75.8)		
High level (equivalent to ISCED 6 to 8)	3 (2.9)	12 (19.4)	<0.001	6/172 (3.5)
Maternal ethnicity, *N*(%)				
White	93 (86.1)	0 (0.0)		
non-White	15 (13.9)	0 (0.0)	N/A	64/172 (37.2)
HUI3 MAU score, mean (SD)	0.77 (0.25)	0.92 (0.12)	<0.001	0/172 (0.0)
HUI3 MAU score, median (min; max)	0.85 (−0.13; 1.00)	0.97 (0.45; 1.00)	<0.001	0/172 (0.0)

Notes: Age of assessment was measured in years. VP/VLBW – very preterm or very low birth weight. ISCED - International Standard Classification of Education. QoL – quality of life. HUI3 MAU – Health Utility Index Mark 3, multi-attribute utility score.

Data availability

The authors do not have permission to share data.

References

Achana, F., Johnson, S., Ni, Y., Marlow, N., Wolke, D., Khan, K., Petrou, S., 2022. Economic costs and health utility values associated with extremely preterm birth: evidence from the EPICure2 cohort study. Paediatr. Perinat. Epidemiol. 36 (5), 696–705.

Adler, N.E., Rehkopf, D.H., 2008. US disparities in health: descriptions, causes, and mechanisms. Annu. Rev. Publ. Health 29 (1), 235–252.

Ahmed, A.M., Pullenayegum, E., McDonald, S.D., Beltempo, M., Premji, S.S., Shoukry, R., Pole, J.D., Bacchini, F., Shah, P.S., Pechlivanoglou, P., 2024. Preterm birth, family income, and intergenerational income mobility. JAMA Netw. Open 7 (6), e2415921.

Allen, M.C., Cristofalo, E., Kim, C., 2010. Preterm birth: transition to adulthood. Developmental Disabilities Research Reviews 16 (4), 323–335.

Arnett, J.J., 2015. The Oxford Handbook of Emerging Adulthood. Oxford University Press.

Arnold, B.F., Khush, R.S., Ramaswamy, P., London, A.G., Rajkumar, P., Ramaprabha, P., Durairaj, N., Hubbard, A.E., Balakrishnan, K., Colford Jr., J.M., 2010. Causal inference methods to study nonrandomized, preexisting development interventions. Proc. Natl. Acad. Sci. 107 (52), 22605–22610.

Assmann, S.F., Pocock, S.J., Enos, L.E., Kasten, L.E., 2000. Subgroup analysis and other (mis)uses of baseline data in clinical trials. Lancet 355 (9209), 1064–1069.

Athey, S., Imbens, G., 2016. Recursive partitioning for heterogeneous causal effects. Proc. Natl. Acad. Sci. 113 (27), 7353–7360.

Audhya, I.F., Szabo, S.M., Bever, A., O'Sullivan, F., Malone, D.C., Feeny, D., Neumann, P., Iannaccone, S.T., Jayasinghe, P., Gooch, K.L., 2023. Estimating health state utilities in Duchenne muscular dystrophy using the Health Utilities Index and EQ-5D-5L. J. Patient-Rep. Outcomes 7 (1), 132.

Auriemma, C.L., O'Donnell, H., Jones, J., Barbati, Z., Akpek, E., Klaiman, T., Halpern, S. D., 2022. Patient perspectives on states worse than death: a qualitative study with implications for patient-centered outcomes and values elicitation. Palliat. Med. 36 (2), 348–357.

Baranger, D.A.A., Finsaas, M.C., Goldstein, B.L., Vize, C.E., Lynam, D.R., Olino, T.M., 2023. Tutorial: power analyses for interaction effects in cross-sectional regressions. Adv. Methods Pract. Psychol. Sci. 6 (3), 25152459231187531.

Beam, A.L., Fried, I., Palmer, N., Agniel, D., Brat, G., Fox, K., Kohane, I., Sinaiko, A., Zupancic, J.A.F., Armstrong, J., 2020. Estimates of healthcare spending for preterm and low-birthweight infants in a commercially insured population: 2008–2016. J. Perinatol. 40 (7), 1091–1099.

Beckr, J.-M., Rai, A., Ringle, C.M., Völckner, F., 2013. Discovering unobserved heterogeneity in structural equation models to avert validity threats. MIS Q. 37 (3), 665–694.

