No evidence for superiority of air or oxygen for neonatal resuscitation: a meta-analysis

Aucune donnée probante n’appuie la supériorité de l’air ou de l’oxygène pour la réanimation néonatale: une méta-analyse

Joanne Guay, MD • Jean Lachapelle, MD

Abstract

Purpose The aim of this meta-analysis was to re-evaluate the evidence in favour of oxygen or room air as the initial gas mixture for neonatal resuscitation in terms of the following outcomes: death, hypoxic/ischemic encephalopathy, need for tracheal intubation, and APGAR score—Appearance (skin color), Pulse (heart rate), Grimace (reflex irritability), Activity (muscle tone), and Respiration—at five minutes.

Methods A search with no language restriction for all available controlled clinical trials (CCT) was conducted in PUBMED, Cochrane Central Register of Controlled Trials, and EMBASE. Data were extracted independently by the two investigators.

Results Eight CCTs were retained for analysis. They included 1,500 patients, 772 in the oxygen group and 728 in the air group. The evidence is based mainly on quasi-randomized studies (1,311/1,500) with unblinded resuscitators (1,421/1,500). The expertise/training of the resuscitators was unspecified for four of the eight studies. The risk ratio (RR) for death was 1.35 (95% confidence intervals [CI] = 0.97 to 1.88; P = 0.08; I-squared 0%). The RR for hypoxic/ischemic encephalopathy was 1.03 (95% CI = 0.86 to 1.23; P = 0.74; I-squared 0%). The RR for requiring a tracheal intubation was 0.85 (95% CI = 0.69 to 1.05 [random effects model]; P = 0.12; I-squared 9.51%).

Conclusions The literature is insufficient to make any statement regarding the superiority of oxygen or room air as the initial gas mixture for neonatal resuscitation.
In their most recent publication, the American Heart Association recommended that neonatal resuscitation should be initiated with air if blended oxygen is not available. This recommendation is based mainly on the results of two meta-analyses. The first meta-analysis included five studies published during 1993 to 2003. The authors concluded, “There is insufficient evidence at present on which to recommend a policy of using room air over 100% oxygen, or vice versa, for newborn resuscitation.” The second meta-analysis included seven studies published during 1993 to 2005. The authors concluded, “Given the sum of the concerns regarding the methodology and patient population, it would be inappropriate to make definitive recommendations in North America based on the pooling of results from these studies.”

The aim of the present meta-analysis was to re-evaluate the evidence in favour of oxygen or room air as the initial gas mixture for neonatal resuscitation with respect to the following outcomes: death, hypoxic/ischemic encephalopathy, need for tracheal intubation, and APGAR score—appearance (skin color), pulse (heart rate), grimace (reflex irritability), activity (muscle tone), and respiration—at five minutes.

**Methods**

A search with no language restriction was conducted in PUBMED on February 20, 2011 for all available controlled clinical trials (CCT), randomized (RCT) or quasi-randomized which compared oxygen vs air as the initial gas mixture for neonatal resuscitation. The following search terms were used: “Randomis* OR Randomized trial* OR Double blind* OR Placebo* OR Clinical trial* OR Randomized OR Controlled Trial[Publication Type] OR Controlled clinical trial[Publication Type] OR Meta-analysis OR Review OR Systematic Review” [limit to human] AND “Air” AND “Infant, Newborn” AND “Asphyxia Neonatorum” [*therapy] OR “Oxygen Inhalation Therapy” [adverse effects *methods] OR “Resuscitation” [*methods] OR “Retinopathy of Prematurity [epidemiology]” OR “Bronchopulmonary Dysplasia [epidemiology]”. We also searched the Cochrane Central Register of Controlled Trials (Clinical Trials) on May 2, 2011 using the following search terms: “Newborn [Search all text]” AND “Oxygen [Search all text]” NOT “Animal [record title]” for any year and EMBASE [1980 to 2011 Week 17] with “Newborn [limit to human]” AND “Oxygen [limit to human]” AND “Resuscitation [limit to human- include all subheadings]”. Reference lists of all studies as well as those of previous meta-analyses on the same topic were also checked.

