


REVIEW ARTICLE

Quality of life of adults born very preterm or very low birth weight: A systematic review

Sylvia van der Pal¹ | Malte Steinhof¹ | Manon Grevinga¹ | Dieter Wolke² |
Gijsbert (Erik) Verrips¹ 

¹TNO Child Health, Leiden, The Netherlands

²Department of Psychology, University of Warwick, Coventry, UK

Correspondence

Gijsbert (Erik) Verrips, TNO, Child Health,
PO Box 3005, NL 2301 DA Leiden, The
Netherlands.
Email: erik.verrips@tno.nl

Funding information

This systematic review was performed in the context of EU-project RECAP (www.recap-preterm.eu) and thus with funding from the European Commission; Horizon 2020; Grant Number: 733280.

Abstract

Aim: To establish differences in health-related quality of life (HRQoL) in adults born term and those born very preterm (VPT) and/or with a very low birth weight (VLBW).

Methods: Our systematic review is preregistered under PROSPERO-ID CRD42018084005. Studies were eligible for inclusion if their authors had stated the HRQoL of adults (18 years or older) born VPT (<32 weeks of gestation) or VLBW (<1500 g of birth weight) had been measured, if written in English, and if they reported a comparison with a control group or valid norms. We searched Pubmed, Scopus, Psycinfo, Web of Science, Embase and contacted experts in this field. Non-response and other bias-related problems were evaluated.

Results: We included 18 studies of 15 unique cohorts from 11 countries. In 11 studies, no differences in HRQoL between VPT or VLBW and term-born adults were found; four studies found lower HRQoL in VPT/VLB adults; and evidence from three studies was inconclusive. Disability, sex and age were associated with HRQoL.

Conclusion: There is no conclusive evidence that HRQoL differs between term-born adults and those born VPT or with a VLBW. The comparability of studies was restricted by differences between HRQoL measurements, age ranges at assessment and definition of disability.

KEYWORDS

health-related quality of life, very low birth weight adults, very preterm adults

Abbreviations: AGA, appropriate for gestational age; APIC, Adults born Preterm International Collaboration; BPD, bronchopulmonary dysplasia; BW, birth weight; ELBW, extremely low birth weight (birth weight < 1000 g); EPT, extremely preterm (<28 weeks GA); GA, gestational age; HRQoL, health-related quality of life; HUI, Health Utilities Index; LBW, low birth weight; NBW, normal birth weight; NSI, neuro-sensory impairments; PT, preterm (<37 weeks GA); RECAP, Research on European Children and Adults born Preterm; SF36/12, 36/12-item Short Form Health Survey; SGA, small for gestational age; VLBW, very low birth weight (birth weight < 1500 g); VPT, very preterm (<32 weeks GA).

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2020 The Authors. *Acta Paediatrica* published by John Wiley & Sons Ltd on behalf of Foundation Acta Paediatrica

1 | BACKGROUND

1.1 | Rationale

Clinical measures such as mortality and morbidity rates are important when evaluating the impact of preterm birth on well-being later in life. In addition, to complement our understanding of the significance of putative impairments and disabilities after preterm birth, the broader comprehensive concept of health-related quality of life (HRQoL) is needed, especially as it incorporates patients' perspectives on their well-being. In 2008, a systematic review by Zwicker and Harris¹ showed that HRQoL differed between individuals born preterm and individuals born full term, with those born preterm reporting a lower average HRQoL. While the same review mainly included assessments of HRQoL in children or adolescents born preterm, it included only four studies in young adults aged 18–23. As it is unknown whether the lower HRQoL reported in individuals born preterm persists into adulthood, our systematic review was intended to close this knowledge gap.

1.2 | Objectives

We reviewed studies that compared the HRQoL of adult individuals (age \geq 18 years) who had been born very preterm (<32 weeks of gestational age) and/or with very low birth weight (<1500 g) with the HRQoL of individuals in the same age range born term.

2 | METHODS

2.1 | Protocol and registration

This review is preregistered using PROSPERO (<https://www.crd.york.ac.uk/prospero/>) under the ID CRD42018084005. It was conducted in accordance with the PRISMA guidelines.²

2.2 | Eligibility criteria

Studies were included if: (a) the study had been published in English; (b) the authors had stated they had measured HRQoL; (c) participants were aged 18 years or older; (d) preterm participants had been compared to a full-term control group or to population norms; and (e) participants had been born very preterm (<32 weeks of gestational age) and/or with very low birth weight (<1500 g). There were no limits on year of publication.

2.3 | Information sources

Publications were retrieved between 08/02/2018 and 15/07/2019 from Pubmed, Scopus, PsycInfo, Web of Science and Embase. In addition to those databases, we also consulted, on the basis of the

Key notes

- There is no conclusive evidence that health-related quality of life (HRQoL) differs between term-born controls and adults born very preterm or with a very low birth weight.
- Standardisation of HRQoL and disability measurement is needed.
- Future research on HRQoL in adults born very preterm should focus on the transition from adolescence into adulthood, on age ranges in adulthood and on disability as a factor that may influence HRQoL.

abstracts, the publication lists of researchers who had been presented at the APIC (Adults Born Preterm International Collaboration, www.apic-preterm.org) Conference in Venice, Italy in November 2017. The publication lists of the experts who had attended this conference produced no additional eligible publications.

2.4 | Search

The search terms used in Scopus were the following: "quality of life" AND adult* AND (preterm OR premature OR "low birthweight" OR "low birth weight").

The search terms used in PsycInfo were the following: exp "Quality of Life"/ AND (exp Premature Birth/ OR exp Birth Weight/ OR preterm.mp.) AND adult*.mp.

The search terms used in Web of Science were the following: (TS = ("quality of life" AND adult* AND (preterm OR premature OR "low birthweight" OR "low birth weight"))) AND LANGUAGE: (English).

The search terms used in Embase were the following: ("quality of life"/exp OR "quality of life") AND ("adult"/exp OR adult) AND (preterm OR "low birthweight"/exp OR "low birthweight" OR "low birth weight"/exp OR "low birth weight").

Although the results from Pubmed are also included in Scopus, it is possible that Scopus does not list publications that have been published less than three months before a search is carried out. For this reason, we included Pubmed, but, to be sure we had included all relevant recent publications, limited our results to studies that had been published after April 2019.

The search terms used in Pubmed were the following: "quality of life" AND adult* AND (preterm OR premature OR "low birthweight" OR "low birth weight"); Limit: From 2019/07/15.

2.5 | Study selection

The abstracts of the publications produced by the database searches were scanned for eligibility by one reviewer (MS) according to our

criteria. Subsequently, articles selected on the basis of the abstracts were read independently by two reviewers (MS, MG) and again evaluated for eligibility. Complete consensus had to be reached on the inclusion of studies; if no agreement could be reached, a third reviewer (SP) would decide.

2.6 | Data collection process

Using a piloted publication evaluation form agreed by the reviewers, one reviewer extracted the relevant data from the selected studies. In one case,³ authors were contacted with the request to provide means and standard deviations of SF-36 scores that they had not reported in their publication; these were duly provided. With this exception, only data available in the publications were used in this review. From one paper,⁴ we calculated combined means and standard deviations, as the results it presented were broken down by subgroups. No decisions were made a priori on how to deal with obtaining additional information from authors.

2.7 | Data items

Data items were derived from a piloted-article evaluation form and can be found in Tables 1 and 2.

