

Preference-based health-related quality of life outcomes associated with preterm birth: A systematic review and meta-analysis

Short running head: Systematic review of utility values associated with preterm birth

Authors: Stavros Petrou,^{a,b} Natnaree Krabuanrat,^a Kamran Khan^a

^a Nuffield Department of Primary Care Health Sciences, University of Oxford
Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK.

^b Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, Coventry
CV4 7AL, UK.

Contact for correspondence:

Professor Stavros Petrou, Nuffield Department of Primary Care Health Sciences, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK.

Tel: 0044 (0)1865 289503

E-mail: stavros.petrou@phc.ox.ac.uk

Abstract

Objectives: Assessments of health-related quality of life outcomes associated with preterm birth provide valuable complementary data to the objective biomedical assessments that have traditionally been reported. The objective of this study was to perform a systematic review and meta-analysis of health utility values associated with preterm birth generated using preference-based approaches to health-related quality of life measurement.

Methods: Systematic searches of Medline, Web of Science, EconLit, Embase, CINAHL, PsycINFO, the Cochrane Library and SCOPUS were performed, covering the literature from inception of the search engines to 26th June 2018. Studies reporting health utility values estimated using either direct or indirect utility elicitation methods and published in the English language were included. Central descriptive statistics and measures of variability surrounding health utility values for each study and control group, and differences between comparator groups, are reported for each included article. The effect of preterm birth on health utility values was estimated using a hierarchical linear model in a linear mixed-effects meta-regression.

Results: Of 2,139 unique articles retrieved, 20 articles met the inclusion criteria. All but one study used the Health Utilities Index (HUI) Mark 2 (HUI2) or Mark 3 (HUI3) measures as their primary health utility assessment method. All studies reporting health utility values for individuals born preterm or at low birthweight and a control group of individuals born at full term or normal birthweight reported lower utility values in the study groups, regardless of age at assessment, respondent type or valuation method. The meta-regression revealed that preterm birth was associated with a mean utility decrement of 0.066 (95% CI: 0.035, 0.098; $p<0.001$) after controlling for valuation method, respondent type, administration mode, year of publication, geographical region of study, study setting and age at assessment.

Conclusion: Evidence identified by this review can act as data inputs into future economic evaluations of preventive or treatment interventions for preterm birth. Future research should focus particularly on estimating health utility values during the various stages of adulthood, and incorporating the effects of preterm birth on the preference-based health-related quality of life outcomes of parents and other family members.

Key Points for Decision Makers

- Although the adverse functional, neurodevelopmental and behavioural sequelae of preterm birth have received considerable attention, there has been no previous systematic review of health utility values associated with the condition.
- This systematic review and meta-analysis demonstrated that preterm birth and extremely low birthweight are associated with significant health utility decrements. Alternative methodological approaches to utility assessment, such as the valuation method and respondent type, had independent effects on health utility values.
- The study has generated a repository of data inputs for future economic evaluations of preventive or treatment interventions for preterm birth.

1 Introduction

Improved survival rates following preterm birth (<37 weeks' gestation) over recent decades have heightened interest in the problems surviving infants and their families face over the life course [1]. The functional, neurodevelopmental and behavioural sequelae of preterm birth are relatively well understood [2-6], and there is increasing interest in the economic challenges that these families face [7]. These data are increasingly used to inform evidence-based policy and research, as well as clinical and budgetary service planning [8].

There is increasing recognition of the need to measure the impact of preterm birth across multiple domains [9]. Assessments of health-related quality of life outcomes provide valuable complementary data to the objective biomedical assessments that have traditionally been reported [9]. Importantly, they also provide a framework for reflecting the multi-dimensional aspects of physical, mental, and social well-being of those born preterm and their families. Several assessments of health-related quality of life outcomes following preterm birth have applied multi-dimensional health profile measures that disaggregate outcomes across dimensions or subscales [10]. An alternative approach to assessing health-related quality of life outcomes following preterm birth involves the estimation of health utility values. These values summarise health-related quality of life outcomes on a cardinal scale typically anchored at 0 (representing death) and 1 (representing full health) [11]. Valuation methods for health utility assessment can be divided into two broad categories: direct and indirect. Direct valuation methods require respondents to value their own or hypothetical health state using scaling or choice-based techniques, such the visual analogue scale, standard gamble approach or time trade-off approach [12]. Indirect valuation methods use multi-attribute utility measures, such as the EuroQol EQ-5D [13] or the Health Utilities Index (HUI) [14], which are health status classification systems with pre-existing utility values or tariffs that

can be attached to each permutation of responses [15]. A particular appeal of health utility values is that they reflect preferences for health-related quality of life outcomes [15]. Furthermore, health utility values can act as data inputs into health economic evaluations and can therefore directly inform cost-effectiveness based decision-making [16].

In this paper, we report the first systematic review and meta-analysis of health utility values associated with preterm birth. In so doing, we generate a repository of data that can be viewed alongside the plethora of other outcomes associated with preterm birth that have been reported, and act as inputs into future economic evaluations of preventive or treatment interventions.

2 Methods

The systematic review and meta-analysis followed the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines [17].

2.1 Identification of studies

The search strategy was developed in collaboration with a local information specialist. It encompassed granulated thesauri terms and free text search terms associated with preterm birth (e.g. preterm or prematur* or low birthweight or gestational age) intersected with terms associated with health utility values (e.g. util* or EQ-5D or standard gamble or time trade-off) and is presented in full in Appendix 1. The electronic databases searched included Medline, Web of Science (including Index to Scientific and Technical Proceedings, Science Citation Index, Social Science Citation Index), EconLit, Embase, CINAHL, PsycINFO, Cochrane Library (including York Database of Abstracts of Reviews of Effectiveness, NHS Economic Evaluation Database), and SCOPUS. Searches of electronic databases were

supplemented by manual reference searching of bibliographies, contacts with experts in the field, citation searching and author searching. The time horizon of the search strategy had no initial date limit and extended to 26th June 2018. All reports identified by the searches, encompassing journal articles, monographs, book chapters and conference abstracts (hereafter ‘articles’ for brevity), were entered into EndNote vX9 and duplicates were removed.

2.2 Inclusion criteria

The inclusion criteria for the systematic review encompassed primary studies reporting health utility values associated with preterm birth (<37 weeks’ gestation) or low birthweight (<2500 g), estimated using either direct or indirect utility elicitation methods, and published in the English language. Studies estimating health utility values associated with low birthweight were included in order to capture potentially relevant information from contexts of clinical uncertainty in, and incomplete recording of, estimates of gestation. Studies were included regardless of age at assessment and regardless of whether the health utility values related to individuals born preterm or at low birthweight or to their parents, carers or siblings.

A two stage screening process was followed; the first stage assessed titles and abstracts, and the second stage assessed full text articles. Assessments at each stage of the screening process were performed independently by two reviewers (NK and KK), with disagreements referred to a third reviewer (SP) for the final assessment. We excluded studies at the full article stage that: (i) did not report health utility values associated with preterm birth or low birthweight; (ii) applied non-preference based approaches to the assessment of health-related quality of life outcomes; or (iii) treated preterm birth or low birthweight as a covariate rather than a primary exposure.

2.3 Data extracting and processing

From each article that met the study selection criteria, we extracted the following information about the characteristics of the study using a bespoke proforma and entered it into an excel database: (i) bibliographic details, including year of publication; (ii) country/geographical jurisdiction; (iii) setting (hospital, educational, community, other); (iv) study design (cross-sectional study, prospective, other); (v) age of target group at assessment; (vi) respondent type (self-assessment, proxy assessment by parents, proxy assessment by others, assessment by children and parent proxies); (vii) data collection method (interview, questionnaire, other); (viii) characteristics and size of study group by gestational age at birth and/or birthweight status; (ix) characteristics and size of control group; (x) comorbidities reported; (xi) direct valuation method applied (if applicable); (xii) indirect valuation method applied (if applicable); (xiii) utility tariff if indirect valuation method was applied; (xiv) utility scores (including central statistics, measures of variability and differences between comparator groups); (xv) statistical methods for analysing utility values; (xvi) response quality (response rate, information on dropouts, reasons for loss to follow-up, etc.); and (xvii) any reported methodological concerns. For studies that applied preference-based measures that generate both multi-attribute and single-attribute utility scores, e.g. HUI Mark 3 (HUI3), both sets of utility values were extracted.

Central descriptive statistics (e.g. means, medians) and measures of variability (e.g. standard deviation (SDs), interquartile ranges (IQRs)) surrounding health utility values for each study and control group, and differences between comparator groups, are reported for each article. Where available, the proportion of children with suboptimal levels of function (defined as below the highest level of function) are reported for study and control groups for individual attributes of multi-attribute utility measures.

2.4 Quality assessment of eligible studies

The quality of each study was assessed using the checklist published by Papaionnou and colleagues, which generates a maximum score of eight [18].

2.5 Statistical analysis

Our primary form of meta-analysis of health utility values took the form of a hierarchical linear model in a linear mixed-effects meta-regression [19]. The aim was to estimate the decrement in health utility score associated with preterm birth or low birthweight relative to that for controls matched for age, after controlling for methodological factors and study-specific random effects not accounted for by explanatory variables. The meta-regression was restricted to studies that applied either the HUI Mark 2 (HUI2) or HUI3 measures as all but one study included in the systematic review used one or both of these measures and it was expected that the valuation method exerts an independent effect on utility score [19].

Moreover, some health care decision-making bodies recommend the use of both the HUI2 and HUI3 for their reference case for economic evaluation purposes [20]. The HUI2 health status classification system covers seven attributes: sensation, mobility, emotion, cognition, self-care, pain and fertility, each with 3-5 levels ranging from normal function (level 1) to severe impairment (levels 3, 4 or 5) [14]. The HUI3 health status classification system covers eight attributes: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain, each with 5-6 levels ranging from normal function (level 1) to severe impairment (levels 5 or 6) [14]. The multi-attribute utility scoring algorithm for the HUI2 health status classification system was based on a survey of 293 parents of school age children in Hamilton, Canada, whilst that for the HUI3 health status classification system was based on a survey of 504 general population adults living Hamilton, Canada, with both algorithms generated using both a visual analogue scaling technique and a standard gamble instrument [15]. There is

evidence to suggest that the HUI3 health status classification system offers wider applicability in both clinical and general population health studies, improvements in a number of definitions, and an increased orthogonality of its attributes for structural independence [14].