Data were extracted from texts, tables, or figures independently by the two investigators as required. Conflicts (n = 1) were resolved by discussion. The latest available outcome up to 28 days was retained (taken earlier if some patients were lost to follow-up). The exact point at which the data were taken is shown in the Table 1. When data were published in more than one report, the available reports were consulted, but the study (not the report) was considered the unit; therefore, no study was considered more than once. Data were analyzed with Comprehensive Meta Analysis version 2.2.044 (www.Meta-Analysis.com) and Review Manager (RevMan) version 5 (for the risk of bias assessment) (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark). Studies were judged on the information contained in the reports without any assumption on the following: adequate sequence generation; allocation concealment (inability of the person who was recruiting the patient to know in advance to what group the patient would be assigned); blinding of the observer for the neurological status examination (hypoxic/ischemic encephalopathy); incomplete outcome data addressed (clear description of the fate of all patients eligible included in the study and of all outcomes); free of selective reporting (outcomes of interest clearly available for all patients included in the study); and free of other bias (any other possible factor that could have influenced the results, including any obvious potential commercial interest). Heterogeneity was assessed by the I-squared value. Numbers needed to treat or harm were calculated on the odds ratios (http://www.nntonline.net/visualrx/). Publication bias was assessed with the classical fail-safe number, the number of missing studies required to bring the P value to 0.05.

**Results**

The flow diagram of the study selection is provided in Fig. 1. The eight studies retained for analysis included 1,500 patients, 772 in the oxygen group and 728 in the air group. The exact distribution per country cannot be given because 18 patients originally reported by Saugstad et al. were counted twice and need to be subtracted; nevertheless, as originally reported, the distribution would be: India = 1,148; Egypt = 121; Philippines = 26; Estonia = 26; Spain = 85; Denmark = 69; Norway = 2; United States = 41. The risk of bias assessment is given in Fig. 2. The characteristics of the studies are detailed in the
### Table 1 Characteristics of included studies