2.8 | Risk of bias in individual studies

Within individual studies, it was checked whether attrition had occurred at random or, if it had not occurred at random, whether attrition could have biased the results. We also established whether there were indications of selective outcome reporting of chance findings. Only after study inclusion had finished did we check and document what the authors had hypothesised for each particular study. This decision was made post hoc, in order to have an additional means of assessing bias in individual studies.

2.9 | Synthesis of results

This systematic review is qualitative in nature. We did not plan to integrate the results quantitatively, as we expected various measurement instruments to have been used that would not allow quantitative comparison without harmonisation protocols.

2.10 | Additional analyses

It was decided a priori to base our choices of subgroup analyses on the features of the publications we included. Such subgroup analyses might have been based on characteristics of groups (eg disability level, sex, age or SES) or on measurement instruments (eg SF-36, HUI-3).

3 | RESULTS

3.1 | Study selection

The search of the databases yielded 2147 articles, all of whose abstracts were screened. After screening, 201 articles remained that seemed to meet the inclusion criteria, 79 of which were non-duplicates. Subsequently, two reviewers read these 79 articles independently and decided independently whether an article should be included. Sixty-one articles were excluded consensually, as they did not meet the inclusion criteria after all. Twenty-three of the 61 were reviews of other articles that did not provide unique data, and nine were not about HRQoL. In 17 articles, the study population did not fit our inclusion criteria; instead, inclusion had been based on BPD, neurological risk or treatment, or participants had not been adults. In six publications, the studies had no control group, three articles were not available in English; one publication was a book without unique data; one article was excluded because, after contact with the authors, we concluded that the results had already been published in another article included in this review (duplicate); and one article was about the HRQoL of parents of children born preterm. Upon request, a listing of references of papers excluded is provided by the authors. Figure 1 shows a flow diagram of the inclusion process.

In this way, we included a total of 18 studies involving 15 unique cohorts in our review. Two cohorts had been included in more than one article: the Trondheim cohort, which was assessed at ages 20⁵ and 23⁶; and the Canadian McMaster Ontario cohort, which was assessed at ages 22-26^{7,8} and 29-36.⁹ For the 22-26 age range, the Ontario cohort had used two measurements for HRQoL: (HUI⁷ and SF-36⁸). These had been published in two separate articles.

3.2 | Study characteristics

3.2.1 | Methods and participants

The 18 articles that were finally selected for our review reported on longitudinal prospective studies and included a comparison of adults born preterm with a control group or, in one case, 'community norms' for the SF-36.¹⁰ Participants in the studies were assessed between ages 18 and 36. The cohorts included were from 11 countries: Australia, Canada, Denmark, Finland, Germany, New Zealand, Norway, Sweden, Switzerland, the United Kingdom and the United States.

3.2.2 | Features of study group and subgroups

In some articles, individuals were classified into groups on the basis of disability and/or sex. Some studies excluded individuals with a disability, others compared VPT/VLBW groups with and without a

TABLE 1 Study characteristics of 18 selected publications

Publication; First author (ref)	Year pub.	Name/origin of cohort	Age at assessment (mean/range)	Study population inclusion criteria	Number of participants in eligible study and control groups	Measure used	Mode of administration of measurement
Batsvik, et al ⁴	2015	Hauckland University Hospital, Bergen, Norway (1982-1985)	24 y	<1000 g or < 28 wk	43 in 51 eligible study group; 43 in 46 eligible controls	SF-36	Self-report
Baumann, et al ¹⁴	2016	Bavarian Longitudinal Study (BLS), Germany (1985-1986)	26 y	<1500 g or < 32 wk	201 in 411 eligible study group; 190 in 308 eligible controls	HUI-3	Self-report and parent report
Bjerager, et al ¹⁵	1995	State University Hospital Copenhagen, Denmark (1971-1974)	18-20 y	<1500 g	85 in 107 eligible study group; 85 in 110 eligible controls	Customised questionnaire ^a	Telephone interviews
Björqvist, et al ¹⁷	2017	Heisinki Study of VLBW adults, Finland (1971-1974)	18-27 y	<1500 g	164 in 255 eligible study group; 172 in 314 eligible controls	15-D	Self-report
Cooke ¹¹	2004	Liverpool Maternity hospital, in UK (1980-1983)	19-22 y	<1500 g or < 31 wk	79 in 138 eligible study group; 71 in 163 eligible controls	SF-36	Self-report
Darlow, et al ³	2013	1986 cohort, (all VLBW) in New Zealand	22-23 y	<1500 g	230 in 251 eligible study group; 69 controls	SF-36	Face-to-face interview
Dinesen & Greisen ¹⁶	2001	Rigshospitalet Copenhagen, in Denmark (1980-1982)	18 y	<1500 g	79 in 114 VLBW study group seen at 2yrs; 69 in 92 controls seen at 4 yrs	Customised questionnaire ^a	Telephone interview
Gaddlin, et al ¹²	2009	South-eastern Sweden (1987-1988)	20 y	<1500 g	77 in 86 surviving study group; 69 in 86 surviving controls	SF-36	Self-report (2 telephone interviews; 2 other participants had help due to disability)
Hack, et al ¹⁸	2007	Rainbow Babies and Children's hospital, Cleveland, Ohio, USA	20 y	<1500 g	241 in 312 surviving study group; 232 in 363 surviving controls	Chip-AE	Self-report (5% read aloud due to reading difficulties)
Hallin & Stjernqvist ¹⁹	2011	Southern Sweden (1985-1986)	18 y	<29 wk	52 in 59 eligible study group and 54 in 61 eligible controls	Visual analogue scale ranging 0-100 ^b	Self-report
Husby, et al ⁶	2016	St. Olavs Hospital Trondheim, in Norway (1986-1988)	23 y	<1500 g	35 in 54 invited study group; 37 in 48 invited controls	SF-36	Self-report
Lund, et al ⁵	2012	St. Olavs Hospital Trondheim, Norway (1986-1988)	20 y	<1500 g	43 in 59 invited study group; 74 in 102 invited controls	SF-36	Self-report

(Continues)

TABLE 1 (Continued)

Publication; First author (ref)	Year pub.	Name/origin of cohort	Age at assessment (mean/range)	Study population inclusion criteria	Number of participants in eligible study and control groups	Measure used	Mode of administration of measurement
Natalucci, et al ¹⁰	2013	(all ELBW) in Switzerland (1983-1985)	22-26 years	<1000g	55 in 65 traced study group, no controls; community norms	SF-36	self-report
Poole, et al ⁸	2017	McMaster cohort, Central-west Ontario, Canada (1977-1982)	22-26 y	<1000 g	70 in 179 surviving study group; 83 in 145 recruited at 8yrs controls	SF-36	Self-report
Roberts, et al ¹³	2013	Victoria, Australia (1991-1992)	18 y	<1000 g or < 28 wk	186 (HUI3)/183 (SF-36) in 298 surviving study group; 138 (HUI3)/144 (SF-36) in 262 matched controls	SF-36 & HUI-3	Self-report
Saigal, et al ⁷	2006	McMaster cohort, Central-west Ontario, Canada (1977-1982)	23 y	<1000 g	143 in 179 surviving study group; 130 in 145 recruited at 8yrs controls	HUI-2; plus hypothetical states	Interview self-report (parents proxy for impaired participants)
Saigal, et al ⁹	2016	McMaster cohort, Central-west Ontario, Canada (1977-1982)	22-26 y AND 29-36 y	<1000 g	141 (22-26 y) and 94 (29-36 y) in 179 surviving study group; 133 & 83 in 145 recruited at 8yrs controls	HUI-3	interview
Vederhus, et al ²⁰	2015	Haukeland University Hospital, Bergen, Norway (1991-1992)	18 y	<1000 g or < 28 wk	31 in 35 eligible study group; 29 in 35 eligible controls	Child Health Questionnaire (CHQ)	Self-report and parent report

