The effect of waterbirth on neonatal mortality and morbidity: a systematic review and meta-analysis

Rowena Davies[1,2]

Deborah Davis[1,2]

Melissa Pearce[2,3]

Nola Wong[2,3]

1 Nursing and Midwifery, Faculty of Health, University of Canberra, Australia
2 The Australian Capital Regional Centre for Evidence Based Nursing and Midwifery Practice: an Affiliate Center of the Joanna Briggs Institute
3 Centenary Hospital for Women and Children, Canberra, Australia

Corresponding author:

Rowena Davies

u3053358@uni.canberra.edu.au

Executive summary

Background

Women have been giving birth in water in many centers across the globe; however, the practice remains controversial. Qualitative studies highlight the benefits that waterbirth confers on the laboring woman, though due to the nature of the intervention, it is not surprising that there are few randomized controlled trials available to inform practice. Much of the criticism directed at waterbirth focuses on the potential impact on the neonate.

Objectives

The objective of this review was to systematically synthesize the best available evidence regarding the effect of waterbirth, compared to landbirth, on the mortality and morbidity of neonates born to low risk women.

Inclusion criteria

Types of participants

This review considered studies that included low risk, well, pregnant women who labor and birth spontaneously, at term (37-42 weeks), with a single baby in a cephalic presentation. Low risk pregnancies are defined as pregnancies with an absence of co-morbidity or obstetric complication, such as maternal diabetes, previous cesarean section, high blood pressure or other illness. Women may be experiencing their first or subsequent pregnancy. The fetus must also be well and without any co-morbidity or complication.

Types of intervention(s)/phenomena of interest

The intervention of interest is waterbirth. The comparator is landbirth. Women and their babies must be cared for by qualified maternity healthcare providers throughout their labor and birth. The birth setting must be clearly described but can include home, hospital or birth center, either freestanding or attached to a hospital.

doi: 10.11124/jbisrir-2015-2105

Types of studies

This review considered randomized controlled trials, quasi-experimental studies and observational prospective and retrospective cohort studies.

Types of outcomes

1. Neonatal death within 28 days of birth
2. Neonatal resuscitation at birth or respiratory distress syndrome
3. Sepsis or infection within seven days of birth
4. APGAR scores at one, five and ten minutes
5. Admission to Neonatal Intensive Care Unit or Special Care Units
6. Cord pH values
7. Cord avulsion
8. Hyponatremia
9. Hypoxic ischemic encephalopathy
10. Birth injury.

Search strategy

A multi-step search strategy was utilized to find published and unpublished studies, in English between January 1999 and June 2014.

Methodological quality

The first author assessed the quality of all eligible studies. The three secondary authors independently assessed six studies each, followed by group discussion using the appropriate Joanna Briggs Institute appraisal checklist.

Data extraction

Data were extracted using a standardized extraction tool from Joanna Briggs Institute.

Data synthesis

Quantitative studies were pooled, where possible, for meta-analysis using software provided by Cochrane. Effect sizes were expressed as odds ratio or relative risk, according to study design, and the 95% confidence intervals were calculated. Heterogeneity was assessed statistically using the standard Chi-square test.

Results

The meta-analyses of 12 studies showed that for the majority of outcomes measured in this review there is little difference between waterbirth and landbirth groups. Meta-analysis was not conducted for mortality within 24 days of birth. Heterogeneity was significant between studies for APGAR (*A*ppearance, *P*ulse, *G*rimace, *A*ctivity, and *R*espiration). scores ≤7 at one minute and admission to Special Care nursery. Sensitivity analysis for case control studies describing infection found results that were not statistically significant (OR 0.74, 95% CI 0.05-11.06). Results of meta-analysis were also not significant for studies describing resuscitation with oxygen (OR 1.12, 95% CI 0.14-8.79) and Respiratory Distress Syndrome (OR 0.81, 95% CI 0.44-1.49). Results comparing APGAR scores ≤7 at five minutes for waterbirth and landbirth groups results for included RCTs demonstrated results that were not statistically significant (OR 6.4, 95% CI 0.63-64.71). However, results for included cohort studies

doi: 10.11124/jbisrir-2015-2105

describing APGAR scores ≤7 at 5 minutes indicate neonates are less likely to have scores ≤7 in the waterbirth group (OR 0.32, 95% 0.15-0.68). Data were not statistically significant for meta-analysis describing admission to NICU (OR 0.51, 95% CI 0.13-1.96) between water and landbirth groups. The differences in arterial (MD 0.02, 95% CI 0.01-0.02) and venous (MD 0.03, 95% CI 0.03-0.03) cord pH, while statistically significant, were clinically negligible.

Conclusions

Analyses of data reporting on a variety of neonatal clinical outcomes comparing land with waterbirth do not suggest that outcomes are worse for babies born following waterbirth. Meta-analysis of results for five-minute APGAR scores ≤7 should be treated with caution due to the different direction of results for meta-analysis of data from randomized controlled trials and cohort studies. Data measuring cord pH (an objective measure of neonatal wellbeing) were robust and showed no difference between groups. Overall this review was limited by heterogeneity between studies and meta-analysis could not be conducted on a number of outcomes. Waterbirth does not appear to be associated with adverse outcomes for the neonate in a population of low risk women.

Implications for practice

There is no evidence to suggest that the practice of waterbirth in a low risk population is harmful to the neonate.

Implications for research

There is a paucity of high level evidence to guide practice in the area of waterbirth. It is unlikely that randomized controlled trials on waterbirth will be acceptable to childbearing women or maternity caregivers. Observational studies are a more appropriate choice for researchers in this field as they offer a more practical and ethical approach.

Keywords

morbidity; mortality; neonatal; outcomes; waterbirth

Background

The first recorded waterbirth occurred in France in 1803.[1] After laboring for 48 hours, an exhausted woman used a warm bath and birthed a healthy baby.[2] In the early 1980s, waterbirth became more popular as water immersion was promoted to help women relax and cope with labor.[3,4]

Waterbirth and water immersion in labor are two distinct phenomena; however they are often confounded. Some women use water immersion in labor as a strategy to manage their labor pain but leave the bath prior to the birth of their baby. As the name implies, however, waterbirth occurs when a baby is born underwater. This can happen either intentionally or accidentally, for example, when a woman uses water immersion during labor and remains in the water to birth her baby. Although the definitions of waterbirth and water immersion are simple to separate, descriptions of their use during a woman's labor are often merged. Given this, research attempting to describe the benefits and risks of both water immersion and waterbirth sometimes confounds these distinct phenomena.[2,5]

Water immersion in labor has been used by many generations of women and is common practice in many birthing suites.[6] Current research describing the benefits for women using water immersion in labor include: increased relaxation[7], pain relief[8,9,10], maximized maternal satisfaction[11] and reduced

doi: 10.11124/jbisrir-2015-2105

length of labor.[5,6,8,12,13] The buoyancy of the water allows women to move more easily during labor and potentially optimizes labor progress.[6,9] Water immersion may also be associated with improved uterine perfusion, less painful contractions and a shorter labor.[1,6,14] A Cochrane systematic review[8] of eight trials comparing water immersion in labor with controls showed that water immersion resulted in a significant reduction in epidural analgesia use (RR 0.90; 95% Confidence Interval [CI] 0.82 to 0.99, six trials), a reduction in duration of the first stage of labor (mean difference -32.4 minutes; 95% CI - 58.7 to -6.13, seven trials) with no difference in assisted vaginal birth (RR 0.86; 95% CI 0.71 to 1.05, seven trials), cesarean sections (RR 1.21, 95% CI 0.87 to 1.68, eight trials), use of oxytocin infusion (RR 0.64, 95% CI 0.32 to 1.28, five trials), perineal trauma (intact perineum 236/678 versus 200/659, RR 1.16, 95% CI 0.99 to 1.35, five trials) or maternal infection (15/647 versus 15/648, RR 0.99, 95% CI 0.50 to 1.96, five trials).[8]

The benefits for women using water immersion in labor are evident and some local or national policies promote the use of water immersion for laboring women. For example, in 1992 baths and birthing pools were integrated into the United Kingdom's mainstream maternity units after the House of Commons Health Committee recommended that all women have access to water for labor and birth.[15] Their national practice guidelines also support the use of baths and birthing pools for women in labor.[19,17] Within Australia, 14 out of 19 birthing centers provide bath facilities.[18]

Waterbirth tends to be supported by midwives as it represents a birthing option congruent with the midwifery philosophy of promoting physiological birth.[11,19] However, the practice of birth underwater remains controversial and due to the nature of the intervention, there are few randomized controlled trials available to inform the issue. As a result of this uncertainty, many birthing units in Australia provide water immersion in labor as an option for women but do not support waterbirth. Implementing waterbirth policies remains a slow and complex process.[20,21]

Concerns about waterbirth

In an uncomplicated pregnancy, water immersion in labor is unlikely to harm the woman or her baby.[8] However, a number of maternity health professionals[2,22-27] and professional groups[28,29] have concerns in relation to waterbirth.

Although the practice of waterbirth has been linked to increased risk to the neonate[2,22-24] there is no high level evidence available to inform this issue.[2,30] The association between waterbirth and adverse neonatal outcomes comes largely from case reports.[24,31,32] These highlight the potential risks to the neonate from waterbirth, including: neonatal respiratory distress, neonatal infection, umbilical cord avulsion, hyponatremia, hypoxic ischemic encephalopathy, fetal thermoregulation and water embolism.[2,22,24,25,33] There are also numerous articles providing commentary about a perceived lack of adequate research and potential disregard for adverse neonatal outcomes following waterbirth.[21,25,26,27,34] Both case reports and commentary are subject to author bias and represent a low level of evidence upon which to build waterbirth policy and protocol.

In a literature review of the evidence informing waterbirth, Young and Kruske (2013)[35] identified five main areas of concern: a perceived risk of neonatal water aspiration, neonatal and maternal infection, neonatal and maternal thermo-regulation, skills of attending midwives and emergency procedures in the event of maternal collapse. They concluded there was no evidence supporting these concerns.[35]

The above neonatal outcomes, as described within the literature, will form the basis of the reviewers' search and discussion concerning potential adverse neonatal outcomes following waterbirth.

doi: 10.11124/jbisrir-2015-2105

Current evidence and policy

A number of systematic reviews have been conducted on waterbirth but all have some limitations. Simpson (2013)[36] conducted a systematic review of neonatal outcomes following waterbirth; however, this review was limited as only two randomized controlled trials (RCT), two systematic reviews and five case reports were reviewed. A number of observational studies have been conducted on waterbirth but these were not included.[36] Also, no meta-analysis was conducted. The Cochrane systematic review by Cluett and Burns (2009)[8] was similarly limited to RCTs, of which only three were on waterbirth, with a total sample size of 186. Another systematic review, conducted in 2004, searched for complications that could be associated with waterbirth and was not limited to RCTs.[2] They reviewed 16 articles and concluded that waterbirth may be associated with complications that are not seen with landbirth; however, outcomes from water immersion and waterbirth are confounded.[2] The quality of this review has also been criticized.[26]

There are two known published RCTs comparing outcomes after waterbirth and landbirth: an Iranian study of 106 women[37] and a pilot study conducted in the UK including 60 women.[30] Both trials were small and therefore too underpowered to identify differences in maternal and neonatal outcomes. Woodward and Kelly (2009) reported that a larger RCT was possible and acceptable to women; however no such trial has been conducted, raising concerns over feasibility.[30]

Published guidelines for health practitioners argue that there is insufficient evidence to guide waterbirth practice. The Royal Australian and New Zealand College of Obstetricians and Gynecologists (2011) offer a cautious review, suggesting that there is little evidence of waterbirth offering any benefits and advise caution when interpreting any current studies due to small sample sizes.[28] The American College of Obstetricians and Gynecologists (2014) state that waterbirth “should be considered an experimental procedure that should only be performed within the context of an appropriately designed clinical trial with informed consent”.[29(p914)] However, it is unlikely that a RCT could ethically be conducted on this practice.

Other protocols agree there may be no benefit, but argue that there is also no adverse effects directly attributable to waterbirth. A joint statement released by the Royal College of Obstetrics and Gynecology (RCOG) and the Royal College of Midwives (RCM) supports women laboring in water while acknowledging the lack of evidence supporting waterbirth and the rarity of complications.[38] The Queensland Normal Birth guidelines also discuss waterbirth, stating there is no evidence of increased adverse effects for the woman or fetus; however they acknowledge there is inadequate evidence to either support or not support a woman’s choice to birth underwater.[39] Young and Kruske (2012) suggest that Australian based policies are not well informed by the existing evidence and are not effective in informing and providing guidance to women or their health care providers.[40]

This lack of evidence results in conflicting opinions and allows anecdotal evidence to influence policy and practice.[3] In the absence of RCTs, observational studies offer the next highest level of evidence and a number of these have been conducted. Therefore a systematic review of this body of research is needed to provide the best available evidence on which to guide clinical decision making in relation to waterbirth.

The objectives, inclusion criteria and methods of analysis for this review were specified in advance and documented in a published protocol.[41]

doi: 10.11124/jbisrir-2015-2105

Objectives

The objective of this review was to systematically synthesize the best available evidence regarding the effect of waterbirth, in comparison to landbirth, on the mortality and morbidity of neonates born to low risk women.

Inclusion criteria

Types of participants

This review considered studies that included low risk, well, pregnant women who labor and birth spontaneously, at term (37-42 weeks), with a single baby in a cephalic presentation. Low risk pregnancies are defined as pregnancies with an absence of co-morbidity or obstetric complication, such as maternal diabetes, previous cesarean section birth, high blood pressure or other illness. Women may be experiencing their first or subsequent pregnancy. The fetus must also be well and without any co-morbidity or complication.

Types of intervention(s)

The intervention of interest is waterbirth. The comparator is landbirth. Women and their babies must be cared for by qualified maternity healthcare providers throughout their labor and birth. The birth setting must be clearly described but can include homebirth, hospital birth or birth center, either freestanding or attached to a hospital.

Types of studies

This review considered randomized controlled trials, quasi-experimental studies, prospective and retrospective cohort studies. Descriptive studies that did not include a comparator were excluded.

Types of outcomes

This review considered studies that included the following outcome measures:

1. Neonatal mortality – neonatal death within 28 days of birth
2. Neonatal resuscitation or Respiratory Distress Syndrome within 24 hours of birth
3. Neonatal sepsis/infection, including fever and other infection markers, as defined within any studies, within seven days of birth
4. APGAR scores at one, five and ten minutes
5. Admission to Neonatal Intensive Care Unit or Special Care Units, including length of stay
6. Cord pH values – arterial and/or venous taken immediately following birth
7. Cord avulsion
8. Hyponatremia
9. Hypoxic ischemic encephalopathy
10. Birth injury

Search strategy

The search strategy aimed to find both published and unpublished studies. A three-step search strategy was utilized in this review. An initial search of MEDLINE and CINAHL was undertaken followed by an analysis of the words contained in the title and abstract, and of the index terms used to describe the article. A second search using all identified keywords and index terms was then undertaken across all included databases (Appendix I). Thirdly, the reference lists of all identified reports and articles were searched for additional studies.

doi: 10.11124/jbisrir-2015-2105

Studies published in English within the last 15 years (from 1999) were considered for inclusion. This will ensure that retrieved studies will provide more recent, up-to-date evidence that will reflect more contemporary practice and policy.

The databases searched included:
CINAHL
MEDLINE
Cochrane Database of Systematic Reviews and Central Register of Controlled Trials
Pubmed Clinical Queries
Web of Science
Scopus
Health Source Nursing

The search for unpublished studies included:
ProQuest Theses and Dissertations
National Library database (Trove)
Google Advanced Search

Initial keywords used:
Waterbirth; birth/s; childbirth; water; underwater; delivery; births in water; outcome; effect; neonate; neonatal; fetal; infant; newborn; morbidity; mortality

Method of the review

Critical appraisal

Papers selected for retrieval were assessed by four independent reviewers for methodological validity prior to inclusion in the review using standardized critical appraisal instruments for randomized controlled/pseudo-randomized trials or comparable cohort/case-control studies as appropriate, the Joanna Briggs Institute Meta-Analysis of Statistics Assessment and Review Instrument (JBI-MAStARI) (Appendix II). Any disagreements that arose between the reviewers were resolved through discussion.

Data extraction

Data were extracted from papers included in the review using the standardized data extraction tool for experimental/observational studies from JBI-MAStARI (Appendix III). The data extracted included specific details about the interventions, populations, settings, study methods and outcomes of significance to the review question and specific objectives.

Five authors where contacted to request further data.[20,42-45] Data were received from three authors and these studies were consequently included in this review.[20,43,45] One author[44] was unable to provide the requested data which excluded women who had a previous cesarean section and the other[42] did not respond. Consequently these studies were excluded from this review.

Data synthesis

Where possible, quantitative data were pooled in statistical meta-analysis using Revman software (Review Manager, version 5.3.4). All results were subject to double data entry. Effect sizes are expressed as odds ratio due to the inclusion of case control studies in this review (for categorical data) and weighted mean differences (for continuous data) and their 95% confidence intervals were calculated for analysis. A random effects model was used for cohort and case control studies and

doi: 10.11124/jbisrir-2015-2105

heterogeneity was assessed statistically using the standard Chi-square test. A p value of < 0.1 was considered statistically significant. Dichotomous data were analyzed using Mantel-Haenszel as the default method or Peto odds ratio when event rates and sample sizes were very small or contained a zero value.