<table>
<thead>
<tr>
<th>Author/Reference</th>
<th>Years of data collection</th>
<th>Country</th>
<th>Population included</th>
<th>Type of study</th>
<th>Resuscitators expertise/training</th>
<th>Resuscitation protocol</th>
<th>Data extracted/Time point</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bajaj 2005&lt;sup&gt;4&lt;/sup&gt;</td>
<td>April 2002-June 2002</td>
<td>India</td>
<td>Oxygen = 97</td>
<td>Quasi-randomized (date of birth)</td>
<td>Unspecified</td>
<td>American Academy of Pediatrics and American Heart Association</td>
<td>APGAR score at five min&lt;br&gt;Death before discharge</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Room air = 107</td>
<td>Presence of hypoxic ischemic encephalopathy and neurological examination confirmed by a blinded observer</td>
<td></td>
<td>Room air group switched to oxygen if HR &lt; 100 beats·min&lt;sup&gt;-1&lt;/sup&gt; or central cyanosis after 90 sec (defined as resuscitation failure)</td>
<td>Any HIE before discharge&lt;br&gt;Need for tracheal intubation during resuscitation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>≥1,000 g plus apnea or gasping respiration and/or HR &lt; 100 beats·min&lt;sup&gt;-1&lt;/sup&gt; requiring positive pressure ventilation; major congenital malformations or hydrops excluded.</td>
<td></td>
<td></td>
<td>Corrugated reservoir connected to a bag (type unspecified), oxygen source 5-6 L·min&lt;sup&gt;-1&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Lundstrom 1995&lt;sup&gt;3&lt;/sup&gt;</td>
<td>November 1, 1990-April 15, 1993</td>
<td>Denmark</td>
<td>Oxygen = 35</td>
<td>RCT</td>
<td>Unspecified</td>
<td>Room air vs oxygen 80%. All infants initially ventilated by bag and face mask (2-4 breaths), followed by CPAP or more face mask and bag ventilation. If abnormal HR at one minute, (\text{FiO}_2) at 30-40% and +10% thereafter as required. Routine vitamin E.</td>
<td>Death within 28 days&lt;br&gt;Death at seven days (the stillbirth of the room air group is included as intention to treat)&lt;br&gt;HIE II or III&lt;br&gt;Need for tracheal intubation not taken because intubation for meconium-stained amniotic fluid was unequally distributed (12 in oxygen group vs 6 in room air group)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Room air = 34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&lt; 33 wks gestational age; known severe malformation and hydrops excluded.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramji 1993&lt;sup&gt;6&lt;/sup&gt;</td>
<td>1993-1995</td>
<td>India</td>
<td>Oxygen = 42</td>
<td>Quasi-randomized (date of birth)</td>
<td>Unspecified</td>
<td>Self-inflated bag and face mask at 60 breaths·min&lt;sup&gt;-1&lt;/sup&gt;, switched to 100% if cyanosed and/or bradycardic &gt; 90 sec. Corrugated reservoir and ≥ 4 L·min&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>APGAR score at five min&lt;br&gt;Death at seven days (the stillbirth of the room air group is included as intention to treat)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Room air = 43</td>
<td>Evaluation of neurological status by a blinded assessor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&gt; 999 g with HR &lt; 80 beats·min&lt;sup&gt;-1&lt;/sup&gt; and/or apnea. Newborn babies with lethal anomalies; hydrops fetalis, or congenital cyanotic heart defects were excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramji 2003&lt;sup&gt;7&lt;/sup&gt;</td>
<td>1995-1997</td>
<td>Four centres from India</td>
<td>Oxygen = 221</td>
<td>Quasi-randomized (date of birth)</td>
<td>Personnel involved in the study recruited by the investigator who also ensured their training and compliance to the treatment protocol during the study period</td>
<td>Self-inflated bag 40-60 breaths·min&lt;sup&gt;-1&lt;/sup&gt; with oxygen reservoir and 4 L·min&lt;sup&gt;-1&lt;/sup&gt;. Air group switched to 100% if HR &lt; 100 beats·min&lt;sup&gt;-1&lt;/sup&gt; and/or cyanosed after 90 sec (defined as treatment failure)</td>
<td>Death at seven days&lt;br&gt;HIE during the first seven days&lt;br&gt;Death at seven days (the stillbirth of the room air group is included as intention to treat)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Room air = 210</td>
<td>Unblinded</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&gt; 1,000 g, HR &lt; 100 beats·min&lt;sup&gt;-1&lt;/sup&gt; and/or apnea, unresponsive to nasopharyngeal suction and tactile stimuli, and having no lethal abnormalities nor congenital pulmonary or cyanotic heart defects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>Reference</td>
<td>Years of data collection</td>
<td>Country</td>
<td>Population included</td>
<td>Type of study</td>
<td>Resuscitators expertise/training</td>
<td>Resuscitation protocol</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------</td>
<td>--------------------------</td>
<td>---------</td>
<td>---------------------</td>
<td>---------------</td>
<td>----------------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Saugstad 1998*</td>
<td></td>
<td>June 1, 1994-May 31, 1996</td>
<td></td>
<td>10 centres from six countries: India, Egypt, Philippines, Estonia, Spain, Norway</td>
<td>Oxygen = 311 Room air = 280</td>
<td>Quasi-randomized (date of birth) Unblinded</td>
<td>Two trained people involved in the study</td>
</tr>
<tr>
<td>Vento 2001</td>
<td></td>
<td>Spain</td>
<td></td>
<td>Oxygen = 21 Room air = 19</td>
<td>Term (37-40 wks) with hypotonia and apnea, nonresponsive to external stimuli, pale skin and mucous colour, and HR &lt; 80 beats-min⁻¹</td>
<td>RCT</td>
<td>Resuscitation team blinded to gas mixture, unclear if assessors were blinded, study designed for measurement of biochemical markers</td>
</tr>
<tr>
<td>Vento 2005</td>
<td></td>
<td>Spain</td>
<td>1999-2002</td>
<td>Oxygen = 22 Room air = 17</td>
<td>Term (37-40 wks) with pale colour, HR &lt; 80 beats-min⁻¹, nonresponsiveness to stimuli, pH ≤ 7.0, APGAR ≤ 5 &gt; five min</td>
<td>RCT</td>
<td>Resuscitation team blinded to gas mixture, unclear if assessors were blinded, study designed for measurement of biochemical markers</td>
</tr>
<tr>
<td>Wang 2008</td>
<td></td>
<td>December 2005-March 2007</td>
<td></td>
<td>Oxygen = 23 Room air = 18</td>
<td>&lt; 32 wks requiring resuscitation; neonates with known congenital malformations or chromosomal anomalies excluded</td>
<td>RCT</td>
<td>Centre A: a pediatric resident, a neonatal fellow, a neonatal nurse, and a respiratory therapist</td>
</tr>
</tbody>
</table>