^aCustomised questionnaire based on Aggenmaes HRQoL theory.^{10,11}

^bVisual analogue scale ranging from 0 to 100 (0 = least desirable HRQoL, and 100 = most desirable)

TABLE 2 Results and bias of 18 selected publications

Publication: First author (ref)	Main conclusion (based on tables or according to authors in abstract/discussion)	Outcomes by subgroups (sex/disability) and/or subscales	Selective non-response bias or other biases/remarks
Batsvik, et al ⁴	<i>Inconclusive:</i> Overall, scores for EPT adults were similar to those of controls on the SF36 at age 24, but the authors only show comparisons for EPT with and without disabilities. EPT without disabilities scored significantly lower on 3 of the 8 SF36 scales	<i>Disability:</i> EPT without disabilities (n = 35) scored lower on social functioning, emotional role, mental health (3 of 8 scales SF36) and psychological complaints. SF36-scales of EPT with disabilities (n = only 8) were comparable to the scale scores of controls	Nothing was reported on non-response bias; response rate was high. However, low sample size, especially in EPT born adults with disabilities (n = 8)
Baumann, et al ¹⁴	<i>Lower HRQoL than controls:</i> At age 26, a higher percentage of VPT/VLBW had more severe levels of disability on the HUI3 (self-report and parent report)	No subgroups reported. <i>Self-reported</i> HRQoL was higher than parent-reported HRQoL. <i>Subscales</i> were reported only with regard to change from 13 to 26 y of age and; parent-reported HRQoL worsened into adulthood	Dropout was not random; socially disadvantaged participants dropped out more. SES and sex were included in the multivariate analyses as covariates. Lower HRQoL was related to economic and social functioning problems in adulthood
Bjerager, et al ¹⁵	<i>No difference in HRQoL:</i> Total objective and subjective HRQoL did not differ between the non-handicapped VLBW adults aged 18-20 and NBW controls.	<i>Disability:</i> Adults born with VLBW reported lower objective and subjective HRQoL than controls on the Elementary Biological Needs subscale	Nothing was reported on non-response bias; authors stated that both groups had comparable dropout rates
Björqvist, et al ¹⁷	<i>No difference in HRQoL:</i> On total 15D score or any of the profile dimensions, HRQoL did not differ between the whole VLBW group and the control group at age 18-27	<i>Sex:</i> Women reported lower a 15D score than men (P 0.001) in all areas except mobility, hearing, eating, speech, usual activities and sexual activities. <i>NSI^a:</i> VLBW participants with NSI did not differ on the total 15D score, but had lower scale scores on mobility, vision, sexual activities and eating. <i>SGA vs AGA:</i> the SGA VLBW group reported a significantly lower total 15D score. Compared to controls, AGA VLBWs scored lower on mobility and higher on depression. Compared to controls, SGA VLBWs scored lower on eating; compared to AGA VLBWs, they scored lower on mental functions, depression and vitality	The authors hypothesised that dropout in participants with NSI might have been high, but emphasised that including the NSI participants in the analyses and reporting separately on this group were strengths of this study>?
Cooke ¹¹	<i>No difference in HRQoL between VPT/VLBW and controls:</i> HRQoL did not differ between groups except for the Physical Functioning scale of the SF36 at ages 19-22	<i>Sex:</i> Male VPT's had lower physical functioning and general health perception than male controls. In females, no differences were found on subscales	50% of the cohort could not be traced, and more females and preterms returned the questionnaires
Darlow, et al ³	<i>No difference in HRQoL:</i> At age 22-23, HRQoL did not differ between the VLBW and control groups on all 8 SF-35 subscales and on the physical and mental component score	No subgroups were reported. No differences were found on subscales (not reported in article; sent after contact with the authors)	The authors reported that although 71% of the VPT/VLBW cohort had participated, there were no differences in the basic demographic characteristics of those who did and did not participate
Dinesen & Greisen ¹⁶	<i>Lower objective HRQoL:</i> Non-handicapped VLBWs scored lower on objective HRQoL, but at age 18 age their subjective HRQoL did not differ from that of controls	<i>Disability:</i> Handicapped VLBW and LBW had lower subjective and objective HRQoL than NBW controls	Five people could not be assessed due to a handicap. There was also a high participation rate in the VLBW group. There was no further report of selective non-response or other bias

(Continues)

TABLE 2 (Continued)

Publication: First author (ref)	Main conclusion (based on tables or according to authors in abstract/discussion)	Outcomes by subgroups (sex/disability) and/or subscales	Selective non-response bias or other biases/remarks
Gaddlin, et al ¹²	<i>No difference in HRQoL:</i> At age 20, HRQoL did not differ between VLBW and controls on all 8 SF-35 subscales and on the physical and mental health scores.	<i>Disability^b:</i> The n = 15 handicapped VLBWs had significantly lower scores on the physical functioning subscale only ($P < .001$)	The authors reported that non-responders did not differ from responders in BW and GA. Univariate and multivariate regression analyses showed an association between physical functioning and some perinatal and neonatal factors
Hack, et al ¹⁸	<i>No difference in poor health profiles:</i> At age 20, similar proportions of VLBW and NBW participants reported excellent, average or poor health profiles on the Chip-AE	<i>20 subdomains of Chip-AE:</i> Although the VLBW group had lower scores on resilience (physical activity and family involvement) and more disorders (acute minor, long-term medical, long-term surgical and psychosocial) than controls, they also had better work performance and less risk behaviour.	A higher IQ in the control group might have led to a higher probability of participation
Hallin & Stjernqvist ¹⁹	<i>No difference in HRQoL:</i> At age 18, HRQoL did not differ between groups: 71.7 in EPTs compared to 74.8 in FT controls	No subgroups reported	High response rate. It was not evident which specific item or items were used to measure HRQoL on a visual scale from 0 to 100
Husby, et al ⁶	<i>Lower HRQoL than controls:</i> VPT/VLBW had lower HRQoL at age 23 (on 6 of 8 SF36 scales and also the physical and mental component scales; no difference after correction for CP and low IQ)	<i>Subscales:</i> VLBW participants (n = 35) reported significantly lower values on the physical component (physical functioning, role-physical and bodily pain) and on the mental component scales (social functioning and role-emotional). <i>Disability:</i> On SF-36 subscales, VLBWs without cerebral palsy and/or low IQ (n = 25) did not show any significant differences from controls	Noting the limited sample size, the authors state that because they did not know the reasons for non-participation, selection bias may have resulted. On perinatal data, however, participants did not differ from non-participants
Lund, et al ⁵	<i>No difference in HRQoL:</i> At age 20, HRQoL did not differ between VLBWs and controls, except on Mental Health on 8 SF36 scales)	<i>Subscales:</i> VLBWs scored lower than controls on mental health	In the VLBW group, non-participants were more often male, no difference on GA, BW and HC. A term SGA group (n = 55) had lower scores than controls on SF36 scales for mental health, social functioning and emotional role
Natalucci, et al ¹⁰	<i>Inconclusive: Lower HRQoL on mental HRQoL but higher HRQoL on physical HRQoL:</i> ELBW's had lower HRQoL total scores on Mental Component Summary, but higher HRQoL on Physical Component Summary on the SF36 at 22-26 y, compared to community norms	<i>Subscales:</i> Compared to community norms, ELBW's had lower HRQoL scores on social functioning, but higher HRQoL scores on bodily pain and general health. <i>Groups at risk:</i> multiple regression analysis showed that females were more at risk for lower mental HRQoL, but ELBW with BPD were less at risk	The authors reported that the high dropout rate and consequent small size of the group studied may have led to non-response bias, even if there was no difference between participants and dropouts with regard to perinatal and socio-demographic variables. It is also unclear where the community norms originated. After contact with the authors, we concluded that there was an overlap with participants in Baumgard et al, where, on all SF36 scales, there were no differences between adults at age 23 with a BW < 1250 g and term controls at the same age. But males did score lower than controls on physical functioning