Benavente-Fernandez, I., Synnes, A., Grunau, R.E., Chau, V., Ramraj, C., Glass, T., Cayam-Rand, D., Siddiqi, A., Miller, S.P., 2019. Association of socioeconomic status and brain injury with neurodevelopmental outcomes of very preterm children. JAMA Netw. Open 2 (5), e192914.

Bhatti, P., Mirick, D.K., Davis, S., 2013. Racial differences in the association between night shift work and melatonin levels among women. Am. J. Epidemiol. 177 (5), 388–393.

Bilgin, A., Wolke, D., 2015. Maternal sensitivity in parenting preterm children: a meta-analysis. Pediatrics 136 (1), e177–e193.

Bilsteen, J.F., Alenius, S., Bråthen, M., Børch, K., Ekstrøm, C.T., Kajantie, E., et al., 2022. Gestational age, parent education, and education in adulthood. Pediatrics 149 (1), e2021051959. https://doi.org/10.1542/peds.2021-051959.

Bolbocean, C., van der Pal, S., van Buuren, S., Anderson, P.J., Bartmann, P., Baumann, N., Cheong, J.L.Y., Darlow, B.A., Doyle, L.W., Evensen, K.A.I., et al., 2022. Health-related quality-of-life outcomes of very preterm or very low birth weight adults: evidence from an individual participant data meta-analysis. Pharmacoeconomics 1–13.

Bolbocean, C., Anderson, P.J., Bartmann, P., Cheong, J.L.Y., Doyle, L.W., Wolke, D., Petrou, S., 2023. Comparative evaluation of the Health Utilities Index Mark 3 and the Short Form 6D: evidence from an individual participant data meta-analysis of very preterm and very low birthweight adults. Qual. Life Res. 1–14.

Bonander, C., Svensson, M., 2021. Using causal forests to assess heterogeneity in cost-effectiveness analysis. Health Econ. 30 (8), 1818–1832.

Braveman, P., Gottlieb, L., 2014. The social determinants of health: it's time to consider the causes of the causes. Publ. Health Rep. 129 (Suppl. 2), 19–31. https://doi.org/10.1177/00333549141291S206.

Braveman, P., Cubbin, C., Marchi, K., Egerter, S., Chavez, G., 2001. Measuring socioeconomic status/position in studies of racial/ethnic disparities: maternal and infant health. Publ. Health Rep.® 116 (5), 449–463.

Brazier, J., Ratcliffe, J., Saloman, J., Tsuchiya, A., 2017. Measuring and Valuing Health Benefits for Economic Evaluation. Oxford University Press.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Brumbaugh, J.E., Bann, C.M., Bell, E.F., Travers, C.P., Vohr, B.R., McGowan, E.C., Harmon, H.M., Carlo, W.A., Hintz, S.R., Duncan, A.F., et al., 2025. The relationship between Hispanic ethnicity and outcomes for infants born extremely preterm. J. Pediatr. 279, 114474.

Call, C., Oran, A., O'Shea, T.M., Jensen, E.T., Frazier, J.A., Vaidya, R., Shenberger, J., Gogcu, S., Msall, M.E., Kim, S., et al., 2024. Social determinants of health rather than