HR = heart rate; RCT = randomized controlled trial; CPAP = continuous positive airway pressure; HIE = Hypoxia/Ischemic encephalopathy; grade I (mild) includes irritability, hyperalertness, mild hypotonia, and poor sucking; grade II (moderate) includes lethargy, seizures, marked abnormalities of tone, and requirement of tube feeding; and grade III (severe) includes coma, prolonged seizures, severe hypotonia, and failure to maintain spontaneous respiration

* 18 Neonates of this study were erroneously counted twice in the data: 10 in the oxygen group and 8 in the room air group, the numbers provided here are the corrected numbers for mortality; the numbers could not be corrected for the other outcomes

NRP (protocol) = Neonatal Resuscitation Program
Table 1. Four studies were quasi-randomized (allocation according to the date of birth [even vs odd]).\textsuperscript{4,6-8} One study mentioned that neonates for whom spontaneous circulation was never achieved were considered stillbirths and were excluded from the analysis.\textsuperscript{8} For that study, deaths included in the analysis occurred only in non-European countries.\textsuperscript{8} Of the four potentially randomized studies (three were said to be randomized, but the exact method was unspecified), three were not designed for clinical outcomes.\textsuperscript{5,9,10} In the study by Bajaj et al.,\textsuperscript{4} the neurological examination at discharge for the survivors in the room air group is missing for 15 patients, while all such data are given for the survivors in the oxygen group. For this study, an abnormal neurological examination at discharge was found in 13.8% of survivors available for follow-up in the oxygen group and 20.0% of the survivors in the room air group. Data retained for this study were any evidence of hypoxic/ischemic encephalopathy before discharge. No obvious methodological problem was detected for this outcome.

In the oxygen group, the risk ratio (RR) bordered the significance level for an increased risk of death (RR = 1.37; 95% confidence interval [CI] = 1.08 to 1.75; P = 0.01; I-squared 0%). The number needed to harm (NNTH) is 20 (95% CI = 93 to 10), classical fail safe number is 2. Excluding the study where the number of true failed resuscitations (spontaneous circulation never achieved) was not counted,\textsuperscript{8} the RR would be 1.35 (95% CI = 0.97 to 1.88; P = 0.08; I-squared 0%) (Fig. 3). There was no difference in the risk for hypoxic/ischemic encephalopathy (RR = 1.03; 95% CI = 0.86 to 1.23; P = 0.74; I-squared 0%) (Fig. 4). Excluding the study by Bajaj et al.,\textsuperscript{4} the RR would be 1.05 (95% CI = 0.79 to 1.39; P = 0.74; I-squared 0%). Excluding both the Bajaj et al.\textsuperscript{4} and Saugstad et al.\textsuperscript{8} studies, the RR for hypoxic/ischemic encephalopathy would be RR = 1.03 (95% CI = 0.81 to 1.32; P = 0.82; I-squared 0%) (n = 516 patients; oxygen = 263 and room air = 253—all derived from the same authors).\textsuperscript{5,7} The risk
of requiring a tracheal intubation was similar for the two treatment groups (RR 0.85; 95% CI = 0.69 to 1.05 [random effects model]; P = 0.12; I-squared = 9.5%) (Fig. 5). Data for the APGAR score at five minutes could be extracted for only two studies (288 patients); one study favoured oxygen and the other favoured room air. An attempt to sum them would give heterogeneity of 83.6% (I-squared value). A follow-up at 18-24 months (not formally blinded) is available for the study of Saugstad et al. representing 69% of the patients resuscitated with oxygen and 62% of those resuscitated with room air. In those infants with no follow-up, the room air group had a significantly lower heart rate at 90 sec of life than those resuscitated with oxygen (median 100 beats·min⁻¹; 5 to 95% CI = 68 to 147 vs 120 beats·min⁻¹; 5 to 95% CI = 60 to 166). The incidence of cerebral palsy was 7% and 10% for oxygen and room air, respectively (RR = 0.75; 95% CI = 0.31 to 1.80; P = 0.52), and the incidence of abnormal development was 10% and 15% for oxygen and room air, respectively (RR = 0.64; 95% CI = 0.31 to 1.32; P = 0.22).
Discussion

