

Strength of association between umbilical cord pH and perinatal and long term outcomes: systematic review and meta-analysis

Gemma L Malin, clinical research fellow,¹ Rachel K Morris, clinical research fellow,¹ Khalid S Khan, professor of obstetrics, gynaecology, and clinical epidemiology^{1,2}

¹Department of Obstetrics, School of Clinical and Experimental Medicine, College of Medical and Dental Sciences, University of Birmingham

²Department of Obstetrics and Gynaecology, Birmingham Women's Foundation NHS Trust, Birmingham

Correspondence to: G L Malin, Academic Department of Obstetrics and Gynaecology, Birmingham Women's Hospital, Edgbaston, Birmingham B15 2TG. g.l.malin@bham.ac.uk

Cite this as: *BMJ* 2010;340:c1471 doi:10.1136/bmj.c1471

ABSTRACT

Objective To evaluate the association between umbilical cord pH at birth and long term outcomes.

Design Systematic review and meta-analysis.

Data sources Medline (1966-August 2008), Embase (1980-August 2008), the Cochrane Library (2008 issue 8), and Medion, without language restrictions; reference lists of selected articles; and contact with authors.

Study selection Studies in which cord pH at birth was compared with any neonatal or long term outcome. Cohort and case-control designs were included.

Results 51 articles totalling 481 753 infants met the selection criteria. Studies varied in design, quality, outcome definition, and results. Meta-analysis carried out within predefined groups showed that low arterial cord pH was significantly associated with neonatal mortality (odds ratio 16.9, 95% confidence interval 9.7 to 29.5, $I^2=0\%$), hypoxic ischaemic encephalopathy (13.8, 6.6 to 28.9, $I^2=0\%$), intraventricular haemorrhage or periventricular leucomalacia (2.9, 2.1 to 4.1, $I^2=0\%$), and cerebral palsy (2.3, 1.3 to 4.2, $I^2=0\%$).

Conclusions Low arterial cord pH showed strong, consistent, and temporal associations with clinically important neonatal outcomes that are biologically plausible. These data can be used to inform clinical management and justify the use of arterial cord pH as an important outcome measure alongside neonatal morbidity and mortality in obstetric trials.

INTRODUCTION

Perinatal asphyxia is a major cause of neonatal and childhood morbidity and mortality and has been associated with neonatal death¹; hypoxic ischaemic encephalopathy and seizures^{2,3}; intraventricular haemorrhage⁴; cerebral palsy⁵; and delayed development.⁵ Perinatal asphyxia is predicated by fetal acidosis, determined by umbilical cord pH at birth.⁶ Cerebral palsy is thought to occur more frequently at an arterial cord pH of <7.00 and a base deficit of ≥ 12 mmol/L.⁷ These criteria, however, have been derived through consensus, not through evaluation of collated summaries of evidence,⁷ leading to clinical uncertainty.⁸ This is because existing observational studies of the association between cord pH and

outcomes have drawn inconsistent inferences, partly as a result of the different thresholds used to define abnormality, the variety of outcomes evaluated, and the different variables measured (arterial cord pH, venous cord pH, or base excess).⁹⁻¹¹ It has been suggested that neonatal complications are associated with metabolic rather than respiratory acidosis.^{12,13} Respiratory acidosis arises in the early stages of impaired blood supply to the fetus; hypoxaemia and hypercapnia occur, leading to a reduction in pH with a normal base excess.⁶ If hypoxia is prolonged, anaerobic metabolism results and base excess rises secondary to the presence of lactic acidosis.

On the basis of nine studies, a review¹⁴ claimed an association between fetal acidosis and neonatal death and cerebral palsy. This review was, however, carried out in the late 1990s, during which time new studies on the subject have been published and guidelines produced on the methodology and reporting of systematic reviews, including quality assessment of included studies, which were not in widespread use at the time of the review's publication.¹⁵ Substantial uncertainty therefore remains about the value clinicians may attach to acidosis in the clinical management of neonates and the long term implications of a low arterial cord pH. However, cord pH is also commonly used as an outcome measure in obstetric clinical trials¹⁶⁻¹⁸ and it is one of the bench marks by which obstetricians judge their performance on the labour ward.¹⁹ A documented low cord pH is a factor that may be used to support medico-legal claims of harm during intrapartum events resulting in long term disability.²⁰ It is therefore imperative that the validity of this association is supported with high quality evidence. We carried out a comprehensive systematic review of the literature to quantitatively establish the strength of association of acidosis at birth with neonatal mortality, morbidity, and long term outcomes and to assess if causal criteria were met.²¹

METHODS

This systematic review was protocol driven using widely recommended methods for reviews²²⁻²⁴ and evaluation of causal associations.^{21,25-28} The review

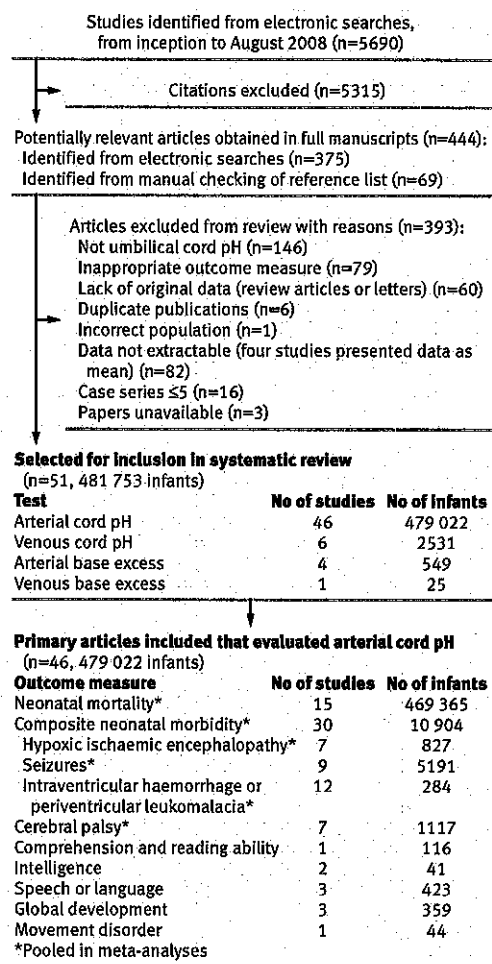


Fig 1 | Study selection process

was carried out and has been reported according to MOOSE guidance.¹⁵

We searched Medline (1966–August 2008), Embase (1980–August 2008), the Cochrane Library (2008 issue 8), and Medion for relevant published articles. To identify “grey” literature we also searched SIGLE, Web of Science, the national research register, and medical conferences register. For the Medline search (also see web extra on bmj.com) we used a combination of MeSH headings, such as umbilical cord, hydrogen-ion concentration, or asphyxia neonatorum; keywords, such as umbilical artery pH and cord pH; and word variants using “OR” for capturing citations of the relevant text. To capture relevant outcomes we combined these using “AND” with a combination of MeSH headings, such as human development and infant mortality; keywords, such as developmental delay and handicap; and word variants. We restricted the search to human studies but applied no language restrictions. We adapted the Medline search strategy for use in other databases. We also hand searched recent major journals. Two investigators (RKM and GLM) carried out the search. Using Reference Manager 11.0, we

constructed a comprehensive database collating all citations.

Study selection and data extraction

Two reviewers (RKM or GLM, partly in duplicate) scrutinised the database and obtained full articles of all citations potentially meeting the predefined selection criteria. Translations of articles in languages other than English were obtained. Two reviewers (GLM and RKM or ZL) included or excluded studies in accordance with the most recent guidance,²³ using strict adherence to the following criteria: population—infants with cord blood obtained at birth; index test—cord blood examined for arterial or venous pH or base excess; outcome—any measure of compromise of neonatal or childhood well-being, such as mortality; neonatal morbidity, including hypoxic ischaemic encephalopathy, seizures, intraventricular haemorrhage, periventricular leukomalacia, and long term outcomes, including cerebral palsy; and study design—observational studies that allowed generation of a 2×2 table (true positives, false positives, false negatives, and true negatives) to compute an estimate of the association between test result and outcomes. We excluded studies with five or fewer cases, because of unreliability.

The manuscripts were examined for duplicated populations. If any were found we selected the most recent and complete version. No language restrictions were applied in study selection. We scrutinised the reference lists of selected studies and review articles and obtained additional relevant articles. Two researchers (GLM and RKM or ZL) used a data collection sheet to extract information in duplicate from the selected articles. Data were extracted on study characteristics, quality, and results, and entered onto an Excel spreadsheet. Data were used to construct 2×2 tables of the association between the cord blood variable at the threshold reported in the paper and the postnatal outcome for each infant. When data were thought to be relevant but 2×2 tables could not be constructed, we contacted the relevant authors. Difficulties in data extraction were resolved by discussion with a third reviewer (KSK).