race impact health-related quality of life in 10-year-old children born extremely preterm. Frontiers in Pediatrics 12, 1359270.
Caron, A., Baio, G., Manolopoulou, I., 2022. Shrinkage Bayesian causal forests for heterogeneous treatment effects estimation. J. Comput. Graph Stat. 31 (4), 1202–1214.
Case, A., Lubotsky, D., Paxson, C., 2002. Economic status and health in childhood: the origins of the gradient. Am. Econ. Rev. 92 (5), 1308–1334.
Cheong, J.L.Y., Anderson, P.J., Burnett, A.C., Roberts, G., Davis, N., Hickey, L., Carse, E., Doyle, L.W., Victorian Infant Collaborative Study Group, 2017. Changing neurodevelopment at 8 years in children born extremely preterm since the 1990s. Pediatrics 139 (6), e20164086.
Cook, D.I., Gebski, V.J., Keech, A.C., 2004. Subgroup analysis in clinical trials. Med. J. Aust. 180 (6), 289.
Cooper, R., Atherton, K., Power, C., 2009. Gestational age and risk factors for cardiovascular disease: evidence from the 1958 British birth cohort followed to mid-life. Int. J. Epidemiol. 38 (1), 235–244.
Currie, J., 2009. Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development. J. Econ. Lit. 47 (1), 87–122.
Currie, J., Hyson, R., 1999. Is the impact of health shocks cushioned by socioeconomic status? The case of low birth weight. Am. Econ. Rev. 89 (2), 245–250. https://doi.org/10.1257/aer.89.2.245.
Cutler, D.M., Lleras-Muney, A., 2010. Understanding differences in health behaviors by education. J. Health Econ. 29 (1), 1–28. https://doi.org/10.1016/j.jhealeco.2009.10.003.
Dandl, S., Bender, A., Hothorn, T., 2022. Heterogeneous Treatment Effect Estimation for Observational Data Using Model-Based Forests arXiv preprint arXiv:2210.02836.
Darlow, B.A., Woodward, L.J., Levin, K.J., Melzer, T., Horwood, L.J., 2020. Perinatal and childhood predictors of general cognitive outcome at 28 years in a very-low-birthweight national cohort. Dev. Med. Child Neurol. 62 (12), 1423–1428.
Davis, J.M.V., Heller, S.B., 2017. Using causal forests to predict treatment heterogeneity: an application to summer jobs. Am. Econ. Rev. 107 (5), 546–550.
Dolatian, M., Sharifi, N., Mahmoodi, Z., 2018. Relationship of socioeconomic status, psychosocial factors, and food insecurity with preterm labor: a longitudinal study. Int. J. Reprod. Biomed. 16 (9), 563.
Doyle, L.W., Cheong, J.L.Y., Burnett, A., Roberts, G., Lee, K.J., Anderson, P.J., Victorian Infant Collaborative Study Group, 2015. Biological and social influences on outcomes of extreme-preterm/low-birth weight adolescents. Pediatrics 136 (6), e1513–e1520.
Doyle, L.W., Spittle, A., Anderson, P.J., Cheong, J.L.Y., 2021. School-aged neurodevelopmental outcomes for children born extremely preterm. Arch. Dis. Child. 106, 834–838.
Drummond, M., 2001. Introducing economic and quality of life measurements into clinical studies. Ann. Med. 33 (5), 344–349.
Drummond, M.F., Sculpher, M.J., Claxton, K., Stoddart, G.L., Torrance, G.W., 2015. Methods for the Economic Evaluation of Health Care Programmes, fourth ed. Oxford University Press.
Duong, B., Senadeera, M., Nguyen, T., Nichols, M., Backholer, K., Allender, S., Nguyen, T., 2024. Utilising causal inference methods to estimate effects and strategise interventions in observational health data. PLoS One 19 (12), e0314761.
D'Souza, N.J., Kuswara, K., Zheng, M., Leech, R., Downing, K.L., Lioret, S., Campbell, K. J., Hesketh, K.D., 2020. A systematic review of lifestyle patterns and their association with adiposity in children aged 5–12 years. Obes. Rev. 21 (8), e13029.
Echevarria, E., Lorch, S.A., 2022. Family educational attainment and racial disparities in low birth weight. Pediatrics 150 (1), e2021052369.
Elek, P., Bíró, A., 2021. Regional differences in diabetes across Europe—regression and causal forest analyses. Econ. Hum. Biol. 40, 100948.
Eves, R., Mendonça, M., Baumann, N., et al., 2021. Association of Very Preterm Birth or Very Low Birth Weight With Intelligence in Adulthood: An Individual Participant Data Meta-analysis. JAMA Pediatr 175 (8), e211058.