Meta-analyses were conducted where included studies were appropriately similar. Clinically, the women included in the studies have similar characteristics, being low risk and without obstetric co-morbidities. Some differences did exist between studies and these are discussed further in the description of studies. Methodologically, RCTs and observational studies were included together in a table of results; however RCTs were subsequently not included in meta-analyses with observational studies. Where possible, and where results indicate significant heterogeneity, case-control and cohort studies are analyzed separately.

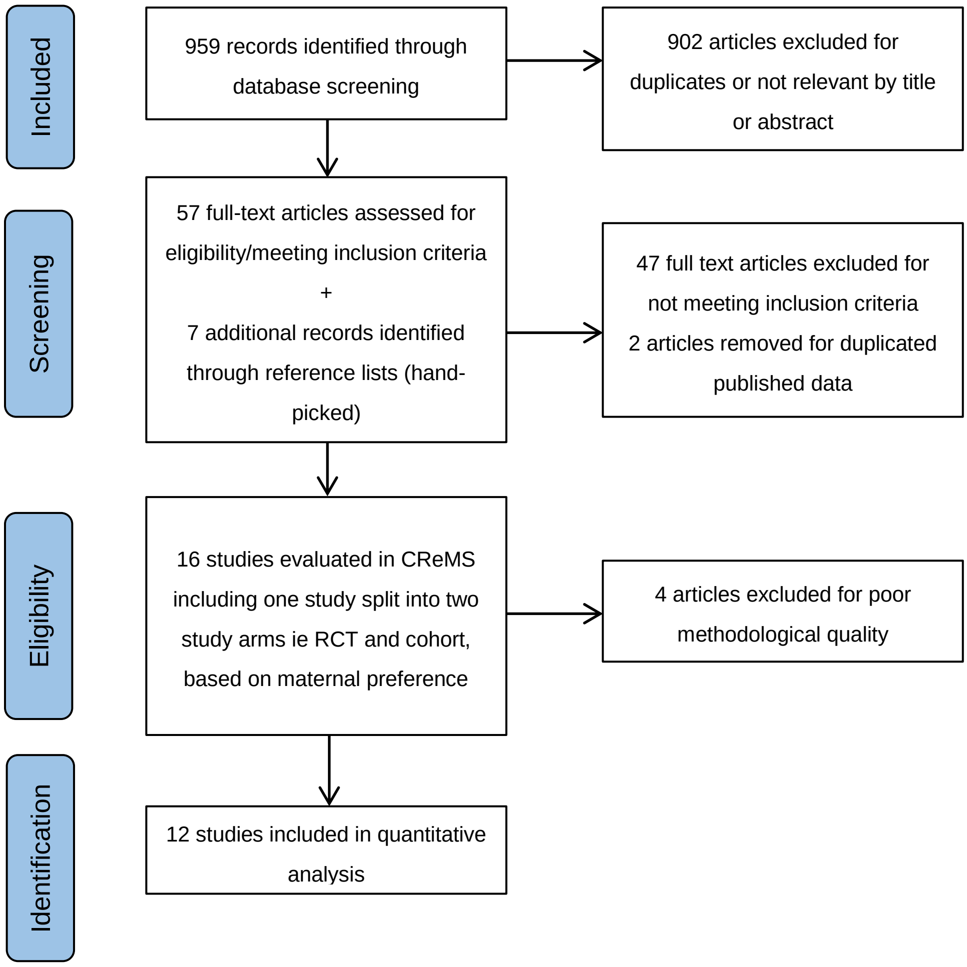
Results

Description of studies

From the search of the databases detailed above, 959 studies were identified as potentially relevant. The titles and abstract were assessed according to the inclusion criteria. The first author read the title and abstract of studies deemed potentially relevant. During this process, 902 studies were excluded as duplicates or because they did not meet the inclusion criteria. The remaining 57 studies were retrieved in full for further examination. An additional seven studies were included at this point after being identified from the reference lists of retrieved studies. After reviewing the 64 full text studies, 47 were further excluded for not meeting the inclusion criteria. At this stage, two studies were excluded for duplicated published data. Consequently, 15 studies were included and added to JBI MAStARI. One study[30] was subsequently split into two separate studies in JBI MAStARI as it described two aspects of one trial (a RCT and a cohort "preference arm"), increasing the number of studies assessed for methodological quality to 16. The preference arm was consequently analyzed with other non-experimental studies and the RCT was analyzed with data from other experimental studies. Four studies were excluded for poor methodological quality. Twelve studies were included in the review: two randomized controlled studies[30,46] and ten observational studies, of which five were cohort studies (three prospective[30,48,50] and two retrospective[20,45]) and five were case-control studies.[6,43,47,49,51]

A flow diagram of this process is presented in Figure 1.

doi: 10.11124/jbisrir-2015-2105



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

Figure 1: Flow diagram of the study inclusion process

All included studies compared neonatal outcomes following waterbirth and landbirth. All studies applied strict inclusion criteria regarding the health of women and fetuses enrolled in their studies. Women were low risk, with no obstetric complications or co-morbidities at entry to the study. Studies were excluded from review if they included women with a previous cesarean section, current gestational diabetes mellitus, macrosomic or intra-uterine growth restricted fetus or other complicating factors.

Included studies were published between 1999 and 2013. The studies were conducted in the following countries: United Kingdom (*n*=3), Switzerland (*n*=2), South Africa (*n*=2), Australia (*n*=2), Turkey (*n*=1), Austria (*n*=1) and Italy (*n*=1).

Further details of the included studies can be found in Appendix IV.

The list of excluded studies appears in Appendix V.

Description of RCTs included in the review

Nikodem[46] conducted a RCT, presented as a thesis, submitted in 1999 at the University of Johannesburg. The study was conducted over two years at two government hospitals, both of which served the same population of women: low to middle income, living in urban Johannesburg. The primary aim of this study was to evaluate maternal outcomes and experience of waterbirth: however

doi: 10.11124/jbisrir-2015-2105

the study also included select neonatal outcomes including: mortality, resuscitation with oxygen, APGAR score at five minutes, admission to SCN and cord pH values. One hundred and twenty women agreed to take part in the trial, and 60 women were randomized to each group. There were no withdrawals. The sample included women who had low risk pregnancies with no complicating factors. English speaking pregnant women in established labor, with a term (37-42 weeks) infant presenting in a cephalic position were approached to take part in the study. Exclusion criteria included grand multiparae women or women aged over 35. Randomization occurred once women were fully dilated or, for multiparous women, when they began to feel the urge to push. The researcher opened consecutively numbered envelopes randomly assigned to waterbirth or landbirth by an independent assistant, thereby removing selection bias of both participant and researcher. There were equal numbers of multiparous and primiparous women in each group. Care protocols for both waterbirth and landbirth are extensively described within the thesis and appear similar. As intention to treat analysis was used within this study, the results include three infants experiencing assisted birth: either cesarean section (*n*=1 control group) or instrumental birth (1 in each group). A standardized birth data form was completed to collect biographical, intrapartum and postnatal data. The researcher completed the data sheet to ensure consistency, collecting birth and postnatal data from patient records.

Woodward and Kelly[30] conducted a pilot randomized controlled trial to assess the feasibility of undertaking a larger multicenter trial. This study randomly allocated 60 women to either waterbirth or landbirth. A ratio of 2:1 was chosen as the authors recognized that not all women randomized to waterbirth would birth in the pool. Consequently, 40 women were randomized to waterbirth and 20 to landbirth at 36-40 weeks gestation. Randomization was blind to the researcher who recruited the women. This trial also included a “preference arm”, which has been separately analyzed as a cohort study within this review.[30] In this part of the trial, the outcomes of 20 women, 10 of whom definitely wanted a waterbirth and 10 who definitely wanted a landbirth, were analyzed separately. Inclusion criteria for both study arms were: term gestation more than 37 weeks, no risk factors or obstetric complications and no previous cesarean section. If required, labor was augmented with oxytocin. Intermittent auscultation of the fetal heart rate was conducted during labor. Within the RCT group, 34 of 40 women in the waterbirth group had a spontaneous vaginal birth (16 of these women did not use the pool for their labor or birth), four had an instrumental birth and two a cesarean section. There were 10 actual waterbirths within the waterbirth group. In the landbirth group, 14 of 20 women had a spontaneous vaginal birth (one woman in this group had a waterbirth), three women experienced an instrumental birth and three a cesarean section. Significant crossover occurred between the groups in the RCT whereby 16 (40%) women randomized to waterbirth did not use the pool for their labor or birth. For the preference arm, each group had 10 participants. Within the waterbirth group, seven women had a spontaneous vaginal birth (of which five were waterbirths), one an instrumental birth and two a cesarean section. In the landbirth group, nine women had a spontaneous vaginal birth and one woman had an instrumental birth. Data were collected on neonatal outcomes relevant to this review: arterial and venous cord pH, APGAR scores at one and five minutes, cord avulsion, admission to Neonatal Intensive Care Unit (NICU) and interventions, infections requiring antibiotics and mortality.

Description of cohort studies included in the review

Mollamahmutoglu et al.[48] prospectively studied a group of 610 women admitted in labor to a regional hospital over a 15-month period. Women elected to waterbirth, landbirth or have epidural analgesia in labor and were consequently split into these three groups for analysis. This review did not include the

doi: 10.11124/jbisrir-2015-2105

outcomes of women choosing to have an epidural. All women were of term gestation (37-42 weeks), with a single pregnancy in cephalic presentation without any obstetric risk factors. The three treatment groups were similar for maternal age, body mass index (BMI) and gestational age. Both primiparous and multiparous women were considered.

Geissbuehler, Stein and Eberhard[45] retrospectively observed 9518 women over a nine-year period at a rural hospital in Switzerland. Their original published data included women with obstetric risk factors; however the authors were contacted and provided further data excluding these women. Women with breech presentations, multiple pregnancies, less than 36 weeks gestation or BMI >40 were excluded from waterbirth. All births were documented on standardized questionnaires from which the data was collected. The participants had spontaneous, cephalic, vaginal births at term. Within this study, landbirths were further divided into maternal positioning and included: bed births, Maia stools, lateral, supported squat, roma wheel, all fours, standing and others. Geissbuehler, Stein and Eberhard[44] stated that neonatal care was provided by pediatricians with neonatal experience who attended births as requested by maternity staff. With suspected or confirmed cases of respiratory distress syndrome, neonates were transferred to an external tertiary referral NICU. No details were provided in relation to criteria for diagnosis of respiratory distress syndrome.

Zanetti-Dallenbach et al.[50] conducted a prospective observational study between April 1998 and May 2002 in a hospital in Switzerland. During routine antenatal check-ups, all women with low risk pregnancies and no obstetric complications where informed about the availability of waterbirth, the inclusion and exclusion criteria and hospital guidelines. Inclusion criteria were a single pregnancy with a cephalic presentation at term (>37 weeks gestation), continuous cardiotocogram (CTG) while laboring, negative tests for HIV, hepatitis B and C, venous access during labor and women understanding the need to leave the bath in the case of suspicious or pathological CTG. Women were also screened for Group B Streptococcus and treated with ampicillin intravenously if the screening was positive. Women entered the bath whenever they desired and the temperature was maintained between 32 and 36 degrees. Intravenous oxytocin was administered to women to augment their labor if cervical dilation was less then 1cm an hour (18% waterbirth group, 44.5% landbirth). Artificial rupture of membranes also occurred at this point (33.7% of women in the waterbirth group, 37% in landbirth group). Labor was managed identically for women in either group. Pain relief was provided to women at their request. Women electing for epidural analgesia were asked to leave the water. All data were collected by midwives or resident doctors and documented on a standardized form in the patient’s notes. Four groups were identified; waterbirth, spontaneous birth with water immersion, spontaneous birth without water immersion and assisted birth with forceps, ventouse or cesarean section. For the purposes of this review, only data from groups one and three were analyzed. The study group, waterbirth, consisted of 89 women. Group three, forming one of the control groups consisted of 146 women who spontaneously birthed and did not use water immersion in labor. Other demographic and obstetric data was similar across both groups. Intravenous oxytocin was used more often in the control group. This study included women with prolonged rupture of membranes (PROM). In the waterbirth group, eight women (9%) had PROM compared to 17 women (11.6%) in the landbirth group. Women presenting with meconium stained liquor also remained within this study (9% in waterbirth group, 13% in landbirth).

Dahlen et al.[20] conducted a cohort study within an Australian birth center to determine rates of perineal trauma, postpartum hemorrhage and APGAR scores five minutes post birth in low risk women. The landbirth data were originally split into six differing positions but were combined for the purposes of this review. In this study data were collected by reviewing handwritten midwifery records

doi: 10.11124/jbisrir-2015-2105

over a 12-year period, from January 1996 to April 2008. During this time frame, 8338 women commenced labor, 6144 women gave birth at this birth center and 2194 were transferred to an alongside labor ward due to obstetric complications. The women included in this study were from the group who birthed at the center. Women attending this birth center were low risk, experiencing normal vaginal births and were attended by midwives and obstetricians experienced in active birth and waterbirth. Continuous monitoring did not occur and epidural analgesia was not available. Of these women, 30% used pharmacological pain relief, the most common of which was nitrous oxide. Midwives were the lead accoucheur in 91% of births, the remainder were led by obstetricians. There were 818 (13%) waterbirths and 5217 (85%) landbirths, with 105 (2%) births where the mode of birth was not recorded. The authors were contacted, and further data received, regarding raw data on the numbers of women with APGAR scores ≤7 at five minutes after birth.

Description of case-control studies included in the review

Ros[49] conducted a case-control study across two different birthing units. Waterbirths occurred in birth units in two private hospitals and were attended by midwives. Landbirths, referred to as traditional bed births, occurred in a labor ward in a government hospital. Women were also attended by midwives and birthed in lithotomy or semi-reclined positions. Although participants were taken purposively from the population of low risk women, selecting women from differing maternity care models (public and private) may expose the study to selection bias. Little information is provided regarding specific protocols of care for either group. The women were healthy, low risk, women with a single pregnancy and a cephalic presentation. Gestational age was limited to 37-42 weeks. Women meeting the inclusion and exclusion criteria were invited to participate in the research. Twenty-seven women were recruited for each group. The author collected and input data. The researcher contacted the women 14 days after the birth of their baby to confirm the occurrence of any further neonatal health issues.

Menakaya and colleagues[47] conducted a retrospective case-control study in Australia comparing waterbirth and landbirth. The location of this study, Bankstown hospital, offers formal training to staff for waterbirth and has an extensive waterbirth protocol. Information on the waterbirth policy is provided to women routinely in the ante-natal time. Women are encouraged to use water immersion in labor if maternal and fetal observations are within normal ranges. Fetal observations are attended 1-2 hourly in first stage and after each contraction in second stage of labor. Entonox (combined nitrous oxide and oxygen) is provided for further pain management. Vaginal examinations are conducted within the bath. The waterbirth group consisted of all women who experienced a waterbirth at Bankstown hospital between 2000 and 2009. This group consisted of 219 low risk women. Inclusion criteria included spontaneous vaginal birth, uncomplicated antenatal care, term gestation, single pregnancy, cephalic presentation, labor without augmentation, clear liquor if membranes had ruptured and established labor upon entry to the pool. Women with prolonged rupture of membranes (PROM) or who were positive for Group B Streptococcus (GBS), were also included in this study and were provided antibiotic cover during labor. The control group was formed by a matching cohort of similar risk women who had a spontaneous vaginal birth not in water within 24 hours of each woman who birthed underwater. The study period extended over 10 years and this introduces potential for bias as practice and the practice context can change over this period. Inclusion criteria for the control group were the same as the waterbirth group and women were matched by gestational age and parity. Women were excluded if they had epidural analgesia, augmentation of labor or opioid pain relief. Data for both groups were extracted from patient records by the researcher.

doi: 10.11124/jbisrir-2015-2105

Bodner et al.[51] conducted a case-control study over 18 months in an Austrian hospital. During the study period 140 women elected for waterbirth. The control group was matched by parity and consisted of low risk women who experienced spontaneous vaginal births. There was no description about the selection of women in the control group leaving this study open to issues of selection bias. Inclusion criteria were restricted to term gestation >37 weeks, normal sized healthy fetus, reactive admission CTG, clear amniotic fluid if membranes had ruptured and a cephalic presentation. Women with medical or obstetric risk factors were excluded. Management of waterbirth and landbirth was similar for both groups of women. Obstetric or pediatric staff attended births if pathological conditions presented, at the request of attending midwives. Midwives received training in waterbirth and management of emergencies during waterbirth. Oxytocin augmentation was instigated in women with slow cervical dilation. Within the waterbirth group, 22 (15.7%) women were augmented with oxytocin and 44 women (31.4%) in the landbirth group. Medical analgesia was given to nine women in the waterbirth group and 42 in the landbirth group. One woman in the landbirth group had an epidural. The pool water was maintained at 37-37.5 degrees Celsius throughout labor and birth. Continuous cardiotocogram (CTG) monitoring was used throughout labor. At the moment of birth, care was taken to ensure the controlled birth of the baby's head and body with the baby quickly and gently bought to the surface. The cord was clamped and cut, the bath emptied and the birth of the placenta was completed within the pool. Data were collected on a pre-coded form. No significant differences were found for neonatal parameters measured in this study between neonates born underwater or on land.