Study quality assessment

We assessed those articles meeting the selection criteria for methodological quality, defined as the confidence that the study design, conduct, and analysis minimised bias in estimation of the association. The articles were assessed using the complete STARD and QUADAS checklists.^{29,30} These are guidelines for reporting, and methodological quality of, studies on diagnostic accuracy. As the focus of this review was to determine association, the checklist items thought to be most relevant were selected for the purpose of defining the overall quality of included studies and to divide the studies according to quality for meta-analysis and metaregression, in accordance with published guidance.²⁸ We did not assign a quality score as this has been shown to give flawed results.³¹ We considered the cohort study design to be superior to the case-control

Table 1 | Studies included in systematic review of association between cord pH at birth and neonatal and long term outcomes

Studies	Risk factors	Index tests	Outcome measures
High risk population:			
Baenziger et al 1999 ⁵⁰	All ventilated neonates; risk factors for HIE, including meconium liquor, abnormalities on cardiotocogram, low Apgar score or pH, gestation >34 weeks	Arterial cord pH <7.00	Death*; hypoxic ischaemic encephalopathy (Samat grade >1)*; neurological optimality score (age 1 year); Griffiths developmental scale (age 1 year)
Beeby et al 199 ⁴⁴	Gestation <32 weeks; excluded congenital anomalies	Arterial cord pH <7.1	Death*; cerebral palsy (age 1 year); intraventricular haemorrhage grade 3 or 4 on cranial ultrasound scan or at autopsy*
Blackwell et al 2001 ⁵²	All neonates requiring ventilation >48 hours for meconium aspiration syndrome; gestation >37 weeks	Arterial cord pH <7.2	Seizures*
Bresadola et al 1995 ⁵¹	Gestation >24 weeks, <37 weeks; birth weight >700 g; excluded congenital anomalies	Arterial cord pH <7.2	Death*
Casey et al 2001 ⁵³	Neonates who developed respiratory symptoms postnatally; gestation >37 weeks	Arterial cord pH <7.2	Death*; seizures*; respiratory distress requiring ventilation*; meconium aspiration syndrome*
Engle 1999 ⁵⁶	Neonates admitted to neonatal unit directly from delivery suite; gestation >37 weeks	Arterial cord pH <7.00	Hypoxic ischaemic encephalopathy including seizures*
Ertan 2001 ⁹¹	Gestation 24-34 weeks	Arterial cord pH <7.10; arterial base excess >16 mmol/l	Intracranial haemorrhage (Papile, all grades) on cranial ultrasound scan*
Gaudier et al 1994 ⁴⁹	Birth weight 500-1000 g; gestation 23-34 weeks; excluded congenital anomalies	Arterial cord pH <7.05	Wechsler intelligence scale for children (IQ <70); cerebral palsy (chronic non-progressive motor disability characterised by abnormal posture and movements), follow-up to age 7 years; motor or mental deficit severe enough to interfere with normal function, including one or more of mental retardation, cerebral palsy, deafness, blindness, or hydrocephaly
Gea 2007 ⁹³	Birth weight <2000 g; gestation <37 weeks; excluded congenital anomalies, maternal diabetes, and rhesus incompatibility	Venous cord pH <7.20; venous base excess >10 mmol/l	Periventricular or intraventricular haemorrhage on cranial ultrasound scan*; ventilation >24 hours*; necrotising enterocolitis (Bell grade 2)*
Gonzalez de Dios et al 2000 ⁵⁹	At least one risk factor for asphyxia (for example, Apgar score <6, cord pH <7.00, antenatal risk factors); excluded if congenital anomaly, sepsis, metabolic disorder, postnatal depression; gestation >37 weeks	Arterial cord pH ≤7.00	Serum creatinine level >1.2 mg/dl*; hypoxic ischaemic encephalopathy (Levine's criteria, all grades)*; abnormal neurological status (Amiel-Tison criteria), age 2 years
Graham et al 2004 ⁶¹	Gestation 23-34 weeks	Arterial cord pH <7.00	Periventricular leucomalacia or ventricular dilatation on cranial ultrasound scan*
Haddad et al 2000 ⁶²	Neonates with Apgar score 0 at 1 and 5 minutes, resuscitated and transferred to neonatal unit; excluded congenital malformations, chromosome abnormalities, birth before arrival; gestation 25.5-42.1 weeks	Arterial cord pH <7.00	Death*; hypoxic ischaemic encephalopathy*
Hernandez et al 1993 ⁶³	Clinical and radiological evidence of meconium aspiration syndrome; excluded if congenital anomalies, cytomegalic inclusion disease, delivery outside hospital	Arterial cord pH <7.00, <7.10, and <7.20	Ventilation required*; ventilation ≥3 days*
Hibbard et al 1991 ⁶⁴	Birth weight 500-1500 g	Arterial cord pH ≤7.05 and ≤7.15	Death*; intraventricular haemorrhage (Papile grade 1-4) on cranial ultrasound scan*; abnormal neurological status defined as seizures, cortical atrophy, need for shunt placement*; hyaline membrane disease (chest radiography)*; bronchopulmonary dysplasia (chest radiography)*; necrotising enterocolitis*
Holmes et al 2001 ⁶⁶	Birth weight 750-2500 g; gestation 25-35 weeks; excluded if caesarean section before labour, congenital anomalies, uninterpretable fetal heart rate trace	Arterial cord pH <7.10	Death*; intraventricular haemorrhage (Papile grade 3 or 4)*; periventricular leucomalacia*
Kato 1997 ⁵	Birth weight <1500 g. Congenital anomalies excluded	Arterial cord pH <7.20	Death*; cerebral palsy or mental retardation (at least 12 months old)
Loh et al 1998 ⁷²	Included if one or more of following risk factors: abnormality on cardiotocogram, scalp pH <7.25, thick meconium or no liquor, cord prolapse or bradycardia, antepartum haemorrhage, estimated fetal weight <1500 g, gestation <34 weeks, breech, poorly controlled type 1 diabetes, pre-eclampsia, suspected fetal anomalies, transverse or oblique lie at caesarean section, multiple pregnancies	Arterial cord pH <7.00	Hypoxic ischaemic encephalopathy
Luthy et al 1987 ⁷⁴	Gestation 26-32 weeks; excluded multiple pregnancy, non-cephalic presentation, malformations, delivery before labour, antenatal haemorrhage	Arterial cord pH ≤7.20	Death; cerebral palsy (age 18 months); intraventricular haemorrhage on cranial ultrasound scan (Papile grade 3 or 4)*
Murphy et al 1995 ⁷⁵	Gestation 23-32 weeks; excluded multiple pregnancy	Arterial cord pH ≤7.10	Cerebral palsy (permanent impairment of voluntary movement or posture), age unreported
Salafia et al 1995 ⁷⁸	Gestation <32 weeks; excluded if congenital anomalies, multiple pregnancy, maternal diabetes mellitus or chronic hypertension, fetal hydrops, placenta praevia, intrauterine growth restriction	Arterial cord pH <7.10; venous cord pH <7.10	Germinal matrix or intraventricular haemorrhage diagnosed on cranial ultrasound scan (grade 1-4)*
Socol 1994 ⁸¹	Apgar score ≤3 at 5 minutes; excluded birth weight <2000 g and gestation <34 weeks	Arterial cord pH <7.00; arterial base excess >12 mmol/l	Cerebral palsy or motor deficit (age 1-7 years); seizures*; renal impairment (serum creatinine level >1.5 mg/dl, or oliguria)*; cognitive impairment on Wechsler scale (cut off <70), age 1-7 years
Spinillo et al 1995 ⁸²	Birth weight <2500 g	Arterial cord pH <7.20	Bayley index abnormal (71-84), age 2 years
Tejani and Verma 1989 ⁸³	Birth weight ≤2000 g; excluded major congenital anomalies	Arterial cord pH ≤7.10	Death*; intraventricular haemorrhage on cranial ultrasound scan (Papile grade 1-4)*; respiratory distress syndrome (radiological evidence of reticulogranular pattern and air bronchograms)*
Yudkin et al 1994 ⁸⁹	Apgar score ≤3 at 1 minute; gestation >37 weeks; excluded multiple pregnancies and death related to congenital anomaly or rhesus disease	Arterial cord pH <7.15	Death*; any impairment (age 5 years); serious impairment (global delay, lateralising signs and severe deficit in a specific area), age 5 years
Yoon et al 1996 ⁸⁸	Gestation 25-36 weeks; excluded major congenital malformations or death before examination	Arterial cord pH <7.15	Periventricular leucomalacia on cranial ultrasound scan or at autopsy*