Fang, M., Oremus, M., Tarride, J.-E., Raina, P., Canadian Willingness-to-pay Study Group, 2016. A comparison of health utility scores calculated using United Kingdom and Canadian preference weights in persons with Alzheimer's disease and their caregivers. Health Qual. Life Outcome 14, 1–8.
Feeny, D., Furlong, W., Boyle, M., Torrance, G.W., 1995. Multi-attribute health status classification systems. Pharmacoeconomics 7 (6), 490–502.
Feeny, D., Furlong, W., Torrance, G.W., Goldsmith, C.H., Zhu, Z., DePauw, S., Denton, M., Boyle, M., 2002. Multiattribute and single-attribute utility functions for the health utilities index Mark 3 system. Med. Care 40 (2), 113–128.
Fuchs, F., Monet, B., Ducruet, T., Chaillet, N., Audibert, F., 2018. Effect of maternal age on the risk of preterm birth: a large cohort study. PLoS One 13 (1), e0191002.
Furlong, W., Feeny, D., Torrance, G., Goldsmith, C., DePauw, S., Zhu, Z., Denton, M., Boyle, M., 1998. Multiplicative Multi-Attribute Utility Function for the Health Utilities Index Mark 3 (HUI3) System: A Technical Report. *Centre for Health Economics And Policy Analysis (CHEPA)*. McMaster University.
Furlong, W.J., Feeny, D.H., Torrance, G.W., Barr, R.D., 2001. The Health Utilities Index (HUI®) system for assessing health-related quality of life in clinical studies. Ann. Med. 33 (5), 375–384.
Grimmer, J., Messing, S., Westwood, S.J., 2017. Estimating heterogeneous treatment effects and the effects of heterogeneous treatments with ensemble methods. Polit. Anal. 25 (4), 413–434.
Grootendorst, P., Feeny, D., Furlong, W., 2000. Health Utilities Index Mark 3: evidence of construct validity for stroke and arthritis in a population health survey. Med. Care 38 (3), 290–299.
Hahn, P.R., Murray, J.S., Carvalho, C.M., 2020. Bayesian regression tree models for causal inference: regularization, confounding, and heterogeneous effects. Bayesian Analysis 15 (3), 865–902.
Hainmueller, J., Mummolo, J., Xu, Y., 2019. How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice. Polit. Anal. 27 (2), 163–192.
Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A., Yavitz, A., 2010. Analyzing social experiments as implemented: a reexamination of the evidence from the HighScope Perry Preschool Program. Quantitative Economics 118 (5), S38–S55. https://doi.org/10.1086/653213.
Horsman, J., Furlong, W., Feeny, D., Torrance, G., 2003. The health utilities index (HUI®): concepts, measurement properties and applications. Health Qual. Life Outcome 1 (1), 1–13.
Horvath, H., Brindis, C.D., Reyes, E.M., Yamey, G., Franck, L., 2017. Preterm birth: the role of knowledge transfer and exchange. Health Res. Pol. Syst. 15 (1), 1–14.
Hovi, P., Vohr, B., Ment, L.R., Doyle, L.W., McGarvey, L., Morrison, K.M., Evensen, K.A. I., van der Pal, S., Grunau, R.E., APIC Adults Born Preterm International Collaboration, 2016. Blood pressure in young adults born at very low birth weight: adults Born Preterm International Collaboration. Hypertension 68 (4), 880–887.
Hoynes, H., Schanzenbach, D.W., Almond, D., 2016. Long-run impacts of childhood access to the safety net. Am. Econ. Rev. 106 (4), 903–934.
Hu, A., 2023. Heterogeneous treatment effects analysis for social scientists: a review. Soc. Sci. Res. 109, 102810.
Imai, K., Ratkovic, M., 2013. Estimating treatment effect heterogeneity in randomized program evaluation. Ann. Appl. Stat. 7 (1), 443–470.
Imbens, G.W., Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. J. Econ. Lit. 47 (1), 5–86.
Jin, X., Al Sayah, F., Ohinmaa, A., Marshall, D.A., Johnson, J.A., 2019. Responsiveness of the EQ-5D-3L and EQ-5D-5L in patients following total hip or knee replacement. Qual. Life Res. 28, 2409–2417.
Khanna, D., Khadka, J., Mpundu-Kaambwa, C., Lay, K., Russo, R., Ratcliffe, J., 2022. Are we agreed? Self- versus proxy-reporting of paediatric health-related quality of life (HRQoL) using generic preference-based measures: a systematic review and meta-analysis. Pharmacoeconomics 40 (11), 1043–1067.