Otigbah and colleagues[6] undertook a case-control study in Scotland, comparing 301 women who had waterbirths and a control group of parity matched, low risk women who experienced a landbirth. The control group was selected from the birth register and also matched maternal age to within five years and induction method where applicable (7.3% of women in the study underwent prostaglandin induction for postdates). Inclusion criteria included normal vaginal birth without oxytocin augmentation. Women presented to birth suite in labor and requested in writing to have a waterbirth. They were reviewed by a senior midwife and assessed for meeting waterbirth guidelines. These include; gestation more than 38 weeks, cephalic presentation, normal sized fetus, normal admission CTG and clear liquor if membranes had ruptured. Women with medical or obstetric risk factors were excluded. Women entered the pool at their request and the water was maintained at 37-37.5 degrees Celsius. Intermittent fetal monitoring occurred routinely throughout labor. Women were provided with pain relief upon request, usually Entonox. In the waterbirth group 58% of women used Entonox, compared to 32% of women in the control group ($p<0.0001$). Transcutaneous electrical nerve stimulation (TENS) was available and used outside of the pool. Pethidine was also administered upon maternal request. In the waterbirth group 1.3% of women requested Pethidine compared to 56% in the landbirth group. Women remained out of the bath for 2-3 hours after Pethidine administration. Epidural analgesia was not administered. Two trained midwives were present at each waterbirth and closely followed written protocols. During second stage, midwives used a hands-on approach and ensured controlled birth of the head and body. The baby was brought to the surface gently, the cord cut and clamped and oxytocin given for third stage. Birth of the placenta was completed outside of the pool.

There was no follow up of cases within this study. As such, outcomes such as infection are unreliable as symptoms and diagnoses may present some time after birth. This is one downfall of the retrospective study design using the birth register as a source for data. Without accessing the full patient notes or contacting women, cases of infection up to seven days after birth will not be found.

doi: 10.11124/jbisrir-2015-2105

Pagano et al.[43] compared maternal and fetal outcomes after waterbirth and landbirth. The study aimed to assess the cost-effectiveness of waterbirth compared to normal birth on land. The authors employed a retrospective case-control study design with a cohort of 110 women in each group. Data were collected over a two-year period and were extracted from an Italian hospital birth register, which may have limited the expanse of data collected. The waterbirth group consisted of primiparous women experiencing a waterbirth within the defined timeframe. For every waterbirth the subsequent woman in the register to have a normal vaginal birth meeting the inclusion criteria was admitted to the study. The inclusion criteria included; normal spontaneous birth, no pregnancy complications, no obstetric complications in labor, no meconium stained liquor, PROM or maternal pyrexia in labor. Women must also be laboring at term, between ages of 18 and 35 years old with a fetus with no diagnosed complications. While this study focused on the economic analysis of waterbirth compared to landbirth, data on APGAR scores were also collected. Mean scores were presented however standard deviation was not supplied. The authors were contacted and subsequently provided raw data on APGAR scores at one and five minutes post-birth allowing calculation of dichotomous data.

Not all studies provided data for all outcomes studied in this review. A full account of the data for individual outcomes derived from each study can be found in Appendix VI.

Labor and birth is a dynamic and complex process and those providing care may influence its course. Newborn assessments made by clinicians may be subjective, specifically when determining outcomes such as APGAR scores, need for differing resuscitation measures and admission into special care nurseries. APGAR scores are somewhat dependent upon the person making the assessment and are at risk of bias, especially as blinding with regards the mode of birth in these studies is not possible. While simple guides exist for assessing a newborn at one and five minutes, in practice, these scores are calculated retrospectively and outcomes, such as neonatal colour, are easily subject to variance. The same applies for simple resuscitation of neonates. Nikodem[45] and Ros[48] both provide data on resuscitation with oxygen, however no explanation of the extent of measures or process is provided. Geissbuehler, Stein and Eberhard[45] and Menakaya et al.[47] provide data on respiratory distress but do not describe their criteria for diagnosis.

Variance between care protocols for both groups, but especially for the women in landbirth groups, is one confounding factor that may influence outcomes included in this review. Most studies clearly describe protocols for waterbirth. Landbirth however is described as 'normal' or 'traditional' care with little further description. Other studies describe care between both groups as equal. Subjective decisions made by healthcare practitioners assisting women during labor and birth, and the protocols that guide their practice, may differ. All studies describe monitoring of the neonate throughout the first and second stages of labor.

The definition and procedures for waterbirth are similar between all studies in this review. Waterbirth is defined as the birth of a baby completely underwater. In all instances the baby was brought gently, and within seconds, to the surface of the water and given to the woman to hold. The umbilical cord was clamped and cut. Women were encouraged to enter the bath at any stage of their labor. Landbirth refers to the birth of an infant, which doesn't occur in the water. This definition allows for variance in maternal positioning. Studies that followed an intention to treat protocol, such as Woodward and Kelly[30] in their RCT, included outcomes after instrumental birth. All studies included women in spontaneous labor, although some also considered women whose labor was augmented with oxytocin if labor progress was slow.[30,48,50,51] This was defined as insufficient cervical dilation over time. Other studies did not discuss this intervention.[44,48] Bath water temperature was described in

doi: 10.11124/jbisrir-2015-2105

most studies. Zanetti-Daellenbach and colleagues[50] stated the water temperature varied between 32 and 36 degrees Celsius in their study. Others state that bath temperature was maintained between 37-37.5 degrees Celsius at the time of birth.[6,47-49,51] All studies, except Pagano et al.[43], also describe using continuous or intermittent fetal heart monitoring throughout labor as described by individual hospital protocols. Birth of the placenta varied among studies, some allowing third stage to complete underwater with others requiring women to leave the bath to birth the placenta.[6,46-48,51] Other studies followed expectant management allowing the woman and her blood loss to guide the care provided for third stage[50] or didn't describe third stage at all.[20,43,45,49] Pagano et al.[43], Woodward and Kelly[30] and Dahlen et al.[20] provide little description of either waterbirth or landbirth protocol for their respective studies.

Research into care options for laboring women does not easily lend itself to random allocation of participants. While both Nikodem[46] and Woodward and Kelly[30] randomized laboring women to waterbirth or landbirth, both studies recruited small numbers. A lack of randomization in the cohort studies found in this review, may suggest that the strength of evidence is less robust and the studies more prone to bias, however, RCTs in this field are uncommon. The observational studies in this review offer high quality data and more practical methods for examining the impact of waterbirth on neonatal outcomes.

Methodological quality

The first author conducted a methodological assessment on each study included in the review. Three secondary authors independently assessed six studies each. Once independent assessments were completed, discussions were held as a group to complete a final assessment. Any disagreements were resolved through discussion, or with a different secondary reviewer conducting a third assessment. Two such re-assessments occurred whereby a third reviewer independently assessed a study and returned for further discussion with the first author. Subsequently, one study was included in the review and one excluded. The final methodological assessment can be found in the following tables. Overall the methodological quality of included studies was high.

Methodological quality of RCTs included in the review

Table 1 details the methodological quality of the two RCTs included in this review. Question two asks if participants were blinded to the intervention, which due to study design and the nature of the intervention, is not possible. This also applies to the researchers providing intrapartum care, which is covered in question five. As such, both these questions are not applicable.[30,46]

Table 1: Methodological quality of RCTs included in this review

Citation	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Woodward J, Kelly SM, 2004[30]	Y	N/A	Y	Y	N/A	Y	Y	Y	Y	Y
Nikodem VC, 1999[46]	Y	N/A	Y	Y	N/A	Y	Y	Y	Y	Y
%	100.0	0.00	100.0	100.0	0.00	100.0	100.0	100.0	100.0	100.0

Methodological quality of cohort and case control studies included in the review

Table 2 below details the methodological quality of cohort studies and case control studies included in the review. Question six concerns appropriate follow-up of participants. In some studies follow-up was not appropriate as only outcomes at the time of birth were studied and did not require longer term monitoring.[6,20,43,47] Other studies provided appropriate follow-up for the outcomes measured.[30,45,48-51]

doi: 10.11124/jbisrir-2015-2105

Question seven details if outcomes from participants who withdrew from the studies are described. The majority of studies did not have any withdrawals due to study design (case-control or retrospective cohort studies).

Question three asks if bias has been minimized in relation to the selection of cases and controls. There is some confusion within the text of Mollamahmutoglu et al.[48] study where they describe women both electing and being assigned to either the control or waterbirth group. This appears to be an issue with the translation of the article. However, as the remainder of the study implies that women have elected waterbirth or landbirth, this study is labeled as unclear in terms of the selection bias of its participants.

Table 2: Methodological quality of cohort and case control studies included in this review

Citation	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Geissbuehler V, Stein S, Eberhard J, 2004[45]	Y	Y	N/A	Y	Y	N/A	N/A	Y	Y
Dahlen HG, Dowling H, Tracy M, Schmied V, Tracy S, 2013[20]	Y	Y	Y	Y	Y	N/A	N/A	Y	Y
Menakaya U, Albayati S, Vella E, Fenwick J, Angstetra D, 2013[47]	Y	Y	Y	Y	Y	N/A	N/A	Y	Y
Mollamahmutoglu L, Moraloglu O, Ozyer S, Su FA, Karayalcin R, Hancerlioglu N, et al. 2012[48]	Y	Y	U	N	Y	N/A	N	Y	Y
Woodward J, Kelly SM, 2004[30]	Y	Y	Y	Y	Y	Y	N/A	Y	Y
Otigbah CH, Dhanjal MK, Harmsworth G, Chard T, 2000[6]	Y	Y	Y	Y	Y	N/A	N/A	Y	Y
Bodner K, Bodner-Adler B, Wierrani F, Mayerhofer K, Fousek C, Niedermayr A, Grunberger W, 2002[51]	Y	Y	Y	Y	Y	Y	N/A	Y	Y
Ros HB, 2009[49]	Y	Y	Y	N	Y	Y	Y	Y	Y
Zanetti-Daellenbach RA, Lapaire O, Maertens A, Holzgreve W, Hosli I, 2006[50]	Y	Y	Y	Y	Y	Y	N/A	Y	Y
Pagano E, De Rota B, Ferrando A, Petrinco M, Merletti F, Gregori D, 2010[43]	Y	Y	Y	Y	Y	N/A	N/A	Y	Y
%	100.0	100.0	88.89	80.0	100.0	100.0	50.0	100.0	100.0

Findings of the review

This review aimed to collate data regarding neonatal outcomes for low risk women after waterbirth and compare these to outcomes after landbirth. Throughout this review, the treatment group refers to neonates born underwater and the control group refers to neonates born on land.

Outcome 1: Mortality of the neonate within 28 days of birth

This review searched for studies that compared mortality of the neonate within 28 days of birth. Three studies were found that documented this outcome. One of these is an RCT and is excluded from meta-analysis. The data from two cohort studies is included in the meta-analysis.

Results from RCTs included in the review

One RCT described mortality as an outcome when comparing waterbirth and landbirth.[46] Nikodem[46] detailed one neonatal death to a woman experiencing a complex pregnancy. The neonatal death occurred after a waterbirth. The woman was 25 years old, HIV positive with one living child. She had been admitted to hospital 16 days prior to birth with purulent vaginal discharge, was treated and discharged home. She was next reviewed when she returned in labor and consequently consented to join the trial. Artificial rupture of membranes was performed and thick meconium stained liquor was observed. She was randomized to waterbirth at full dilation and birthed her baby five minutes later in the pool. APGAR scores were 9:10:10 at one, five and ten minutes. The woman and her baby were transferred to postnatal ward 2.5 hours after birth where the baby was observed to be not breathing. Resuscitation commenced but was unsuccessful. Autopsy was not performed due to parental request. Upon further examination, blood tests and chest x-rays, independent physicians diagnosed the most likely cause of death to be intrauterine infection. Most waterbirth protocols exclude women who develop complications in labor and the presence of thick meconium (a potential sign of fetal distress) would exclude women from waterbirth. Nikodem[46] concludes that maternal and neonatal morbidity is increased after waterbirth and that further, more extensive randomized trials are required.

Results from cohort studies included in the review

Mollamahmutoglu et al.[48] found there were no adverse perinatal outcomes or neonatal deaths in their study. This outcome has been included in the analysis below; however as it is a zero value, it does not register within the forest plot.

Geissbuehler, Stein and Eberhard[45] reported two neonatal deaths after waterbirth and two after landbirth within seven days of birth. One of the neonates in the waterbirth group was born at 42 weeks gestation, weighing 3500g and with APGAR scores of 7:9:10. This infant was not transferred to NICU. No other details were provided. The other death in the waterbirth group was a neonate at 41 weeks gestation, 3360g with APGAR scores of 9:9:10. This infant was transferred to NICU and diagnosed with respiratory distress syndrome and other complications including a malformation. In the landbirth group, one neonate was 40 weeks gestation, 3860 g with APGARs of 2:0:0. This infant suffered adaption problems and malformation. The other neonate in this group was born at 41 weeks gestation, weighing 3450g with APGAR 9:10:10. This infant was transferred to NICU for complications. No further detail was provided.

No case-control study reported data on this outcome.

Meta-analyses of results for mortality of the neonate within 28 days of birth

Meta-analysis was not conducted for this outcome as only results from one study[45] could be included as one study is an RCT and the other has zero results which are not able to be analyzed. The data table below presents the data described for studies included in this review.

As this outcome is rare, a very large sample size would be required to adequately compare mortality between waterbirth and landbirth groups.

doi: 10.11124/jbisrir-2015-2105

Table 3: Results for included studies describing neonatal mortality

Study	Waterbirth event	Waterbirth total	Landbirth event	Landbirth total
Mollamahmutoglu et al 2012 [48]	0	207	0	207
Gueissbeuhler et al 2004[45]	2	3296	2	4074
Nikodem 1999[46]	1	60	0	60
Total	3	3563	2	4341

Outcome 2: Neonatal resuscitation and respiratory distress syndrome

This review searched for studies comparing the incidence of neonatal resuscitation after waterbirth and landbirth. Three studies provided data on neonatal resuscitation[46,47,49], one of which was an RCT[46] and was analyzed separately. Two studies provided data on neonates diagnosed with respiratory distress syndrome.[45,47] Neither of these studies provided detailed explanation of guiding protocols or clear definition of these outcomes. Authors were contacted for further details on these outcomes but no responses were received.[46,49]

Resuscitation with oxygen at birth

Results from RCTs included in the review

One RCT described neonates requiring resuscitation with oxygen after birth.[46] Nikodem[46] stated that 6% of neonates born after waterbirth and 3% after landbirth required resuscitation with oxygen.[46] This was not statistically significant. A pediatrician was called to attend 18 births (eight waterbirth and 10 control group). Reasons for attendance included meconium stained liquor (waterbirth *n*=7, control *n*=5), respiratory distress (control group *n*=1), decreased fetal heart rate before birth (control group *n*=2), and caring for infants born after assisted birth or cesarean section (waterbirth *n*=1, control group *n*=2). These neonates are included in this review due to the intention to treat analysis employed by this RCT. After assessment these neonates were admitted to the nursery for observation (waterbirth *n*=3, control group *n*=5).

No cohort studies reported on this outcome.

Results from case-control studies included in the review

Ros[49] conducted a case-control study across two different birthing units. This study collected data on the use of oxygen as a resuscitation measure as well as other resuscitation measures. Within the waterbirth group, five neonates were suctioned through mouth and nose, three received oxygen through a funnel and two received intramuscular Narcan (an opioid antagonist). From the control group, six neonates were suctioned through mouth and nose, one neonate received bagged oxygen and oxygen through a funnel a few minutes later. Four other neonates also required oxygen through a funnel and four neonates received Narcan intramuscularly.

Menakaya and colleagues[47] conducted a retrospective case-control study in Australia comparing waterbirth and landbirth. Two neonates in the waterbirth group and none in the landbirth group required active resuscitation and were admitted to the hospital Special Care Nursery (SCN) for observation. One neonate was 41 weeks gestation, born to a multiparous woman, birth weight of 3670g, with APGAR scores of 2:9 and was admitted to SCN for two days following Intermittent Positive Pressure Ventilation (IPPV) for two minutes and oxygen resuscitation. The other neonate was 41 weeks gestation, 2990g birth weight and APGARS of 4:5 requiring active resuscitation with an oxygen mask.

doi: 10.11124/jbisrir-2015-2105

Meta-analysis of results for resuscitation with oxygen at birth

Figure 1 presents data from two case-control studies and one RCT. Chi^2 test for heterogeneity is not significant. Results from the RCT are included within the table but excluded from meta-analysis. The results are not statistically significant (OR 1.12, 95% CI 0.14-8.79) suggesting there is no clinical difference between neonates requiring resuscitation after waterbirth or landbirth.

	Waterbirth		Land birth			Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI
Nikodem 1999	4	60	2	60		Not estimable
Ros 2009	3	27	5	27	67.7%	0.55 [0.12, 2.58]
Menakaya et al 2013	2	219	0	219	32.3%	5.05 [0.24, 105.71]
Total (95% CI)		**246**		**246**	**100.0%**	**1.12 [0.14, 8.79]**
Total events	5		5			

Heterogeneity: $Tau^2 = 1.01$; $Chi^2 = 1.66$, df = 1 (P = 0.20); $I^2 = 40\%$
Test for overall effect: Z = 0.11 (P = 0.91)

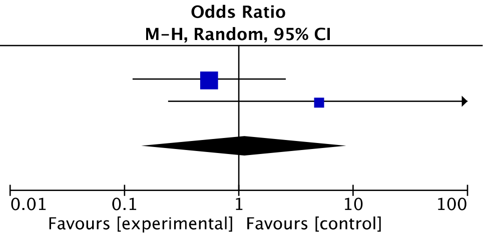


Figure 1: Resuscitation with oxygen at birth – observational studies only

Respiratory distress syndrome

Results regarding diagnosed neonatal respiratory distress syndrome was provided by two studies.[45,47]

No RCTs reported on this outcome.