Separate models were estimated for gestational age at birth and birthweight status. Mean utility scores were weighted by the inverse of their standard error, whilst the individual studies were weighted by the total number of their respondents. The explanatory variables incorporated into the meta-regression took categorical forms and included the valuation method (reference HUI2 vs HUI3), respondent type (reference self-response vs parent response), administration mode (reference self-administered vs interviewer vs self and interviewer), year of publication (after 2000 vs 2000 or earlier), geographical region of study (reference North America vs Europe), study setting (reference community vs educational vs hospital) and age at assessment (reference 8-12 years vs 13-18 vs 19 years and above). Separate analyses used fixed effect meta-analysis and random effects meta-analysis to further explore the relationship between gestational age at birth /birthweight status and HUI2 or HUI3 health utility scores [19]. All statistical analyses were conducted using STATA software, version 14 (Stata-Corp, College Station, TX).

3 Results

3.1 Data source and selection

A total of 4,190 articles were identified by the combined literature searches (Figure 1). Following removal of duplicates, 2,139 articles were screened at the title and abstract stage. Of these, only 24 articles satisfied the first stage inclusion criteria, whilst a further 4 articles

were excluded at the second stage of the review process. A total of 20 articles met the inclusion criteria at both stages and were included in the systematic review [21-40].

3.2 Study characteristics

The study characteristics reported by these articles are summarised in Table 1. Only two of the articles were published before the year 2000 [31, 32]. The main country settings for the studies were Canada [22, 31-35, 38], the Netherlands [24, 36-39], Germany [21, 38, 40], the United Kingdom [23, 26] and Australia.[29, 30] The target ages ranged from 5 years [28] to 36 years [33] with most studies focussing on adolescents. All but one study [28] used the HUI2 or HUI3 as their primary health utility assessment method. Furthermore, the study groups in all but one study [25] focussed on children born very preterm (<32 weeks) or at very low birthweight (<1500g). No studies reported health utility values for parents, carers or siblings of individuals born preterm or at low birthweight.

3.3 Descriptive statistics

Table 2 summarises descriptive statistics for the health utility outcomes reported by the 20 articles included in the systematic review. All 12 studies that report health utility values for individuals born preterm or at low birthweight and a control group of individuals born at full term or normal birthweight reported lower utility values in the study groups, regardless of age at assessment, respondent type or valuation method [21-23, 26, 28, 30-35, 40]. Mean health utility value differences ranged from 0.02 when the HUI3 was completed by two groups of adolescents born very preterm or very low birthweight versus full term or normal birthweight [21], and 0.73 when the HUI3 was completed for two groups of adolescents born very preterm or very low birthweight with moderate to severe disability versus full term [40]. Two studies reported changes in health utility outcomes based on the same Dutch cohort of

individuals born very preterm or very low birthweight. There was a small decline in the HUI3 multi-attribute utility score between ages 14 and 19 [39], which contrasted with a later significant increase in utility score between ages 19 and 28 following multiple imputation for missing values [37]. Overall quality scores for contributing studies ranged between five [25, 31] and eight [24, 26, 30, 37-39].

Four of the 20 articles included in the systematic review reported levels of function for study and control groups for individual attributes of the HUI3 multi-attribute utility measure [21, 23, 26, 40]. Increased levels of suboptimal (below level 1) function were observed in the study groups in four of the eight attributes, namely vision, speech, ambulation and dexterity, regardless of the gestational age and birthweight composition of the comparator groups, age at assessment and respondent type (Appendix 2).

3.4 Quantitative synthesis

Table 3 shows the results of the mixed-effects meta-regression, based on the studies that used either the HUI2 or HUI3 measure to estimate health utility values in a study group of individuals born preterm and a control group born at term. Preterm birth was associated with a mean utility decrement of 0.066 (95% confidence interval (CI): 0.035, 0.098; $p < 0.001$) after controlling for covariates. The use of the HUI2 measure, parents as respondents and assessment at 13-18 years of age were each associated with significantly higher health utility values in comparison to their respective referents, whilst assessment that did not rely solely on self-administration and assessment in a hospital setting were each associated with significantly lower health utility values in comparison to their respective referents. Separate mixed-effects meta-regression models that specified birthweight as the main exposure revealed that very low birthweight was associated with a mean utility decrement of 0.030 (95% CI: 0.030, 0.030; $p < 0.001$) and extremely low birthweight was associated with a mean

utility decrement of 0.068 (95% CI: 0.038, 0.098; $p < 0.001$) after controlling for covariates (Appendix 3).

The fixed-effect estimate of mean HUI2/HUI3 utility score was 0.868 (95% CI: 0.863, 0.887) for the study groups and 0.930 (95% CI: 0.925, 0.935) for the control groups. There was significant variability between individual mean utility values overall in both the study groups ($Q=238.93$; $p < 0.001$; I^2 statistic=92.9%) and control groups ($Q=145.53$; $p < 0.001$; I^2 statistic=93.1%), and in both sets of groups there was greater variability between mean utility values published after 2000 than those published in 2000 or earlier. The random effects estimate of mean HUI2/HUI3 utility score was 0.838 (95% CI: 0.818, 0.858) for the study groups and 0.919 (95% CI: 0.899, 0.940) for the control groups. Graphical representations of the results of the fixed effect meta-analysis and random effects meta-analysis are shown in Figure 2.

4 Discussion

Although the adverse functional, neurodevelopmental and behavioural sequelae of preterm birth have received considerable attention, there is increasing recognition of the importance of capturing the health and well-being of those born preterm and their families across multiple domains [9, 41]. Alternative approaches to health-related quality of life measurement provide vehicles for assessing outcomes across multi-dimensional aspects of physical, mental, and social well-being. Two previous reviews of the health-related quality of life of individuals born preterm have been published. The first review only included four studies that estimated health utility values using preference-based health-related quality of life measures, and did not perform a meta-analysis [10]. The second review focussed on assessment approaches and information sources only [42]. This is the first paper, to our

knowledge, to report a systematic review and meta-analysis of health utility values associated with preterm birth. In order to capture potentially relevant information from contexts of clinical uncertainty in and incomplete recording of estimates of gestation, the review also encompasses evidence surrounding health utility values associated with low birthweight.

Our meta-analyses revealed that preterm birth and extremely low birthweight are associated with reductions in mean HUI utility scores of 0.066 (95% CI: 0.035, 0.098) and 0.068 (95% CI: 0.038, 0.098), respectively. Notably, the lower limits of the confidence intervals surrounding these marginal disutility estimates exceed the 0.03 minimally important difference in utility score postulated in the literature as clinically important for evaluative purposes [43, 44]. Our estimates also exceed the marginal decrements in mean HUI3 utility scores associated with a wide range of childhood conditions, including obesity or type II diabetes (-0.032), allergies (-0.054) and upper respiratory infections and diseases (-0.038), revealed by our recent systematic review and meta-analysis of childhood utility values reported by 272 studies [45]. This suggests that our results are likely to be of interest and relevance to both clinical and policy decision-makers.

This systematic review revealed that alternative methodological approaches to utility assessment had independent effects on health utility values. Notably, the use of the HUI2 measure was associated with significantly higher health utility values in comparison to those generated by the HUI3, which may be explained, at least in part, by the differences in its dimension and item structures and the broader range of its utility scale. The use of parents as respondents was similarly associated with significantly higher health utility values. Empirical evidence of the concordance between parent and child ratings of attributes of children's health-related quality of life suggests that parents are able to accurately rate observable behaviors, such as physical functioning and physical symptoms, but are less

successful at identifying social or emotional impairments [46], highlighting the importance of capturing both perspectives when describing the health-related quality of life of individuals born preterm.

The main strength of this study is that it was conducted in accordance with recent methodological guidance for systematic reviews and meta-analyses of health utility values [19]. There are, however, a number of features of this study that merit further discussion. First, given the presentation of data in the contributing studies, gestational age at birth is treated as a categorical variable in our meta-analyses. The relationship between gestational age at birth and health utility values may be more complex in form, suggesting a need to treat gestational age as a continuous variable in future iterations of this research. Second, all but one study included in the systematic review used the HUI2 or HUI3 as their primary health utility assessment method and applied Canadian tariffs, generated in the 1990s, to responses to their descriptive systems. These tariffs may not necessarily reflect preference structures for HUI health states in other jurisdictions, nor reflect changes in preference structures for HUI health states that might have occurred over the past twenty years [47]. Moreover, recent national guidelines, from the UK for example, have recommended that alternative multi-attribute utility measures, such as the EQ-5D, valued using locally-derived tariff sets, are used for health technology assessment purposes [16]. A separate assessment of the relevance of contributing studies to the requirements of local health technology assessment agencies may therefore be required if data from this review are to act as inputs into future economic evaluations of preventive or treatment interventions for preterm birth [18]. Third, the review provides limited evidence that the utility decrements associated with preterm birth and low birthweight persist into early adulthood [21, 30, 33, 35]. Recently established collaborative research platforms with a focus on outcomes through adulthood, such as the ‘Adults Born Preterm International Collaboration’ (<http://www.apic-preterm.org/>) and the ‘Research on

European Children and Adults born Preterm' programme (<https://recap-preterm.eu/>), should provide opportunities for assessments of preference-based health-related quality of life outcomes at later ages. Fourth, there is limited evidence in the broader literature that suggests that preterm birth and low birthweight may also affect the health-related quality of life of parents when the offspring are children [48] and even when the offspring are adults [49]. Broader effects of preterm birth and low birthweight on the preference-based health-related quality of life outcomes of parents and other family members should therefore be assessed in future research studies.

How might the results of our study be used? At the most basic level, our health utility estimates can act as complementary data to the objective biomedical assessments of the effects of preterm birth and low birthweight that have been reported. More fundamentally, however, we have generated a significant new resource that can inform economic evaluations of preventive or treatment interventions for preterm birth. Specifically, in the context of decision-analytic modelling based economic evaluations, analysts generally lack the time and resources to estimate primary utility values for all health states of interest, and previous economic evaluations focussed on preterm birth and low birthweight have resorted to a limited pool of secondary utility evidence [50-53]. We therefore anticipate that economic analysts will draw upon our catalogue when parametrising their decision analytic models. The process of selecting health utility values from our catalogue should necessarily be informed by a number of factors. This should include an assessment of whether contributing studies applied the preferred or an accepted valuation method of the analyst's local health technology assessment agency or derived utility values on the basis of responses from the preferred or an accepted population group or, indeed, whether contributing studies were conducted in the age group(s) of interest. A particular methodological challenge may arise for analysts requiring estimates of the lifelong consequences of preterm birth or low birthweight

for their models. The mean adjusted utility decrements of 0.066 and 0.068 for preterm birth and extremely low birthweight generated by our mixed effect meta-regressions may not hold through the life course. Indeed, we provide tentative evidence that utility decrements ameliorate during adolescence. Until firmer evidence on the life course trajectories in health utility values becomes available on the basis of well-designed longitudinal studies, analysts are encouraged to be cautious when applying our pooled estimates across life stages and to reflect potential ameliorations in utility decrements within their sensitivity analyses.