Studies	Risk factors	Index tests	Outcome measures
Unselected or low risk population:			
Dennis et al 1989 ⁵⁴	Gestation >37 weeks; singletons surviving to age 4.5 years	Arterial cord pH ≤7.10; arterial base excess >12 mmol/l	Griffiths developmental scales age 4.5 years: locomotor, personal or social, hearing and speech, performance, and overall impairment
Dijkhoorn et al 1986 ⁶¹	Gestation >37 weeks; excluded caesarean delivery or breech presentation	Arterial cord pH ≤7.10 and ≤7.20	Neurological status (abnormal if one of hyperkinesias or hypokinesia, hypertonia or hypotonia, hemisindrome, apathy syndrome, or hyperexcitability syndrome)*
D'Souza et al 1983 ⁵⁵	Normal pregnancy, vaginal delivery 39-42 weeks	Venous cord pH <7.27	Neurological status (abnormal if one of hypotonia, lethargy, feeding difficulties, jittery)*
Ghosh et al 2003 ⁵⁷	Gestation >37 weeks; singletons excluded if rhesus isoimmunisation, maternal anaemia, or diabetes mellitus	Arterial cord pH ≤7.15	Hypoxic ischaemic encephalopathy*; death*
Gilstrap et al 1989 ⁵⁸	Gestation >37 weeks, cephalic presentation, birth weight >2500 g; excluded congenital anomalies	Arterial cord pH <7.00	Hypotonia >24-48 hours*; respiratory distress requiring oxygen*; seizures*
Graham 2002 ⁶⁰	Excluded congenital anomalies	Arterial cord pH <7.00	Seizures*
Heller et al 2003 ¹	Excluded congenital anomalies	Arterial cord pH ≤7.00, ≤7.10, and ≤7.20	Death*
Hogan 2007 ⁶⁵	Gestation >37 weeks	Arterial cord pH ≤7.15	Composite reference standard hypoxic ischaemic encephalopathy (Sarnat, all grades) or hypoxic death*
Huisjes and Aamoudse 1979 ⁶⁷	Population characteristics unreported	Arterial cord pH ≤7.09	Abnormal neonatal neurological status (Prechtl method)*
Ingemansson et al 1997 ⁶⁸	Population characteristics unreported	Arterial cord pH <7.05 and <7.00	Death*; cerebral palsy (age 4 years); impact of events scale (Sarnat grade 1-3)*; attention deficit (age 4 years); speech difficulty (age 4 years); motor delay (age 4 years)
Jurgens-van der Zee et al 1979 ⁶⁹	Neonates; excluded if neonatal death or parents refused examination	Venous cord pH <7.20	Abnormal neurological examination result (Prechtl method): 1 or more of increased or decreased excitability including seizures, apathy or coma; abnormal motility or tone; peripheral or central nervous system lesions*
Lama et al 2007 ⁷⁰	Gestation ≥24 weeks	Arterial cord pH <7.00	Seizures*; periventricular leucomalacia*; intraventricular haemorrhage*; respiratory dysfunction*; renal dysfunction* (thresholds unreported)
Litschgi et al 1974 ⁷¹	Population characteristics unreported	Arterial cord pH <7.09	Abnormal neurological examination result (24 hours after birth)*
Low 1997 ⁷³	Gestation ≥37 weeks	Arterial base excess >12 mmol/l	Composite outcome score*; neurological abnormality (lethargy, abnormal tone, coma, or seizures)*; respiratory dysfunction (continuous positive airway pressure or ventilation required)*; cardiovascular dysfunction (hypotension, hypertension, abnormal electrocardiogram or echocardiogram)*; renal dysfunction (serum creatinine level >100 μmol/l, anuria, or oliguria <1 ml/kg/h)*
Perlman and Risser 1996 ⁷⁶	Gestation ≥37 weeks	Arterial cord pH ≤7.00	Seizures*
Sakuraba and Saling 1989 ⁷⁷	Population characteristics unreported	Arterial cord pH ≤7.19 and ≤7.24	Intracranial haemorrhage diagnosed on cranial ultrasound scan*
Schneider and Tanner 1985 ⁷⁹	Twins only included	Arterial cord pH <7.20	Binet Simon Kramer (intelligence); language test; emotional intelligence test; Raven intelligence test (non-spoken); neurological status (all tests carried out at age 5-7 years)
Silva et al 2008 ⁸⁰	Gestation ≥34 weeks; excluded congenital malformations and chromosome anomalies	Arterial cord pH <7.00 and <7.10	Hypotonia at birth necessitating admission to neonatal unit*; hypoxic ischaemic encephalopathy*; seizures*
Svirko et al 2008 ⁹²	Gestation ≥36 weeks; one of reference tests done and cord pH available; excluded if delivered by prelabour elective caesarean section	Arterial cord pH <7.10	Wechsler objective reading dimensions (age 6-8 years); test for comprehension of grammar (age 5-7 years); Naglieri non-verbal ability (age 6-8 years); cut off <100 age standardised score used for all tests
Thoulon et al 1972 ⁸⁴	Population characteristics unreported	Venous cord pH <7.20	Death*; abnormal neurological examination (age 18 months)
Valentin et al 1993 ⁸⁵	Anomalies not excluded	Arterial cord pH ≤7.00 and ≤7.10; venous cord pH ≤7.10 and ≤7.20	Composite measure of neonatal sequelae, including severe symptoms requiring treatment (for example, ventilation, intravenous fluids), death, or survival with sequelae*
Van den Berg 1996 ¹⁰	Excluded chromosomal or major congenital anomalies or intrauterine infection	Arterial cord pH <7.00 (compared with group of neonates with cord pH >7.24)	Seizures*; intracranial haemorrhage on cranial ultrasound scan*; periventricular leucomalacia on cranial ultrasound scan*; renal impairment (serum creatinine level >90 μmol/l*); abnormal liver function (aspartate transaminase level >33 U/L, alanine aminotransferase level >25 U/l)*; necrotising enterocolitis (criteria unreported)*
Vintzileos et al 1993 ⁹⁰	Gestation ≥26 weeks; excluded known congenital or chromosomal anomalies	Arterial cord pH <7.10	Death*
Wildshut et al 2005 ¹¹	Gestation 37-42 weeks; neonates included if growth between 2.3 and 97.7th centiles, vertex position, stay in hospital at least 3 days after birth; excluded if hypoxia ischaemia other than caused by perinatal adverse conditions. Exclusions included meconium aspiration, respiratory distress syndrome, infection, birth after complicated pregnancy, congenital malformations, maternal medication, alcohol or drug misuse, metabolic disorders	Arterial cord pH <7.10	Movement ABC test (score <16 or unable to perform due to movement disorder), age 4 years
Winkler 1991 ⁸⁶	Gestation ≥37 weeks	Arterial cord pH <7.20	Composite reference standard: seizures or neonatal death*
Wu et al 1998 ⁸⁷	Gestation ≥37 weeks; singletons only	Arterial cord pH <7.19	Motor delay; speech delay; difficulty concentrating (all assessed age 4 years)

*Outcomes within neonatal period.

Table 2 | Methodological quality of studies in systematic review of association between low cord pH at birth and neonatal and long term morbidity

Quality item	No (%) of studies (n=51)		
	Yes	No	Unclear
Cohort study design	42 (82)	9 (18)	0
Population adequately described	40 (78)	2 (4)	9 (18)
Consecutive recruitment	2 (4)	1 (2)	48 (94)
Prospective recruitment	15 (29)	29 (57)	7 (14)
Appropriate outcome measure	51 (100)	0	0
Outcome measure blinded	7 (14)	0	44 (86)
>90% of infants had outcome measure	41 (80)	10 (20)	0
Index test and outcome measure described	8 (16)	12 (24)	31 (61)
Intervention between index test and outcome	0	0	51 (100)
Quality classification:			
High	23 (45)	0	0
Medium	0	20 (39)	0
Low	0	0	8 (16)

design, and we considered the cohort design as a covariate in metaregression. Because of the small number of studies included in each meta-analysis we did not think it appropriate to consider each remaining quality item as an individual covariate for metaregression,³² and we therefore divided the studies into two categories on the basis of several criteria. We rated a study as being of high quality if it had at least four of the following items: adequate description of population, adequate description of the test and outcome measure, consecutive recruitment, prospective recruitment, more than 90% completion of follow-up, appropriate outcome measurement, blinding of the investigators carrying out the outcome measure, and a statement on the use of intervention between the index test and outcome. When a study adhered to three or fewer of these criteria we considered it to be of medium or low quality.

Data synthesis

We used 2×2 tables to compute odds ratios and 95% confidence intervals for each pair of index test and outcome. The odds ratio was selected as the summary statistic as it is relatively constant regardless of the test threshold³³ and is often used to show an epidemiological association.³⁴ It represents the effect of the exposure on the odds in an unbiased fashion and enables linkage between the results of case-control and cohort studies.³⁴ As the focus of this review was to establish the association between cord pH and adverse outcome rather than the value of cord pH as a prognostic test, we did not calculate other summary statistics (including sensitivity, specificity, and likelihood ratios). We calculated the estimated predictive interval for each meta-analysis, which relates to the effect of a new study that would be eligible for inclusion in meta-analysis and therefore allows the full uncertainty around inferences to be calculated, including both magnitude and consistency.³⁵

When tables contained cells with a value of 0, we added 0.5 to these cells to allow the calculation of

variances.³⁶ Analyses were done for groups defined a priori according to index test (arterial cord pH, venous cord pH, and base excess). The main outcome measures considered were neonatal mortality, a composite measure of neonatal morbidity, and cerebral palsy. To maximise the number of events that could be included in the analysis and to avoid the need to select a single morbidity as a primary outcome measure, we used a composite outcome measure for neonatal morbidity. However, a hazard of composite outcome measures is the assumption that the significance of the result applies to all components.³⁷ To deal with this, we analysed the component outcomes of hypoxic ischaemic encephalopathy, seizures, intraventricular haemorrhage, or periventricular leucomalacia. When the composite outcome measure was used, we took care to ensure that each individual was counted only once in each analysis, particularly when studies reported multiple outcomes for a single population. When multiple outcomes and test thresholds were reported, we attempted to select the most consistent threshold and outcome across the composite analysis—for example, where studies reported on hypoxic ischaemic encephalopathy, seizures, and intraventricular haemorrhage, we included hypoxic ischaemic encephalopathy and seizures in the meta-analysis as we believed this to be more clinically consistent and overall the more common outcome across the included studies.

We plotted odds ratios data in forest plots, assessing heterogeneity visually and statistically using the I^2 statistic, a scale that measures the degree of heterogeneity across studies in a meta-analysis.³⁸ Within the largest groups (arterial cord pH paired with neonatal mortality and neonatal morbidity) we explored the reasons for heterogeneity using metaregression, planned a priori in keeping with published recommendations.^{39,40} We considered study design (cohort *v* case-control), study quality (high *v* medium or low), and population risk (high *v* low/unselected or unreported) as potential sources of heterogeneity. We defined a population at high risk for complications on the basis of reported characteristics, including abnormalities on cardiotocograms, meconium liquor, a low Apgar score at birth, gestation less than 37 weeks, and birth weight less than 2500 g. It was not possible to use a low risk population as comparator because only one included study met the criteria for this definition.⁴¹ Given the reporting quality of the primary studies it was also not possible to explore other potential sources of heterogeneity. Metaregression analysis was univariate throughout, in accordance with published guidance that recommends allowing at least 10 studies per covariate explored.⁴² When a reason for heterogeneity was identified, we carried out subgroup analyses. As complications such as cerebral palsy and intraventricular haemorrhage have been shown to be increased in pre-term and low birthweight infants we analysed studies with these populations separately and compared the results with that of a term population.

As clinical heterogeneity was present between studies, we used a random effects model throughout.

Table 3 | Exploration of heterogeneity in estimation of association between low arterial cord pH at birth and neonatal mortality and morbidity

Factor	Univariable analysis	
	Odds ratio (95% CI)	P value
Neonatal mortality:		
Cohort v case-control design	0.4 (0.03 to 6.5)	0.52
High quality v medium or low quality study	0.8 (0.3 to 2.4)	0.63
High risk v unselected population	0.3 (0.1 to 1.0)	0.048*
Composite neonatal morbidity:		
Cohort v case-control design	1.1 (0.2 to 5.3)	0.90
High quality v medium or low quality study	1.4 (0.5 to 3.7)	0.51
High risk v unselected or low risk population	0.4 (0.2 to 0.9)	0.03*

Dummy variables were used to set the reference category as medium or low quality, case-control design, and unselected or low risk population.

*P<0.05.

The presence of publication bias was determined using the Harbord test.^{42,43} To assess the effect of the use of different thresholds to define low pH, we repeated the meta-analyses using the bivariate method, which takes into account any threshold effect.⁴⁴ We additionally analysed subgroups within the outcome measures of neonatal mortality and neonatal morbidity according to the arterial pH threshold reported (<7.00, <7.10, and <7.20). We statistically analysed data using Meta-Disc 1.3 software,⁴⁵ and Stata version 10.0 using the metan, metandi, and meta-bias commands.⁴⁶⁻⁴⁸

To assess a causal association for each outcome we considered Hill's criteria.^{21,25} We examined strength of association using point estimates of odds ratios and consistency using the direction of association in the forest plots and heterogeneity statistics.