(Continues)

TABLE 2 (Continued)

Publication: First author (ref)	Main conclusion (based on tables or according to authors in abstract/discussion)	Outcomes by subgroups (sex/disability) and/or subscales	Selective non-response bias or other biases/remarks
Poole, et al ⁸	No difference in HRQoL: HRQoL did not differ between ELBW and NBW adults at 22-26 years on all 8 SF36 scales	Subscales: there were no significant differences on SF35 subscales. Subgroups: no subgroups were reported. In ELBWs, motor coordination was significantly associated with Physical and Social Functioning. This was also the case in the NBW group	The authors did not report non-response bias, but there was a low response rate and a higher percentage of females. The authors warn about the broad age range, which made it difficult to base conclusions on a specific ages.
Roberts, et al ¹³	No difference in HRQoL: HRQoL did not differ between EPT/ELBW and controls at 18 years (not on overall HUI3 score and only on Physical Functioning of 8 SF36 scales).	Subscales: there was only one significant difference on physical functioning. Disability: Additional analyses without individuals with a major disability did not alter the results (data not shown). Low GA/BW: Regression analyses revealed no evidence of an association between GA and BW on HUI3 and SF36 physical functioning scale scores.	Non-responders had had more problems than responders, such as PVL after birth, and disabilities and lower IQ at age 8
Saigal, et al ⁷	Inconclusive: At age 23, HUI2 mean utility scores did not differ between the ELBW and control groups, but did differ after substitution of missing values. No differences were found on the preferences for hypothetical health states based on the HUI2.	NSI ^a : Within ELBW, there were no differences between those with NSI and those without NSI (n = 38). Sex: No effect of sex. Subscales: HUI2 attributes Sensory and Cognition (out of 7 HUI2 attributes) were significantly more affected in the ELBW group than in controls.	Except for lower maternal education, ELBW non-responders and ELBW responders had a similar prevalence of NSIs and parental socio-demographic features. HRQoL decreased over time (from age 12-16 to age 23), but this happened in both the ELBW group and the control group, and there was no group effect
Saigal, et al ⁹	Lower HRQoL than controls: ELBWs had lower HRQoL on the HUI3 both in their 20s and 30s than controls and the HUI3 score decreased more in the ELBW groups (with and without NSI) from their 20s to 30s	NSI ^a : The HUI3 score was even worse for ELBW with NSI (n = 37) than for ELBW without NSI and NBW controls, and more HUI3 attributes were affected. A participant's sex or socio-economic status had no significant impact on HUI3 trajectories. HUI attributes: at age 22-26, there were differences between VPT/LBW adults and controls on 5 subscales of HUI3 and at 29-36 y on 3 attributes	More males and more people whose mother had lower SES or lower maternal education had missing data on one of the measurements
Vederhus, et al ²⁰	No difference in HRQoL: At age 18, HRQoL did not differ on the Child Health Questionnaire between the EPT/ELBW group and term controls.	Sex and subscales: EPT/ELBW boys significantly improved on 3 subscales from age 10-18 y and at 18 y HRQoL on subscales did not differ from controls anymore. While differences between the EPT/ELBW and control girls remained significantly lower in the EPT/ELBW study group on 3 subscales (mental health, physical functioning and general health), compared to controls	The authors reported nothing on non-response bias, but, due to the small sample sizes, warned about any interpretation of the results

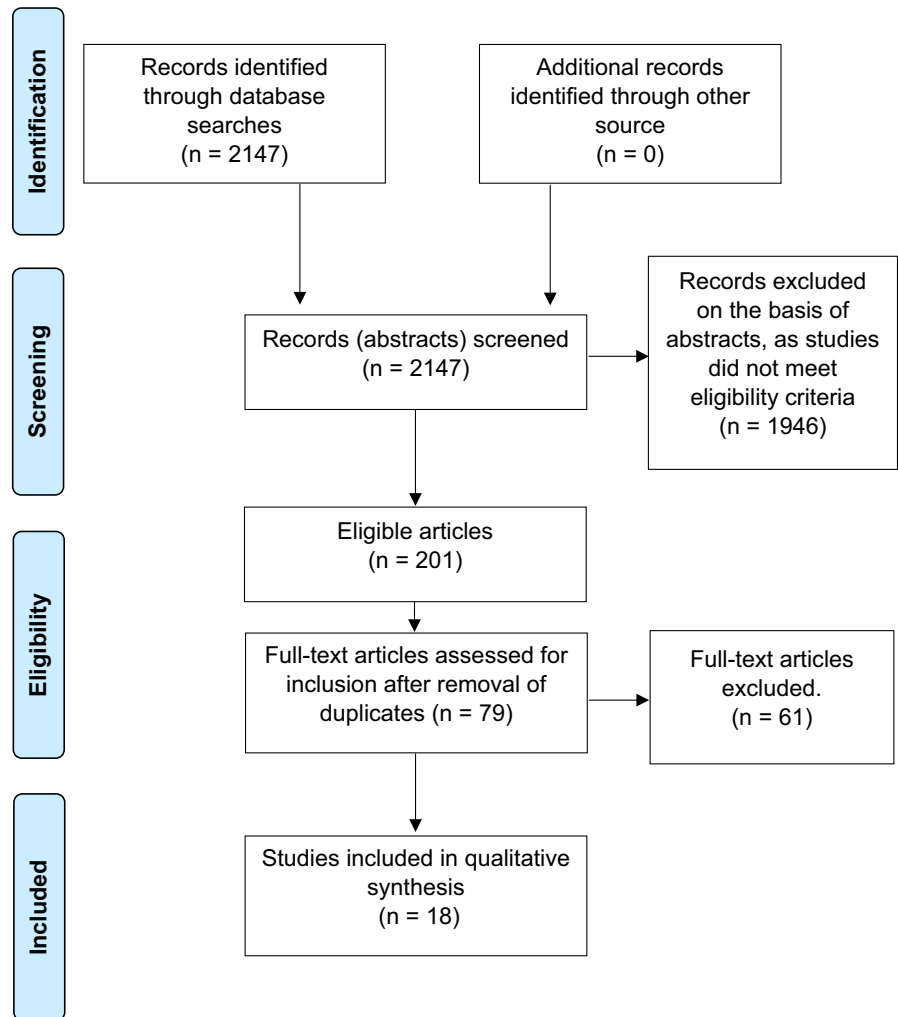
^aNSI (neuro-sensory impairments) definitions. Björqvist, et al¹⁷: cerebral palsy, developmental disorders and severe visual impairment. Saigal, et al⁹: cerebral palsy, blindness, deafness and microcephaly. (Saigal, et al⁷ report no definition of NSI, but it is assumed to be the same as Saigal, et al⁹).