Results of cohort studies included in the review

Geissbuehler, Stein and Eberhard[45] found 16 cases of neonates diagnosed with respiratory distress at birth in the waterbirth group and 25 cases in the landbirth group. No further details were provided.

Results of case control studies included in the review

Menakaya and colleagues[47] found one neonate from each group diagnosed with respiratory distress syndrome. The neonate from the waterbirth group was 41 weeks gestation born with APGAR scores of 8:9 and a birth weight of 4620g. This neonate was subsequently diagnosed with meconium aspiration and was transferred to a tertiary center. The neonate from the landbirth group was diagnosed with respiratory distress and admitted to SCN for nine days. This neonate was 38 weeks gestation with APGAR scores of 9:9 and a birth weight of 3160g.

Meta-analysis of results for respiratory distress syndrome

Figure 2 presents the meta-analysis of data from two studies. Chi^2 test for heterogeneity is not significant. The results are not statistically significant suggesting there is no difference between cohorts on the outcome of respiratory distress syndrome.

	Waterbirth		Land birth			Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI
Geissbeuhler et al 2004	16	3023	25	3785	95.1%	0.80 [0.43, 1.50]
Menakaya et al 2013	1	219	1	219	4.9%	1.00 [0.06, 16.09]
Total (95% CI)		**3242**		**4004**	**100.0%**	**0.81 [0.44, 1.49]**
Total events	17		26			

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 0.02$, df = 1 (P = 0.88); $I^2 = 0\%$
Test for overall effect: Z = 0.68 (P = 0.50)

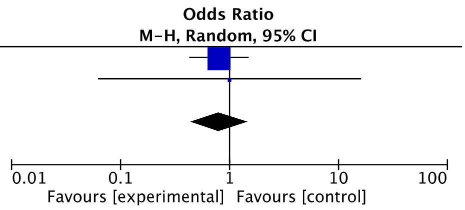


Figure 2: Respiratory distress syndrome within 24 hours of birth

doi: 10.11124/jbisrir-2015-2105

Outcome 3: Sepsis or infection within seven days of birth

Eight studies were found in this review that presented data on neonatal sepsis or infection.[6,30,45,48-51] Infection was commonly defined as any illness requiring treatment with antibiotics. This review placed the limit of diagnosis of infection within seven days of birth; however the majority of studies did not provide a detailed time frame for follow-up.

Results of RCTs included in the review

Woodward and Kelly[30] collected data on neonatal infection for both the RCT and cohort groups within their trial. Infection was defined as neonates requiring treatment with antibiotics. Swabs were taken of the neonate's umbilicus, behind the ear and mouth for culture, sensitivity and microscopy; however no further results were provided. The number of neonates treated with antibiotics is shown in the forest plot below. No further description of these outcomes is provided within the study. Within the RCT, three out of 40 neonates in the waterbirth group developed an infection and one out of 20 in the landbirth group.

Results of cohort studies included in the review

Geissbuehler and colleagues[45] presented data on neonatal infection according to four types; pulmonary, urogenital, ocular and other. The criteria for diagnosing infection were not described. Ocular infections occurred more frequently after landbirths (p=0.022). The results are combined for the purposes of this review. Significantly fewer infections were seen among neonates born underwater. There were a total of 19 infections in the waterbirth group and 42 in the landbirth group.

The cohort group within Woodward and Kelly's[30] trial found one neonate treated with antibiotics from the landbirth group only. No further details were provided.

Zanetti-Dallenbach et al.[50] did not describe their measures for determining infection in neonates post-birth. The authors describe monitoring the neonate for "signs of postpartum infection" without further clarification.[50(p357)] Results found five neonates in the waterbirth group and two in the landbirth group that met their criteria for infection.

Mollamahmotoglu et al.[48] found no evidence of infections among the neonates within their study. Criteria for assessing and diagnosing neonatal infection were not described.

Results of case control studies included in the review

Otigbah and colleagues[6] found no documented cases of neonatal infection and no statistically significant difference for maternal or neonatal outcomes between waterbirths and controls.

Bodner et al.[51] described neonatal infection as a clinical diagnosis requiring antibiotic treatment. The authors stated that this was due to laboratory parameters being insensitive in a newborn. A time frame for diagnoses was not provided. This study found two cases of neonatal antibiotic treatment for neonates born on land. Neither of these cases was detailed within the study. There were no cases in the waterbirth group.

Ros[49] described one neonate who was diagnosed with sepsis after a waterbirth. The infant was diagnosed with congenital pneumonia ten hours after birth and was admitted to the NICU. This neonate is also included in meta-analysis of admission to NICU. No cause for the infection was found. Details on criteria for diagnosis of infection were not described.

doi: 10.11124/jbisrir-2015-2105

Meta-analysis of results for neonatal infection within seven days of birth

Eight studies provided data on neonatal infection after birth including one RCT, four cohort and three case-control studies. Figure 3 shows the results of meta-analysis. Data from Woodward and Kelly's[30] RCT has been included within the below table, but excluded from forest plot. Geissbuehler et al.[45] is the only study to independently reach statistical significance. The Chi^2 test for heterogeneity is significant indicating that results cannot be drawn from this meta-analysis. Figures 4 and 5 show the results of sensitivity analysis. Heterogeneity was improved when case control studies were analyzed together but the Chi^2 test for heterogeneity remained significant for cohort studies. Meta-analysis of case control studies suggests no significant difference in neonatal infection between landbirth and waterbirth groups.

Study or Subgroup	Waterbirth Events	Waterbirth Total	Land birth Events	Land birth Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Bodner et al 2002	0	140	2	140	2.8%	0.13 [0.01, 2.16]
Geissbeuhler et al 2004	19	3017	42	3778	85.2%	0.58 [0.35, 0.97]
Mollamahmutoglu et al 2012	0	207	0	204		Not estimable
Otigbah et al 2000	0	301	0	301		Not estimable
Ros 2009	1	27	0	27	1.4%	7.39 [0.15, 372.38]
Woodward and Kelly 2004a	3	40	1	20		Not estimable
Woodward and Kelly 2004b	0	10	1	10	1.4%	0.14 [0.00, 6.82]
Zanetti-Dallenbach et al 2006	5	89	2	146	9.2%	4.32 [0.92, 20.30]
Total (95% CI)		**3791**		**4606**	**100.0%**	**0.68 [0.43, 1.09]**
Total events	25		47			

Heterogeneity: $Chi^2 = 9.24$, $df = 4$ ($P = 0.06$); $I^2 = 57\%$
Test for overall effect: $Z = 1.61$ ($P = 0.11$)

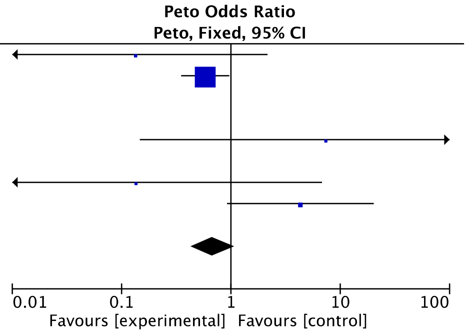


Figure 3: Infection within 7 seven days of birth

Study or Subgroup	Waterbirth Events	Waterbirth Total	Land birth Events	Land birth Total	Weight	Odds Ratio M-H, Random, 95% CI
Geissbeuhler et al 2004	19	3017	42	3778	51.6%	0.56 [0.33, 0.97]
Mollamahmutoglu et al 2012	0	207	0	204		Not estimable
Woodward and Kelly 2004b	0	10	1	10	15.2%	0.30 [0.01, 8.33]
Zanetti-Dallenbach et al 2006	5	89	2	146	33.2%	4.29 [0.81, 22.58]
Total (95% CI)		**3323**		**4138**	**100.0%**	**1.01 [0.22, 4.59]**
Total events	24		45			

Heterogeneity: $Tau^2 = 1.09$; $Chi^2 = 5.41$, $df = 2$ ($P = 0.07$); $I^2 = 63\%$
Test for overall effect: $Z = 0.01$ ($P = 0.99$)

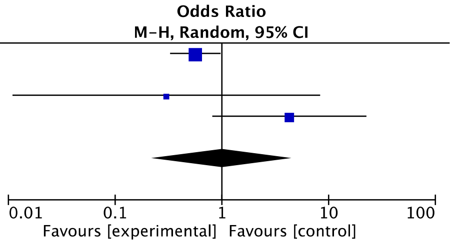


Figure 4: Infection within 7 seven days of birth – cohort studies

Study or Subgroup	Waterbirth Events	Waterbirth Total	Land birth Events	Land birth Total	Weight	Odds Ratio M-H, Random, 95% CI
Bodner et al 2002	0	140	2	140	52.1%	0.20 [0.01, 4.14]
Otigbah et al 2000	0	301	0	301		Not estimable
Ros 2009	1	27	0	27	47.9%	3.11 [0.12, 79.87]
Total (95% CI)		**468**		**468**	**100.0%**	**0.74 [0.05, 11.06]**
Total events	1		2			

Heterogeneity: $Tau^2 = 1.24$; $Chi^2 = 1.48$, $df = 1$ ($P = 0.22$); $I^2 = 33\%$
Test for overall effect: $Z = 0.22$ ($P = 0.83$)

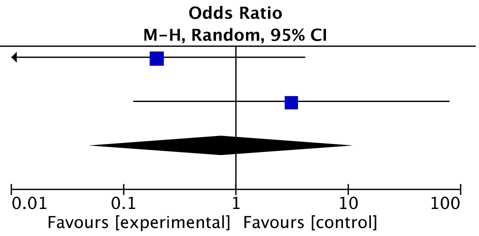


Figure 5: Infection within 7 days of birth – case control studies

Outcome 4: APGAR score at one, five and ten minutes

Results from APGAR scores have been divided according to the method of reporting data. Some studies describe dichotomous data of APGAR scores ≤7 at one or five minutes. Other studies provided data for APGAR scores less than, but not equal to, seven at one and five minutes. Forest plots and meta-analysis has been undertaken for these results. Data was also presented as mean APGAR scores at one and five minutes after birth. In this case, narrative analysis is provided. No study provided data for APGAR scores at ten minutes.

doi: 10.11124/jbisrir-2015-2105

Results for mean APGAR scores at one and five minutes

No RCTs reported results on this outcome.

Results of cohort studies included in the review

Zanetti-Dallenbach et al.[50] provided mean APGAR scores at one and five minutes post-birth and found significantly lower APGAR scores at one minute in a control group of women experiencing instrumental births. However they found no statistically significant difference between the low risk women in the waterbirth group or other control groups (landbirth, with or without temporary water immersion) on the same APGAR measure at one or five minutes post-birth.

Results of case control studies included in the review

Otigbah et al.[6] also calculated mean APGAR scores at one and five minutes post birth and further analyzed these results by separating primiparous and multiparous women. Mean data is provided; however the standard deviation was not available. All the neonates were documented to take their first breath outside of the water.

The above results for mean APGAR scores at one and five minutes are presented below, in Table 3.

Table 3: Mean APGAR scores (standard deviation) at one and five minutes post-birth

Author	1 minute		5 minutes	
	Waterbirth	Landbirth	Waterbirth	Landbirth
Zanetti-Dallenbach et al. (2006)[50]	8.7 (0.08) *n*=89	8.7 (1.0) *n*=146	9.8 (0.5) *n*=89	9.8(0.5) *n*=146
Otigbah et al. (2000)[6] Primiparous *n*=133	8.37	8.47	9.54	9.56
Multiparous *n*=168	8.43	8.55	9.59	9.6

Results for APGAR scores at one minute ≤7

No RCTs reported results on this outcome.

Results of cohort studies included in the review

Geissbuehler et al.[45] provided extensive data to the author of this review to allow calculation of the APGAR scores at one minute for the neonates in their study. The results were not statistically significant; however they indicate a trend towards favoring the waterbirth group. Results show 107/3290 neonates in the waterbirth group and 161/4057 in the landbirth group had an APGAR score ≤7 at one minute.

Results of case control studies included in the review

Menakaya and colleagues[47] found more neonates with APGAR scores of ≤7 at one minute in the waterbirth compared to the landbirth group (RR 3.40, 95% CI 1.5- 7.71, $p<0.05$). There were 25/219 waterbirth neonates and 8/219 neonates in the landbirth group with an APGAR ≤7 at one minute.

Pagano et al.[43] provided further data to that published allowing the authors of this review to calculate the following results. One neonate in the waterbirth group and five neonates in the landbirth group were found to have APGARS ≤7 at one minute post birth.

doi: 10.11124/jbisrir-2015-2105

Meta-analysis of results for APGAR score ≤7 at one minute

Figure 6 presents data from studies measuring APGAR scores ≤ 7 at one minute. The results demonstrate significant heterogeneity, therefore conclusions cannot be drawn from this meta-analysis.

	Waterbirth		Land birth			Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI
Geissbeuhler et al 2004	107	3290	161	4057	43.8%	0.81 [0.63, 1.04]
Menakaya et al 2013	25	219	8	219	37.3%	3.40 [1.50, 7.71]
Pagano 2010	1	110	5	110	18.9%	0.19 [0.02, 1.68]
Total (95% CI)		**3619**		**4386**	**100.0%**	**1.06 [0.31, 3.64]**
Total events	133		174			

Heterogeneity: $Tau^2 = 0.89$; $Chi^2 = 12.69$, df = 2 (P = 0.002); $I^2 = 84\%$
Test for overall effect: Z = 0.09 (P = 0.93)

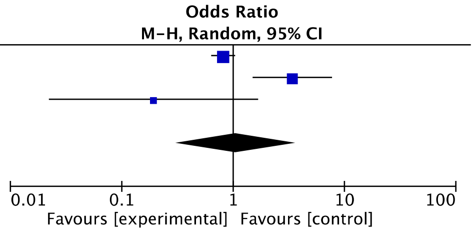


Figure 6: APGAR score ≤7 at 1 minute

Results for APGAR score <7 at one minute

The following studies provided data on APGAR scores <7 at one minute after birth.

No RCTs reported on this outcome.

Results of cohort studies included in the review

Mollamahmutoglu et al.[48] found that APGAR scores were lower in the waterbirth group compared to the conventional landbirth group (without epidural analgesia). They found 26 neonates in the waterbirth group and three in the landbirth group with APGAR scores <7 at one minute.

Results of case control studies included in the review

Bodner et al.[51] studied 140 waterbirths and landbirths in their case-control study. They found two neonates in the waterbirth group and six in the landbirth group had APGAR scores <7 at one minute. The significance was p>0.05.

Ros[49] found the landbirth neonates had lower APGAR scores than those in the waterbirth group at one minute. In the landbirth group, 11% (3/27) had APGAR scores <7 at one minute. No neonates in the waterbirth group had APGAR scores <7.

Meta-analysis of results for APGAR scores <7 at one minute

Figure 7 includes data from one cohort and two case-control studies. The Chi^2 test for heterogeneity is significant indicating that results cannot be drawn from this meta-analysis. Sensitivity analysis was conducted by analyzing only case control studies (figure 8). The Chi^2 test for heterogeneity is not significant and results suggest that there is no difference in this outcome between landbirth and waterbirth groups.

	Waterbirth		Land birth			Peto Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	Peto, Fixed, 95% CI
Bodner et al 2002	2	140	6	140	20.7%	0.36 [0.09, 1.46]
Mollamahmutoglu et al 2012	26	207	3	204	71.7%	5.40 [2.54, 11.48]
Ros 2009	0	27	3	27	7.7%	0.13 [0.01, 1.26]
Total (95% CI)		**374**		**371**	**100.0%**	**2.31 [1.22, 4.37]**
Total events	28		12			

Heterogeneity: $Chi^2 = 17.78$, df = 2 (P = 0.0001); $I^2 = 89\%$
Test for overall effect: Z = 2.57 (P = 0.01)

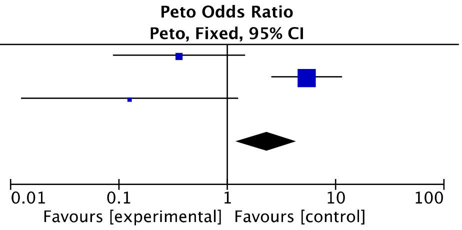


Figure 7: APGAR score <7 at 1 minute

doi: 10.11124/jbisrir-2015-2105

Study or Subgroup	Waterbirth Events	Waterbirth Total	Land birth Events	Land birth Total	Weight	Odds Ratio M-H, Random, 95% CI
Bodner et al 2002	2	140	6	140	77.6%	0.32 [0.06, 1.63]
Ros 2009	0	27	3	27	22.4%	0.13 [0.01, 2.59]
Total (95% CI)		**167**		**167**	**100.0%**	**0.26 [0.06, 1.09]**
Total events	2		9			

Heterogeneity: Tau² = 0.00; Chi² = 0.29, df = 1 (P = 0.59); I² = 0%
Test for overall effect: Z = 1.84 (P = 0.07)

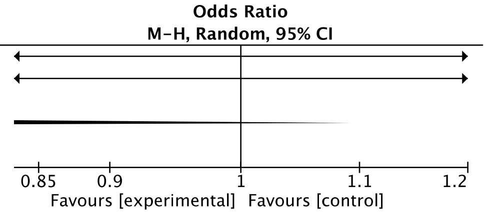


Figure 8. APGAR score <7 at 1 minute – case control studies

Results for APGAR score at five minutes after birth

APGAR scores attended five minutes after birth offer a clearer picture of the neonatal condition after adaptation to extrauterine life. Four observational studies[20,43,45,47] and two RCT[30,46] provide data on this outcome.