5 Conclusion

In conclusion, this paper presents the first systematic review and meta-analysis of health utility values associated with preterm birth. Evidence identified by this review can act as data inputs into future economic evaluations of preventive or treatment interventions for preterm birth. Future research should focus particularly on estimating health utility values during the various stages of adulthood, and assessing the effects of preterm birth on the preference-based health-related quality of life outcomes of parents and other family members.

Acknowledgments

We would like to thank Samantha Johnson for her assistance with the development and piloting of the literature search strategies. We would also like to thank colleagues at the University of Warwick for comments on earlier drafts of the paper.

Author's Contributions

SP designed the study, extracted and reviewed the evidence from all contributing studies, wrote the article and acts as guarantor. NK and KK performed the literature searches and screened all titles and abstracts and full text articles. NK also extracted data from all full text articles. All authors read and approved the final version of the article. SP is the guarantor of the overall content.

Data Availability Statement

The data generated or analysed during the current study are available from the corresponding author on reasonable request.

Funding

SP receives support as a NIHR Senior Investigator (NF-SI-0616-10103). SP also receives funding from the European Union's 2020 research and innovation programme, 'Research on European Children and Adults born Preterm' (RECAP), under grant agreement number 733280. The views contained within this paper are those of the authors and not necessarily of the funders.

Conflict of interests

SP, NK and KK have no conflicts of interest.

Details of ethics approval

Ethics approval was not required.

Supplementary material

Appendix 1. Search terms applied in literature search strategy

Appendix 2. Number (%) of children with suboptimal levels of function by study and multi-attribute utility measure attributes.

Appendix 3. Mixed-effect meta-regression by hierarchical linear model of multi-attribute utility values for low birthweight measured by HUI2 and HUI3.

References

1. Harrison MS, Goldenberg RL. Global burden of prematurity. *Semin Fetal Neonatal Med.* 2016;21(2):74-9. doi:10.1016/j.siny.2015.12.007.
2. Larroque B, Ancel PY, Marret S, Marchand L, Andre M, Arnaud C et al. Neurodevelopmental disabilities and special care of 5-year-old children born before 33 weeks of gestation (the EPIPAGE study): a longitudinal cohort study. *Lancet.* 2008;371(9615):813-20. doi:10.1016/S0140-6736(08)60380-3.
3. Moore T, Hennessy EM, Myles J, Johnson SJ, Draper ES, Costeloe KL et al. Neurological and developmental outcome in extremely preterm children born in England in 1995 and 2006: the EPICure studies. *BMJ.* 2012;345:e7961. doi:10.1136/bmj.e7961.
4. Serenius F, Kallen K, Blennow M, Ewald U, Fellman V, Holmstrom G et al. Neurodevelopmental outcome in extremely preterm infants at 2.5 years after active perinatal care in Sweden. *JAMA.* 2013;309(17):1810-20. doi:10.1001/jama.2013.3786.
5. Wood NS, Marlow N, Costeloe K, Gibson AT, Wilkinson AR. Neurologic and developmental disability after extremely preterm birth. EPICure Study Group. *N Engl J Med.* 2000;343(6):378-84. doi:10.1056/NEJM200008103430601.
6. Johnson S, Wolke D. Behavioural outcomes and psychopathology during adolescence. *Early Hum Dev.* 2013;89(4):199-207. doi:10.1016/j.earlhumdev.2013.01.014.
7. Petrou S, Yiu HH, Kwon J. Economic consequences of preterm birth: a systematic review of the recent literature (2009-2017). *Arch Dis Child.* 2018. doi:10.1136/archdischild-2018-315778.
8. Horvath H, Brindis CD, Reyes EM, Yamey G, Franck L, Knowledge T et al. Preterm birth: the role of knowledge transfer and exchange. *Health Res Policy Syst.* 2017;15(1):78. doi:10.1186/s12961-017-0238-0.

9. Saigal S, Doyle LW. An overview of mortality and sequelae of preterm birth from infancy to adulthood. *Lancet*. 2008;371(9608):261-9. doi:10.1016/S0140-6736(08)60136-1.
10. Zwicker JG, Harris SR. Quality of life of formerly preterm and very low birth weight infants from preschool age to adulthood: a systematic review. *Pediatrics*. 2008;121(2):e366-76. doi:10.1542/peds.2007-0169.
11. Torrance GW, Feeny D. Utilities and quality-adjusted life years. *Int J Technol Assess Health Care*. 1989;5(4):559-75.
12. Arnold D, Girling A, Stevens A, Lilford R. Comparison of direct and indirect methods of estimating health state utilities for resource allocation: review and empirical analysis. *BMJ*. 2009;339:b2688. doi:10.1136/bmj.b2688.
13. Brooks R. EuroQol: The current state of play. *Health Policy*. 1996;37(1):53-72. doi:10.1016/0168-8510(96)00822-6.
14. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems. *Health Utilities Index. Pharmacoeconomics*. 1995;7(6):490-502.
15. Brazier J, Ratcliffe, J., Salomon, J., Tsuchiya, A. *Measuring and valuing health benefits for economic evaluation*. Oxford: Oxford University Press; 2007.
16. National Institute for Health and Care Excellence (NICE). *Guide to the methods of technology appraisal 2013*. London, UK: NICE;2013.
17. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009;339:b2535. doi:10.1136/bmj.b2535.
18. Papaioannou D, Brazier J, Paisley S. Systematic searching and selection of health state utility values from the literature. *Value Health*. 2013;16(4):686-95. doi:10.1016/j.jval.2013.02.017.

19. Petrou S, Kwon J, Madan J. A Practical Guide to Conducting a Systematic Review and Meta-analysis of health state utility values. *Pharmacoeconomics*. 2018;36(9):1043-61. doi:10.1007/s40273-018-0670-1.
20. Canadian Agency for Drugs and Technologies in Health (CADTH). Guidelines for the economic evaluation of health technologies. 4th Edition. . Ottawa, Canada: CADTH; 2017.
21. Baumann N, Bartmann P, Wolke D. Health-Related Quality of life into adulthood after very preterm birth. *Pediatrics*. 2016;137(4). doi:10.1542/peds.2015-3148.
22. Feeny D, Furlong W, Saigal S, Sun J. Comparing directly measured standard gamble scores to HUI2 and HUI3 utility scores: group- and individual-level comparisons. *Soc Sci Med*. 2004;58(4):799-809.
23. Gray R, Petrou S, Hockley C, Gardner F. Self-reported health status and health-related quality of life of teenagers who were born before 29 weeks' gestational age. *Pediatrics*. 2007;120(1):e86-93. doi:10.1542/peds.2006-2034.
24. Hille ET, Weisglas-Kuperus N, van Goudoever JB, Jacobusse GW, Ens-Dokkum MH, de Groot L et al. Functional outcomes and participation in young adulthood for very preterm and very low birth weight infants: the Dutch Project on Preterm and Small for Gestational Age Infants at 19 years of age. *Pediatrics*. 2007;120(3):e587-95. doi:10.1542/peds.2006-2407.
25. James JD. Health status and psychological adjustment in low birth weight and normal birth weight Jamaican preadolescents. Northwestern University; 2003.
26. Petrou S, Abangma G, Johnson S, Wolke D, Marlow N. Costs and health utilities associated with extremely preterm birth: evidence from the EPICure study. *Value Health*. 2009;12(8):1124-34. doi:10.1111/j.1524-4733.2009.00580.x.

27. Quinn GE, Dobson V, Saigal S, Phelps DL, Hardy RJ, Tung B et al. Health-related quality of life at age 10 years in very low-birth-weight children with and without threshold retinopathy of prematurity. *Arch Ophthalmol.* 2004;122(11):1659-66. doi:10.1001/archopht.122.11.1659.
28. Rautava L, Hakkinen U, Korvenranta E, Andersson S, Gissler M, Hallman M et al. Health-related quality of life in 5-year-old very low birth weight infants. *J Pediatr.* 2009;155(3):338-43 e1-3. doi:10.1016/j.jpeds.2009.03.061.
29. Roberts G, Anderson PJ, Cheong J, Doyle LW, Victorian Infant Collaborative Study G. Parent-reported health in extremely preterm and extremely low-birthweight children at age 8 years compared with comparison children born at term. *Dev Med Child Neurol.* 2011;53(10):927-32. doi:10.1111/j.1469-8749.2011.04025.x.
30. Roberts G, Burnett AC, Lee KJ, Cheong J, Wood SJ, Anderson PJ et al. Quality of life at age 18 years after extremely preterm birth in the post-surfactant era. *J Pediatr.* 2013;163(4):1008-13 e1. doi:10.1016/j.jpeds.2013.05.048.
31. Saigal S, Feeny D, Furlong W, Rosenbaum P, Burrows E, Torrance G. Comparison of the health-related quality of life of extremely low birth weight children and a reference group of children at age eight years. *J Pediatr.* 1994;125(3):418-25.
32. Saigal S, Feeny D, Rosenbaum P, Furlong W, Burrows E, Stoskopf B. Self-perceived health status and health-related quality of life of extremely low-birth-weight infants at adolescence. *JAMA.* 1996;276(6):453-9.
33. Saigal S, Ferro MA, Van Lieshout RJ, Schmidt LA, Morrison KM, Boyle MH. Health-Related Quality of Life Trajectories of Extremely Low Birth Weight Survivors into Adulthood. *J Pediatr.* 2016;179:68-73 e1. doi:10.1016/j.jpeds.2016.08.018.
34. Saigal S, Rosenbaum PL, Feeny D, Burrows E, Furlong W, Stoskopf BL et al. Parental perspectives of the health status and health-related quality of life of teen-aged children

- who were extremely low birth weight and term controls. *Pediatrics*. 2000;105(3 Pt 1):569-74.
35. Saigal S, Stoskopf B, Pinelli J, Streiner D, Hoult L, Paneth N et al. Self-perceived health-related quality of life of former extremely low birth weight infants at young adulthood. *Pediatrics*. 2006;118(3):1140-8. doi:10.1542/peds.2006-0119.
 36. van Dommelen P, van der Pal SM, Bennebroek Gravenhorst J, Walther FJ, Wit JM, van der Pal de Bruin KM. The effect of early catch-up growth on health and well-being in young adults. *Ann Nutr Metab*. 2014;65(2-3):220-6. doi:10.1159/000364884.
 37. van Lunenburg A, van der Pal SM, van Dommelen P, van der Pal-de Bruin KM, Bennebroek Gravenhorst J, Verrips GH. Changes in quality of life into adulthood after very preterm birth and/or very low birth weight in the Netherlands. *Health Qual Life Outcomes*. 2013;11:51. doi:10.1186/1477-7525-11-51.
 38. Verrips E, Vogels T, Saigal S, Wolke D, Meyer R, Hoult L et al. Health-related quality of life for extremely low birth weight adolescents in Canada, Germany, and the Netherlands. *Pediatrics*. 2008;122(3):556-61. doi:10.1542/peds.2007-1043.
 39. Verrips G, Brouwer L, Vogels T, Taal E, Drossaert C, Feeny D et al. Long term follow-up of health-related quality of life in young adults born very preterm or with a very low birth weight. *Health Qual Life Outcomes*. 2012;10:49. doi:10.1186/1477-7525-10-49.
 40. Wolke D, Chernova J, Eryigit-Madzwamuse S, Samara M, Zwierzynska K, Petrou S. Self and parent perspectives on health-related quality of life of adolescents born very preterm. *J Pediatr*. 2013;163(4):1020-6 e2. doi:10.1016/j.jpeds.2013.04.030.
 41. McCormick MC, Litt JS. The Outcomes of Very Preterm Infants: Is It Time to Ask Different Questions? *Pediatrics*. 2017;139(1). doi:10.1542/peds.2016-1694.