RESULTS

An initial search of 5690 citations identified 51 primary articles^{14,5,10,11,41,49-91}, two after obtaining information from the authors (fig 1).^{92,93} The included studies totalled 481 753 infants. Overall, 190 2×2 tables were produced; 43 of the studies providing these data were eligible for inclusion in meta-analyses according to predefined outcome measures, totalling 479 383 infants and 70 2×2 tables. One included study was thought to contain duplication⁶⁰; however, the outcomes reported were different and therefore both papers^{60,61} were included in the review in separate meta-analyses. When more than one 2×2 table was produced per study, double counting of infants in any one analysis was avoided.

The included studies used various index tests (table 1): most papers (n=46) reported arterial cord pH, with thresholds ranging from 7.00 to 7.24. Arterial base excess was reported in four studies, with thresholds of 12-16 mmol/l.^{61,73,81,91} Five studies reported venous cord pH; three using a threshold of 7.20 and one a threshold of 7.10.^{55,69,78,85,93} Only one study

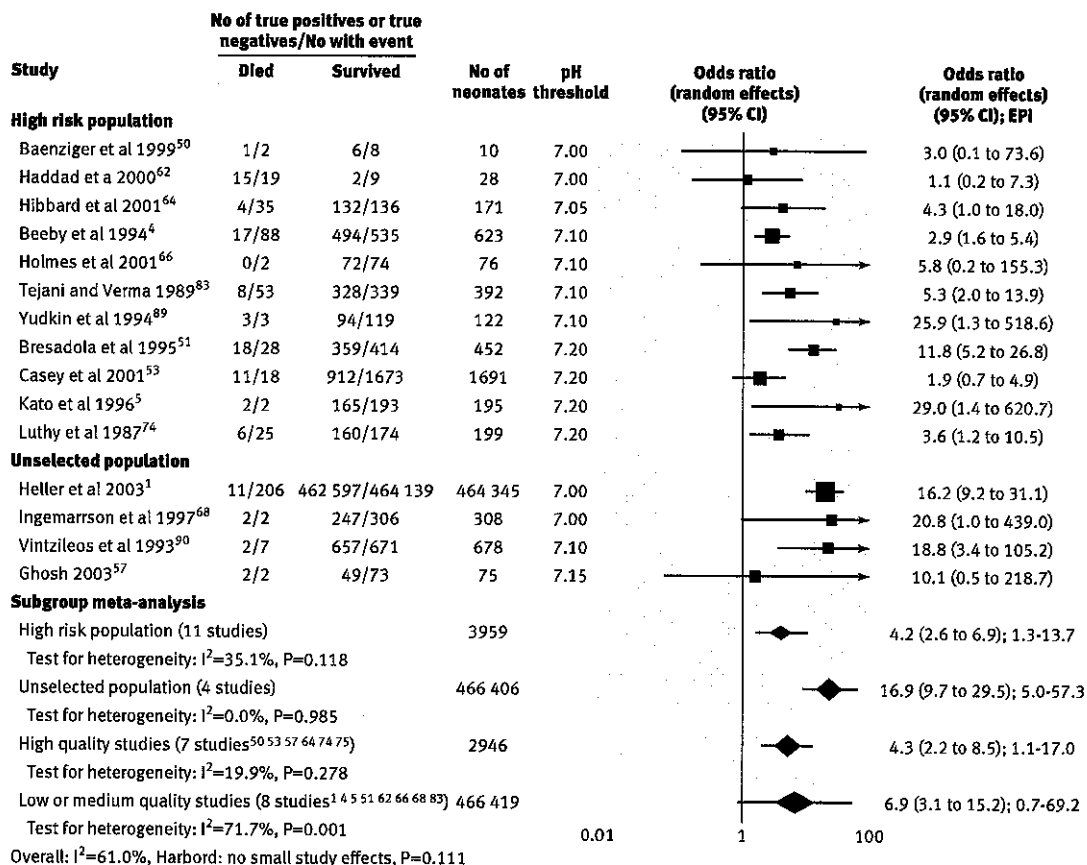
**Fig 2** | Association of low arterial cord pH with neonatal mortality. EPI=estimated predictive interval

Table 4 | Effect of using varying thresholds of arterial cord pH at birth on accuracy to predict neonatal morbidity and mortality

Variables	Odds ratio (95% CI)	Overall odds ratio (95% CI); estimated predictive interval
Neonatal mortality		
pH threshold 7.00:		
Baenziger et al 1999 ⁵⁰	3.00 (0.1 to 73.6)	6.1 (0.90 to 41.6); 0.0-20 406.6
Haddad et al 2000 ⁶²	1.07 (0.2 to 7.3)	
Heller et al 2003 ¹	16.9 (9.2 to 31.1)	
Ingemarrson et al 1997 ⁶⁸	20.8 (1.0 to 439.0)	
pH threshold 7.10:		
Beeby et al 1994 ⁴	2.9 (1.6 to 5.4)	7.1 (3.3 to 15.3); 0.8-64.3
Heller et al 2003 ¹	10.6 (7.4 to 15.1)	
Holmes et al 2001 ⁶⁶	5.8 (0.2 to 155.3)	
Vintzileos et al 1993 ⁹⁰	18.8 (3.4 to 105.2)	
Yudkin et al 1994 ⁸⁹	25.9 (1.3 to 518.6)	
pH threshold 7.20:		
Bresadola et al 1995 ⁵¹	11.8 (5.2 to 26.8)	4.3 (2.2 to 8.7); 0.5-40.7
Casey et al 2001 ⁵³	1.9 (0.7 to 4.9)	
Heller et al 2003 ¹	3.1 (2.3 to 4.1)	
Kato et al 1996 ⁵	29.0 (1.4 to 10.5)	
Luthy et al 1987 ⁷⁴	3.6 (1.2 to 10.5)	
Neonatal morbidity		
pH threshold 7.00:		
Baenziger et al 1999 ⁵⁰	2.7 (0.2 to 45.1)	12.5 (6.1 to 25.6); 1.7-89.9
Engle et al 1999 ⁵⁶	42.4 (2.3 to 782.1)	
Graham et al 2002 ⁶⁰	22.9 (1.1 to 494.6)	
Gonzalez de Dios 2000 ⁵⁹	9.2 (3.2 to 26.5)	
Haddad et al 2000 ⁶²	8.0 (0.3 to 255.8)	
Hernandez et al 1993 ⁶³	2.4 (0.8 to 7.4)	
Ingemarrson et al 1997 ⁶⁸	18.5 (3.8 to 89.6)	
Lama et al 2007 ⁷⁰	8.6 (1.1 to 69.7)	
Loh et al 1998 ⁷²	92.1 (3.7 to 2309.1)	
Perlman 1996 ⁷⁶	50.3 (2.7 to 955.6)	
Silva et al 2008 ⁸⁰	47.4 (2.2 to 1030.3)	
Socol 1994 ⁸¹	14.4 (0.7 to 311.8)	
Valentin et al 1993 ⁸⁵	1.8 (0.2 to 15.8)	
Van den Berg 1996 ¹⁰	10.0 (1.2 to 80.5)	
Gilstrap et al 1989 ⁵⁸	169.9 (22.5 to 1281.4)	
pH threshold 7.10:		
Beeby et al 1994 ⁴	3.9 (1.9 to 7.9)	2.4 (1.3 to 4.2); 0.4-12.7
Dijkshoorn et al 1986 ⁴¹	0.3 (0.04 to 13.4)	
Ertan et al 2006 ⁹¹	2.0 (0.9 to 4.5)	
Hernandez et al 1993 ⁶³	1.6 (0.6 to 4.1)	
Holmes et al 2001 ⁶⁶	3.1 (0.1 to 75.5)	
Huisjes 1979 ⁶⁷	4.8 (2.6 to 8.7)	
Silva et al 2008 ⁸⁰	14.7 (4.3 to 50.6)	
Tejani 1989 ⁸³	1.2 (0.3 to 4.3)	
Valentin et al 1993 ⁸⁵	0.8 (0.2 to 3.9)	
Salafia et al 1995 ⁷⁸	0.7 (0.04 to 13.4)	
pH threshold 7.20:		
Blackwell et al 2001 ⁵²	1.0 (0.3 to 4.2)	2.2 (1.3 to 3.7); 0.5-9.9
Casey et al 2001 ⁵³	3.1 (1.8 to 5.3)	
Dijkshoorn et al 1986 ⁴¹	1.4 (0.6 to 3.0)	
Hernandez et al 1993 ⁶³	1.6 (0.7 to 3.9)	
Luthy et al 1987 ⁷⁴	5.5 (1.9 to 15.7)	

utilised venous base excess.⁹³ The cord pH level was obtained before the occurrence of the adverse outcome. A wide variety of outcome measures were reported, ranging from neonatal mortality and morbidity to long term outcomes, including cerebral palsy, unspecified neurological abnormality, IQs, and developmental assessments (for example, Bayley and Griffiths scores). When the same outcome was reported, the thresholds and ages at assessment varied between studies, therefore meta-analysis was not possible for long term outcomes other than cerebral palsy. The odds ratio point estimate varied widely. Thirty six studies reported on outcomes within the neonatal period only. Fifteen studies had long term follow-up; one did not report the age of ascertainment of the outcome. The age range reported in other studies was 1 to 8 years; the median across studies was 5 years.

Quality assessment

Overall, study quality was variable (table 2). Most studies were retrospective but with a cohort design, allowing inferences on temporality of association. Over 80% of studies met the following quality items: appropriate outcome measure, more than 90% verification with outcome measure, and cohort design. Studies scored poorly for adequate description of index test and outcome measure, and consecutive recruitment. All included studies were unclear about the use of medical intervention between the index test and outcome measure. Overall, 45% (n=23) of studies were of high quality, 39% (n=20) of medium quality, and 16% (n=8) of low quality. Study design and quality did not seem to influence results on metaregression (table 3).