Results of RCTs included in this review

Woodward and Kelly[30] described in their methods section their intention to collect data on APGAR scores at one, five and ten minutes after birth; however only data on five minutes after birth was presented. They found one neonate from the waterbirth group with an APGAR score less than 8 in the RCT arm of their trial.

Nikodem[46] found two neonates in the waterbirth group with an APGAR score ≤7 at five minutes after birth.

Meta-analysis of RCTs describing APGAR scores ≤7 at five minutes

Figure 9 presents the meta-analysis of data relating to neonates with APGAR scores ≤7 at five minutes after birth in both RCTs included in this review.[30,46] The wide confidence interval reflects the small sample size of both of these studies and points to the difficulty of undertaking such trials within a maternity setting. The Chi^2 test for heterogeneity is not significant. The results of this meta-analysis are not statistically significant and demonstrate no difference in the outcome of APGAR score ≤7 at five minutes between babies in the waterbirth group compared to the landbirth group.

Study or Subgroup	Waterbirth Events	Waterbirth Total	Land birth Events	Land birth Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Nikodem 1999	2	60	0	60	69.1%	7.52 [0.46, 121.57]
Woodward and Kelly 2004a	1	40	0	20	30.9%	4.48 [0.07, 286.49]
Total (95% CI)		**100**		**80**	**100.0%**	**6.40 [0.63, 64.71]**
Total events	3		0			

Heterogeneity: Chi² = 0.04, df = 1 (P = 0.84); I² = 0%
Test for overall effect: Z = 1.57 (P = 0.12)

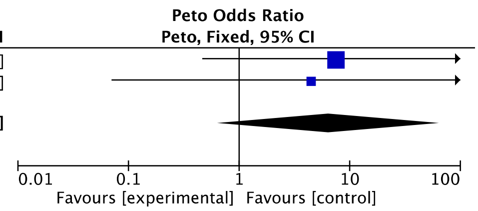


Figure 9: APGAR scores ≤7 at 5 minutes – RCTs

Results of cohort studies included in the review

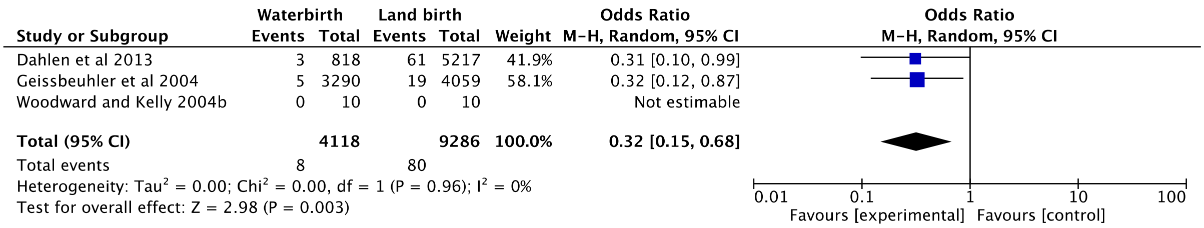
Dahlen et al.[20] provided further data on request relating to APGAR scores ≤7 at five minutes after birth. This study found no difference in five minute APGAR scores ≤7 between waterbirth and any of the other birth positions, excepting recumbent which showed a significantly higher incidence (AOR 4.61, CI 1.29-16.52). This study found 3/818 of neonates born after waterbirth and 61/5217 landbirth neonates had an APGAR score of ≤7 at five minutes. The data included combines the six positions on land into one group.

Geissbuehler et al.[45] provided data and contributed a significant number of participants to the meta-analysis of the APGAR scores of neonates at five minutes. This study found 5/3290 waterbirth neonates and 19/4059 landbirth neonates had an APGAR score ≤7 at 5 minutes. The results show a statistically significant trend to fewer neonates in the waterbirth group with APGAR ≤7 at five minutes post-birth (OR 0.38, 95% CI 0.17-0.85).

Woodward and Kelly[30], in the cohort preference arm of their trial, found no neonates with an APGAR score ≤7 at 5 minutes.

Meta-analysis of results from cohort studies for APGAR score ≤7 at 5 minutes

Meta-analysis has been conducted for cohort studies describing this outcome. Figure 10 presents this meta-analysis. The Chi^2 test for heterogeneity is not significant. Results are statistically significant suggesting neonates born after waterbirth are less likely to have an APGAR score ≤7 at five minutes compared to neonates born after landbirth.



Study or Subgroup	Waterbirth Events	Waterbirth Total	Land birth Events	Land birth Total	Weight	Odds Ratio M-H, Random, 95% CI
Dahlen et al 2013	3	818	61	5217	41.9%	0.31 [0.10, 0.99]
Geissbeuhler et al 2004	5	3290	19	4059	58.1%	0.32 [0.12, 0.87]
Woodward and Kelly 2004b	0	10	0	10		Not estimable
Total (95% CI)		**4118**		**9286**	**100.0%**	**0.32 [0.15, 0.68]**
Total events	8		80			

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 0.00$, df = 1 (P = 0.96); $I^2 = 0\%$
Test for overall effect: Z = 2.98 (P = 0.003)

Figure 10: APGAR scores ≤7 at 5 minutes – cohort studies

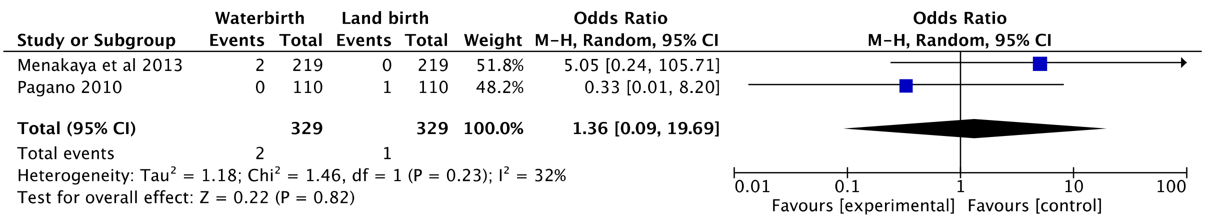
Results of case control studies included in the review

Menakaya et al.[47] found there was no statistical difference between APGAR scores at five minutes post-birth between the two groups. They found two neonates in the waterbirth group and zero in the landbirth group had APGAR scores ≤7 at five minutes.

Pagano et al.[43] also provided additional data to the authors of this review on request. They described one neonate with a five-minute APGAR score ≤7 from the landbirth group within their study. Their results are not statistically significant (OR 0.14, 95% CI 0.00-6.82).

Meta-analysis of results from case control studies included in this review

Figure 11 shows results of meta-analysis of case control studies. The Chi^2 test for heterogeneity is not significant. Individual studies demonstrate effects in different directions and the meta-analysis shows no difference in this outcome between landbirth and waterbirth groups.



Study or Subgroup	Waterbirth Events	Waterbirth Total	Land birth Events	Land birth Total	Weight	Odds Ratio M-H, Random, 95% CI
Menakaya et al 2013	2	219	0	219	51.8%	5.05 [0.24, 105.71]
Pagano 2010	0	110	1	110	48.2%	0.33 [0.01, 8.20]
Total (95% CI)		**329**		**329**	**100.0%**	**1.36 [0.09, 19.69]**
Total events	2		1			

Heterogeneity: $Tau^2 = 1.18$; $Chi^2 = 1.46$, df = 1 (P = 0.23); $I^2 = 32\%$
Test for overall effect: Z = 0.22 (P = 0.82)

Figure 11: APGAR scores ≤7 at 5 minutes – case control studies

Results for APGAR scores <7 at five minutes

Three further studies described APGAR scores <7 at five minutes post-birth, one cohort study[48] and two case control studies.[49,51] No RCTs included in this review described this outcome.

doi: 10.11124/jbisrir-2015-2105

Results of cohort studies included in this review

Mollamahmutoglu and colleagues[48] did not find any neonates with APGARs <7 at this time after birth from either study group.

Results of case control studies included in this review

Ros[49] stated that 4% (1/27) of neonates in the landbirth group had APGAR scores <7 at five minutes post-birth. There were no neonates in the waterbirth group with scores less than 7.

Bodner et al.[51] found one neonate from each group with an APGAR score <7 at five minutes.

Meta-analysis of observational studies describing APGAR scores <7 at five minutes

Figure 12 presents meta-analysis conducted on the outcome of APGAR < 7 at five minutes. The Chi^2 test for heterogeneity is not significant. Results were not statistically significant.

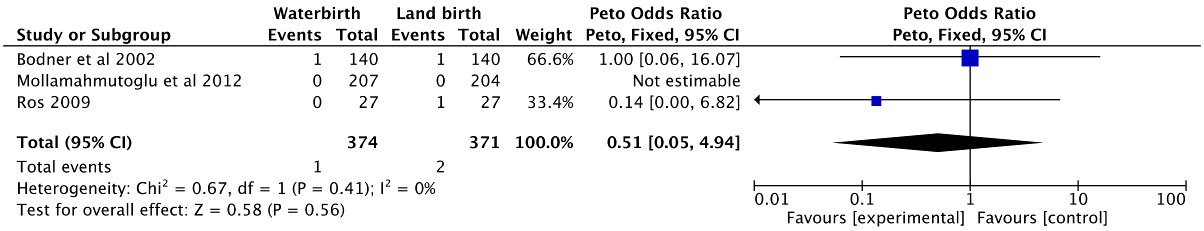


Figure 12: APGAR <7 at 5 minutes

Outcome 5: Admission to Neonatal Intensive Care Unit or Special Care Unit

Special Care Units (SCN or SCBU) and Neonatal Intensive Care Units (NICU) provide varied levels of support to neonates. Special care, or Level II, nurseries provide care to neonates born after 32 weeks or weighing more than 1500g and have access to pediatric specialists providing care to neonates with short term issues that do not need sub-specialist services.[52] Neonatal Intensive Care Units, or Level III, nurseries provide a higher level of care to infants requiring more complex, life sustaining or care for critical illness and provide access to medical and pediatric sub-specialists.[52]

Studies in this review tended to record data on either SCN or NICU according to the facilities available at the place the study was undertaken. This was also determined by individual hospital protocols for each study. Six studies collected data on NICU admission including one RCT which is not included in the meta-analysis.[30,45,46,49,50] One RCT and two case control studies presented data on SCN admission.[6,46,47] Meta-analysis was conducted for both outcomes.

Results for admission to NICU

Results of RCTs included in the review

Woodward and Kelly[30] described five neonates admitted to NICU after waterbirth and one neonate in the landbirth group of this RCT. No further detail or discussion was provided regarding the circumstances or length of stay for these admissions.

Results of cohort studies included in the review

Geissbuehler, Stein and Eberhard found there was "uncomplicated neonatal care" seen equally between waterbirth and landbirth (89.5% compared to 89.1%).[45(p311)] This study consisted of 3/3278 neonates in the waterbirth group and 13/4050 neonates in the landbirth group. Complications that did arise were seen more frequently after landbirth. Admission to NICU was defined as a neonate

doi: 10.11124/jbisrir-2015-2105

requiring transfer to an offsite tertiary NICU center, either from the on-site SCN or birthing room. The authors reported 8.5% of missing data from waterbirth and 6.9 % after landbirth, on this outcome due to mothers returning home within four to six hours after birth. The most frequent causes for NICU admission were suspected or established respiratory distress syndrome, cardiac abnormalities, hyperbilirubinaemia, infection or renal malformation.

Mollamahmutoglu et al.[48] found there was no statistically significant difference to rates of admission to the NICU between groups. This study found five neonates in the waterbirth group and two in the landbirth group were admitted to NICU. Further details of admissions were not provided.

Zanetti-Dallenbach et al.[50] described rates of admission to NICU between the study group and three control groups as comparable and not statistically significant. Data is included here from the waterbirth study group and the group of women experiencing a landbirth without using temporary immersion in water. This study found no neonates in the waterbirth group and five in the landbirth group requiring admission to NICU. No definition or descriptions of the instances of NICU admission were described in this study.

Woodward and Kelly[30] in the cohort arm of their trial describe one neonate from the landbirth group requiring admission to NICU. No further description was provided.

Results of case control studies included in the review

Ros[49] stated that no babies from either group were admitted to NICU immediately after birth; however, one infant was admitted ten hours after birth with sepsis and congenital pneumonia after a waterbirth. Investigations were conducted but no cause for the infection was found.

Meta-analysis of cohort studies for admission to NICU

Meta-Analysis was conducted for the cohort studies describing this outcome. Figure 13 presents this meta-analysis. Chi[2] test for heterogeneity is not significant. Results are not statistically significant suggesting that there is no difference in this outcome between landbirth and waterbirth groups.

	Waterbirth		Land birth			Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI
Geissbeuhler et al 2004	3	3278	13	4050	39.6%	0.28 [0.08, 1.00]
Mollamahmutoglu et al 2012	5	207	2	204	31.7%	2.50 [0.48, 13.04]
Woodward and Kelly 2004b	0	10	1	10	12.9%	0.30 [0.01, 8.33]
Zanetti-Dallenbach et al 2006	0	89	5	146	15.8%	0.14 [0.01, 2.63]
Total (95% CI)		**3584**		**4410**	**100.0%**	**0.51 [0.13, 1.96]**
Total events	8		21			

Heterogeneity: $Tau^2 = 0.78$; $Chi^2 = 5.24$, df = 3 (P = 0.16); $I^2 = 43\%$
Test for overall effect: Z = 0.98 (P = 0.33)

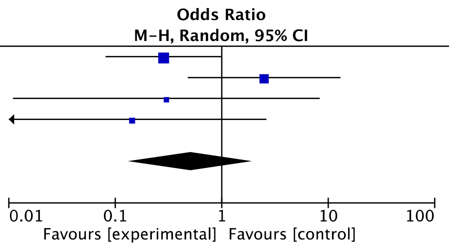


Figure 13: Admission to NICU – cohort studies

Results for admission to SCN

Results for RCTs included in the review

Nikodem[46] did not find any statistically significant differences between the intervention or control group regarding admission to SCN. This study describes three admissions to SCN for the waterbirth group and five for the landbirth group. Details of these admissions were not provided. The results of this RCT are not included in the meta-analysis.

No cohort studies provided results of this outcome.

Results of case control studies included in the review

doi: 10.11124/jbisrir-2015-2105

Menakaya and colleagues[47] describe nine admissions to the special care nursery in their case-control study: eight from the waterbirth group and one from the landbirth group. This was statistically significant (p=0.023). The authors conducted a review of the clinical notes to determine the reasons for admissions to SCN. In the waterbirth group, four neonates were admitted to the SCN for observation: two after active resuscitation, one after an apnoea event and with a birth weight of 2560g, and one after a mild shoulder dystocia described in the notes as “tight shoulders”. These neonates had an average length of stay of one night. One other neonate was admitted to the SCN with respiratory distress, was consequently diagnosed with meconium aspiration and transferred to a tertiary center. The remaining three neonates were admitted for feeding difficulties. One neonate from the landbirth group was admitted to SCN with diagnosed respiratory distress and remained in the SCN for nine days. The data for the number of infants requiring resuscitation and diagnosed with respiratory distress syndrome are also included in the corresponding outcome analysis.

Otigbah and colleagues[6] reported six neonatal admissions to SCN within their study, two from waterbirth and four from the landbirth group. Admissions to the Special Care Baby Unit (SCBU) were prompted by low APGAR scores. One of the neonates in the waterbirth group experienced a difficult birth due to a compound presentation. The other was associated with a true knot in the cord. There was no description for the outcomes for the landbirth control group or length of stay.

Meta-analysis of results describing admission to SCN

Figure 14 includes one RCT, which is not included in the meta-analysis, and two case-control studies, incorporating a random effects model and a Mantel-Haenszell odds ratio. The Chi^2 test for heterogeneity is significant, thus conclusions cannot be drawn from this meta-analysis.

Study or Subgroup	Waterbirth Events	Waterbirth Total	Land birth Events	Land birth Total	Weight	Odds Ratio M-H, Random, 95% CI
Nikodem 1999	3	60	5	60		Not estimable
Otigbah et al 2000	2	301	4	301	52.3%	0.50 [0.09, 2.73]
Menakaya et al 2013	8	219	1	219	47.7%	8.27 [1.02, 66.66]
Total (95% CI)		**520**		**520**	**100.0%**	**1.90 [0.11, 31.45]**
Total events	10		5			

Heterogeneity: $Tau^2 = 3.16$; $Chi^2 = 4.35$, df = 1 (P = 0.04); $I^2 = 77\%$
Test for overall effect: Z = 0.45 (P = 0.65)

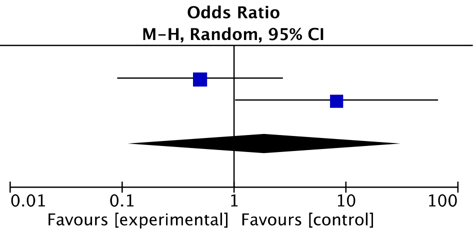


Figure 14: Admission to SCN – observational studies only

Outcome 6: Cord pH values

Cord blood gas analysis is recommended after all high risk births as it can provide valuable information regarding the neonatal condition.[53] Data about the cord pH is one value available as part of this analysis. Cord pH provides an objective measure of neonatal wellbeing, compared to APGAR scores and admission to SCN, both of which are subjective measures and vulnerable to bias.