42. Vieira ME, Linhares MB. Quality of life of individuals born preterm: a systematic review of assessment approaches. *Qual Life Res.* 2016;25(9):2123-39. doi:10.1007/s11136-016-1259-9.
43. Drummond M. Introducing economic and quality of life measurements into clinical studies. *Ann Med.* 2001;33(5):344-9.
44. Marra CA, Woolcott JC, Kopec JA, Shojania K, Offer R, Brazier JE et al. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. *Soc Sci Med.* 2005;60(7):1571-82. doi:10.1016/j.socscimed.2004.08.034.
45. Kwon J, Kim SW, Ungar WJ, Tsiplova K, Madan J, Petrou S. A Systematic Review and Meta-Analysis of Childhood Health Utilities. *Med Decis Making.* 2018;38(3):277-305. doi:10.1177/0272989X17732990.
46. Eiser C, Morse R. Quality-of-life measures in chronic diseases of childhood. *Health Technol Assess.* 2001;5(4):1-157.
47. McCabe C, Stevens K, Roberts J, Brazier J. Health state values for the HUI 2 descriptive system: results from a UK survey. *Health Econ.* 2005;14(3):231-44. doi:10.1002/hec.925.
48. Lakshmanan A, Agni M, Lieu T, Fleegler E, Kipke M, Friedlich PS et al. The impact of preterm birth <37 weeks on parents and families: a cross-sectional study in the 2 years after discharge from the neonatal intensive care unit. *Health Qual Life Outcomes.* 2017;15(1):38. doi:10.1186/s12955-017-0602-3.
49. Wolke D, Baumann N, Busch B, Bartmann P. Very preterm birth and parents' quality of life 27 years later. *Pediatrics.* 2017;140(3). doi:10.1542/peds.2017-1263.
50. Honest H, Forbes CA, Duree KH, Norman G, Duffy SB, Tsourapas A et al. Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness

- literature with economic modelling. *Health Technol Assess.* 2009;13(43):1-627.
doi:10.3310/hta13430.
51. McCurdy RJ, Baxter JK. Universal cervical length screening with a cervicometer to prevent preterm birth <34 weeks: a decision and economic analysis. *J. Matern.-Fetal Neonatal Med.* 2019:1-10. doi:10.1080/14767058.2019.1583202.
 52. Werner EF, Hamel MS, Orzechowski K, Berghella V, Thung SF. Cost-effectiveness of transvaginal ultrasound cervical length screening in singletons without a prior preterm birth: an update. *Am J Obstet Gynecol.* 2015;213(4):554 e1-6.
doi:10.1016/j.ajog.2015.06.020.
 53. Moya MP, Goldberg RN. Cost-effectiveness of prophylactic indomethacin in very-low-birth-weight infants. *Ann Pharmacotherapy.* 2002;36(2):218-24. doi:10.1345/aph.10347.

Table 1. Description of study characteristics

Authors	Year	Country	Setting	Study Design	Age at assessment (years)	Sample Size: Baseline		Valuation Method (Country Source of Tariff Set)	Respondent Type	Data-Collection Method
						Study group	Control Group			
Baumann <i>et al</i> [21]	2016	Germany	Community	Prospective cohort study	13 26	260 VP/VLBW (<32 weeks or <1500 g)	229 FT (37-41 weeks/ >2500 g)	HUI3 (Canada)	Self and parents	Questionnaires
Feeny <i>et al</i> [22]	2004	Canada	Community	Prospective cohort study	Mean (SD) Cases: 14.0 (1.6) Controls: 14.4 (1.3)	140 ELBW (<1000 g)	124 FT (37-41 weeks)	SG, HUI2 (Canada) HUI3 (Canada)	Self	Interviews
Gray <i>et al</i> [23]	2007	UK	Community	Prospective cohort study, contemporaneous controls	15-16	140 ELGA (<29 weeks)	108 FT (37-41 weeks)	HUI3 (Canada)	Self	Questionnaires
Hille <i>et al</i> [24]	2007	The Netherlands	Hospital	Prospective cohort study	Mean (SD) 19.3 (0.2)	705 VP/VLBW (<32 weeks or <1500g)	N/A	HUI3 (Canada)	Self or parents when individuals incapable	Questionnaires
James [25]	2003	Jamaica	Educational	Prospective cohort study	11–12	115 LBW (<2500g)	119 NBW (>2500 g)	HUI2 (Canada)	Self	Questionnaires
Petrou <i>et al</i> [26]	2009	UK & Republic of Ireland	Community	Prospective cohort study	11	219 EP (20-25 weeks)	153 FT (37-41 weeks)	HUI3 (Canada)	Parents	Questionnaires
Quinn <i>et al</i> [27]	2004	USA	Hospital	Prospective cohort study, contemporaneous controls	10	255 VLBW (<1251 g) with ROP	104 VLBW (< 1251 g) without ROP	HUI3 (Canada)	Parents	Interviews
Rautava <i>et al</i> [28]	2009	Finland	Community	National population register study	5	588 VP/VLBW (<32 weeks or ≤1500 g)	176 FT (37-41 weeks)	17D (Finland)	Parents	Questionnaires
Roberts <i>et al</i> [29]	2011	Australia	Hospital	Prospective cohort study	8	189 EP/ELBW (<28 weeks or <1000 g)	173 Term (>36 weeks or >2499 g)	HUI2 (Canada)	Parents	Questionnaires
Roberts <i>et al</i> [30]	2013	Australia	Community	Prospective cohort study	18	194 EP/ELBW (<28 weeks or <1000 g)	148 NBW (>2499 g)	HUI3 (Canada)	Self	Questionnaires

Saigal <i>et al</i> [31]	1994	Canada	Hospital	Prospective cohort study, contemporaneous controls	8	156 ELBW (501-1000 g)	145 students matched for age, sex, and socioeconomic status	HUI2 (Canada)	Parents	Questionnaires
Saigal <i>et al</i> [32]	1996	Canada	Educational	Prospective cohort study, contemporaneous controls	12-16	141 ELBW (<1000 g)	124 Term (37-41 weeks and >2500 g)	HUI2 (Canada) SG VAS	Self; parents for 9 severely impaired teenagers	Interviews
Saigal <i>et al</i> [33]	2016	Canada	Hospital	Prospective cohort study, contemporaneous controls	12-16, 22-26, 29-36	153 ELBW (<1000 g)	137 NBW (> 2500 g)	HUI3 (Canada)	Self	Interviews and questionnaires
Saigal <i>et al</i> [34]	2000	Canada	Educational	Prospective cohort study, contemporaneous controls	12-16	149 ELBW (\leq 1000 g)	126 Term/NBW (37-41 weeks and >2500 g)	HUI2 (Canada) SG VAS	Parents	Interviews
Saigal <i>et al</i> [35]	2006	Canada	Educational	Prospective cohort study, contemporaneous controls	ELBW: 23.3 (\pm 1.2) NBW: 23.6 (\pm 1.1)	143 ELBW (501-1000 g)	130 NBW (> 2500 g)	HUI2 (Canada) SG	Self	Interviews
Van Dommelen <i>et al</i> [36]	2014	The Netherlands	Hospital	Prospective cohort study	19	334 individuals born <32 weeks and/or at <1500g and SGA and without severe complications	None	HUI3 (Canada)	Self	Questionnaires
Van Lunenburg <i>et al</i> [37]	2013	The Netherlands	Community	Prospective cohort study	19 28	314 VP/VLBW (<32 weeks or <1500 g)	None	HUI3 (Canada)	Self	Online and postal questionnaires
Verrips <i>et al</i> [38]	2008	Canada, Germany, The Netherlands	Community	Multinational, prospective cohort studies, contemporaneous controls	12-16	341 ELBW (\leq 1000 g)	1737 Term or drawn from normative population samples	HUI3 (Canada)	Self and parents	Interviews or questionnaires
Verrips <i>et al</i> [39]	2012	The Netherlands	Community	Prospective cohort study	14 19	630 VP/VLBW (<32 weeks or <1500 g)	None	HUI3 (Canada)	Self and parents	Questionnaires
Wolke <i>et al</i> [40]	2013	Germany	Community	Prospective cohort study, contemporaneous controls	13	272 VP/VLBW (<32 weeks or <1500 g)	282 Term	HUI3 (Canada)	Self and parents	Questionnaires

ELBW, Extremely low birthweight; ELGA, Extremely low gestational age; EP, Extremely preterm; FT, Full term; HUI2, Health Utilities Index Mark 2; HUI3, Health Utilities Index Mark 3; LBW, Low birthweight; NBW, Normal birthweight; ROP, Retinopathy of prematurity; SD, Standard deviation; SG, Standard Gamble; SGA, small for gestational age; VAS, Visual analogue scale; VLBW, Very low birthweight; VP, very preterm; 17D, 17-Dimensional Health Related Measure; N/A, Not applicable.