Kim, S.W., Andronis, L., Seppanen, A.-V., Aubert, A.M., Barros, H., Draper, E.S., Sentenac, M., Zeitlin, J., Petrou, S., 2023. Health-related quality of life of children born very preterm: a multinational European cohort study. Qual. Life Res. 32 (1), 47–58.
Kokandakar, A.H., Kang, H., Deshpande, S.K., 2023. Bayesian causal forests & the 2022 ACIC data challenge: scalability and sensitivity. Observational Studies 9 (3), 29–41.
Kreif, N., DiazOrdaz, K., Moreno-Serra, R., Mirelman, A., Hidayat, T., Suhrcke, M., 2022. Estimating heterogeneous policy impacts using causal machine learning: a case study of health insurance reform in Indonesia. Health Serv. Outcome Res. Methodol. 1–36.
Kwon, J., Kim, S.W., Ungar, W.J., Tsiplova, K., Madan, J., Petrou, S., 2018. A systematic review and meta-analysis of childhood health utilities. Med. Decis. Mak. 38 (3), 277–305.
Kwon, J., Kim, S.W., Ungar, W.J., Tsiplova, K., Madan, J., Petrou, S., 2019. Patterns, trends and methodological associations in the measurement and valuation of childhood health utilities. Qual. Life Res. 28, 1705–1724.
Kwon, J., Freijser, L., Huynh, E., Howell, M., Chen, G., Khan, K., Daher, S., Roberts, N., Harrison, C., Smith, S., et al., 2022. Systematic review of conceptual, age, measurement and valuation considerations for generic multidimensional childhood patient-reported outcome measures. Pharmacoeconomics 40 (4), 379–431.
Kwon, J., Bolbocean, C., Onyimadu, O., Roberts, N., Petrou, S., 2023. Psychometric performance of generic childhood multi-attribute utility instruments in preterm and low birthweight populations: a systematic review. Children 10 (11), 1798.
Larsen, J., Kochhar, P., Wolke, D., Draper, E.S., Marlow, N., Johnson, S., 2024. Comparing behavioural outcomes in children born extremely preterm between 2006 and 1995: the EPICure studies. Eur. Child Adolesc. Psychiatr. 33 (5), 1517–1528.
Linsell, L., Malouf, R., Morris, J., Kurinczuk, J.J., Marlow, N., 2015. Prognostic factors for poor cognitive development in children born very preterm or with very low birth weight: a systematic review. JAMA Pediatr. 169 (12), 1162–1172.
Linsell, L., Johnson, S., Wolke, D., O'Reilly, H., Morris, J.K., Kurinczuk, J.J., Marlow, N., 2018. Cognitive trajectories from infancy to early adulthood following birth before 26 weeks of gestation: a prospective, population-based cohort study. Arch. Dis. Child. 103 (4), 363–370.
Ludwig, J., Phillips, D.A., 2007. Long-term effects of Head Start on low-income children. Ann. N. Y. Acad. Sci. 1136 (1), 257–268.
Lund, L.K., Vik, T., Lydersen, S., Løhaugen, G.C.C., Skranes, J., Brubakk, A.-M., Indredavik, M.S., 2012. Mental health, quality of life and social relations in young adults born with low birth weight. Health Qual. Life Outcome 10 (1), 1–10.
Madzwamuse, S.E., Baumann, N., Jaekel, J., Bartmann, P., Wolke, D., 2015. Neuro-cognitive performance of very preterm or very low birth weight adults at 26 years. JCPP (J. Child Psychol. Psychiatry) 56 (8), 857–864.
Marlow, N., Ni, Y., Lancaster, R., Suonpera, E., Bernardi, M., Fahy, A., Larsen, J., Trickett, J., Hurst, J.R., Morris, J., et al., 2021. No change in neurodevelopment at 11 years after extremely preterm birth. Arch. Dis. Child. Fetal Neonatal Ed. 106 (4), 418–424.
Marmot, M., 2005. Social determinants of health inequalities. Lancet 365 (9464), 1099–1104. https://doi.org/10.1016/S0140-6736(05)71146-6.
Marmot, M., Allen, J., Goldblatt, P., Boyce, T., McNeish, D., Grady, M., Geddes, I., 2010. Fair Society, Healthy Lives: the Marmot Review. Institute of Health Equity.
McHale, P., Fahy, K., Pennington, A., Schluter, D.K., Barr, B., Taylor-Robinson, D., 2024. How do socioeconomic inequalities and preterm birth interact to modify health and education outcomes? A narrative systematic review. BMJ Open 15 (1), e084147. https://doi.org/10.1136/bmjopen-2024-084147.
McMaughan, D.J., Oloruntoba, O., Smith, M.L., 2020. Socioeconomic status and access to healthcare: interrelated drivers for healthy aging. Front. Public Health 8, 231.