Cord pH values differ according to whether samples are obtained from the umbilical artery or vein. Accurate pH values can indicate fetal compromise during labor and/or birth, but greater accuracy can be achieved when these results are paired with other data such as base excess or deficit and APGAR scores at birth.[53] Obtaining accurate arterial and venous cord pH is difficult and subject to time constraints and issues with technique.[54] Paired samples from both arterial and venous specimens are recommended for thorough analysis: however some studies in this review do not provide data on both.[53] In general, normal arterial pH values range from 7.10-7.38 and 7.20-7.44 for venous samples from term neonates.[55,56] Values less than this can indicate a compromised neonate.

doi: 10.11124/jbisrir-2015-2105

The manner in which data were provided for umbilical cord pH values varied greatly between studies found in this review. Both continuous and dichotomous data were provided and consequently, direct comparison was not possible. Meta-analysis has been conducted where possible. Otherwise narrative descriptions and a table of results are provided.

Results for cord pH value

Results of RCTs included in the review

Nikodem[46] analyzed umbilical cord gases and presented dichotomous data according to number of cases where arterial cord pH was below 7.2. After birth, the umbilical cord was double clamped and a section of cord was removed. Arterial cord blood was taken from this section. Within the waterbirth group 57 samples were obtained from 60 participants. Twelve samples had a pH value less than 7.2. Within the landbirth group, 54 samples from 60 participants were taken, and 14 had values less than 7.2. This represents 21% of infants in the waterbirth and 24% of landbirth infants. This shows slightly fewer infants had arterial cord blood pH less than 7.2 when infants were born underwater.

In the Woodward and Kelly[30] study, both the RCT and preference arm collected data on cord pH. This study did have some concerns over some midwives not collecting cord gas samples as this was not routinely completed after birth at the time of the study. Cord pH values were provided in a median and range format. The authors could not be contacted to obtain further data. The data provided within the study is described below. The findings did not reach statistical significance.

Results of cohort studies included in the review

No cohort studies provided results on this outcome.

Results of case control studies included in the review

Ros[49] provided data on umbilical cord blood pH; however details were not provided within the study. The author was contacted for further details but did not respond. The author did not clarify if the samples were arterial or venous or a combination of both. The process for collection of umbilical cord gases was not described, apart from stating that the collection of cord blood was delayed within the waterbirth group as the water was drained to below the level of the infant on the woman's chest before the umbilical cord was clamped and cut. This may have influenced results. Also, Ros[49] stated that cord pH was not measured for all participants due to samples clotting or other collection issues, which is a common occurrence when collecting cord blood samples. This impacted on the sample size in relation to this outcome. The mean umbilical cord pH of both groups was the same at 7.26.

Bodner et al.[51] provided dichotomous data on cord pH less than 7.1 for both groups in their case-control study. Within the waterbirth group, three (2.1%) neonates had a cord pH<7.1 as did four (2.9%) neonates in the landbirth group. These results did not reach statistical significance.

Table 4 summarizes the above results.

doi: 10.11124/jbisrir-2015-2105

Table 4: Summary of pH results of included RCT and case control studies included in the review

Study	Waterbirth	Landbirth	Statistical Significance
Ros [48]	Average 7.26 Median 7.27 Range 7.09-7.43 *n*=21	Average 7.26 Median 7.3 Range 7.00-7.36 *n*=18	
Bodner et al.[50]	<7.1 3/140	<7.1 4/140	
Nikodem[45]	Arterial <7.2 12/57	Arterial <7.2 14/54	
Woodward and Kelly[30] RCT arm	Arterial Median 7.23 Range7.04-7.40 *n*=35 Venous Median 7.32 Range 7.15-7.52 *n*=36	Arterial Median 7.18 Range7.04-7.26 *n*=13 Venous Median7.33 Range 7.15-7.42 *n*=16	p=0.15 p=0.59
Woodward and Kelly[30] Preference arm	Arterial Median 7.25 Range 7.16-7.37 *n*=7 Venous Median 7.28 Range 7.25-7.28 *n*= 5	Arterial Median 7.2 Range 7.12-7.3 *n*=7 Venous Median 7.33 Range 7.27-7.49 *n*=10	p=0.336 p=0.72

Results of cohort studies included in the review

Further data on mean arterial and venous cord pH was provided by two studies. Geissbuehler, Stein and Eberhard[45] found that arterial umbilical cord pH was significantly higher for waterbirth than landbirth groups. Mean umbilical cord pH was 7.29 after waterbirth and 7.27 after landbirth. There was no description of the procedure for obtaining umbilical cord blood samples.

Zanetti-Dallenbach[50] also described cord pH levels. Analysis here includes data from the waterbirth group (*n*=89) and the landbirth group not using water immersion in labor (*n*=146). Cord pH analysis was attended immediately after birth, however the process was not described. Zanetti-Dallenbach et al.[50] found significantly lower venous pH measurements in all three control groups compared to the waterbirth group (p<0.05). The authors also described the waterbirth group having no arterial cord pH <7.13. They explained that this might be due to the rigorous inclusion and exclusion criteria employed for women wanting to waterbirth.

Meta-analysis of cohort studies describing cord pH values

Figures 15 and 16 present data on the mean differences for arterial and venous cord pH respectively between neonates born after waterbirth and landbirth incorporating two cohort studies. Both forest plots use a random effects model calculating weighted mean difference. Chi^2 tests for heterogeneity are not significant. Figure 15 shows a statistically significant result (0.02, 95% CI, 0.01-0.02) indicating negligible difference for arterial cord pH between neonates born after waterbirth or landbirth. Similarly figure 16 examining venous cord pH shows a statistically significant though small

doi: 10.11124/jbisrir-2015-2105

difference between mean venous pH for water and landbirth groups (0.03, 95% CI, 0.03-0.03). These differences are not clinically significant.

	Waterbirth			Land birth				Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Random, 95% CI**
Geissbeuhler et al 2004	7.29	0.07	3254	7.27	0.08	4017	90.6%	0.02 [0.02, 0.02]
Zanetti-Dallenbach et al 2006	7.26	0.06	89	7.25	0.08	146	9.4%	0.01 [-0.01, 0.03]
Total (95% CI)			**3343**			**4163**	**100.0%**	**0.02 [0.01, 0.02]**
Heterogeneity: Tau² = 0.00; Chi² = 1.14, df = 1 (P = 0.28); I² = 13%								
Test for overall effect: Z = 6.53 (P < 0.00001)								

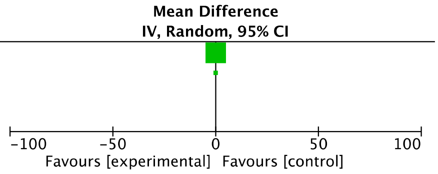


Figure 15: Arterial cord pH

	Waterbirth			Land birth				Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Random, 95% CI**
Geissbeuhler et al 2004	7.4	0.07	3281	7.37	0.07	4048	96.7%	0.03 [0.03, 0.03]
Zanetti-Dallenbach et al 2006	7.38	0.07	89	7.35	0.06	146	3.3%	0.03 [0.01, 0.05]
Total (95% CI)			**3370**			**4194**	**100.0%**	**0.03 [0.03, 0.03]**
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 1.00); I² = 0%								
Test for overall effect: Z = 18.55 (P < 0.00001)								

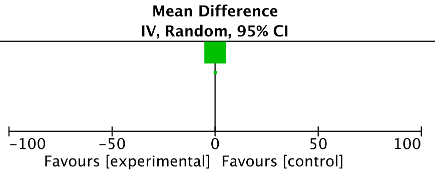


Figure 16: Venous cord pH

Outcomes 7, 8, 9 and 10

Cord avulsion, hyponatremia, hypoxic ischemic encephalopathy and birth injury

Zanetti-Dallenbach et al.[50] described in their discussion that no other negative fetal outcomes parameters such as aspiration, hypoxemia, pulmonary edema or hyponatremia were observed within their study. However, these outcomes were not expressly measured within their study.

Similarly, Woodward and Kelly[30] included cord avulsion as an outcome within both arms of their study. They noted in their discussion that no cord avulsions occurred during the course of their study possibly explained by small sample sizes.

Cord avulsion, hyponatremia and hypoxic ischemic encephalopathy are rare neonatal outcomes and would require large sample sizes in order to both detect and measure appreciable differences according to mode of birth.

Summary

The meta-analyses of 12 studies showed that for the majority of outcomes measured in this review there is little difference between waterbirth and landbirth groups. Meta-analysis was not conducted for mortality within 24 days of birth. Heterogeneity was significant between studies for APGAR scores ≤7 at one minute and admission to Special Care nursery. Sensitivity analysis for case control studies describing infection found results that were not statistically significant (OR 0.74, 95% CI 0.05-11.06). Results of meta-analysis were also not significant for studies describing resuscitation with oxygen (OR 1.12, 95% CI 0.14-8.79) and respiratory distress syndrome (OR 0.81, 95% CI 0.44-1.49). Results comparing APGAR scores ≤7 at five minutes for waterbirth and landbirth groups results for included RCTs demonstrated results that were not statistically significant (OR 6.4, 95% CI 0.63-64.71). However, results for included cohort studies describing APGAR scores ≤7 at five minutes indicate neonates are less likely to have scores ≤7 in the waterbirth group (OR 0.32, 95% 0.15-0.68). Data were not statistically significant for meta-analysis describing admission to NICU (OR 0.51, 95% CI 0.13-1.96) between water and landbirth groups. The differences in arterial (MD 0.02, 95% CI 0.01-0.02) and venous (MD 0.03, 95% CI 0.03-0.03) cord pH, while statistically significant, are clinically negligible.

doi: 10.11124/jbisrir-2015-2105

Discussion

Waterbirth and neonatal outcomes

This review aimed to synthesize the best available evidence regarding neonatal outcomes following waterbirth versus landbirth to advance our understanding of the safety of waterbirth for the neonate. Although small, the randomized control trials contributed to the evidence base informing this clinical issue, and the majority of the research evidence came from observational studies. Given the nature of the intervention it is unlikely that adequately powered randomized controlled trials will ever be possible, thus a meta-analysis of observational studies is appropriate. A recently published systematic review,[57] though comprehensive, did not conduct a methodological assessment or meta-analysis, making this review necessary.

There was significant heterogeneity between studies in this review, limiting the possibility of meta-analysis and justifying the need for further research to expand knowledge in this area. This is supported by Nutter and colleagues[57] who had similar findings. Through a meta-analysis, this review has generated significant sample sizes for some of the outcomes measured. This increases the applicability of results as this study is more likely to be powered to detect a number of important outcomes with more precision, which is a significant limitation of many of the smaller individual studies included here.

Meta-analysis was not conducted for mortality within 24 days of birth due to heterogeneity.

Meta-analysis for neonates requiring resuscitation with oxygen after birth produced results that were not statistically significant (OR 1.12, 95% CI 0.14-8.79) suggesting there is no clinical difference between neonates requiring resuscitation after waterbirth or after landbirth.

APGAR scores were the most commonly reported outcome measure in this review; however there was no consistency in the reporting of this measure. Consequently, it was not possible to combine all datasets. Data on mean APGAR scores at one and five minutes demonstrates no significant difference between waterbirth and landbirth groups.

Meta-Analysis was not conducted for APGAR scores at one minute post-birth due to heterogeneity. However, sensitivity analysis of case control studies suggested meta-analysis was appropriate and produced results that were not statistically significant (OR 0.26, 95% CI 0.06-1.09), suggesting there is no difference between both study groups. Study results across all design methods showed contradictory results, with different effect sizes and direction of results. This may be due to subjectivity in assigning APGAR scores, varied study designs, differences in study participants or chance. Anecdotally, maternity care providers suggest that babies born in water are stimulated less after birth and tend to make a slower transition to establishing respiration, possibly due to the warmth of the water and less handling by caregivers. These results appear to confirm this variability, suggesting that differences in care and stimulation after birth may affect one-minute post-birth APGAR scores.

Varying results are also evident for APGAR scores at five minutes post-birth. Meta-analysis of five-minute APGAR scores of ≤7 in cohort studies provided a statistically significant result favoring waterbirth (OR 0.32, 95% CI 0.15-0.68), while meta-analysis of data from RCTs was not statistically significant (OR 6.40, 95% CI 0.63-64.71). With a total sample size of 14,062 in the cohort studies and 180 in the RCTs, it may be that the RCTs were too under-powered to detect differences in this outcome.

Evaluating umbilical cord pH provides a robust measure of neonatal wellbeing. Meta-analysis in this review provided good sample sizes indicating a negligible difference between neonates born after

doi: 10.11124/jbisrir-2015-2105

waterbirth or landbirth. This is a significant and reassuring result suggesting that waterbirth does not compromise the neonate.

Analyses of data reporting on a variety of neonatal clinical outcomes comparing land with waterbirth do not suggest that outcomes are worse for babies born following waterbirth. Overall this review was limited by heterogeneity between studies and conducting meta-analysis on a number of outcomes was not possible. Waterbirth does not appear to be associated with adverse outcomes for the neonate in a population of low risk women.

Women's experiences

Waterbirth is increasing in popularity though it is still the choice of a small proportion of women. Zanetti-Dallenbach and colleagues[50] found waterbirth represented 1.3% of all births occurring during their study, which is consistent with studies done in other tertiary centers.[17,42] Surveys of birth mode in the United Kingdom suggest waterbirth represents between 0.6-30% of all births, with higher percentages occurring in midwifery-led units.[23,44]

Qualitative studies suggest that waterbirth eases women's pain during labor and birth[11], promotes relaxation and freedom of movement[7,11,58] and provides for a more gentle birth.[11] Some women choose waterbirth to enhance their sense of control and autonomy, and improve their chances for a physiological birth as it is more difficult for caregivers to intervene in a labor and birth that is occurring in a bath.[11,59] Women using water immersion in labor are less likely to use analgesia for labor pain[1,5,6,8,57] and are more satisfied with their labor and birth than other women.[11,46,60] They are also less likely to have an episiotomy although it is not clear how waterbirth might impact on perineal outcomes.[1,6,8,48,50] A large proportion of women experiencing waterbirth enjoy the experience and would choose it again for subsequent births.[60]

While research has not identified any harms associated with water immersion for the mother or baby and some potential benefits[5,15,34,40,57], there is a paucity of high level evidence to guide practice and informed decision making. This has led some professional groups[28,29] to warn against waterbirth, citing lack of evidence, particularly with neonatal outcomes. Waterbirth is more likely to be supported by midwives[6,17,14,20,43,45,47,50,61] as it aligns with the midwifery philosophy of promoting physiological birth, though midwives too have an important role to play in promoting safe practice and informed decision making.[62-65] This review provides evidence that waterbirth is not associated with poor neonatal outcomes, but, as with all studies, this study has some limitations.

Critique

Limitations of the review

This review was limited to studies published in English. The authors are aware of a number of observational studies conducted within this field that have been published in languages other than English. Waterbirth is also not currently described as a Medical Subject Heading (MeSH) term leading to a number of issues during the search process. The subject 'waterbirth' was poorly indexed by databases and despite the reviewers' best efforts some research studies may have been missed by the search strategy.

As with all systematic reviews, this review is limited by the quality of evidence upon which the review is based. Given the thorough search strategy covering multiple databases, this review demonstrates the limited evidence available to inform waterbirth practice. Other systematic reviews confirm this.[2,8,36,57] In using this systematic review as a source to inform practice, it must be noted that this

doi: 10.11124/jbisrir-2015-2105

review draws on data from studies providing evidence comprising mainly cohort (JBI evidence level 3c) and case control (JBI evidence level 3d) designs.[66] Due to the nature of the study design, these studies cannot demonstrate a cause and effect relationship between intervention and outcome.[67]

There are a number of limitations associated with retrospective case control studies, the primary study design of the studies included in this review, the fundamental disadvantage being the lack of control over data collected. Retrospective studies in this review accessed information from hospital birth registers, which provide a selective picture of birth that may be limited, inaccurate, inconsistently measured between subjects and subject to missing data, thereby decreasing the internal validity of individual studies.

Internal validity of included studies

Internal validity refers to the degree to which an outcome may be attributed to the intervention and not to any other causes.[68] Validity is an evaluation of any systematic or random errors within a study. Two systematic issues were identified that may impact upon the internal validity of some of the studies in this review. Firstly, systematic errors may occur where different care is provided to each group generating performance bias.[67] As outlined within the results, studies included in this review provided more or less detailed information about the care and support provided to laboring women according to their study group. One significant difference between study groups is that women experiencing a waterbirth were more likely to be provided one-on-one care by a dedicated midwife which can lead to better outcomes and reduces the likelihood of women using pain medication.[69] Further treatment differences are exemplified by Ros[49] whereby women electing to waterbirth attended a private hospital, while women undertaking landbirth were admitted to a government hospital labor ward.