Table 2. Description of health utility outcomes

Authors/Study	Measure	Unit of analysis	Mean (SD), [median (range)] utility value		Utility score differences	Quality score
			Study Group	Control Group		
Baumann <i>et al</i> [21]	HUI3	Self-reports				7
		<i>Adolescence</i>	0.86	0.88	Mean utility difference of 0.02	
		<i>Adulthood</i>	0.86	0.89	Mean utility difference of 0.03	
		Parent reports				
		<i>Adolescence</i>	0.88	0.92	Mean utility difference of 0.04	
Feeny <i>et al</i> [22]	SG	<i>Adulthood</i>	0.86	0.94	Mean utility differences of 0.08	7
		All participants	0.90 (0.20), [1.00 (-0.05 to 1.00)]	0.93 (0.11), [0.95 (0.55 to 1.00)]	Mean utility difference of 0.03	
		All participants	0.89 (0.14), [0.95 (0.38 to 1.00)]	0.95 (0.09), [1.00 (0.34 to 1.00)]	Mean utility difference of 0.06 ($p<0:0001$)	
		All participants	0.80 (0.22), [0.87 (-0.07 to 1.00)]	0.89 (0.13), [0.93 (0.38 to 1.00)]	Mean utility difference of 0.09 ($p<0:0002$)	
Gray <i>et al</i> [23]	HUI3	All participants	0.86 (0.19), [0.93 (0.07 to 1.00)]	0.89 (0.12), [0.93 (0.45 to 1.00)]	Mean utility difference of 0.03 ($p=0.123$)	7
Hille <i>et al</i> [24]	HUI3	All participants	Not reported	Not reported	Not reported	8
James [25]	HUI2	All participants	Not reported	Not reported	Not reported	5
Petrou <i>et al</i> [26]	HUI3	All participants	0.789 (0.264)	0.956 (0.102)	Mean utility difference of 0.167 ($p<0:001$)	8
		≤ 23 weeks	0.772 (0.291)		Mean utility difference of 0.184 ($p=0:016$)	
		24 weeks	0.717 (0.333)		Mean utility difference of 0.239 ($p<0:001$)	
		25 weeks	0.830 (0.208)		Mean utility difference of 0.126 ($p<0:001$)	
Quinn <i>et al</i> [27]	HUI3	All participants	0.59 (0.39), [0.72 (-0.36 to 1.00)]	0.90 (0.16), [0.97 (-0.01 to 1.00)]	Median utility value for ROP participants with blindness/low vision in better eye significantly lower than the median utility value for non-ROP participants ($p<0.001$).	6
		<i>ROP participants with blindness/low vision in better eye</i>	0.25 (0.37), [0.27 (-0.36 to 1.00)]		Median utility value for ROP participants with blindness/low vision in better eye significantly lower than the median utility value for the sighted ROP participants ($p<0.001$).	
		<i>ROP participants sighted in better eye</i>	0.78 (0.25), [0.87 (-0.24 to 1.00)]			
Rautava <i>et al</i> [28]	17D	23 weeks	0.92	0.98	The adjusted 17D utility score difference (following adjustments for sex, family income, mother's employment situation and family structure) between all study group participants	7
		24-25 weeks	0.94			
		26-27 weeks	0.96			
		28-29 weeks	0.96			
		30-31 weeks	0.97			

		>31 weeks	0.97		and control group participants was 0.02 (95% CI (0.01 to 0.03)).	
Roberts <i>et al</i> [29]	HUI2	All participants	Not reported	Not reported	Not reported	7
Roberts <i>et al</i> [30]	HUI3	All participants	[0.93, IQR of 0.82 to 1.0]	[0.95, IQR of 0.86 to 1.0]	Non-significant (<i>p</i> =0.18) difference in median utility scores between comparator groups.	8
Saigal <i>et al</i> [31]	HUI2	All participants	0.82 (0.21), [0.88 (0.09 to 1.00)]	0.95 (0.07), [1.00 (0.59 to 1.00)]	Mean utility difference of 0.13	5
Saigal <i>et al</i> [32]	HUI2	All participants	0.87 (0.26), [1.00 (-0.50 to 1.00)]	0.93 (0.11), [0.95 (0.55 to 1.00)]	Mean utility difference of 0.06 (<i>p</i> =0.02)	6
	VAS	All participants	87 (20), [94, (-1 to 100)]	91 (8), [93, (46 to 100)]	Mean VAS score difference of 4 (<i>p</i> =0.01)	
Saigal <i>et al</i> [33]	HUI3	Adolescents (12-16 years)	0.831 (without NSI), 0.680 (with NSI)	0.884	Mean utility differences of 0.053 and 0.204 between NBW individuals and those born at ELBW without and with NSI, respectively. Mean utility differences of 0.064 and 0.239 between NBW individuals and those born at ELBW without and with NSI, respectively. Mean utility differences of 0.087 and 0.253 between NBW individuals and those born at ELBW without and with NSI, respectively.	7
		Young adults (22-26 years)	0.829 (without NSI), 0.654 (with NSI)	0.893		
		Adults (29-36 years)	0.766 (without NSI), 0.600 (with NSI)	0.853		
Saigal <i>et al</i> [34]	HUI2	All participants	0.91 (0.20), [1.00 (-0.10,1.0)]	0.97 (0.08), [1.00 (-0.45,1.0)]	Mean utility difference of 0.06 (95% CI: 0.02 to 0.09; <i>p</i> =0.002)	7
Saigal <i>et al</i> [35]	SG	All participants	0.85 (0.26) [0.95]	0.88 (0.24) [0.95]	Mean utility differences of 0.03 (95% CI: -0.03 to 0.09; <i>p</i> =0.32)	7
		<i>ELBW participants without NSIs</i>	0.85 (0.27)			
		<i>ELBW participants with NSIs</i>	0.85 (0.26)			
Van Dommelen <i>et al</i> [36]	HUI3	All participants	Not reported	Not reported	Not reported	7
Van Lunenburg <i>et al</i> [37]	HUI3	All participants			Mean utility score change over time of 0.01 (SD: 0.15).	8
		<i>Age 19</i>	0.89 (0.16)	N/A		
		<i>Age 28</i>	0.88 (0.16)	N/A		
		All participants (after multiple imputation)			Mean utility score change over time of 0.02 (SD: 0.17; 95% CI: 0.01 to 0.03; <i>p</i> <0.05).	
		<i>Age 19</i>	0.83 (0.22)	N/A		
<i>Age 28</i>	0.85 (0.20)	N/A				

Verrips <i>et al</i> [38]	HUI3	Canadian participants	0.76 (0.27)	Not reported	Statistically significant differences in utility scores between the three national cohorts ($p<0.001$). No comparisons with controls reported.	8
		German participants	0.75 (0.29)	Not reported		
		Dutch participants	0.87 (0.20)	Not reported		
Verrips <i>et al</i> [39]	HUI3	Participants at age 14	0.87 (0.18)	N/A	Mean utility score change over time of 0.01 (SD: 0.18).	8
		Participants at age 19	0.86 (0.20)	N/A		
Wolke <i>et al</i> [40]	HUI3	Parent reports		0.91 (0.11)	Mean difference in parent reported utility score of: (i) 0.03 between term adolescents and VLBW/VP adolescents with no or mild disability; and (ii) 0.73 between term adolescents and VLBW/VP adolescents with moderate or severe disability	7
		VLBW/VP I: no or mild disability	0.88 (0.18)			
		VLBW/VP II: moderate to severe disability	0.18 (0.35)			
		Adolescent reports		0.87 (0.15)	Mean difference in self-reported utility score of 0.03 between term adolescents and VLBW/VP adolescents with no or mild disability	
		VLBW/VP I: no or mild disability	0.84 (0.18)			
		VLBW/VP II: moderate to severe disability				

CI, confidence interval; ELBW, Extremely low birthweight; HUI2, Health Utilities Index Mark 2; HUI3, Health Utilities Index Mark 3; IQR, Inter-quartile range; NBW, Normal birthweight; NSI, neurosensory impairments; ROP, Retinopathy of prematurity; SD, Standard deviation; SG, Standard Gamble; Very low birthweight; VP, very preterm; 17D, 17-Dimensional Health Related Measure; N/A, Not applicable.

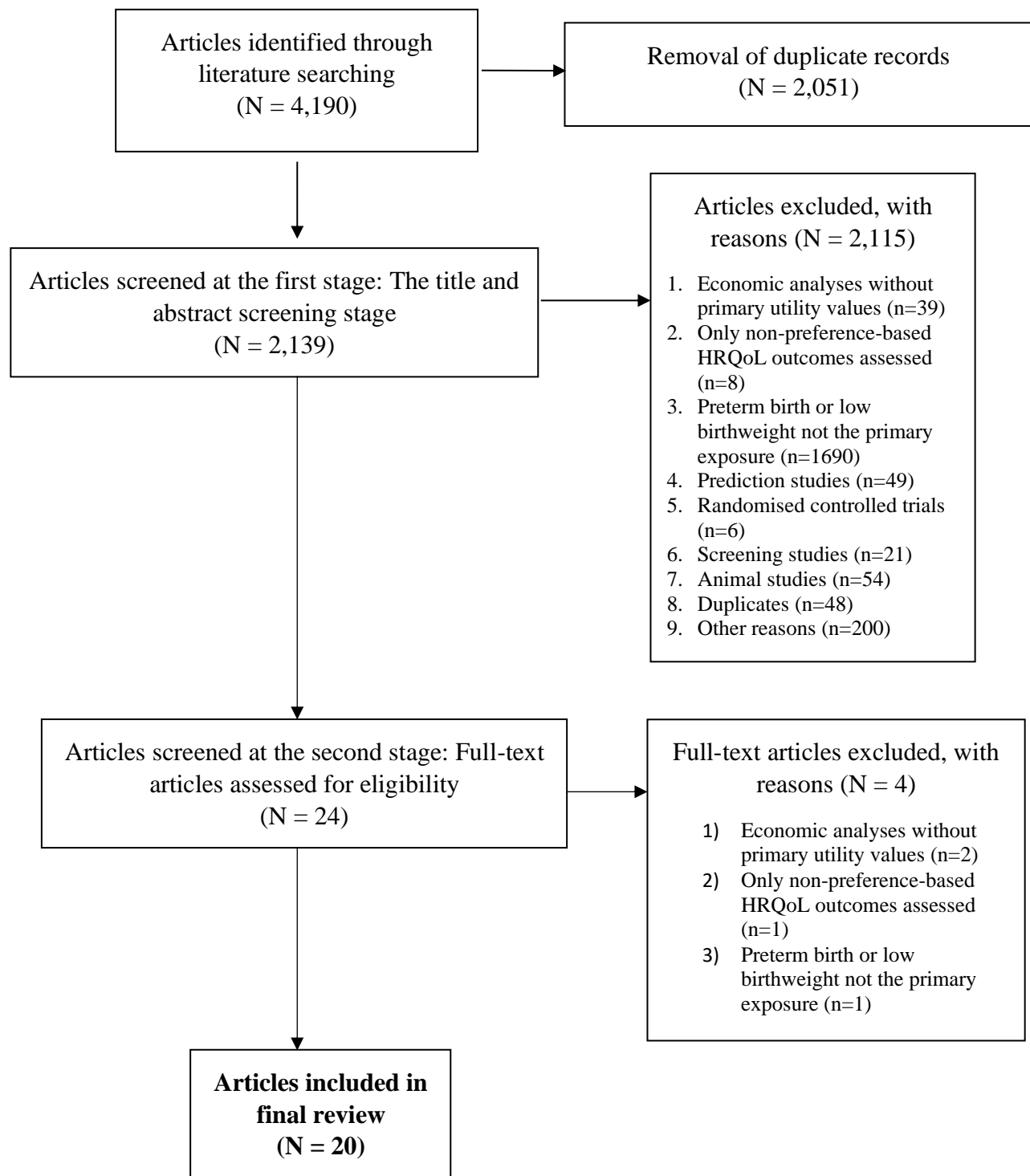
Table 3. Mixed-effect meta-regression by hierarchical linear model of multi-attribute utility values for preterm birth measured by HUI2 and HUI3