Association between cord pH and neonatal mortality

Fifteen studies (13 cohort, two case-control) totalling 469 395 infants reported on the association between arterial cord pH and mortality. The Harbord test was not significant, suggesting no influence of small study effect (see web extra on bmj.com). All studies had an odds ratio point estimate greater than 1.0. Heterogeneity was significant ($I^2=61.0\%$) overall. Metaregression (table 3) identified population risk as a significant explanatory factor. Within subgroups (fig 2) the association was consistent across studies. The association of low arterial cord pH with neonatal mortality was stronger in the unselected population (odds ratio 16.9, 95% confidence interval 9.7 to 29.5, estimated predictive interval 5.0-57.3, $I^2=0.0\%$) than in the high risk population (4.2, 2.6 to 6.9, 1.3-13.7, $I^2=35.1\%$). The analysis for the unselected population was dominated by one large study,¹ which was based on data from a regional perinatal register where cord pH was determined at most of the births. A sensitivity analysis excluding this study this did not significantly alter the point estimate (17.0, 4.4 to 65.5); however, the estimated predictive interval became wide (0.0-106 299.0).

Examination of the high risk population further showed that when studies reporting on a population of infants born at less than 32 weeks' gestation, or with a birth weight of less than 2000 g, were analysed

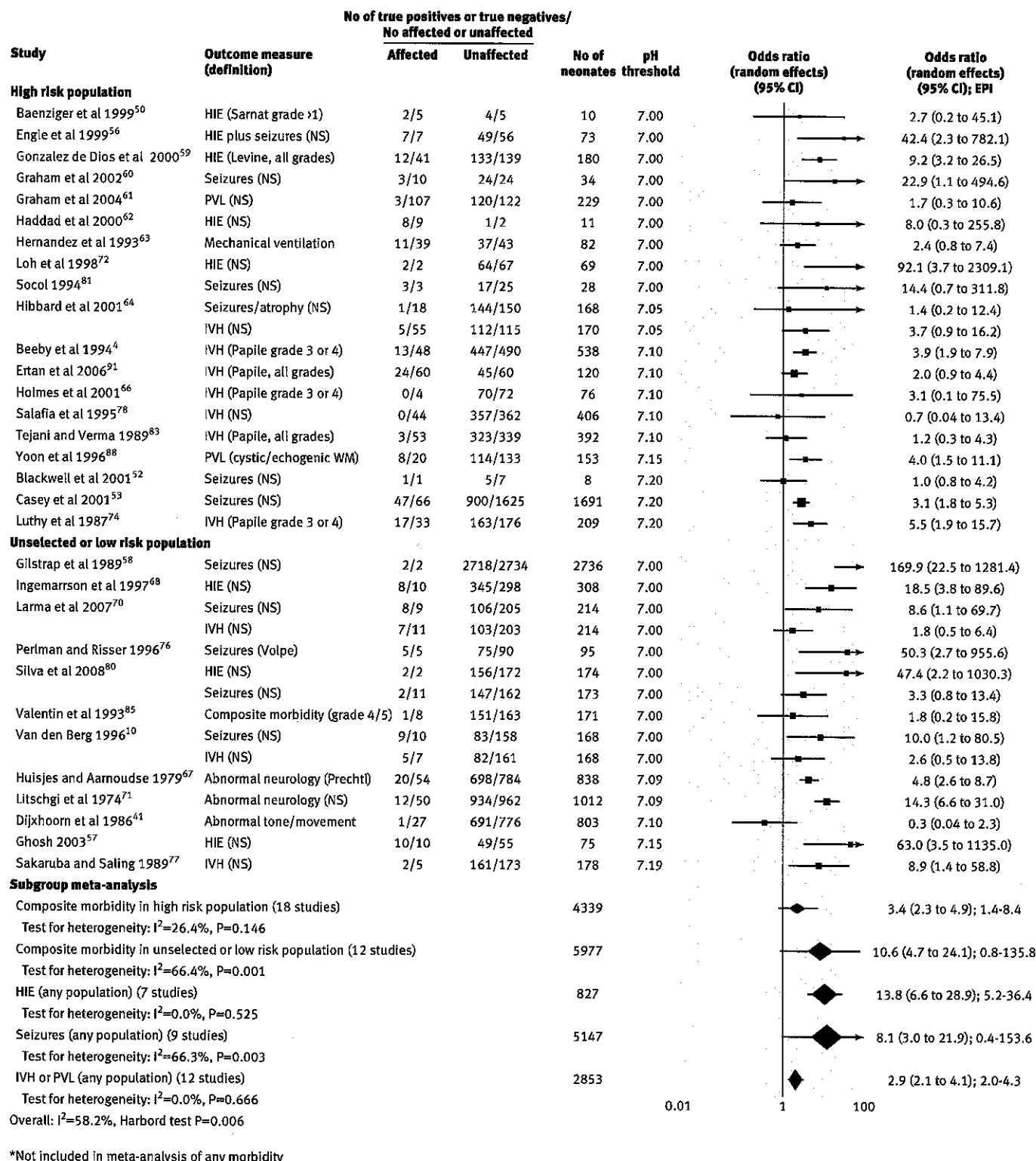


Fig 3 Association of low arterial cord pH with neonatal morbidity. HIE=hypoxic ischaemic encephalopathy; IVH=intraventricular haemorrhage; PVL=periventricular leucomalacia; NS=not stated; WM=cerebral white matter; EPI=estimated predictive interval

separately (seven studies), a significant association was found between cord pH and mortality (3.5, 2.3 to 5.4, $I^2=0.0\%$). When limited to a term (≥ 37 weeks' gestation) population (four studies), the association was also

strong (9.3, 1.4 to 63.2, $I^2=84.0\%$); however, heterogeneity was significant and the predictive interval broad (0.0-38 169.8). Grouping the studies according to quality did not affect the significance or direction of

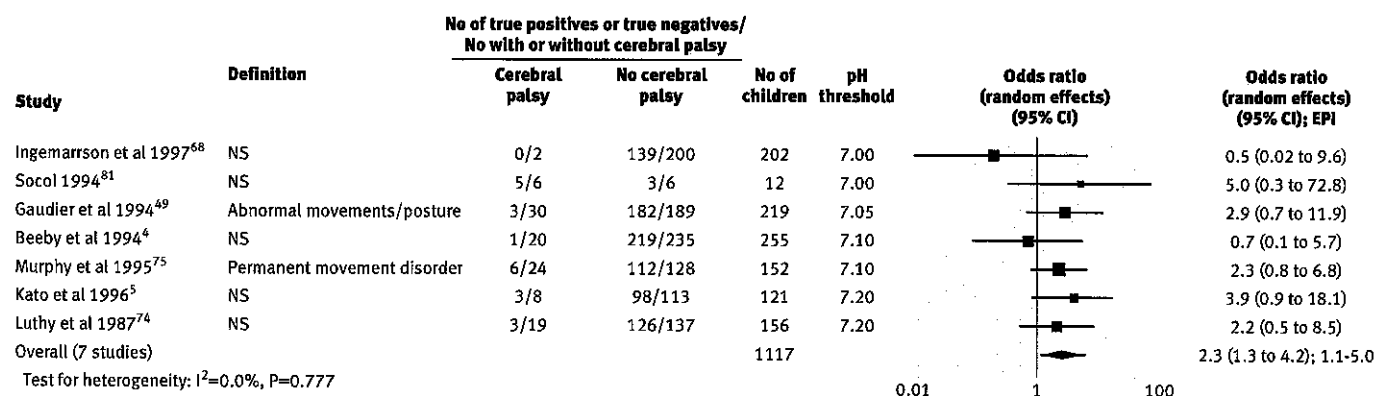


Fig 4 | Association of low arterial cord pH with cerebral palsy. NS=not stated; EPI=estimated predictive interval

association of the pooled result (fig 2). Exploration for a threshold effect (table 4) showed that for a cut-off pH threshold of 7.00 the association did not reach significance overall (6.1, 0.9 to 41.6) and the predictive interval was broad (0.0-20 406.6). The results for a threshold of 7.10 gave a similar point estimate but achieved significance (7.1, 3.3 to 15.3); however, the predictive interval remained broad (0.8-64.3) and crossed the line of no effect. For a threshold of 7.20, the odds ratio was lower (4.3, 2.2 to 8.7), with a broad predictive interval (0.5-40.6). Only one study examined all three thresholds,¹ with the strongest association at threshold 7.00 (16.9, 9.2 to 31.1) and the weakest at 7.20 (3.1, 2.3 to 4.1).

Association between cord pH and neonatal morbidity

Thirty one studies compared arterial cord pH with a variety of neonatal outcomes (fig 3). One of these⁶¹ was excluded from meta-analysis of composite neonatal morbidity because it contained a duplicate population to another included study,⁶⁰ therefore the dataset to assess the association of arterial pH with a composite measure of morbidity included 30 studies with 10 904 infants. The Harbord test was significant, suggesting the presence of a small study effect (see web extra on bmj.com). Only two studies had an odds ratio point estimate of less than 1.0, the rest showing an association. Significant heterogeneity was, however, present ($I^2=58.2\%$). Metaregression showed population to be an explanatory variable (table 3). Subgroup meta-analysis for a high risk population (fig 3) showed a weaker association (3.4, 2.3 to 4.9, 1.4-8.4, $I^2=26.4\%$) than in an unselected or undefined population (10.6, 4.7 to 24.1, 0.8-135.8, $I^2=66.4\%$). When analysed in subgroups according to quality, the direction of effect remained consistent and significant between the group of 12 high quality studies (6.6, 3.7 to 11.8, 1.3-32.6, $I^2=51.2\%$) and 18 low or medium quality studies (4.6, 2.6 to 8.0, 0.7-29.2, $I^2=61.5\%$). Exploration by subgroup analysis for a threshold effect (table 4) showed the most substantial association at a pH threshold of 7.00 (12.5, 6.1 to 25.6). The predictive interval was broad but did not cross the line of no effect (1.7-89.9). The results were similar for a

threshold of 7.10 (2.4, 1.3 to 4.3) and 7.20 (2.2, 1.3 to 3.7).

When components of the composite outcome were analysed (fig 3), the odds ratio for an association between arterial cord pH and hypoxic ischaemic encephalopathy was 13.8 (6.6 to 28.9, 5.2-36.4, $I^2=0.0\%$), between arterial cord pH and seizures was 8.1 (3.0 to 21.9, 0.4-153.6, $I^2=66.3\%$), and between intraventricular haemorrhage or periventricular leucomalacia was 2.9 (2.1 to 4.1, 2.0-4.3, $I^2=0.0\%$). Only two of nine studies reporting this outcome were not limited to a preterm (<32 weeks) or low birthweight (<2000 g) population. Excluding these two studies did not affect the strength or significance of the association.