Secondly, the retrospective study design of the majority of studies in this review may inadvertently select women who successfully complete a waterbirth without any adverse neonatal outcomes. Women who develop obstetric risk profiles would be excluded from observational studies as waterbirth was no longer considered a safe option for birth, whereas women who completed a landbirth and may have had complications could continue to be in the study. This is likely to skew results towards positive outcomes for waterbirth. The RCTs included in this review used an intention to treat design, meaning that women who developed complications and were unable to waterbirth remained in the study, as appropriate, and their outcomes contributed to waterbirth data, even though a waterbirth was not completed.[30,46] Retrospective study designs however would exclude any women who attempt a waterbirth, develop complications and consequently have a landbirth.

Complementary to this, practitioners may experience a level of tolerance to signs of neonatal compromise for women birthing on land, which they would not tolerate for women in the bath. Practitioners are more likely to ask women to leave the bath at any sign of fetal compromise, compared to having more tolerance for these signs in women laboring on land. This again may distort adverse neonatal outcomes to landbirth.

Internal validity of the systematic review

Heterogeneity existed between studies regarding labor augmentation with oxytocin and meconium stained liquor. Four studies[30,48,50,51] included women whose labor was augmented with oxytocin which is associated with uterine hyperstimulation leading to further intervention.[70] Studies suggest fewer women need augmentation with oxytocin when using water in labor.[51] Other studies included women with meconium stained liquor[45,46,50] which is associated with increased neonatal morbidity and mortality.[15] All other included studies named meconium stained liquor in their exclusion criteria,

doi: 10.11124/jbisrir-2015-2105

except Woodward and Kelly[30], which included women with meconium stained liquor; however they were not permitted to use the pool.

There were also few studies assessing some of the most important outcomes of interest. Outcomes such as mortality, resuscitation and other respiratory issues at birth are poorly researched and require further evidence to inform practice. There were no studies reporting on cord avulsion, hyponatremia and hypoxic ischemic encephalopathy. A number of studies reviewed as full text articles included cord avulsion as an outcome; however these studies were excluded from this review for not meeting inclusion criteria.[17,32,42,71]

External validity

While internal validity may be of concern due to the types of studies included in this review, the external validity of this study, or the ability to generalize results of this study to the wider population, is possible.[67] Women included in studies on waterbirth were recruited while pregnant and options for birth were being considered, usually in the third trimester. Women participated in studies at the same time in their pregnancy, at term and in the same condition with low obstetric risks and well babies. While small differences in care protocols may exist between studies, the women were essentially offered the same waterbirth intervention. Also, the studies included in this review were sourced from a wide range of countries.

This review also focused on neonatal outcomes for low risk women only, therefore the results are not applicable to other groups, including women with identified obstetric risk factors in pregnancy.

Implications for practice

This meta-analysis demonstrates there is limited research from which to draw conclusions about a number of outcomes measured in this review. However, the measurement of umbilical cord pH provides an objective measure of neonatal wellbeing and meta-analysis in this review, which indicates there is a negligible difference in cord pH between neonates born after waterbirth or landbirth. Meta-Analysis of APGAR scores ≤7 at five minutes also favor neonates born after waterbirth. This provides reassuring data about the validity of waterbirth as a safe option for birth. There is no evidence to suggest that the practice of waterbirth in a low risk population is harmful to the neonate.

Current research makes arguments both advocating for and discouraging the use of a pool for birth; however this review suggests that waterbirth is a safe and viable option for birth for low risk women. In this light, the choice is dependent on what individual practitioners are willing to provide to the women they care for during labor. Practitioners should be guided by the choices of the women they work with and workplace policy. Women should be encouraged to make informed decisions about waterbirth and be supported by their maternity care provider in this decision.

Implications for research

Current research demonstrates the minimal risks and potential benefits for women who labor and birth in the water; however the picture is less clear for neonatal outcomes. The paucity of evidence in this area calls for additional comprehensive studies to generate more robust data to inform practice. However, while RCTs are both necessary and theoretically possible for women in labor, future RCTs assessing outcomes following waterbirth are unlikely as they are not readily acceptable to laboring women. Observational studies are a more appropriate choice for researchers in this field as they offer a more practical and ethical approach. Selection bias is an inherent risk to observational studies;

doi: 10.11124/jbisrir-2015-2105

however, as a woman's choice and sense of control, or lack thereof, over her experience can impact outcomes, random allocation appears unethical and impractical, and can potentially subject the study to further confounding factors.[67] Despite these limitations, observational studies provide a useful insight and the best level of evidence on outcomes after waterbirth in a manner that is more conducive to the maternity setting. Studies with large sample sizes, preferably prospectively designed, are required and would offer valuable data and further understanding in this field. Further research should also carefully consider criteria for excluding women from the bath if complications develop, such as meconium stained liquor or augmentation with oxytocin.

Conclusion

Waterbirth continues to be offered as a choice to women in labor in response to maternal demand, despite warnings by some professional bodies. Advocates of waterbirth emphasize the advantages of waterbirth including increased maternal autonomy, freedom of movement, pain management and no increased risk to the neonate. On the other hand, others argue against waterbirth citing the potential risks, especially those associated with the neonate, i.e. aspiration, hypoxemia, pulmonary edema and hyponatremia. Most of these objections have not been appropriately evaluated. A lack of sufficient data and a proliferation of opinion further fuel the debate regarding the safety and appropriateness of providing waterbirth as an option for laboring women.

This review incorporates a comprehensive search and accessed good quality studies generating meta-analyses with significant sample sizes thereby increasing the applicability of results. However, data is limited for some outcomes, suggesting further research is required. Note must be made that the women included in this study are pre-selected, low risk women and waterbirth results are only generalizable to this group.

Analyses of data reporting on a variety of neonatal clinical outcomes comparing land with waterbirth do not suggest that outcomes are worse for babies born following waterbirth. Meta-analysis of results for five-minute APGAR scores ≤7 should be treated with caution due to the different direction of results for meta-analysis of data from RCTs and cohort studies.Data measuring cord pH (an objective measure of neonatal wellbeing) are robust and show no difference between groups. Overall this review is limited by heterogeneity between studies and meta-analysis on a number of outcomes was not possible. Waterbirth does not appear to be associated with adverse outcomes for the neonate in a population of low risk women.

Conflict of interest

The reviewers have no conflict of interest to declare.

doi: 10.11124/jbisrir-2015-2105

References

1. Geissbuehler V, Eberhard J. Waterbirths: a comparative study on more than 2000 water births. Fetal Diag Ther. 2000; 15(5): 291-300.
2. Pinette MG, Wax J, Wilson E. The risks of underwater birth. Am J Obstet Gynecol. 2004; 190(5): 1211-5.
3. MIDIRS. The use of water during childbirth. MIDIRS. 2005. [Internet]. [Cited on March 4, 2014] Available from: www.waterbirth.org/assest/documents/MIDIRS%20waterbirth.pdf.
4. Odent M. Birth under water. Lancet. 1982; (2): 1476-7.
5. Thoeni A, Zech N, Moroder L, Ploner F. Review of 1600 water births. Does water birth increase the risk of neonatal infection. J Matern Fetal Neonatal Med. 2005; 17(5): 357-61.
6. Otigbah CH, Dhanjal MK, Harmsworth G, Chard T. A retrospective comparison of water births and conventional vaginal deliveries. Eur J Obstet Gynecol Reprod Biol. 2000; 91(1): 15-20.
7. Maude R, Foureur M. It's beyond water: stories of women's experience of using water for labor and birth. Women and Birth. 2007; 20(1): 17-24.
8. Cluett ER, Burns, E. Immersion in water in labor and birth. Cochrane Database for Systematic Reviews. 2009;(2):CD000111.
9. Jones L, Othman M, Dowswell T, Alfirevic Z, Gates S, Newburn M. Pain management for women in labor: an overview of systematic reviews. Cochrane Database of Systematic Reviews. 2012;3:CD009234.
10. Lukasse M, Rowe R, Townsend J, Knight M, Hollowell J. Immersion in water for pain relief and the risk of intrapartum transfer among low risk nulliparous women: secondary analysis of the Birthplace national prospective cohort study. BMC: Pregnancy Childbirth. 2014;14:60.
11. Richmond H. Women's experience of waterbirth. Pract Midwife. 2003;6(3): 26-31.
12. Richmond H. Theories surrounding waterbirth. Pract Midwife. 2003;6(2): 10-3.
13. Cortes E, Basra R, Kelleher CJ. Waterbirth and pelvic floor injury: a retrospective study and postal survey using ICIQ modular long form questionnaires. Euro J of Obstet Gynecol Reprod Biol. 2011;155(1): 27-30.
14. Garland, D. Waterbirth in the United Kingdom. Midwifery Today Int Midwife. 2000;54(1): 26.
15. National Institute of Clinical Excellence. Intrapartum Care: care of healthy women and their babies during childbirth. London. 2007; NICE. [Internet]. [Cited on March 4, 2014]. Available from: www.nice.org.ul/CGo55.
16. Embry, M. Observation sur un accouchement termaine dans le bain. Ann Soc Med Prat Montpellier. 1805;5(1): 13.
17. Burns E, Boulton M, Cluett E, Cornelius V, Smith L. Characteristics, interventions and outcomes of women who used a birthing pool; a prospective observational study. Birth. 2012;39(3): 192-202.
18. Laws PJ, Lim C, Tracy S, Sullivan EA. Characteristics and practices of birth centres in Australia. Aust N Z J Obstetr Gynaecol. 2009;49(3): 290-295.
19. Meyer SL, Weible CM, Woeber K. Perceptions and practice of waterbirth: a survey of Georgia midwives. J Midwifery Womens Health. 2010;55(1): 55-9.
20. Dahlen HG, Dowling H, Tracy M, Schmied V, Tracy S. Maternal and perinatal outcomes amongst low risk women giving birth in water compared to six birth positions on land. A descriptive cross sectional study in a birth centre over 12 years. Midwifery. 2013;29(7): 759-764.
21. Veltman L, Doherty D. Safety and underwater birth? What every risk manager should know. J Healthc Risk Manage. 2013;32(4): 16-24.
22. Carpenter L, Weston P. Neonatal respiratory consequences from water birth. J Paediatr Child Health. 2012;48(5): 419-423.

doi: 10.11124/jbisrir-2015-2105

23. Gilbert RE, Tookey PA. Perinatal mortality and morbidity among babies delivered in water; surveillance study and postal survey. BMJ. 1999;319(7208): 483-487.
24. Kassim Z, Sellars M, Greenough A. Underwater birth and neonatal respiratory distress. BMJ.2005;330(1): 1071-2.
25. Davies MW. Water births and the research required to assess the benefits versus the harms. J Paediatr Child Health. 2012;48(9): 726-729.
26. Keirse M. Challenging Water Birth? How Wet Can It Get? Birth. 2005;32(4): 318-322.
27. Schroeter K. Water births: a naked emperor. Pediatrics. 2004;114(3): 855-8.
28. Royal Australian and New Zealand College of Obstericians And Gynaecologists. Warm water immersion in labor and birth. College Statement. 2011. [Internet]. [Cited on April 24, 2014]. Available from www.ranzcog.edu.au/college-statements-guidelines.html#obstetrics.
29. American College of Obstetricians and Gynecologists. Immersion in water during labor and delivery. Committee opinion No. 594. Obstet Gynecol. 2014;123(4): 912-5.
30. Woodward J, Kelly SM. A pilot study for randomised controlled trial of water birth versus landbirth. BJOG. 2004;111(6);537-45.
31. Cohain JS. Waterbirth and GBS. Midwifery Today Int Midwife. 2010-2011 Winter;(96): 9-10.
32. Schafer R. Umbilical cord avulsion in waterbirth. J Midwifery Womens Health. 2014;59(1): 91-4.
33. Eckert K, Turnbull D, MacLennan A. Immersion in water in the first stage of labor: a randomised controlled trial. Birth. 2001;28(2): 84-93.
34. Batton DG, Blackmon LR, Adamkin DH, Bell EF, Denson SE, Engle WA. Underwater Births. Pediatrics. 2005;115(1): 1413.
35. Young K, Kruske S. How valid are the common concerns raised against water birth? A focused review of the literature. Women Birth. 2013;26(2): 105-9.
36. Simpson KR. Underwater birth. J Obstet Gynecol Neonatal Nurs. 2013;42(5): 588-94.
37. Chaichian S, Akhlaghi A, Rousta F, Safavi M. Experience of water birth delivery in Iran. Arch Iran Med. 2009;12(5): 469-471.
38. Professor Z Alfirevic FRCOG, Liverpool and Ms D Gould RM. Royal College of Obstetricians and Gynaecologists/Royal College of Midwives Joint Statement No.1: Immersion in water during labor and birth. 2006.
39. Queensland Maternity and Neonatal Clinical Guidelines Program. Normal Birth. 2012. [Internet] [Cited March 3, 2014]. Available from: https://www.health.qld.gov.au/qcg/documents/g_normbirth.pdf.
40. Young K, Kruske S. Water immersion in Queensland. Evidence, Access and Uptake. 2012; Queensland Centre for Mothers and Babies: University of Queensland. [Internet]. [Cited on May 21, 2014]. Available from www.qcmb.org.au/media/pdf/Water%20Immersion%20Report.pdf.
41. Davies R, Davies D, Pearce M, Wong N. The effect of waterbirth on neonatal mortality and morbidity: a systematic review protocol. JBI Syst Rev Impl Reps. 2014;12(7): 89-100
42. Henderson J, Burns EE, Regalia AL, Casarino G, Boulton MG, Smith LA. Laboring women who used a birthing pool in obstetric units in Italy; prospective observational study. BMC Pregnancy Childbirth. 2014;14: 17.
43. Pagano E, De Rota B, Ferrando A, Petrinco M, Merletti F, Gregori D. An economic evaluation of water birth: the cost-effectiveness of mother well-being. J Eval Clin Pract. 2010;16(5): 916-9.
44. Garland D. Collaborative waterbirth audit- 'supporting practice with audit': a collaborative audit between ten units with data collected during 2001. MIDIRS Midwifery digest. 2002;12(4): 508-11.
45. Geissbuehler V, Stein S, Eberhard J. Waterbirths compared with landbirths: an observational study of nine years. J Perinat Med. 2004;32(4): 308-14.
46. Nikodem VC. The effects of water on birth: a randomised controlled trial (PhD. thesis). 1999; Johannesburg, South Africa: Rand Afrikaans University.

doi: 10.11124/jbisrir-2015-2105

47. Menakaya U, Albayati S, Vella E, Fenwick J, Angstetra D. A retrospective comparison of water birth and conventional vaginal birth among women deemed to be low risk in a secondary level hospital in Australia. Women and Birth. 2013;26(1): 114-118.

48. Mollamahmutoglu L, Moraloglu O, Ozyer S, Su FA, Karayalcin R, Hancerlioglu N et.al. The effects of immersion in water on labor, birth and newborn and comparison with epidural analgesia and conventional vaginal delivery. J Turk Ger Gynecol Assoc. 2012;13(1): 45-9.

49. Ros HB. Effects of waterbirths and traditional bedbirths on outcomes for neonates. Curationis. 2009;32(2): 46-52.

50. Zanetti-Daellenbach RA, Lapaire O, Maertens A, Holzgreve W, Hosli I. Water birth, more than a trendy alternative; a prospective, observational study. Arch Gynecol Obstet. 2006;274(6): 355-65.

51. Bodner K, Bodner-Adler B, Wierrani F, Mayerhofer K, Fousek C, Niedermayr A, Grunberger W. Effects of water birth on maternal and neonatal outcomes. Wien Klin Wochenschr. 2002;114(10): 391-5.

52. Committee on the Fetus and Newborn. Levels of Neonatal Care: policy statement. Pediatrics. 2012;130(3): 587-597.

53. Armstrong L, Stenson BJ. Use of umbilical cord blood gas analysis in the assessment of the newborn. Arch Dis Child Fetal Neonatal Ed. 2007;92: F430-34.

54. White CR, Mok T, Doherty DA, Henderson JJ, Newnham JP, Pennell CE. The effect of time, temperature and storage device on umbilical cord blood. J Maternal Fetal Neonatal Med. 2012;25(6): 587-94.

55. King Edward Memorial Hospital. Cord Blood Gases- protocol. [Internet]. [Cited on Nov 15, 2014]. Available from www.kemh.health.wa.gov.au/development/manuals/O&G.guidelines/section6/5/65.14.1.pdf.

56. American College of Obstetricians and Gynecologists. Committee Opinion- Umbilical Cord Blood Gas and Acid-Base Analysis. ACOG Committee Opinion No. 348. Obstet Gynecol. 2006;108:1319-22.

57. Nutter E, Meyer S, Shaw-Battista J, Marowitz A. Waterbirth: an integrative analysis of peer-reviewed literature. J Mid Womens Heal. 2014;59(3): 286-319.

58. Hall S, Holloway I. Staying in control; women's experiences of labor in water. Midwifery. 1998;14: 30-6.

59. Wu CJ, Chung UL. The decision making experience of mothers selecting waterbirth. J Nurs Res. 2003;11(4): 261-8.

60. Forde C, Creighton S, Batty A, Hawdon J, Summers-Ma S, Rideway G. Labor and delivery in the birthing pool. BJM. 1999;7(3): 165-171.

61. Shaw-Battista JC. Optimal outcomes of labor and birth in water compared to standard maternity care (PhD thesis). 2009; San Francisco: University of California.