	Coefficient (SE)	(95% CI)	p-value
Constant [±]	0.820 (0.034)	(0.752, 0.887)	<0.001
Gestational age at birth			
Term	Reference		
Preterm	-0.066 (0.016)	(-0.098, -0.035)	<0.001
Valuation method			
HUI3	Reference		
HUI2	0.133 (0.034)	(0.067, 0.200)	<0.001
Respondent type			
Self-response	Reference		
Parent response	0.047 (0.008)	(0.031, 0.063)	<0.001
Administration mode			
Self-administered	Reference		
Interviewer-administered	-0.091 (0.028)	(-0.146, -0.035)	0.001
Self and interviewer administered	-0.088 (0.042)	(-0.170, -0.006)	0.036
Year of publication			
After 2000	Reference		
2000 or earlier	-0.016 (0.018)	(-0.051, 0.020)	0.390
Region study is based			
North America	Reference		
Europe	0.058 (0.040)	(-0.020, 0.137)	0.146
Setting			
Community	Reference		
Educational	-0.031 (0.020)	(-0.070, 0.007)	0.113
Hospital	-0.071 (0.018)	(-0.105, -0.036)	<0.001
Age at assessment (years)			
Ages 8-12	Reference		
Ages 13-18	0.075 (0.025)	(0.025, 0.124)	0.003
Age 19+	0.043 (0.029)	(-0.014, 0.101)	0.139

SE denotes standard error; CI denotes confidence interval; HUI denotes Health Utilities Index.

[±] Constant represents utility values for individuals born at term, measured using the HUI3, using self-responses and self-administered questionnaires, after 2000, in North American community settings, between ages 8-12 years.

Figure 1. PRISMA[#] Flow Diagram



[#] PRISMA denotes preferred reporting items for systematic reviews and meta-analyses; HRQoL denotes health-related quality of life.

Figure 2(i). Fixed-effect meta-analysis for the mean utility values for the study group populations

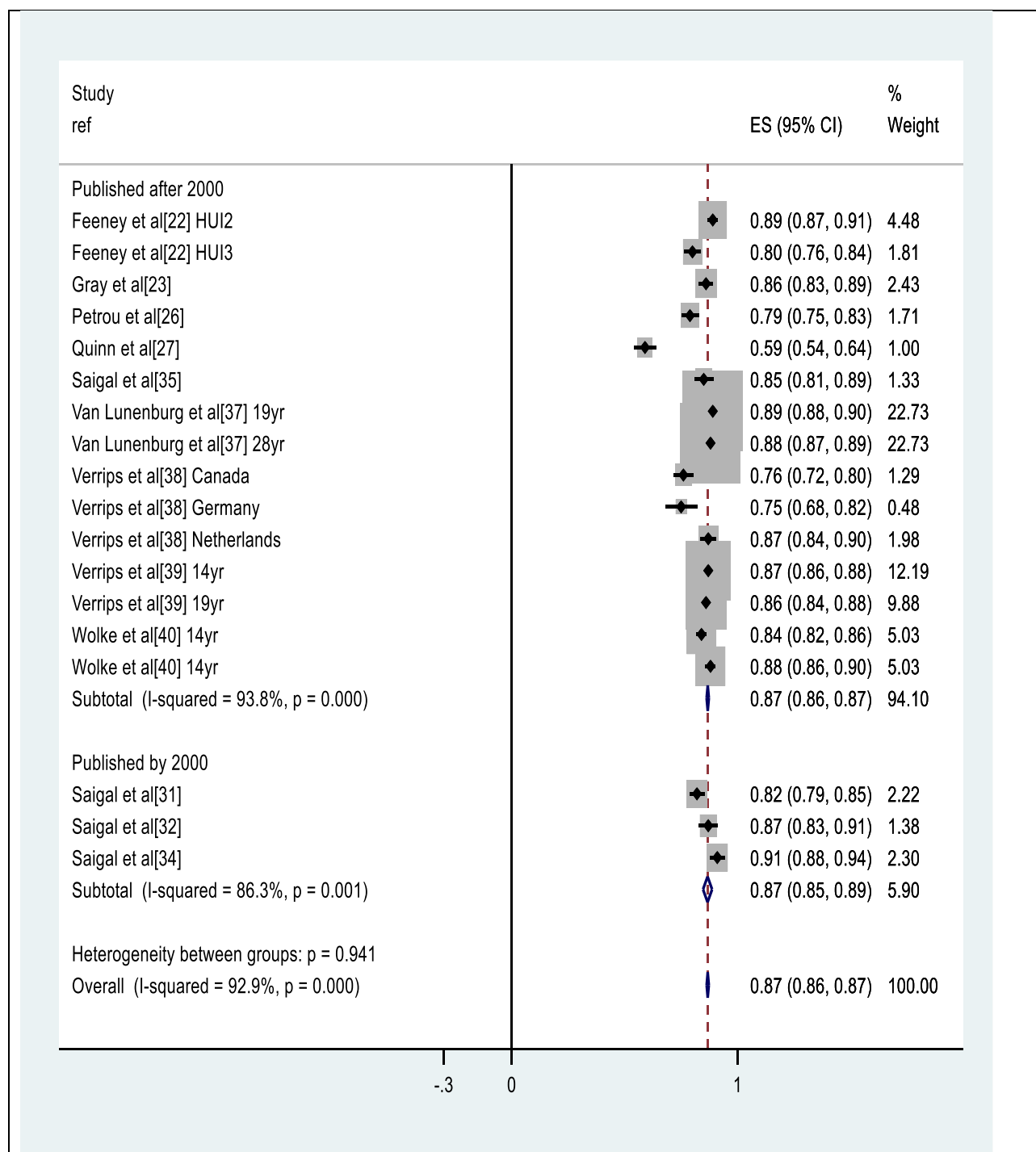


Figure 2(ii). Fixed-effect meta-analysis for the mean utility values for the control group populations

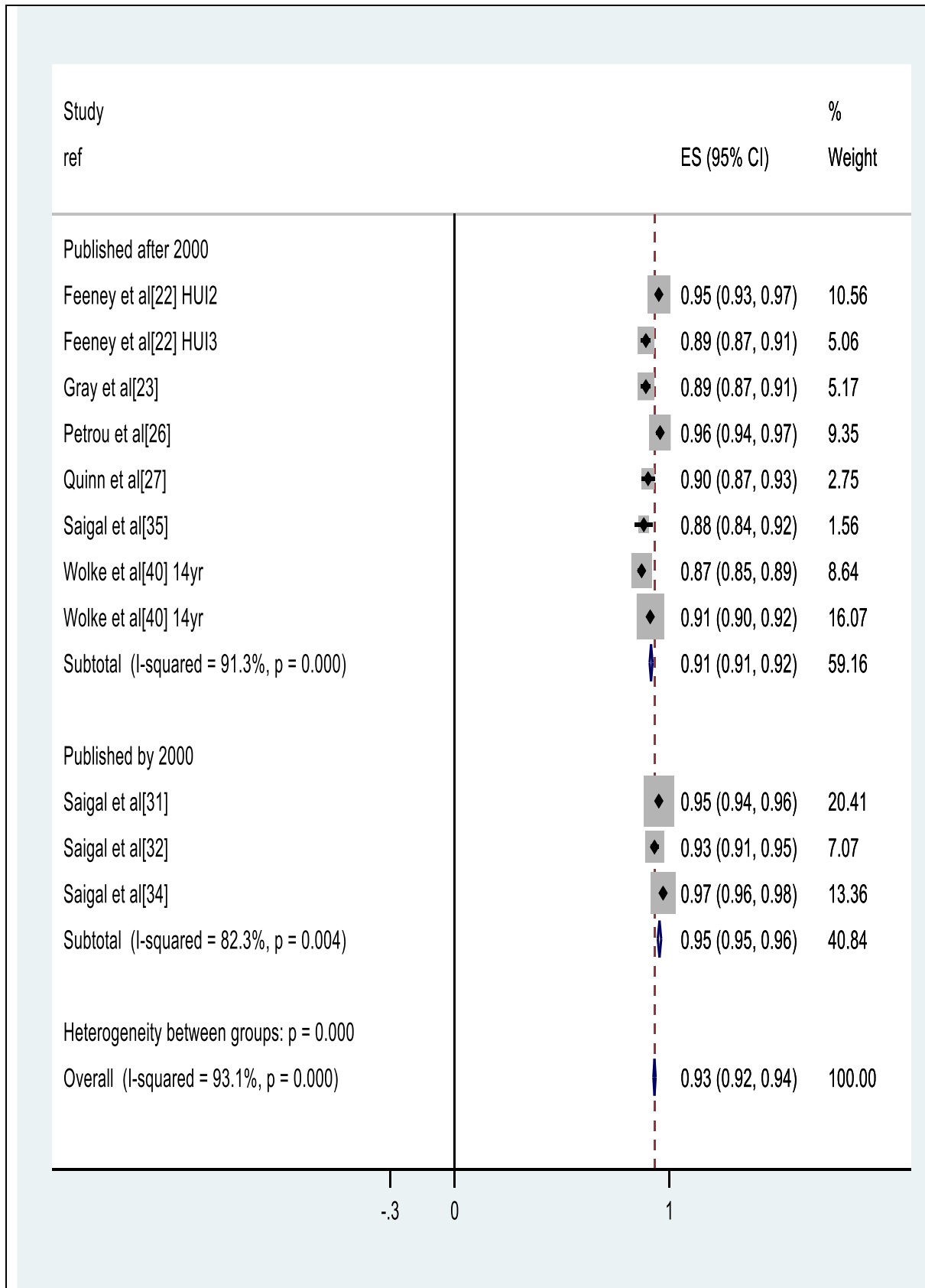


Figure 2(iii). Random-effect meta-analysis for the mean utility values for the study group populations