^bDefinition of handicap by Gaddlin, et al¹²: moderate or severe cerebral palsy, moderate or severe attention deficit hyperactivity disorder or mental retardation (IQ < 70).

disability, and still, other did not make this distinction (see Table 1 for details per study and Table 2 for results per subgroup). The definition of disability varied considerably, thereby hampering comparability between the studies.

3.2.3 | Outcome measurements

Nine studies measured HRQoL using the SF-36 questionnaire.^{3-6,8,10-13} The data of these studies are presented in Table 3.

FIGURE 1 Flow-diagram outlining final inclusion of studies

Three studies used the Health Utilities Index mark 3 (HUI3),^{9,13,14} and one study used the HUI2 instrument.⁷ One study used both the SF36 and the HUI3.¹³ Two studies on two different cohorts from Copenhagen used the same customised questionnaire.^{15,16} Four other instruments were used only once: the 15D questionnaire,¹⁷ the Chip-AE,¹⁸ a visual analogue scale ranging from 0 to 100¹⁹ and the Child Health Questionnaire (CHQ).²⁰

3.3 | Results of the studies included

Table 2 presents the results of the studies we included. The SF36 was the most used HRQoL measure. Authors who used the SF36 had concluded that their results showed lower overall HRQoL in one of the two following cases: (a) if one of the two summary component scales (mental or physical) showed a statistically significant lower score for the preterm group than the controls or community norms; or (b) if three or more of the eight scales showed this. Otherwise, they concluded that there was no difference.

Table 3 presents detailed results of the studies that used the SF-36. We decided not to analyse the SF36 results in a quantitative meta-analysis: the study populations were too dissimilar, and there

was no consistency in statistically significant differences across domains. In addition, even cursory inspection of Table 3 shows that any differences were small or non-existent. Authors who had used the HUI or other instruments concluded that their results indicated lower HRQoL in the preterm group than in the control group if the statistical significance of the overall score was lower. Otherwise, it was concluded as above that there was no difference. Results were qualified as inconclusive if lower HRQoL had been found only in subgroups⁴; if a lower mental HRQoL had been found but a higher physical HRQoL¹⁰; or if a lower HRQoL had been found only after substitution of missing values.⁷

3.3.1 | Main conclusions regarding the HRQoL of adults born VPT or with a VLBW

Table 2 shows that 11 of the 18 articles concluded that there was no difference between the HRQoL of adults born very preterm or with a very low birth weight, and those born at term.^{3,5,8,11-13,15,17-20} While the results of three articles were inconclusive,^{4,7,10} four articles found that HRQoL was lower in adults born very preterm or with a very low birth weight.^{6-7,14,16} Four of the articles that concluded

TABLE 3 SF-36 scale scores per study

Publication ^a Group (n)	Physical functioning		Role-physical		Bodily pain		General health		Vitality		Social functioning		Role-emotional		Mental health	
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Batsvik et al ⁴																
EPT disab (n = 8)	56.9 (43.0)	68.8 (37.2)	73.9 (28.1)	65.9 (22.2)	53.1 (14.4)	87.5 (17.7)	83.3 (35.6)	75.5 (12.7)								
EPT healthy (n = 35)	92.1 (15.0)	80.5 (30.4)	68.4 (25.0)	74.1 (25.0)	50.7 (20.5)	74.3 (30.2) ^b	61.4 (42.2) ^c	69.5 (17.8) ^b								
FT controls (n = 43)	94.2 (11.3)	86.9 (25.9)	80.2 (23.0)	77.1 (20.0)	60.0 (13.8)	90.0 (17.3)	88.3 (31.6)	80.2 (12.2)								
Cooke ¹¹																
VPT (n = 79)	91.4 (14.3) ^b	87.0 (30.2)	82 (23.2)	70.3 (23.3)	60.6 (22.2)	86.8 (23.7)	80.6 (32.3)	70.3 (19.9)								
Controls (n = 71)	95.4 (10.0)	89.8 (26.2)	86.1 (17.6)	73.0 (20.5)	59.2 (22.5)	86.2 (21.8)	75.6 (39.0)	70.6 (21.1)								
Darlow, et al ³																
VLBW (n = 230)	91.8 (18.9)	89.8 (19.7) ^b	82.9 (24.0)	71.2 (22.7)	60.8 (18.8)	87.1 (22.3)	90.8 (19.6)	78.6 (17.5)								
Controls (n = 69)	93.6 (15.2)	83.0 (27.3)	79.3 (26.3)	73.6 (20.7)	62.7 (15.6)	86.6 (19.8)	92.1 (17.0)	77.4 (13.1)								
Gaddlin et al ¹²																
VLBW (n = 76)	92.2 (19.6)	87.5 (26.3)	83.6 (21.4)	81.4 (20.3)	64.7 (21.3)	89.5 (17.2)	82.0 (32.9)	57.8 (20.2)								
Controls (n = 68)	96.0 (9.1)	86.6 (26.9)	78.4 (22.7)	77.4 (21.7)	58.4 (24.3)	87.7 (16.8)	82.1 (31.4)	75.3 (17.4)								
Husby, et al ⁶																
VLBW (n = 35)	90.4 (13.6) ^b	80 (30.8) ^c	68.7 (28.3) ^b	72.1 (18.9)	49.2 (14.2)	86.1 (16.5) ^b	78.1 (33.3) ^b	70.6 (16.8)								
VLBW disability (n = 25)	94.6 (8.3)	86.0 (24.0)	74.5 (23.8)	70.5 (18.5)	50.5 (13.7)	88.5 (13.0)	82.7 (29.1)	73.4 (14.1)								
NBW controls (n = 35)	96.6 (5.9)	96.4 (10.7)	82 (18.3)	66.8 (20)	54.9 (13.2)	94.3 (13.3)	95.2 (20)	77.4 (13.2)								
Lund, et al ⁵																
VLBW (n = 43)	90.2 (20.4)	89.0 (19.1)	80.2 (22.6)	79.3 (17.8)	50.1 (19.1)	91.0 (12.6)	87.6 (24.2)	73.6 (15.0) ^b								
Controls (n = 73)	95.5 (10.1)	91.1 (22.2)	80.2 (22.5)	78.8 (19.8)	56.2 (14.2)	92.6 (13.1)	90.9 (23.7)	79.2 (11.9)								
Natalucci, et al ¹⁰																
ELBW (n = 55)	94.3 (12.8)	89.1 (23.0)	83.4 (23.3) ^{dc}	78.4 (19.5) ^{db}	57.4 (19.1)	80.2 (24.0) ^b	84.6 (30.2)	68.7 (18.7)								
Community norms	94.5 (11.4)	90.4 (23.6)	73.6 (23.7)	73.1 (15.8)	60.5 (16.4)	88.7 (17.5)	92.3 (21.1)	73.3 (15.2)								
Poole, et al ⁸																
ELBW (n = 70)	54.7 (4.1)	53.8 (7.1)	54.6 (9.1)	54.0 (9.3)	52.6 (11.0)	50.6 (10.5)	50.0 (9.9)	48.6 (13.0)								
NBW (n = 83)	52.4 (8.3)	52.4 (8.3)	55.2 (6.8)	53.8 (10.4)	53.5 (9.0)	50.9 (7.9)	50.8 (8.3)	51.2 (9.2)								

(Continues)

TABLE 3 (Continued)

	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)
Roberts et al ¹³									
EPT/ELBW (n = 183)	95 (85-100) ^c	100 (75-100)	84 (74-100)	77 (62-90)	70 (50-80)	100 (75-100)	100 (67-100)	80 (64-88)	
Controls (n = 144)	100 (95-100)	100 (100,100)	84 (72-100)	77 (62-89)	65 (55-75)	100 (100,100)	100 (100-100)	80 (68-88)	

^aSee Table 1 for the basic characteristics of the studies.