62. Nursing and Midwifery Board of Australia. Code of Ethics for Midwives in Australia. 2008; Australia; NMBA.

63. Australian College of Midwives. National Midwifery Guidelines for Consultation and Referral. 3rd edition. 2013. [Internet]. [Cited on Dec 1, 2014]. Available from www.midwives.org.au/scripts/cglip.exe/WService=MIDW/ccms.r?pageid=10196.

64. Nursing and Midwifery Board of Australia. Code of Professional Conduct for Midwives in Australia. 2008; Australia; NMBA.

65. International Confederation of Midwives. Philosophy and Model of Care. [Internet]. [Cited on Dec 1, 2014]. Available from www.internationalmidwives.org/assest/uploads/documents/CoreDocuments/CD2005_001%.

66. JBI Levels of Evidence and Grades of Recommendation Working Party. Supporting documentation for the JBI levels of evidence and grades of recommendation. 2014. The Joanna

doi: 10.11124/jbisrir-2015-2105

Briggs Institute. [Internet]. [Cited on Nov 10, 2014]. Available from http://www.joannabriggs.com/assests/docs/approach/Levels-of-Evidence-SupportingDocuments-v2.pdf.

67. Webb P, Bain C. Essential Epidemiology: An introduction for students and health professionals. 2nd Edition. 2013; Cambridge University Press: Cambridge.

68. Carlson MDA, Morrison RS. Study Design, Precision and Validity in Observational Studies. J Palliative Med. 2009;12(1): 77-82

69. Hodnett ED, Gates S, Hofmeyer GJ, Sakala C. Continuous support for women during childbirth. Cochrane Database of Systematic Reviews. 2012.

70. Bugg GJ, Siddiqui F, Thornton JG. Oxytocin versus no treatment or delayed treatment for slow progress in the first stage of spontaneous labor. Cochrane Database of Systematic Reviews. 2013.

71. Burns E. Waterbirth. MIDIRS Midwifery Digest. 2001; 11(supp 2):S10-13.

72. Thoni A, Mussner K, Moroder L. The risk of infection during waterbirths. Midwifery Today. 2004; summer: 46-7.

doi: 10.11124/jbisrir-2015-2105

Appendix I: Search strategy

Database	Date	Search and Limits	Findings	Relevant by title and/or abstract	Included new studies
		(waterbirth* OR "water birth*" OR "underwater delivery" OR "births in water") AND (outcome* OR neonatal outcome* OR fetal OR infant OR newborn OR mortality OR morbidity OR resuscitation OR infection OR APGAR OR admission* OR "cord avulsion" OR hyponatremia OR hypoxi* OR injury)			
CINAHL	5.6.14	Full algorithm, basic search. Limits- 1/1/99-5/6/14	94	17	17
Scopus	5.6.14	Full algorithm keywords in abstract, title or key words. Limits- 1999-2014	21	3	3
MEDLINE	12.6.14	Full algorithm. Limits- 1999-2014, English	71	33	27
PubMed	12.6.14	Full algorithm, NOT drinking. Limits- 1999, English	256	37	6
Health Source Nursing	12.6.14	Full algorithm. Limits- 1999-2014, English	25	4	0
Web of Science	12.6.14	Full algorithm, basic search. Limits- 1999-2014	81	32	0
ProQuest Dissertations and Theses	12.6.14	Limits- English, 1999	174	1	0
Google Scholar	19.6.14	waterbirth OR water birth OR underwater delivery OR immersion AND outcome OR neonatal OR mortality OR morbidity OR resuscitation OR infection OR APGAR OR admission OR avulsion OR hyponatremia OR hypoxia OR injury. Advanced search. Limits- 1999-2014	44	14	4
Trove	19.6.14	Limits- 1999-2014, theses	0	0	0
Cochrane Central Register of Controlled Trials	19.6.14	Limits- 1999-2014, trials	169	2	0
Cochrane Library	19.6.14	Full algorithm. Limits- 1999-2014, all reviews	24	2	0
Hand searched Reference lists	28.6.14			7	7
TOTAL			**959**	**152**	**64**

doi: 10.11124/jbisrir-2015-2105

Appendix II: MAStARI appraisal instruments

JBI Critical Appraisal Checklist for Comparable Cohort/ Case Control

Reviewer _ Date _

Author _ Year _ _ _ _ _ _ _ Record Number _ _ _ _ _ _ _

	Yes	No	Unclear	Not Applicable
1. Is sample representative of patients in the population as a whole?	☐	☐	☐	☐
2. Are the patients at a similar point in the course of their condition/illness?	☐	☐	☐	☐
3. Has bias been minimised in relation to selection of cases and of controls?	☐	☐	☐	☐
4. Are confounding factors identified and strategies to deal with them stated?	☐	☐	☐	☐
5. Are outcomes assessed using objective criteria?	☐	☐	☐	☐
6. Was follow up carried out over a sufficient time period?	☐	☐	☐	☐
7. Were the outcomes of people who withdrew described and included in the analysis?	☐	☐	☐	☐
8. Were outcomes measured in a reliable way?	☐	☐	☐	☐
9. Was appropriate statistical analysis used?	☐	☐	☐	☐

Overall appraisal: Include ☐ Exclude ☐ Seek further info. ☐

Comments (Including reason for exclusion)

___

___

___

doi: 10.11124/jbisrir-2015-2105

JBI Critical Appraisal Checklist for Randomised Control / Pseudo-randomised Trial

Reviewer _ Date _

Author _ Year _ _ _ _ _ _ Record Number _ _ _ _ _ _

	Yes	No	Unclear	Not Applicable
1. Was the assignment to treatment groups truly random?	☐	☐	☐	☐
2. Were participants blinded to treatment allocation?	☐	☐	☐	☐
3. Was allocation to treatment groups concealed from the allocator?	☐	☐	☐	☐
4. Were the outcomes of people who withdrew described and included in the analysis?	☐	☐	☐	☐
5. Were those assessing outcomes blind to the treatment allocation?	☐	☐	☐	☐
6. Were the control and treatment groups comparable at entry?	☐	☐	☐	☐
7. Were groups treated identically other than for the named interventions	☐	☐	☐	☐
8. Were outcomes measured in the same way for all groups?	☐	☐	☐	☐
9. Were outcomes measured in a reliable way?	☐	☐	☐	☐
10. Was appropriate statistical analysis used?	☐	☐	☐	☐
	☐	☐	☐	☐

Overall appraisal: Include ☐ Exclude ☐ Seek further info. ☐

Comments (Including reason for exclusion)

____

____

____

doi: 10.11124/jbisrir-2015-2105

Appendix III: MAStARI data extraction instrument

JBI Data Extraction Form for Experimental / Observational Studies

Reviewer ________________________ Date _________________________

Author _________________________ Year _________________________

Journal_________________________ Record Number__________________

Study Method

RCT	☐	Quasi-RCT	☐	Longitudinal	☐
Retrospective	☐	Observational	☐	Other	☐

Participants

Setting ____

Population ______________________________________

Sample size

Group A __________________ Group B __________________

Interventions

Intervention A ____________________________________

Intervention B ____________________________________

Authors Conclusions:

Reviewers Conclusions:

doi: 10.11124/jbisrir-2015-2105

Study results

Dichotomous data

Outcome	Intervention ( ) number / total number	Intervention ( ) number / total number

Continuous data

Outcome	Intervention ( ) number / total number	Intervention ( ) number / total number

doi: 10.11124/jbisrir-2015-2105

Appendix IV: Included studies

Study	Methods	Participants	Intervention A	Intervention B	Notes
Bodner K, Bodner-Adler B, Wierrani F, Mayerhofer K, Fousek C, Niedermayr A, 2002[50]	case-control study	well women with spontaneous, low risk pregnancies who had a waterbirth matched to next similarly low risk, spontaneous vaginal landbirth	waterbirth	landbirth	focus on maternal outcomes, however measures neonatal outcomes- APGAR scores at 1 and 5 minutes, infection and cord pH <7.1.
Dahlen HG, Dowling H, Tracy M, Schmied V, Tracy S, 2013[20]	descriptive cohort analysing midwifery handwritten notes from each birth over 12 years	6144 low-risk women who birthed at birth centre between January 1996 and April 2008. 60% multiparous, 40% primiparous	waterbirth	landbirth- described data split over six different positions, including all fours, semi-recumbent, lateral, standing, birth stool and squatting	Author's correspondence received for further raw data on APGAR scores.
Geissbuehler V, Stein S, Eberhard J, 2004 [45]	observational cohort study extracting data from standardized questionnaires and documentation	Women with singleton pregnancies with cephalic presentation, at term, without operative births in spontaneous labor from 1991-2000	waterbirth	landbirth	Original data includes women with obstetric complications. Data requested and received by author correspondence for women with no obstetric risk factors
Menakaya U, Albayati S, Vella E, Fenwick J, Angstetra D, 2013[47]	retrospective case-control comparing waterbirth and a matched cohort landbirth	low-risk, well women who had a waterbirth between 2000-2009, and a matched cohort of similar risk women who had a spontaneous vaginal birth on land.	waterbirth	spontaneous vaginal birth on land	Good quality, well designed study. Describes maternal outcomes (PPH and length of labor) and neonatal outcomes (APGAR, SCN admission).

doi: 10.11124/jbisrir-2015-2105

Study	Methods	Participants	Intervention A	Intervention B	Notes
Mollamahmutoglu L, Moraloglu O, Ozyer S, Su FA, Karayalcin R, Hancerlioglu N et al. 2012[48]	prospective cohort of laboring women split into 3 groups according to mode of birth	610 well women with no obstetric risk factors laboring at term (37-42 weeks) between June 2007 and Sept 2008	waterbirth	conventional vaginal landbirth	conventional vaginal births not clearly described, details throughout study. Data for women with epidural analgesia (study group 2, n=191) are not included in this review.
Nikodem VC, 1999[46]	RCT	Women in established labor at term with no obstetric risk factors randomly allocated to either group A or B once fully dilated	waterbirth	spontaneous vaginal landbirth	Thesis-detailed methodology described, thorough analysis, midwifery care carried out by one person. Only second stage conducted in bath.
Otigbah CH, Dhanjal MK, Harmsworth G,Chard T, 2000[6]	Retrospective case-control study comparing waterbirth and normal vaginal birth	Low risk women electing for a water birth, parity matched to low risk women who had a normal vaginal birth on land between 1989 - 1994	waterbirth	Normal vaginal birth on land	Good quality, well described study. Assesses maternal outcomes (perineal trauma, infection, PPH, analgesia use, length of labor) and neonatal outcomes (SCN admission, infection, mean APGAR scores at 1 and 5 minutes).
Pagano E, De Rota B, Ferrando A, Petrinco M, Merletti F, Gregori D, 2010[43]	retrospective case-control study	low-risk women with a spontaneous labor and nil complications.	waterbirth	spontaneous vaginal birth on land	Raw APGAR data requested by personal correspondence. Economic assessment of waterbirth compared to landbirth.

doi: 10.11124/jbisrir-2015-2105

Study	Methods	Participants	Intervention A	Intervention B	Notes
Ros HB, 2009[49]	case control study, purposive sample	purposive sample- well, low risk women, 37-42 weeks pregnant, with no obstetric complications asked to participate in study, self-selecting water birth or landbirth.	waterbirths occurring at two private hospitals	landbirths that occurred in a government hospital	Study group and control group located at different hospital units- high risk of confounding. Author contacted for further clarification of cord pH values, correspondence not received.
Woodward J and Kelly SM, 2004[30]	pilot RCT randomizing women to waterbirth or landbirth.	low-risk women recruited through community midwives, 36-40 weeks gestation	waterbirth	landbirth	considerable cross-over between randomized groups, only 10 women allocated to WB experienced a WB, and 13 women in LB group had a LB.
Woodward J and Kelly SM, 2004[30]	non-randomized preference arm of an RCT trial, women self-selecting either water birth or landbirth, prospective cohort	Women with no obstetric risk factors or pregnancy complications	waterbirth	landbirth	Part of RCT-this data represents 20 women who pre-selected waterbirth or landbirth before labor
Zanetti-Daellenbach RA, Lapaire O, Maertens A, Holzgreve W, Hosli I, 2006[50]	prospective observational study	low-risk, well pregnant women wanting waterbirth between 1998-2002, with no obstetric complications	waterbirth	spontaneous landbirth without temporary immersion	Data reproduced in other articles, this study used as most thorough and containing all relevant information. Data for women experiencing an operative birth excluded from this review

doi: 10.11124/jbisrir-2015-2105

Appendix V: Excluded studies

Burns E. Waterbirth[70]

> **Reason for exclusion:** does not include data from the full study term, confounds water immersion and waterbirth, not clearly described landbirth group participants

Chaichian S, Akhlaghi A, Rousta F, Safavi M.[37] Experience of water birth delivery in Iran

> **Reason for exclusion:** no description of methods, unclear recruitment process, data not supplied for all outcomes studied.

Thoeni A, Zech N, Moroder L, Ploner F.[5] Review of 1600 water births. Does water birth increase the risk of neonatal infection

> **Reason for exclusion:** methods not clearly described, describes 1600 waterbirths but only 737 primiparae reported, not all ourtcomes reported.

Thoni A, Mussner K, Moroder L.[72] The risk of infection during waterbirths

> **Reason for exclusion:** scant information about participants, methods and results provided, not comparable study groups.

doi: 10.11124/jbisrir-2015-2105

Appendix VI: Outcomes reported in included studies

Author Year	**Participants**		**Mortality**		**Resuscitation**		**Sepsis/ Infection**		**APGAR Score ≤7 1 min**		**APGAR score ≤7 5 min**		**Mean APGAR Score**		**NICU Admission**		**SCN Admission**		**Cord pH values**	
	WB	**LB**	**WB**	**LB**	**WB**	**LB**	**WB**	**LB**	**WB**	**LB**	**WB**	**LB**	**WB**	**LB**	**WB**	**LB**	**WB**	**LB**	**WB**	**LB**
Nikodem VC, 1999[46]	60	60	1	0	4	2 RDS n=1					2	0					3	5	<7.2 arterial 12/57	<7.2 arterial 14/54
Woodward J and Kelly SM, 2004[30] (RCT)	40	20					Ab's n=3	Ab's n= 1			1	0			3	1			Arterial-Med 7.23 Range 7.04-7.40 N=35 Venous-Med 7.32 Range 7.15-7.52 N=36	Arterial Med 7.18 Range 7.04-7.26 N=13 Venous-Med 7.33 Range 7.15-7.42 N=16
Woodward J and Kelly SM, 2004[30] (cohort)	10	10					Ab's n= 0	Ab's n=1			0	0			0	1			Arterial-Med 7.25 Range 7.16-7.37 n=7 Venous-Med 7.28 Range 7.25-7.28 n=5	Arterial-Med 7.2 Range 7.12-7.3 n=7 Vanour-Med 7.33 Range 7.27-7.49 n=10.

Dahlen HG, Dowling H, Tracy M, Schmied V, Tracy S, 2013[20]	818	5217									3	61								
Geissbuehler V, Stein S, Eberhard J, 2004 [45]	3299	4076	3/ 3296	5/ 4074	16/ 3023	25/ 3785	19/ 3017	42/ 3778	107/ 3290	161/ 4057	5/ 3290	19/ 4059			3/ 3278	13/ 4050			Arterial mean 7.29 (0.07) n=3254 Venous Mean 7.4 (0.07) n=3281	Arterial mean 7.27 (0.08) n=4017 Venous mean 7.37 (0.07) n=4048
Mollamahmutoglu L, Moraloglu O, Ozyer S, Su FA, Karayalcin R, Hancerlioglu N et al. 2012[48]	207	204	0	0			0	0	<7 n=26	<7 n=3	<7 n=0	<7 n=0			5	2				
Zannetti-Daellenbach RA, Lapaire O, Maertens A, Holzgreve W, Hosli I, 2006[50]	89	146					5	2					1 min 8.7 (0.8) 5 min 9.8 (0.5)	1 min 8.7 (1.0) 5 min 9.8 (0.5)	0	5			Arterial Mean 7.26 (SD 0.06) Venous mean 7.38 (SD 0.07)	Arterial mean 7.25 (SD 0.08) Venous Mean 7.35 (SD 0.06)
Bodner K, Bodner-Adler B, Wierrani F, Mayerhofer K, Fousek C, Niedermayr A, 2002[50]	140	140					0	2	<7 n=2	<7 n=6	<7 n=1	<7 n=1							<7.1 n=3	<7.1 n=4
Menakaya U, Albayati S, Vella E, Fenwick J, Angstetra D, 2013[47]	219	219			2 RDS n=1	0 RDS n=1			25	8	2	0					8	1		

Otigbah CH, Dhanjal MK, Harmsworth G, Chard T, 2000[6]	301	301					0	0					Primip 1 in 8.37 5 min 9.54 Multip 1 min 8.43 5 min 9.59	Primip 1 min 8.47 5 min 9.56 Multip 1 min 8.55 5 min 9.6			2	4		
Pagano E, De Rota B, Ferrando A, Petrinco M, Merletti F, Gregori D, 2010[43]	110	110							1	5	0	1								
Ros HB, 2009[49]	27	27			3	5	1	0	<7 n=0	<7 n=3	<7 n=0	<7 n=0			0	0			Aver 7.26 Med 7.27 Range 7.09-7.43 n=21	Aver 7.26 Med 7.3 Range 7.00-7.36 n=18