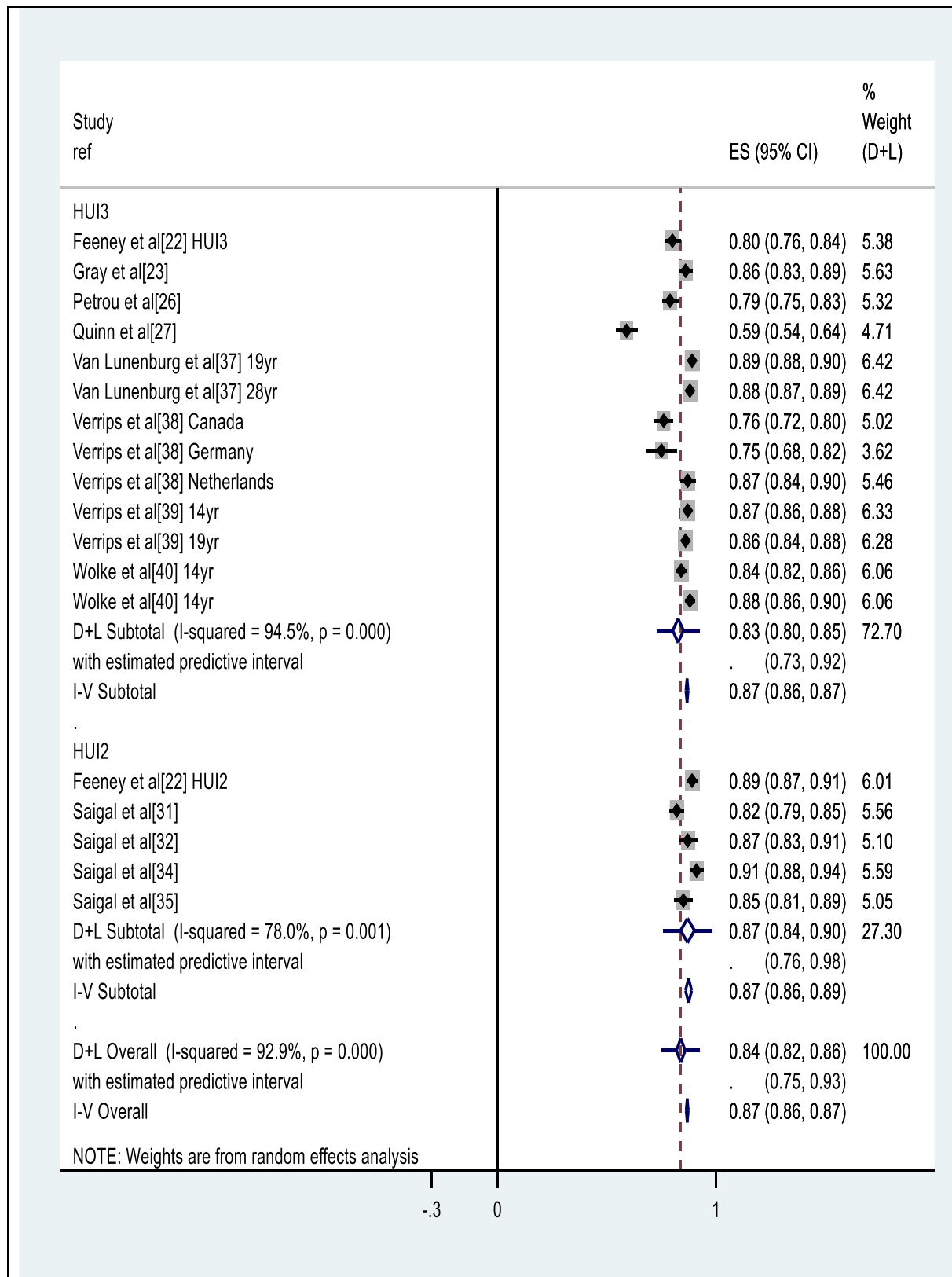
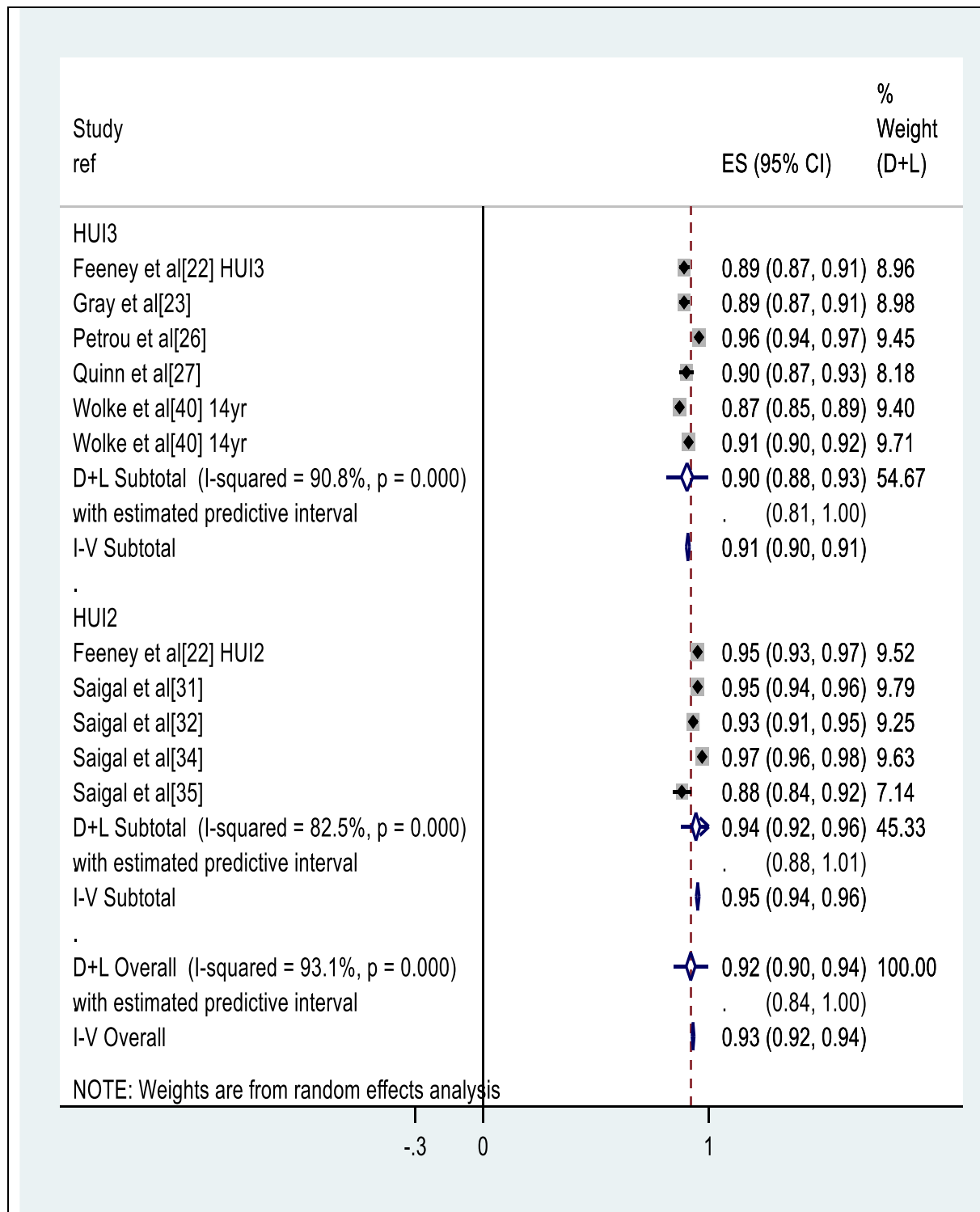


Figure 2(iv). Random-effect meta-analysis for the mean utility values for the control group populations



Appendix 1. Search terms applied in literature search strategy

Search Category	#	Search Terms
Preterm Birth and Low Birthweight	1	“gestational age”
	2	“low birthweight” OR “low birth weight” OR “low-birthweight” OR “low-birth-weight” OR “low-birth weight”
	3	(pre-term OR preterm OR “pre term” OR prematur* OR early) π (birth OR delivery OR parturition OR labour OR labor OR “labour onset” OR “labor onset”)
	4	OR (1 to 3)
Utility Terms	5	Utilit* or disutilit* or HSUV
	6	“quality adjusted life year*” or QALY or “quality-adjusted life year*” or “quality-adjusted life-year*”
	7	OR (5 to 6)
Indirect Valuation Method Terms	8	EQ-5D or “EQ 5D” or EQ5D or Euroqol or “Euro qol” or EQ-5D-Y or “EQ 5D Y”
	9	Short-form survey-6D or short form 6D or SF-6D or “SF 6D” or SF6D
	10	“health utilities index” or HUI
	11	“quality of well being” or “quality of well-being” or QWB
	12	16D Health-Related Quality of Life or 16D HRQoL or 17D Health-Related Quality of Life or 17D HRQoL
	13	AQoL-6D or Assessment of Quality of Life-6D
	14	“Child Health Utility 9 Dimension” or CHU9D or CHU-9D or “CHU 9D”
	15	Adolescent Health Utility Measure or AHUM
	16	15-dimensional instrument or 15 dimensional instrument or 15D
	17	preference-based measure of HRQoL or preference based measure of HRQoL
	18	multi-attribute utility instrument or multiattribute utility instrument
Direct Valuation Method Terms	19	OR (8 to 18)
	20	Standard Gamble or standard-gamble
	21	Time trade off or time trade-off
	22	best worst scaling or best-worst scaling
	23	Discrete choice experiment or discrete-choice experiment
	24	person trade off or person trade-off
	25	scoring algorithm or scoring-algorithm
	26	utility elicitation or direct elicitation
	27	OR (20 to 26)
	28	7 OR 19 OR 27
Main Search	29	4 AND 28
English Filter	30	Remove non-English Title and/or Abstract
Non-Duplicated	31	Remove Duplicates

Appendix 2. Number (%) of children with suboptimal levels of function by study and multi-attribute utility measure attributes