Five studies reported on the association between venous cord pH and neonatal morbidity.^{55 69 78 85 93} Meta-analysis showed a significant association (4.0, 1.2 to 13.3, $I^2=44.5\%$). However the predictive interval crossed the line of no effect (0.1-125.3). Four studies examined the association between arterial base excess and neonatal morbidity, which was similar to venous cord pH and also had an estimated predictive interval that crossed the line of no effect (2.5, 1.3 to 4.8, 0.6-10.4, $I^2=0.0\%$).^{61 73 81 91} Owing to the small number of studies reporting this variable, it was not possible to combine the cord pH and base excess to compare the difference between respiratory and metabolic acidosis.

Association between cord pH and cerebral palsy

Seven studies totalling 1117 infants examined the association between arterial cord pH and cerebral palsy.^{45 49 68 74 75 81} Of those, two had an odds ratio point estimate of less than 1.0, but the rest showed an association of low pH with cerebral palsy (table 4). Only two studies gave the criteria used to diagnose cerebral palsy.^{49 75} The overall association was odds ratio 2.3 (1.3 to 4.2, 1.1-5.0, $I^2=0.0\%$) (fig 4). It was not possible to explore the threshold effect by subgroup analysis owing to the small number of studies reporting this outcome, therefore bivariate meta-analysis was carried out, which did not change the overall results (2.1, 1.2 to 4.1).

DISCUSSION

Low arterial umbilical cord pH had a strong, consistent, and temporal association with neonatal mortality and morbidity (composite of hypoxic ischaemic encephalopathy, seizures, and intraventricular haemorrhage or periventricular leucomalacia) and long term outcome (cerebral palsy). In all of the associations between arterial cord pH and outcome explored, with the exception of composite morbidity in a low risk population and seizures, the estimated predictive interval suggested that a future study would have the same direction and significance of effect observed. The associations observed are biologically plausible.⁶⁹⁻⁹⁶

Strengths and limitations of the review

This review provides the best available evidence, at the time of publication, of the association between cord pH at birth and outcomes. The strength of our review and the validity of our inferences lie in the methods used. We complied with existing guidelines for the reporting of systematic reviews of diagnostic^{22,97} and observational studies evaluating causal association.^{15,25} We used the most up to date techniques for carrying out and interpreting meta-analysis.^{35,44} An extensive literature search was done in relevant databases with no language restrictions applied. The studies we pooled had heterogeneity in terms of quality, population risk, threshold of cord pH used, and ascertainment of neonatal outcome. We carried out recommended analyses to tackle this problem, including bivariate and subgroup meta-analysis to take into account the threshold effect,⁴⁴ metaregression analysis to explore reasons for heterogeneity,³⁹ and component outcome analysis to examine the suitability of the composite morbidity outcome. This did not significantly affect our results. Recent work suggesting the differences between standard and bivariate techniques may be small.⁹⁸ The observed associations were qualitatively in the same direction, and statistical heterogeneity, when present, arose from variation in strength of associations from study to study rather than opposition in direction of association. We are therefore confident that our inference concerning a causal association between low arterial cord pH at birth and neonatal mortality and a variety of neonatal morbidities merits consideration.

This review has several limitations. Despite our efforts to search for published and unpublished literature, the Harbord test suggested that the meta-analysis of the relation of arterial cord pH and neonatal morbidity may be affected by small study bias. The quality of the primary studies varied. We used both meta-regression and subgroup analysis to explore for the effect of this problem, which showed that study quality did not affect our inferences. The poor reporting of population characteristics limited the subgroup analysis according to risk factors: a large number of the papers with an "unselected" population did not fully report characteristics such as birth weight or gestational age, making it difficult to extrapolate our findings to the general obstetric population. Only one paper included in the meta-analysis specified that it

was limited to a low risk term population. With regard to the index test examined, only a small number of papers reported base excess, which meant that it was not possible to compare metabolic acidosis with respiratory acidosis. We found that a high base excess was associated with neonatal morbidity, but in clinical practice the pH and base excess level would be considered together in any infant, and we are unable to comment on this. Additionally, we were only able to base our analysis on the thresholds reported in the primary studies, which limited our exploration of the effect of varying pH levels. Some of these issues may be dealt with using meta-analysis of individual patient data.⁹⁹

Venous cord pH and arterial base excess showed weaker associations with neonatal morbidity than that of arterial cord pH, with estimated predictive intervals that crossed the line of no effect. This analysis was, however, based on a small number of studies, and as no other subgroup analysis could be done, a direct comparison with arterial cord pH was not possible. It seems likely from our results that arterial cord pH has a stronger association with outcome than venous cord pH.

Evidence supporting causality of association of low arterial cord pH with adverse outcomes

Hypoxic ischaemia initiates energy depletion, the accumulation of extracellular glutamate, and activation of receptors, leading to a deleterious cascade of events resulting in neuronal death.¹⁰⁰ However, different areas of the brain are susceptible to injury at various stages in development, and the consequences of injury are unpredictable.¹⁰¹ This would support our findings of a strong association with hypoxic ischaemic encephalopathy but a weaker association with cerebral palsy. Only 10% of infants with evidence of hypoxic ischaemic encephalopathy develop cerebral palsy, but the reasons for this are not yet fully understood.¹⁰² This issue is further highlighted by the findings of the term breech trial, where long term follow-up did not show any difference in neurodevelopmental delay, despite an apparent increase in neonatal morbidity in the vaginal breech group.¹⁰³

It is difficult to comment on the specificity with which cord pH is associated with the outcomes examined within our review. All of the outcomes may arise from a variety of causes—for example, in addition to hypoxia, neonatal seizures may be caused by congenital brain anomaly, infection, and metabolic disorders. Although some of the included studies excluded infants with congenital anomaly, others did not specify whether they had excluded from their analysis infants with an outcome related to another cause. This may explain why the strength of association of arterial cord pH with neonatal mortality and morbidity was stronger in the unselected group than in the high risk group. The infants in the high risk group had other factors (for example, prematurity, low birth weight) that may have contributed to the adverse outcome, thus diluting the strength of association. This is supported by the higher number of deaths in the high

WHAT IS ALREADY KNOWN ON THIS TOPIC

Umbilical cord pH at birth is frequently used to measure perinatal asphyxia

Neonatal and childhood mortality and morbidity, including cerebral palsy, are often attributed to fetal acidosis, as defined by a low cord pH at birth

Existing reports of the association between cord pH and adverse outcome are conflicting

WHAT THIS STUDY ADDS

Low cord pH is substantially associated with neonatal mortality and morbidity and cerebral palsy in childhood

These outcomes justify the increased surveillance of infants born with a low cord pH

Further research is, however, needed to explore the cost effectiveness of doing this test in all neonates

risk population, with a higher specificity for low arterial cord pH and death in the unselected population.

We explored for the presence of a dose-response relation by examining the strength of association between arterial cord pH and neonatal mortality and morbidity at different pH thresholds. Although the meta-analysis did not show a clear dose-response relation, and the estimated predictive intervals were broad, the association for both neonatal mortality and neonatal morbidity was weakest at the highest threshold. Only one study¹ explored the association of arterial cord pH and neonatal mortality at all three thresholds examined. A clear dose-response relation was apparent within this study, with the strongest association at a threshold of 7.00 and the weakest at a threshold of 7.20. Comparison within a study may be more valid than comparisons between studies as it avoids confounding by other study level factors.

Conclusions and practice implications

Our review did not explore the use of cord pH as a prognostic test, and therefore the extent to which our results can be used to counsel parents and target interventions for infants with a low cord pH at birth is limited. To explore this matter, further research is required. A recent review highlighted the current deficits in prognosis research and suggested that when high quality primary studies exist, meta-analysis of individual patient data enables the prognostic value of a test to be assessed at an individual level.¹⁰⁴ Our review, however, highlights the lack of high quality primary studies, and therefore a large prospective cohort study with long term follow-up and accounting for potential confounding factors is required. Such a study must include evaluation of outcomes relevant to the individual and society, such as impaired development and use of healthcare and educational resources, before the evaluation of use of cord pH as a prognostic test and its cost effectiveness. This will enable further exploration of the threshold effect and the use of combined cord pH and base excess to predict outcome.

Cord pH is currently assessed in infants believed to be at high risk for neonatal asphyxia. Our results suggest, however, that the strength of association with cord

pH and outcome is not limited to this high risk population. Therefore future research should assess the use of cord pH across neonatal populations, particularly exploring the cost effectiveness of testing all neonates. Systematic reviews to establish the association of other neonatal wellbeing measures with long term outcomes are ongoing in our department.

Based on our review, increased initial surveillance of neonates born with a low arterial cord pH, regardless of their clinical condition, is warranted as the odds of complications have been shown to be higher in this group. The avoidance of a low cord pH at birth should continue to be a target for day to day obstetric practice. Our findings justify the use of arterial cord pH as an important outcome measure alongside neonatal morbidity and mortality in obstetric clinical trials. It is difficult to draw strong conclusions on the necessity of long term follow-up for babies with low arterial cord pH, as the observed association with cerebral palsy was in a limited number of primary studies and the association, although statistically significant, was only moderately strong and based on studies of varying quality. We therefore support the use of long term follow-up with childhood assessments in obstetric trials, as this may show developmental anomalies and allow evaluation of cord pH as a predictive test for these outcomes.

Zainab Laftah helped with data extraction. Richard Riley provided statistical expertise and advice.

Contributors: GLM designed the review; extracted, analysed, and interpreted the data, drafted the article; and is responsible for the integrity of the work. She is the guarantor. RKM extracted and interpreted the data, revised the article critically for intellectual content, and approved the final draft for publication. KSK conceived the review, helped analyse and interpret the data, revised the article critically for intellectual content, and approved the final draft for publication.

Funding: GLM is funded by the Mary Crosse fellowship, Birmingham Women's Foundation Trust. RKM is funded by a Medical Research Council/Royal College of Obstetrics and Gynaecology clinical research training fellowship (RRAK12923).

Competing interests: All authors have completed the unified competing interest form at www.icmje.org/col_disclosure.pdf (available on request from the corresponding author) and declare (1) no financial support for the submitted work from anyone other than their employer; (2) no financial relationships with commercial entities that might have an interest in the submitted work; (3) no spouses, partners, or children with relationships with commercial entities that might have an interest in the submitted work; and (4) no non-financial interests that may be relevant to the submitted work.