Misund, A.R., Nerdrum, P., Diseth, T.H., 2014. Mental health in women experiencing preterm birth. BMC Pregnancy Childbirth 14 (1), 1–8.
Moster, D., Lie, R.T., Markestad, T., 2008. Long-term medical and social consequences of preterm birth. N. Engl. J. Med. 359 (3), 262–273.
Ni, Y., Johnson, S., Marlow, N., Wolke, D., 2022. Reduced health-related quality of life in children born extremely preterm in 2006 compared with 1995: the EPICure studies. Arch. Dis. Child. Fetal Neonatal Ed. 107 (4), 408–413.
Nie, X., Wager, S., 2021. Quasi-oracle estimation of heterogeneous treatment effects. Biometrika 108 (2), 299–319.
Norman, G.R., Sloan, J.A., Wyrwich, K.W., 2003. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med. Care 41 (5), 582–592.
Noto, S., Uemura, T., 2020. Japanese health utilities index Mark 3 (HUI3): measurement properties in a community sample. J. Patient-Rep. Outcomes 4, 1–19.
Olds, D.L., Eckenrode, J., Henderson Jr., C.R., Kitzman, H., Powers, J., Cole, R., Sidora, K., Morris, P., Pettitt, L.M., Luckey, D., 1997. Long-term effects of home visitation on maternal life course and child abuse and neglect: fifteen-year follow-up of a randomized trial. JAMA 278 (8), 637–643. https://doi.org/10.1001/jama.1997.03550080047038.
Parkinson, J.R.C., Hyde, M.J., Gale, C., Santhakumaran, S., Modi, N., 2013. Preterm birth and the metabolic syndrome in adult life: a systematic review and meta-analysis. Pediatrics 131 (4), e1240–e1263.
Pascal, A., Govaert, P., Oostra, A., Naulaers, G., Ortibus, E., Van den Broeck, C., 2018. Neurodevelopmental outcome in very preterm and very-low-birthweight infants born over the past decade: a meta-analytic review. Dev. Med. Child Neurol. 60 (4), 342–355.
Petrou, S., 2003. Economic consequences of preterm birth and low birthweight. BJOG Int. J. Obstet. Gynaecol. 110, 17–23.
Petrou, S., Sach, T., Davidson, L., 2001. The long-term costs of preterm birth and low birth weight: results of a systematic review. Child Care Health Dev. 27 (2), 97–115.
Petrou, S., Yiu, H.H., Kwon, J., 2019. Economic consequences of preterm birth: a systematic review of the recent literature (2009–2017). Arch. Dis. Child. 104 (5), 456–465.
Petrou, S., Krabuanrat, N., Khan, K., 2020. Preference-based health-related quality of life outcomes associated with preterm birth: a systematic review and meta-analysis. Pharmacoeconomics 1–17.
Petticrew, M., Tugwell, P., Kristjansson, E., Oliver, S., Ueffing, E., Welch, V., 2012. Damned if you do, damned if you don't: subgroup analysis and equity. J. Epidemiol. Community Health 66 (1), 95–98.
Puthussery, S., Chutiyami, M., Tseng, P.-C., Kilby, L., Kapadia, J., 2018. Effectiveness of early intervention programs for parents of preterm infants: a meta-review of systematic reviews. BMC Pediatr. 18, 223. https://doi.org/10.1186/s12887-018-1182-3.
Rothwell, P.M., 2005. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. Lancet 365 (9454), 176–186.
Rubin, E.B., Buehler, A.E., Halpern, S.D., 2016. States worse than death among hospitalized patients with serious illnesses. JAMA Intern. Med. 176 (10), 1557–1559.
Sadique, Z., Grieve, R., Diaz-Ordaz, K., Mouncey, P., Lamontagne, F., O'Neill, S., 2022. A machine-learning approach for estimating subgroup-and individual-level treatment effects: an illustration using the 65 trial. Med. Decis. Mak. 42 (7), 923–936.
Saigal, S., Streiner, D.L., Stoskopf, B., Pinelli, J., Hoult, L., 2000. Self-perceptions and predictors of health-related quality of life in adolescents born very preterm and term born controls. Acta Paediatr. 89 (9), 1087–1096. https://doi.org/10.1111/j.1651-2227.2000.tb00617.x.