^b $p < .05$.

^c $p < .01$.

^dHigher HRQoL in the group studied than in controls.

^eNorm-based scores, with a mean of 50 and SD of 10.

that a very preterm or VLBW birth did not affect overall HRQoL nonetheless found a difference on one or two SF36 scales.^{5,11,13,18} In these instances, we concurred with the author's conclusion of no difference.

Table 3 shows that lower HRQoL was found mainly on the SF36 scales of Physical Functioning^{6,11,13} and Social Functioning.^{4,6,10}

3.3.2 | Disability or neuro-sensory impairments

As Table 2 shows, some of the articles found that, relative to controls, HRQoL was lower in the group born VPT or with a VLBW who in addition had a disability or neuro-sensory impairments (NSI).^{12,15-17} However, this association was not found in other articles.^{6,13} Batsvik et al⁴ presented scores separately for extremely preterm (EPT) adults with a disability and for those without, showing that those without a disability had a lower HRQoL on three 'mental' SF-36 scales. However, after our own calculation of weighted means from the reports in this study, the HRQoL differences between EPT adults and controls were not statistically significant.

3.3.3 | Sex

Table 2 shows that two out of the five articles that reported on associations between sex and the HRQoL of VPT or VLBW adults found females to be at risk for lower HRQoL.^{10,17} Two other studies found that differences in HRQoL between VPT or VLBW adults and the control group differed between females and males.^{11,20} Of these, one study showed physical functioning and general health perception to be poorer in males born EPT than in male controls, while females did not differ.¹¹ The other study showed the HRQoL of EPT/ELBW males improved significantly on three subscales from ages 10 to 18, but differed no longer from controls at age 18, whereas HRQoL in EPT/ELBW females remained significantly lower on three scales (mental health, physical functioning and general health) than in controls.²⁰ One study found no effect of sex.¹²

4 | DISCUSSION

'Quality of life: What is it? How should it be measured?' is the title of the paper written by Aaronson in 1988²¹ in an effort to define Quality of Life Research. Almost three decades later, in 2016, Karimi and Brazier²² pointed out that the terms 'Health', 'Health-Related Quality of Life' and 'Quality of Life' are still often used interchangeably, causing conceptual confusion and involving a great variety of measurement instruments. This issue is certainly important to our systematic review, which found no conclusive evidence of differences between the HRQoL of term-born adults aged 18 and those born very preterm or with a very low birth weight. Although it is perfectly clear how VPT or VLBW are to be defined and accurately measured, the same cannot be said of HRQoL. Between them, in the

18 follow-up studies of preterms included in our review as many as nine different HRQoL instruments had been used. In essence, results across measures were incomparable, as different domains of HRQoL are covered by different measures. If two studies conclude that there is a HRQoL difference between adults born VPT/VLBW and those born term, one study may refer to differences in physical domains of HRQoL and the other to differences in psychological domains.

Comparability was problematic even in studies that used the same HRQoL instrument. The HUI3, for instance, has eight separate single-attribute scores, as well as a total utility score that ranges from -0.36 to 1 . Although a 0.03 difference in total HUI3 utility score was reported to be clinically significant,²³ such a total difference may result from differences in any one of the eight HUI3 attributes, which range from the clinical attributes of vision or hearing to psychological attributes such as emotion or pain. Differences may be clinically significant, but if the domains in which such differences were found vary between studies, these are not comparable. With this caveat in mind, it is remarkable that two out of four studies in our review that did find lower HRQoL in VPT/VLBW subjects (ie involving the McMaster and the Bavarian cohort), had used the HUI3 as an outcome measure. As Wolke²⁴ pointed out, a large nationwide Dutch VPT/VLBW cohort²⁵—which was not included in our review because it lacks a control group—showed similar HUI3 results in adulthood. It may well be that the HUI is more sensitive than other instruments to HRQoL fluctuations in VPT/VLBW subjects. It should be noted that the McMaster cohort included only extremely low birth weight babies (<1000 g). Such small babies are very vulnerable and more of them are disabled and have a lower HUI3 score in adulthood than VPT/VLBW subjects.^{7,9}

Health-related quality of life questionnaires may be administered in various modes, such as computer-assisted mode, by paper-and-pencil, by telephone, or in a face-to-face interview. One study comparing modes of administration of HUI3 in VPT/VLBW patients showed that patients reported more morbidity in a face-to-face interview than in the paper-and-pencil mode.²⁶ Different modes of administration may have added to the incomparability of the studies in our review.

In HRQoL research, there is at least a consensus that the source of information should be the patients themselves. This is reflected in the growing popularity of patient-reported outcome measures (PROMs) in evaluating the effects of clinical trials and care.²⁷ A proxy such as a parent, a spouse, a nurse or physician may be used only when patients are not capable of providing information themselves, for instance due to severe illness or disability. Proxies should be instructed to behave as 'true' proxies, reporting what they think the patient's report would have been, and not their own opinion and perspective on the patient's functioning and well-being. Selective attrition and bias by variation in the source of information were two major issues in our evaluation of the association between VPT/VLBW and HRQoL. High numbers of non-responders were severely disabled, male, from an ethnic minority, or from a lower socio-economic stratum.^{28,29}

If it is assumed that disability results in lower HRQoL, it is possible that any underrepresentation of those most disabled, and other risk groups, leads to the overestimation of HRQoL. This may in turn lead to the false conclusion that there are no differences between controls and those born VPT or with a VLBW. This phenomenon was also reported by Wolke and colleagues³⁰ with respect to HRQoL in 13-year-old VPT/VLBW subjects in the Bavaria cohort.

In order to have data on all participants, some studies included in this review used parent proxies for participants with physical or mental health difficulties.²⁴ As parents tend to underestimate the HRQoL of their children—especially in the psychological domains^{24,31}—these studies are likely to have underestimated HRQoL. Overestimation and underestimation effects may have occurred simultaneously: as well as disabled patients who have only a parent proxy report, a typical VPT/VLBW adult cohort will have disabled patients who drop out altogether. One sophisticated way to solve this differential attrition issue is to do as Lunenburg and colleagues³² did: statistically impute values for missing observations on the basis of the information from previous assessments of the patients themselves.

As a final note on the proxy problem, we should indicate that although a patient is, by definition, never biased in HRQoL reports, while a parent probably is, a parent report is probably more accurate for purposes of predicting societal performance than self-report is. While patients compare their performance with that of their direct peers, parents are likely to report from a more comprehensive societal point of reference.³³

Some studies in our review reported on results from regional or even hospital-based cohorts. As most of these had small sample sizes, they provided little statistical power for evaluation of differences between VPT/VLBW subjects and controls, especially in subgroup analyses. Recently established collaborative research platforms, such as RECAP (Research on European Children and Adults born Preterm; www.recap-preterm.eu), greatly increase statistical power by making it possible to consolidate results over several cohorts. When studying the effects of disabilities on HRQoL, such power is clearly needed. Some studies in our review broke down their results to two levels of disability and consequently subsample sizes were very small. As well as a lack of statistical power, differences in conceptualisation of disability make these studies difficult to compare. Not only should HRQoL be harmonised and standardised, the concept of disability should be clearly defined and thereby measured more meaningfully.