Ethical approval: Not required.

Data sharing: The odds ratios for outcomes included in table 1 but not reported in the paper are available from the author on request.

- Heller G, Schnell RR, Misselwitz B, Schmidt S. Umbilical blood pH, Apgar scores, and early neonatal mortality. *Z Geburtshilfe Neonatol* 2003;207:84-9.
- Goodwin TM, Belai I, Hernandez P, Durand M, Paul RH. Asphyxial complications in the term newborn with severe umbilical acidemia. *Am J Obstet Gynecol* 1992;167:1506-12.
- Williams KP, Singh A. The correlation of seizures in newborn infants with significant acidosis at birth with umbilical artery cord gas values. *Obstet Gynecol* 2002;100:557-60.
- Beeby PJ, Elliott EJ, Henderson-Smart DJ, Rieger ID. Predictive value of umbilical artery pH in preterm infants. *Arch Dis Child* 1994;71:F93-6.
- Kato EH, Yamada H, Matsumoto Y, Hattori S, Makinoda S, Fujimoto S. Relation between perinatal factors and outcome of very low birth weight infants. *J Perinat Med* 1996;24:677-86.
- Fahy J, King TL. Intrauterine asphyxia: clinical implications for providers of intrapartum care. *J Midwifery Women's Health* 2005;50:498-506.

- 7 MacLennan A. A template for defining a causal relation between acute intrapartum events and cerebral palsy: international consensus statement. *BMJ* 1999;319:1054-9.
- 8 Dear P, Newell S, Rosenbloom L, Rennie JM, MacLennan A. Establishing probable cause in cerebral palsy [letter]. *BMJ* 2000;320:1075.
- 9 Dijkshoorn MJ, Visser GH, Huisjes HJ, Fidler V, Touwen BC. The relation between umbilical pH values and neonatal neurological morbidity in full term appropriate-for-dates infants. *Early Hum Dev* 1985;11:33-42.
- 10 Van den Berg PP, Nelen WL, Jongsma HW, Nijland R, Kollée LA, Nijhuis JG, et al. Neonatal complications in newborns with an umbilical artery pH<7.00. *Am J Obstet Gynecol* 1996;175:1152-7.
- 11 Wildschut J, Feron FJ, Hendriksen JG, van Hall M, Gavilanes-Jimenez DW, Hadders-Algra M, et al. Acid-base status at birth, spontaneous motor behaviour at term and 3 months and neurodevelopmental outcome at age 4 years in full-term infants. *Early Hum Dev* 2005;81:535-44.
- 12 Goldaber KG, Gilstrap LC 3rd, Leveno KJ, Dax JS, McIntire DD. Pathologic fetal acidemia. *Obstet Gynecol* 1991;78:1103-7.
- 13 Ross MG, Gala R. Use of umbilical artery base excess: algorithm for the timing of hypoxic injury. *Am J Obstet Gynecol* 2002;187:1-9.
- 14 Van de Riet JE, Vandenbussche FP, Le Cessie S, Keirse MJ. Newborn assessment and long-term adverse outcome: a systematic review. *Am J Obstet Gynecol* 1999;180:1024-9.
- 15 Stroup DF, Berlin JA, Morton SC, Olkin I. Meta-analysis of observational studies in epidemiology. *JAMA* 2000;283:2008-12.
- 16 Shennan AH, Smith R, Browne D, Edmonds DK, Morgan B. The elective use of oxytocin infusion during labour in nulliparous women using epidural analgesia: a randomised double-blind placebo-controlled trial. *Int J Obstet Anesth* 1995;4:78-81.
- 17 Cluett ER, Pickering RM, Getliffe K, St George Saunders NJ. Randomised controlled trial of labouring in water compared with standard of augmentation for management of dystocia in first stage of labour. *BMJ* 2004;328:314-8.
- 18 Hansen SL, Clark SL, Foster JC. Active pushing versus passive fetal descent in the second stage of labor: a randomized controlled trial. *Obstet Gynecol* 2002;99:29-34.
- 19 Cottee C, Harding K. Risk management in obstetrics. *Obstet, Gynaecol Reprod Med* 2008;18:155-62.
- 20 Milkhu v NorthWest hospitals NHS Trust 2003 EWHC 94 (QB). http://www.lexisnexis.com/uk/legal/search/runRemoteLink.do?service=citator&csi=279841&remotekey2=%5B2003%5D+All+ER+%280%29+333+%28Jan%29&remotekey1=REPORT-CITATION&risb=21_T8890416457&citatorCC=GB.
- 21 Hill AB. The environment and disease: association or causation? *Proc R Soc Med* 1965;58:295-300.
- 22 Khan KS, Dinnes J, Kleijnen J. Systematic reviews to evaluate diagnostic tests. *Eur J Obstet Gynecol Reprod Biol* 2001;95:6-11.
- 23 Akers J, Aguiar-Ibanez R, Sari ABA, Benyon S, Booth A, Burch J, et al. Systematic reviews. CRD's guidance for undertaking reviews in health care. Centre for Reviews and Dissemination, University of York, 2009.
- 24 Deville W, Buntinx F. Guidelines for conducting systematic reviews of studies evaluating the accuracy of diagnostic studies. In: Knotterus JA, ed. The evidence base of clinical diagnosis. BMJ Publishing Group, 2002:145-65.
- 25 Weed DL. Interpreting epidemiological evidence: how meta-analysis and causal inference methods are related. *Int J Epidemiol* 2000;29:387-90.
- 26 Weed DL. On the use of causal criteria. *Int J Epidemiol* 1997;26:1137-41.
- 27 Weed DL, Hursting SD. Biological plausibility in causal inference: current method and practice. *Am J Epidemiol* 1998;147:415-25.
- 28 Fox C, Mignini L, Khan KS. Systematic reviews of research to assess causation: a guide to methods and application. *Eur Clinics Obstet Gynecol* 2006;1:251-6.
- 29 Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003;326:41-4.
- 30 Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
- 31 Whiting P, Harbord R, Kleijnen J. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Med Res Methodol* 2005;5:19.
- 32 Juni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. *BMJ* 2001;323:42-6.
- 33 Deeks J. Systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 2001;323:157-62.
- 34 Greenland S. Interpretation and choice of effect measures in epidemiologic analyses. *Am J Epidemiol* 1987;125:761-8.
- 35 Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Statist Soc A* 2009;172:137-59.
- 36 Sankey S, Weissfeld L, Fine M, Kapoor W. An assessment of the use of the continuity correction for sparse data in meta-analysis. *Commun Stat Simul Comput* 1996;25:1031-56.
- 37 Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C. Composite outcomes in randomized controlled trials: greater precision but with greater uncertainty? *JAMA* 2003;289:2554-9.
- 38 Higgins JP, Thompson SG, Deeks JJ. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557-60.
- 39 Lijmer JG, Bossuyt PM, Heisterkamp SH. Exploring sources of heterogeneity in systematic reviews of diagnostic tests. *Stat Med* 2002;21:1525-37.
- 40 Sterne JA, Egger M, Smith GD. Systematic reviews in health care: investigating and dealing with publication and other biases in meta-analysis. *BMJ* 2001;323:101-5.
- 41 Dijkshoorn MJ, Visser GH, Fidler V, Touwen BC, Huisjes HJ. Apgar score, meconium and acidemia at birth in relation to neonatal neurological morbidity in term infants. *BJOG* 1986;93:217-22.
- 42 Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions. Cochrane Collaboration, 2009.
- 43 Harbord RM, Egger M, Sterne JA. A modified test for small study effects in meta-analysis of controlled trials with binary end points. *Stat Med* 2006;25:3443-57.
- 44 Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate meta-analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58:982-90.
- 45 Zamora J, Abraira V, Muriel A, Khan KS, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol* 2003;12:31.
- 46 Harbord RM, Harris RJ, Sterne JAC, Steichen T. METABIAS: Stata module to test for small-study effects in meta-analysis. Statistical Software Components S404901, Boston College Department of Economics, 2009.
- 47 Harris R, Bradburn M, Deeks J, Harbord R, Altman D, Steichen T, et al. METAN: Stata module for fixed and random effects meta-analysis. Statistical Software Components S456798, Boston College Department of Economics, 2009.
- 48 Harbord R. METANDI: Stata module to perform meta-analysis of diagnostic accuracy. Statistical Software Components S456932, Boston College Department of Economics, 2008.
- 49 Gaudier FL, Goldenberg RL, Nelson KG, Peralta-Carcelen M, Johnson SE, DuBard MB, et al. Acid-base status at birth and subsequent neurosensory impairment in surviving 500 to 1000 gm infants. *Am J Obstet Gynecol* 1994;170:48-53.
- 50 Baenziger O, Moenkhoff M, Morales CG, Waldvogel K, Wolf M, Bucher H, et al. Impaired chemical coupling of cerebral blood flow is compatible with intact neurological outcome in neonates with perinatal risk factors. *Biol Neonate* 1999;75:9-17.
- 51 Bresadola M, Lo Mastro M, Arena V, Bellavaglia L, Scarpellini F. Preterm labour and neonatal parameters. *Clin Exp Obstet Gynecol* 1995;3:235-9.
- 52 Blackwell SC, Moldenhauer J, Hassan SS, Redman ME, Refuerzo JS, Berry SM, et al. Meconium aspiration syndrome in term neonates with normal acid-base status at delivery: is it different? *Am J Obstet Gynecol* 2001;184:1422-5.
- 53 Casey BM, Goldaber KG, McIntire DD. Outcomes amongst term infants when two-hour postnatal pH is compared with pH at delivery. *Am J Obstet Gynecol* 2001;184:447-50.
- 54 Dennis J, Johnson A, Mutch L, Yudkin P, Johnson P. Acid-base status at birth and neurodevelopmental outcome at four and one-half years. *Am J Obstet Gynecol* 1989;161:213-20.
- 55 D'Souza SW, Black P, Cadman J, Richards B. Umbilical venous blood pH: a useful aid in the diagnosis of asphyxia at birth. *Arch Dis Child* 1983;58:15-9.
- 56 Engle D, Laptook AR, Perlman J. Acute changes in arterial carbon dioxide tension and acid-base status and early neurologic characteristics in term infants following perinatal asphyxia. *Resuscitation* 1999;42:11-7.
- 57 Ghosh BM. Prediction of perinatal asphyxia with nucleated red blood cells in cord blood of newborns. *Int J Gynecol Obstet* 2003;81:267-71.
- 58 Gilstrap LC 3rd, Leveno KJ, Burris J, Williams ML, Little BB. Diagnosis of birth asphyxia on the basis of fetal pH, Apgar score, and newborn cerebral dysfunction. *Am J Obstet Gynecol* 1989;161:825-30.
- 59 Gonzalez de Dios J, Moya M, Carratala F. Neurological evolution of asphyctic full-term newborns with severe umbilical acidosis (pHUA <7.00). *Rev Neurol* 2000;31:107-13.
- 60 Graham EM, Holcroft CJ, Blakemore KJ. Evidence of intrapartum hypoxia-ischemia is not present in the majority of cases of neonatal seizures. *J Matern Fetal Neonatal Med* 2002;12:123-6.
- 61 Graham EM, Holcroft CJ, Karishma KR, Donohue PK, Allen MC. Neonatal cerebral white matter injury in preterm infants is associated with culture positive infections and only rarely with metabolic acidosis. *Am J Obstet Gynecol* 2004;191:1305-10.