Study	Outcome Measure	Attribute	Source of Measurement	Study Group Total N; n (%)	Control Group Total N; n (%)	p-value [±]
Baumann <i>et al</i> [21] (13 year follow-up)	HUI3	Vision	Self-reported	189, 71 (37.6)	201, 45 (22.4)	Not reported
		Hearing		190, 3 (1.6)	201, 2 (1.0)	Not reported
		Speech		186, 44 (23.7)	196, 38 (19.4)	Not reported
		Ambulation		190, 7 (3.7)	201, 1 (0.5)	Not reported
		Dexterity		190, 10 (5.3)	201, 1 (0.5)	Not reported
		Emotion		189, 62 (32.8)	201, 62 (30.8)	Not reported
		Cognition		190, 66 (34.7)	201, 77 (38.3)	Not reported
		Pain		190, 59 (31.1)	201, 79 (39.3)	Not reported
Baumann <i>et al</i> [21] (13 year follow-up)	HUI3	Vision	Parent-reported	188, 68 (36.0)	201, 46 (22.9)	Not reported
		Hearing		189, 3 (1.6)	201, 1 (0.5)	Not reported
		Speech		188, 14 (7.4)	201, 11 (5.5)	Not reported
		Ambulation		190, 8 (4.2)	201, 0 (0.0)	Not reported
		Dexterity		190, 13 (6.8)	201, 0 (0.0)	Not reported
		Emotion		189, 59 (31.2)	201, 66 (32.8)	Not reported
		Cognition		190, 38 (20.0)	201, 27 (13.4)	Not reported
		Pain		189, 44 (23.2)	201, 63 (31.3)	Not reported
Baumann <i>et al</i> [21] (26 years follow-up)	HUI3	Vision	Self-reported	189, 114 (60.3)	201, 83 (41.3)	Not reported
		Hearing		190, 3 (1.6)	201, 4 (2.0)	Not reported
		Speech		186, 21 (11.3)	196, 16 (8.2)	Not reported
		Ambulation		190, 7 (3.7)	201, 0 (0.0)	Not reported
		Dexterity		190, 9 (4.7)	201, 1 (0.5)	Not reported
		Emotion		189, 68 (36.0)	201, 54 (26.9)	Not reported
		Cognition		190, 60 (31.6)	201, 53 (26.4)	Not reported
		Pain		190, 57 (30.0)	201, 65 (32.3)	Not reported
Baumann <i>et al</i> [21] (26 years follow-up)	HUI3	Vision	Parent-reported	188, 117 (62.2)	201, 90 (44.8)	Not reported
		Hearing		189, 3 (1.6)	201, 2 (1.0)	Not reported
		Speech		188, 12 (6.4)	201, 2 (1.0)	Not reported
		Ambulation		190, 9 (4.7)	201, 0 (0.0)	Not reported
		Dexterity		190, 10 (5.3)	201, 0 (0.0)	Not reported
		Emotion		189, 80 (42.3)	201, 47 (23.4)	Not reported

		Cognition		190, 27 (14.2)	201, 9 (4.5)	Not reported
		Pain		189, 60 (31.7)	201, 45 (22.4)	Not reported
Gray <i>et al</i> [23]	HUI3	Vision	Self-reported	140, 38 (27.1)	108, 25 (23.1)	0.557
(Children with any level of functional impairment)		Hearing		140, 1 (0.7)	108, 0 (0.0)	0.999
		Speech		140, 19 (13.6)	108, 13 (12.0)	0.849
		Ambulation		140, 5 (3.6)	108, 1 (0.9)	0.237
		Dexterity		140, 6 (4.3)	108, 0 (0.0)	0.037
		Emotion		140, 40 (28.6)	108, 37 (34.3)	0.406
		Cognition		140, 57 (40.7)	108, 27 (25.0)	0.010
		Pain		140, 31 (22.1)	108, 39 (36.1)	0.022
Gray <i>et al</i> [23]	HUI3	Vision	Self-reported	140, 3 (2.1)	108, 0 (0.0)	0.260
(Children with severe functional impairment)		Hearing		140, 0 (0.0)	108, 0 (0.0)	N/A
		Speech		140, 1 (0.7)	108, 0 (0.0)	0.999
		Ambulation		140, 3 (2.1)	108, 0 (0.0)	0.260
		Dexterity		140, 3 (2.1)	108, 0 (0.0)	0.260
		Emotion		140, 1 (0.7)	108, 2 (1.9)	0.999
		Cognition		140, 15 (10.7)	108, 9 (8.3)	0.666
		Pain		140, 1 (0.7)	108, 2 (1.9)	0.582
Petrou <i>et al</i> [26]	HUI3	Vision	Parent-reported	190, 61 (32.1)	141, 11 (7.8)	<0.001
		Hearing		190, 11 (5.8)	141, 1 (0.7)	0.015
		Speech		190, 37 (19.5)	141, 5 (3.5)	<0.001
		Ambulation		190, 16 (8.4)	141, 0 (0.0)	<0.001
		Dexterity		190, 19 (10.0)	141, 2 (1.4)	0.001
		Emotion		190, 35 (18.4)	141, 9 (6.4)	0.002
		Cognition		190, 99 (52.1)	141, 24 (17.0)	<0.001
		Pain		190, 40 (21.0)	141, 14 (9.9)	0.007
Wolke <i>et al</i> [40]	HUI3	Vision	Self-reported	260, 90 (34.6)	282, 63 (22.3)	0.002
VLBW/VP I: no or mild disability		Hearing		260, 1 (0.4)	282, 2 (0.7)	N/A
		Speech		260, 72 (27.9)	282, 55 (19.5)	0.025
		Ambulation		260, 11 (4.2)	282, 1 (0.4)	0.002
		Dexterity		260, 14 (5.4)	282, 2 (0.7)	0.001
		Emotion		260, 83 (31.9)	282, 93 (33.0)	0.793
		Cognition		260, 104 (39.9)	282, 108 (38.2)	0.624

		Pain		260, 78 (30.0)	282, 110 (39.1)	0.028
Wolke <i>et al</i> [40]	HUI3	Vision	Parent-reported	260, 94 (36.0)	282, 67 (24.0)	0.002
VLBW/VP I: no or mild		Hearing		260, 1 (0.4)	282, 2 (0.7)	N/A
Disability		Speech		260, 26 (10.0)	282, 22 (7.8)	0.368
		Ambulation		260, 12 (4.6)	282, 0 (0.0)	<0.001
		Dexterity		260, 18 (6.9)	282, 1 (0.4)	<0.001
		Emotion		260, 78 (30.0)	282, 98 (34.8)	0.238
		Cognition		260, 60 (23.0)	282, 40 (14.1)	0.005
		Pain		260, 57 (21.9)	282, 91 (32.3)	0.007
Wolke <i>et al</i> [40]	HUI3	Vision	Parent-reported	12, 8 (61.5)	282, 67 (24.0)	0.005
VLBW/VP II: moderate		Hearing		12, 0 (0.0)	282, 2 (0.7)	N/A
or severe disability		Speech		12, 11 (84.6)	282, 22 (7.8)	<0.001
		Ambulation		12, 6 (46.2)	282, 0 (0.0)	<0.001
		Dexterity		12, 8 (61.5)	282, 1 (0.4)	<0.001
		Emotion		12, 8 (61.5)	282, 98 (34.8)	0.073
		Cognition		12, 12 (92.3)	282, 40 (14.1)	<0.001
		Pain		12, 5 (38.5)	282, 91 (32.3)	0.763

[‡] Calculated using Fisher's exact test where reported. Very low birthweight; VP, very preterm; N/A, not applicable.

Appendix 3. Mixed-effect meta-regression by hierarchical linear model of multi-attribute utility values for low birthweight measured by HUI2 and HUI3

	Model 1 [±]			Model 2 [±]		
	Coefficient (SE)	(95% CI)	<i>p</i> -value	Coefficient (SE)	(95% CI)	<i>p</i> -value
Constant (Model 1) [#]	0.873 (0.039)	(0.796, 0.950)	<0.001			
Constant (Model 2) [#]				0.756 (0.022)	(0.712, 0.799)	<0.001
Birthweight status						
Normal birthweight	Reference			Reference		
Very low birthweight	-0.030 (0.000)	(-0.030, -0.030)	<0.001	-0.030 (0.000)	(-0.030, -0.030)	<0.001
Extremely low birthweight	-0.068 (0.015)	(-0.098, -0.038)	<0.001	-0.068 (0.015)	(-0.098, -0.038)	<0.001
Valuation method						
HUI3	Reference			Reference		
HUI2	0.244 (0.020)	(0.204, 0.284)	<0.001	0.142 (0.029)	(0.085, 0.199)	<0.001
Respondent type						
Self-response	Reference			Reference		
Parent response	0.040 (0.000)	(0.040, 0.040)	<0.001	0.029 (0.008)	(0.012, 0.045)	<0.001
Administration mode						
Self-administered	Reference			Reference		
Interviewer-administered	-0.201 (0.039)	(-0.278, -0.124)	<0.001	-0.072 (0.020)	(-0.112, -0.033)	<0.001
Self and interviewer administered	-0.045 (0.042)	(-0.128, 0.038)	0.285	-0.032 (0.043)	(-0.117, 0.053)	0.463
Year of publication						
After 2000	Reference			Reference		
2000 or earlier	0.016 (0.018)	(-0.019, 0.052)	0.363	0.008 (0.021)	(-0.033, 0.049)	0.706
Region study is based						
North America	Reference			Reference		
Europe	0.030 (0.039)	(-0.047, 0.106)	0.444	0.030 (0.039)	(-0.047, 0.106)	0.444
Setting						

Community	Reference	-	-	-
Educational	-0.240 (0.039)	(-0.317, -0.162)	<0.001	-
Hospital	-0.033 (0.006)	(-0.044, -0.022)	<0.001	-
Age at assessment (years)				
Ages 8-12	-	-	-	Reference
Ages 13-18	-	-	-	0.104 (0.020)
Age 19+	-	-	-	0.085 (0.030)

Model 1 controls for all covariates with the exception of age of child at assessment. Model 2 controls for all covariates with the exception of study setting.

SE denotes standard error; CI denotes confidence interval; HUI denotes Health Utilities Index.

Constant for model 1 represents utility values for individuals born at normal birthweight, measured using the HUI3, using self-responses and self-administered questionnaires, after 2000, in North American community settings. Constant for model 2 represents utility values for individuals born at normal birthweight, measured using the HUI3, using self-responses and self-administered questionnaires, after 2000, in North American settings between ages 8-12 years.