Zwicker and Harris' systematic review¹ on HRQoL from pre-school age to adulthood concluded that the impact of a VPT or VLBW birth on HRQoL is greatest during the younger years, but starts to decrease at the onset of adolescence. Seeing the results of our review, it is justified to hypothesise that differences continue to diminish later in adulthood HRQoL, and may gradually fade out altogether. Such fading is in line with recent findings^{34,35} showing that mental health diagnoses decreased as the age of VP or EP patients increased. This may have resulted from adaptation, coping and resilience by adults who—considering what might have happened

after such a vulnerable start in life—perceived themselves to be doing well. However, as Zwicker and Harris¹ speculated, fading of HRQoL differences may also have been the result of a measurement artefact, in that parents' proxy reports were in the process of being replaced in due course by adolescents/adults' self-report, who often rate their own HRQoL higher than parent proxies do.¹

In adulthood, developmental milestones and societal challenges may vary from young adulthood to more advanced adulthood.³⁶ HRQoL differences between VPT/VLBW adults and term-born controls that fade away in early adulthood may well recur in these subjects' late twenties or early thirties. This may account for a lower HRQoL in VPT/VLBW adults than in term controls later in adulthood compared to less profound differences at an earlier adult age, as found longitudinally in the Norwegian Trondheim cohort⁵ and Canadian McMaster cohort.⁷ Batsvik et al⁴ hypothesised specifically that, when dealing with the demands and responsibilities of later adulthood, disabled adults born EPT will be more troubled by physical signs and symptoms, while non-disabled adults born EPT will experience less than optimal mental HRQoL.

How is the external validity of our results? To which populations may the results be generalised? Our review consolidates results from 18 studies that met our strict inclusion criteria. We feel our results can be generalised at least to the samples of the cohorts involved in these studies and very likely beyond the samples, in fact to the populations from which these had been carefully drawn. These populations include only adults born preterm in the pre-surfactant and pre-antenatal-steroid era. We report that adults born VP/VLBW differed little in HRQoL from term-born adults. We cannot really know whether the same findings will in the future apply for infants born since the mid 1990s. Current evidence suggests that increased survival is not matched by improved outcome, and sometimes by slightly worse outcome.³⁷⁻⁴⁰ In addition, in some areas there may be an increase in what were rarer outcomes. For example, in the Swedish EXPRESS cohort study, some 25% of infants born < 27 weeks of gestation were reported to test positive for autism spectrum disorders (www.ncbi.nlm.nih.gov/pubmed/26689588). The emergence of social cognition disorders in the most immature infants emphasises the need to monitor HRQoL in later cohorts of very preterm adults.

4.1 | Limitations

Some articles included in this systematic review^{4-5,7} stated that, due to the multiple subscales and subgroups and the large number of comparisons being made, multiple testing can increase the risk of false-positive findings. This issue is apparent in Table 3, which shows the results of the SF36 studies; these present a total of 72 tests for statistical significance.

While there is always the possibility of some publication bias—reflecting an interest on the part of some groups of researchers and neonatologists in showing that no HRQoL problems result from the treatment of VPT or VLBW infants—results are usually more likely to get published if they show 'significant differences'.⁴¹ But for the

purposes of this systematic review, we contacted all known cohorts for publications on HRQoL through APIC. The risk of publication bias is therefore low.

As six studies had no control group, we had to exclude them from our review. To be able to include such articles in a future systematic review or meta-analysis, it would be useful to all those in the field if controls or reliable and comparable population norms were identified.

5 | CONCLUSIONS

5.1 | Implications for research

Due to various dissimilarities—of disability, definitions and measurements of HRQoL, and of study characteristics such as age ranges and inclusion criteria—it was difficult to compare studies. Standardisation of measurement in preterms research is highly recommendable.

Many studies in our review had used the 36-item Short Form Health Survey (SF-36), a profile measure that provides a comprehensive appraisal of functioning in several domains of health. Other studies had used the preference-based utility measure of the Health Utility Index, which, through a preference-based utility function, produces a single health utility score that can be used to calculate quality-adjusted life years, thereby enabling health economic evaluations. As both approaches have advantages and drawbacks, we propose to combine them. By using the SF-12, a parsimonious combination is possible.^{42,43} Of course, HRQoL measurement may be extended with other measures as needed.

Further research is needed to better understand the causes of HRQoL differences between adults born preterm and adults born term. More systematic account should be taken of the role of disability, perinatal and neonatal factors. It would also be interesting to study why, despite their evident neonatal medical struggles and proven educational and social problems, there is no conclusive evidence that the HRQoL of those born VPT or VLBW in adulthood is affected. This is all the more challenging in the light of recent findings that VPT/VLBW subjects lagged behind their term-born peers in related areas of long-term psychosocial outcome, such as wealth⁴⁴ and engagement in romantic partnership and sexual intercourse.⁴⁵

The selective non-response of males and those most disabled can be corrected for in data analyses, for example by multiple imputation techniques.⁴⁶ In long-term follow-up studies, it is also important to explore why certain groups respond less than others and how their continued participation can be facilitated as it is in the context of the RECAP collaboration.

If the data of several cohorts are combined, the power of the analyses will increase, thereby allowing multivariate and subgroup analyses and the interpretation of cultural and treatment differences. Research collaboration and data sharing are facilitated by international consortia such as the APIC research group (www.apic-preterm.org) and data platforms such as the RECAP platform initiative.

5.2 | Implications for practice

Although the inconclusiveness of our results makes it impossible to derive any implications for clinical practice, they indicate that the HRQoL of individuals born preterm seems generally satisfactory in these cohorts' home countries. However, our review also shows that disabled groups may be at risk for lower HRQoL and that, later in adulthood, even those without a disability may experience impacts on mental HRQoL for which treatment or coaching might be beneficial.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Gijsbert (Erik) Verrips  <https://orcid.org/0000-0003-0861-945X>