- 62 Haddad B, Mercer BM, Livingston JC, Talati A, Sibai BM. Outcome after successful resuscitation of babies born with Apgar scores of 0 at both 1 and 5 minutes. *Am J Obstet Gynecol* 2000;182:1210-4.
- 63 Hernandez C, Little BB, Dax JS, Gilstrap LC 3rd, Rosenfeld CR. Prediction of the severity of meconium aspiration syndrome. *Am J Obstet Gynecol* 1993;169:61-70.
- 64 Hibbard JU, Hibbard MC, Whalen MP. Umbilical cord blood gases and mortality and morbidity in the very low birth weight infant. *Obstet Gynecol* 1991;78:768-73.
- 65 Hogan LI, Ingemarsson I, Thorngren-Jemeck K, Herbst A. How often is a low 5-min Apgar score in term newborns due to asphyxia? *Eur J Obstet Gynecol Reprod Biol* 2007;130:169-75.
- 66 Holmes P, Oppenheimer LW, Gravelle A, Walker M, Blayney M. The effect of variable heart rate decelerations on intraventricular hemorrhage and other perinatal outcomes in preterm infants. *J Matern Fetal Med* 2001;10:264-8.
- 67 Huisjes HJ, Aamoudse JG. Arterial or venous umbilical pH as a measure of neonatal morbidity? *Early Hum Dev* 1979;3:155-61.
- 68 Ingemarsson I, Herbst A, Thorngren-Jemeck K. Long term outcome after umbilical artery acidemia at term birth: influence of gender and duration of fetal heart rate abnormalities. *BJOG* 1997;104:1123-7.
- 69 Jurgens-van der Zee AD, Bierman-van Eendenburg MEC, Fidler V, Olinga AA, Visch JH, Touwen BC, et al. Preterm birth, growth retardation and acidemia in relation to neurological abnormality of the newborn. *Early Hum Dev* 1979;32:141-54.
- 70 Larma JD, Silva AM, Holcroft CJ, Thompson RE, Donohue PK, Graham EM. Intrapartum electronic fetal heart rate monitoring and the identification of metabolic acidosis and hypoxic-ischemic encephalopathy. *Am J Obstet Gynecol* 2007;197:301-8.
- 71 Litschgi M, Benz JJ, Glatthaar E. Actual and prognostic value of arterial cord pH for the newborn infant. *Z Geburtshilfe Perinatol* 1974;178:23-9.
- 72 Loh SF, Woodworth A, Yeo GS. Umbilical cord blood gas analysis at delivery. *Singapore Med J* 1998;39:151-5.
- 73 Low JA. Threshold of metabolic acidosis associated with newborn complications. *Am J Obstet Gynecol* 1997;177:1391-4.
- 74 Luthy DA, Shy KK, Strickland D, Wilson J, Bennett FC, Brown ZA, et al. Status of infants at birth and risk for adverse neonatal events and long-term sequelae: a study in low birth weight infants. *Am J Obstet Gynecol* 1987;157:676-9.
- 75 Murphy DJ, Sellers S, MacKenzie IZ, Yudkin P, Johnson A. Case-control study of antenatal and intrapartum risk factors for cerebral palsy in very preterm singleton babies. *Lancet* 1995;346:1449-54.
- 76 Perlman JM, Risser R. Can asphyxiated infants at risk for neonatal seizures be rapidly identified by current high-risk markers? *Pediatrics* 1996;97:456-62.
- 77 Sakuraba M, Saling E. Umbilical cord blood coagulability, acidosis and intracranial hemorrhage. *J Perinat Med* 1989;17:99-106.
- 78 Salafia CM, Minior VK, Rosenkrantz TS, Pezzullo JC, Popek EJ, Cusick W, et al. Maternal, placental, and neonatal associations with early germinal matrix/intraventricular hemorrhage in infants born before 32 weeks' gestation. *Am J Perinatol* 1995;12:429-36.
- 79 Schneider R, Tanner R. Perinatal umbilical artery pH and cerebral function disorders in twins starting school. *Z Kinder Jugendpsychiatr* 1985;13:24-30.
- 80 Silva AM, Cootauco AC, Aina-Mumuney A, Donohue PK, Graham EM. The association of hypotonia and depression in the term and near-term neonate with metabolic acidemia. *J Perinat Med* 2008;36:151-6.
- 81 Socol ML. Depressed Apgar score, acid-base balance and neurologic outcome. *Am J Obstet Gynecol* 1994;170:991-9.
- 82 Spinillo A, Fazzi E, Orcesi S. Perinatal factors and 2 year minor neurodevelopmental impairment in low birth weight infants. *Biol Neonate* 1995;67:39-46.
- 83 Tejani N, Verma U. Correlation of Apgar scores and umbilical artery acid-base status to mortality and morbidity in the low birth weight neonate. *Obstet Gynecol* 1989;73:597-600.
- 84 Thoulon JM, Vamier C, Faure M. Prognostic value of umbilical blood pH measurement in newborn infants at birth. *Lyon Med* 1972;227:699-702.
- 85 Valentin L, Ekman G, Isberg PE, Polberger S, Marsál K. Clinical evaluation of the fetus and neonate. Relation between intra-partum cardiotocography, Apgar score, cord blood acid-base status and neonatal morbidity. *Arch Gynecol Obstet* 1993;253:103-15.
- 86 Winkler CL, Hauth JC, Tucker JM, Owen J, Brumfield CG. Neonatal complications at term as related to the degree of umbilical artery acidemia. *Am J Obstet Gynecol* 1991;164:637-41.
- 87 Wu L, Thorngren-Jemeck K, Ingemarsson I. Different types of acidemia at birth, fetal heart rate patterns and infants outcome at four years of age. *Chung-Hua Fu Chan Ko Tsa Chih* 1998;33:462-5.
- 88 Yoon BH, Romero R, Yang SH, Jun JK, Kim IO, Choi JH, et al. Interleukin-6 concentrations in umbilical cord plasma are elevated in neonates with white matter lesions associated with periventricular leukomalacia. *Am J Obstet Gynecol* 1996;174:1433-40.
- 89 Yudkin P, Johnson A, Clover LM, Murphy KW. Clustering of perinatal markers of birth asphyxia and outcome age 5 years. *BJOG* 1994;101:774-81.
- 90 Vintzileos AM, Antsaklis A, Varvarigos I, Papas C, Sofatzis I, Montgomery JT. A randomized trial of intrapartum electronic fetal heart rate monitoring versus intermittent auscultation. *Obstet Gynecol* 1993;81:899-907.
- 91 Ertan AK, Tanriverdi HA, Meier M, Schmidt W. Perinatal risk factors for neonatal intracerebral hemorrhage in preterm infants. *Eur J Obstet Gynecol Reprod Biol* 2006;127:29-34.
- 92 Svirkó E, Mellanby J, Impey L. The association between cord pH at birth and intellectual function in childhood. *Early Hum Dev* 2008;84:37-41.
- 93 Gea YA. Clinical value of lactate measurement and nucleated red blood cell counts in the placental segment of the umbilical vein of premature newborns for diagnosis of hypoxia-ischemia. *J Pediatr (Rio J)* 2007;83:186-90.
- 94 Thorp JA, Rushing R. Umbilical cord blood gas analysis. *Obstet Gynecol Clin North Am* 1999;26:695-709.
- 95 Ferriero DM. Neonatal brain injury. *N Engl J Med* 2004;351:1985-95.
- 96 Johnston MV, Trescher WH, Ishida A, Nakajima W. Neurobiology of hypoxic ischemic injury in the developing brain. *J Child Neurol* 2001;15:588-91.
- 97 Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B What are the results and will they help me in caring for my patients? *JAMA* 1994;271:703-7.
- 98 Simel D, Bossuyt PM. Differences between univariate and bivariate models for summarizing diagnostic accuracy may not be large. *J Clin Epidemiol* 2009;62:1292-300.
- 99 Khan KS, Bachmann LM, ter Riet G. Systematic reviews with individual patient data meta-analysis to evaluate diagnostic tests. *Eur J Obstet Gynecol Reprod Biol* 2003;108:121-5.
- 100 Volpe JJ. Perinatal brain injury: from pathogenesis to neuroprotection. *Ment Retard Dev Disabil Res Rev* 2001;7:56-64.
- 101 Hossain MA. Hypoxic ischemic injury in neonatal brain: involvement of a novel neuronal molecule in neuronal cell death and potential target for neuroprotection. *Int J Dev Neurosci* 2008;26:93-101.
- 102 Blumenthal I. Cerebral palsy—medicolegal aspects. *J R Soc Med* 2001;94:624-6.
- 103 Whyte H, Hannah ME, Saigal S, Hannah WJ, Hewson S, Amankwah K, et al. Outcomes of children at 2 years after planned cesarean birth versus planned vaginal birth for breech presentation at term: the international randomized term breech trial. *Am J Obstet Gynecol* 2004;191:864-71.
- 104 Hemingway H, Riley RD, Altman D. Ten steps towards improving prognosis research. *BMJ* 2009;339:b4184.

Accepted: 14 February 2010