Saigal, S., Stoskopf, B., Streiner, D., Paneth, N., Pinelli, J., Boyle, M., 2006. Growth trajectories of extremely low birth weight infants from birth to young adulthood: a longitudinal, population-based study. Pediatr. Res. 60 (6), 751–758.
Samsa, G., Edelman, D., Rothman, M.L., Williams, G.R., Lipscomb, J., Matchar, D., 1999. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. Pharmacoeconomics 15, 141–155.
Serenius, F., Källén, K., Blennow, M., Ewald, U., Fellman, V., Holmström, G., Lindberg, E., Lundqvist, P., Marsál, K., Norman, M., et al., 2013. Neurodevelopmental outcome in extremely preterm infants at 2.5 years after active perinatal care in Sweden. JAMA 309 (17), 1810–1820.
Spittle, A.J., Cameron, K., Doyle, L.W., Cheong, J.L., Victorian Infant Collaborative Study Group, 2018. Motor impairment trends in extremely preterm children: 1991–2005. Pediatrics 141 (4), e20173458.
Steyerberg, E.W., 2009. Applications of Prediction Models. Springer.
Tamura, N., Hanaoka, T., Ito, K., Araki, A., Miyashita, C., Ito, S., Minakami, H., Cho, K., Endo, T., Sengoku, K., et al., 2018. Different risk factors for very low birth weight, term-small-for-gestational-age, or preterm birth in Japan. Int. J. Environ. Res. Publ. Health 15 (2), 369.
Torrance, G.W., Feeny, D.H., Furlong, W.J., Barr, R.D., Zhang, Y., Wang, Q., 1996. Multiattribute utility function for a comprehensive health status classification system: health Utilities Index Mark 2. Med. Care 34 (7), 702–722.
Treyvaud, K., Lee, K.J., Doyle, L.W., Anderson, P.J., 2014. Very preterm birth influences parental mental health and family outcomes seven years after birth. J. Pediatr. 164 (3), 515–521.
UNESCO Institute for Statistics, 2012. International standard classification of education: ISCED 2011. Comp. Soc. Res. 30.
van der Pal, S., Steinhof, M., Grevinga, M., Wolke, D., Verrips, G., 2020. Quality of life of adults born very preterm or very low birth weight: a systematic review. Acta Paediatr. 109 (10), 1974–1988.
van Lunenburg, A., van der Pal, S.M., van Dommelen, P., van der Pal-de, K.M., Bennebroek Gravenhorst, J., Verrips, G.H.W., et al., 2013. Changes in quality of life into adulthood after very preterm birth and/or very low birth weight in The Netherlands. Health Qual. Life Outcome 11 (1), 1–8.
Wager, S., Athey, S., 2018. Estimation and inference of heterogeneous treatment effects using random forests. J. Am. Stat. Assoc. 113 (523), 1228–1242.
Wallace, M.E., Mendola, P., Chen, Z., Hwang, B.S., Grantz, K.L., 2016. Preterm birth in the context of increasing income inequality. Matern. Child Health J. 20, 164–171.
Williams, D.R., Mohammed, S.A., Leavell, J., Collins, C., 2010. Race, socioeconomic status, and health: complexities, ongoing challenges, and research opportunities. Ann. N. Y. Acad. Sci. 1186 (1), 69–101.
Wolke, D., 2019. Is social inequality in cognitive outcomes increased by preterm birth–related complications? JAMA Netw. Open 2 (5), e192902.
Wong, H.S., Edwards, P., 2013. Nature or nurture: a systematic review of the effect of socioeconomic status on the developmental and cognitive outcomes of children born preterm. Matern. Child Health J. 17 (9), 1689–1700. https://doi.org/10.1007/s10995-012-1183-8.
Xie, Y., Brand, J.E., Jann, B., 2012. Estimating heterogeneous treatment effects with observational data. Sociol. Methodol. 42 (1), 314–347.
Zimmerman, E.B., Woolf, S.H., Haley, A., 2015. Understanding the relationship between education and health: a review of the evidence and an examination of community perspectives. Population health: behavioral and social science insights. Rockville (MD): Agency for Health-care Research and Quality 22 (1), 347–384.