REFERENCES

- 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-e376.
- Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med*. 2009;6(7):e1000100.
- Darlow BA, Horwood LJ, Pere-Bracken HM, Woodward LJ. Psychosocial outcomes of young adults born very low birth weight. *Pediatrics*. 2013;132(6):e1521-e1528.
- Båtsvik B, Vederhus BJ, Halvorsen T, Wentzel-Larsen T, Graue M, Markestad T. Health-related quality of life may deteriorate from adolescence to young adulthood after extremely preterm birth. *Acta Paediatr*. 2015;104(9):948-955.
- Lund LK, Vik T, Lydersen S, et al. Mental health, quality of life and social relations in young adults born with low birth weight. *Health Qual Life Outcomes*. 2012;10(1):146.
- Husby IM, Stray KMT, Olsen A, et al. Long-term follow-up of mental health, HRQoL and associations with motor skills in young adults born preterm with very low birth weight. *Health Qual Life Outcomes*. 2016;14(1):56.
- Saigal S, Stoskopf B, Pinelli J, et al. Self-perceived HRQoL of former extremely low birth weight infants at young adulthood. *Pediatrics*. 2006;118(3):1140-1148.
- Poole KL, Islam UA, Schmidt LA, et al. Childhood motor function, health related quality of life and social functioning among emerging adults born at term or extremely low birth weight. *J Dev Phys Disabil*. 2017;29(3):369-383.
- Saigal S, Ferro MA, Van Lieshout RJ, Schmidt LA, Morrison KM, Boyle MH. HRQoL trajectories of extremely low birth weight survivors into adulthood. *J Pediatr*. 2016;179:68-73.
- Natalucci G, Becker J, Becher K, Bickle GM, Landolt MA, Bucher HU. Self-perceived health status and mental health outcomes in young adults born with less than 1000 g. *Acta Paediatr*. 2013;102(3):294-299.
- Cooke RWI. Health, lifestyle, and quality of life for young adults born very preterm. *Arch Dis Child*. 2004;89(3):201-206.
- Gäddlin PO, Finnström O, Sydsjö G, Leijon I. Most very low birth weight subjects do well as adults. *Acta Paediatr*. 2009;98(9):1513-1520.
- Roberts G, Burnett AC, Lee KJ, et al. Quality of life at age 18 years after extremely preterm birth in the post-surfactant era. *The Journal of Pediatrics*. 2013;163(4):1008-1013.
- Baumann N, Bartmann P, Wolke D. HRQoL into adulthood after very preterm birth. *Pediatrics*. 2016;137(4):e20153148.
- Bjerager M, Steensberg J, Greisen G. Quality of life among young adults born with very low birth weights. *Acta Paediatr*. 1995;84(12):1339-1343.
- Dinesen SJ, Greisen G. Quality of life in young adults with very low birth weight. *Arch Dis Child Fetal Neonatal Ed*. 2001;85(3):F165-F169.
- Björkqvist J, Hovi P, Pesonen AK, et al. Adults who were born preterm with a very low birth weight reported a similar health-related quality of life to their term-born peers. *Acta Paediatr*. 2017;107(2):354-357.
- Hack M, Cartar L, Schluchter M, Klein N, Forrest CB. Self-perceived health, functioning and well-being of very low birth weight infants at age 20 years. *J Pediatr*. 2007;151(6):635-641.
- Hallin AL, Stjernqvist K. Adolescents born extremely preterm: behavioral outcomes and quality of life. *Scand J Psychol*. 2011;52(3):251-256.
- Vederhus BJ, Eide GE, Natvig GK, Markestad T, Graue M, Halvorsen T. HRQoL and emotional and behavioral difficulties after extreme preterm birth: developmental trajectories. *PeerJ*. 2015;3:e738.
- Aaronson NK. Quality of life: what is it? How should it be measured? *Oncology*. 1988;2:69-74.
- Karimi M, Brazier J. Health, health-related quality of life and quality of life: what is the difference? *Pharmacoeconomics*. 2016;34(7):645-649.
- Horseman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI): concepts, measurement properties and applications. *Health Qual Life Outcomes*. 2003;1:54.
- Wolke D. Born extremely low birth weight and health related quality of life into adulthood. *J Pediatr*. 2016;179:11-12.
- Verloove-Vanhorick SP, Pauline, Verwey RA, Brand R, Bennebroek Gravenhorst J, Keirse M, Ruys JH. Neonatal mortality risk in relation to gestational age and birthweight. Results of a national survey of preterm and very low birthweight infants in the Netherlands. *Lancet*. 1986;327(8472):55-57.
- Verrips GHW, Vogels AGC, den Ouden AL, Paneth N, Verloove-Vanhorick SP. Measuring health-related quality of life in adolescents: agreement between raters and between methods of administration. *Child Care Health Dev*. 2000;26(6):457-469.
- Cella D, Hahn EA, Jensen SE, et al. *Patient-Reported Outcomes in Performance Measurement*. Research Triangle Park (NC): RTI Press; 2015.
- Hille ETM, Elbertse MD, Bennebroek Gravenhorst J, Brand R, Verloove-Vanhorick SP. Nonresponse bias in a follow-up study of 19-year-old adolescents born as preterm infants. *Pediatrics*. 2005;116(5):662-666.
- Hille ETM, den Ouden AL, Stuifbergen MC, et al. Is attribution bias a problem in neonatal follow-up? *Early Hum Dev*. 2005;81:901-908.
- 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;163:1020-1026.
- Verrips GHW, Stuifbergen MC, den Ouden AL, et al. Measuring health status using the health utilities index: agreement between raters and between modalities of administration. *J Clin Epidemiol*. 2001;54:475-481.
- van Lunenburg A, van der Pal SM, van Dommelen P, van der Pal - de Bruin KM, Bennebroek Gravenhorst J, Verrips GHW. 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.
- Wolke D, Chernova J, Eryigit-Madzwamuse S, Samara M, Zwierzynska K, Petrou S. Self and parents perspectives on health-related quality of life of adolescents born very preterm. *J Pediatr*. 2013;163:1020-1026.

34. Jaekel J, Baumann N, Bartmann P, Wolke D. Mood and anxiety disorders in very preterm/very low-birth weight individuals from 6 to 26 years. *J Child Psychol Psychiatry*. 2018;59(1):88-95.
35. Johnson S, O'Reilly H, Ni Y, Wolke D, Marlow N. Psychiatric symptoms and disorders in extremely preterm young adults at 19 years of age and longitudinal findings from middle childhood. *J Am Acad Child Adolesc Psychiatry*. 2019;58(8):820-826.e6.
36. van der Pal SM, Maurice-Stam H, Grootenhuis MA, van Wassenaer-Leemhuis AG, Verrips GHW. Psychosocial developmental trajectory of young adults born very preterm in the Netherlands. *J Patient Rep Outcomes*. 2019;3(1):17.
37. Wolke D, Johnson S, Mendonça M. The life course consequences of very preterm birth. *Ann Rev Dev Psychol*. 2019;1(1):69-92.
38. Spittle AJ, Cameron K, Doyle LW, Cheong JL. Motor impairment trends in extremely preterm children: 1991-2005. *Pediatrics*. 2018;141:e20173410.
39. Twilhaar ES, Wade RM, de Kieviet JF, van Goudoever JB, van Elburg RM, Oosterlaan J. Cognitive outcomes of children born extremely or very preterm since the 1990s and associated risk factors: a meta-analysis and meta-regression. *JAMA Pediatr*. 1990;172:361-367.
40. Cheong JLY, Anderson PJ, Burnett AC, et al. Changing neurodevelopment at 8 years in children born extremely preterm since the 1990s. *Pediatrics*. 1990;139:e20164086.
41. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Medicine*. 2009;6(7):e1000100
42. Gandek B, Ware J, Aaronson N, et al. Cross validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA project. *J Clin Epidemiol*. 1998;51(11):1171-1178.
43. Van den Berg B. SF-6D population norms. *Health Econ*. 2012;21(12):1508-1512.
44. Bilgin A, Mendonca M, Wolke D. Preterm birth/low birth weight and markers reflective of wealth in adulthood: a meta-analysis. *Pediatrics*. 2018;142(1):e20173625.
45. Mendonça M, Bilgin A, Wolke D. Association of preterm birth and low birth weight with romantic partnership, sexual intercourse, and parenthood in adulthood: a systematic review and meta-analysis. *JAMA Network Open*. 2019;2(7):e196961.
46. Van Buuren S, Groothuis-Oudshoorn K. MICE: multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45(3):1-67.

How to cite this article: van der Pal S, Steinhof M, Grevinga M, Wolke D, Verrips G(E). Quality of life of adults born very preterm or very low birth weight: A systematic review. *Acta Paediatr*. 2020;00:1-15. <https://doi.org/10.1111/apa.15249>