The available literature is clearly insufficient to make any statement regarding the superiority of air or oxygen as the initial gas mixture for neonatal resuscitation. The evidence is based mainly on quasi-randomized studies (1,311/1,500) with unblinded resuscitators (1,421/1,500). The expertise/training of the resuscitators was unspecified for four of the eight studies.4–6,10 One study11 excluded from the count of deaths those neonates for whom a spontaneous circulation was never achieved (true failed resuscitation). This ought to be considered as a major flaw. In an intention-to-treat analysis, those deaths should have been retained in their allocated group for the final analysis. It is possible here that the sicker neonates, those who needed a more vigorous resuscitation, were simply excluded from the results. Excluding this study, no difference in risk of death can be demonstrated between the two treatments (oxygen vs room air). For this study, all recorded deaths occurred in non-European countries.5 In the present meta-analysis, a high percentage of the patients were recruited in India. The perinatal (stillbirth plus early neonatal mortality) death rate in India is 49 per 1,000 pregnancies.13 The perinatal death rate in Canada from 1999 to 2003 was 6.3 per 1,000 (95% CI = 6.2 to 6.4).14 Clearly, there are factors other than the initial gas mixture used for neonatal resuscitation (room air or oxygen) that can make a substantial difference in neonatal mortality. Therefore it is doubtful that the data collected in India apply to Canada and similar countries.

The present meta-analysis contains studies that mix premature and term babies, clearly mixing two different populations with different physiology, different morbidity, and different mortality rate.15 An unstratified analysis may lead to spurious results (Simpson’s paradox). Finally, if the two large studies containing major flaws4,8 are excluded, we then encounter the issue of the “small-study effect”.16 The small-study effect has two major aspects: contribution to heterogeneity and publication bias. Heterogeneity was not a problem in the present meta-analysis. It is well recognized, however, that small studies are more likely to be published when there are positive (statistically significant) rather than negative (not statistically significant) results; this is part of the “publication bias”. To assess the potential influence of this problem, several techniques have been proposed, e.g., the “Trim and Fill” technique. This technique simply imputes values to the potentially missing studies (for instance the small negative ones), thus allowing an estimate of what the effect size would be if these small studies were published and included. Since this technique requires at least about 20 studies, it could not be applied here. Instead, we have used the classical fail-safe number on the effect on mortality. In the room air group, the RR for death was statistically significantly lower only when all studies were included (including those with major flaws). We then see that only two missing studies would be enough to bring this P value to 0.05 (not statistically significant). This shows the extreme uncertainty of this alleged difference in mortality between room air and 100% oxygen.

An appropriate study on the effect of oxygen vs room air on mortality should therefore be conducted in countries with a low perinatal death rate using a stratified analysis (term vs preterm neonates) and a sample size large enough to provide an appropriate power. Considering a perinatal mortality rate of 0.6% in Canada, 39,046 neonates (19,523 per group) would be required to prove or disprove a one-third reduction in mortality (from 0.6% to 0.4%) with an alpha value of 0.05 and a beta value of 0.2 (two-sided test) (http://www.stat.ubc.ca/~rollin/stats/ssize/b2.html).

In the present meta-analysis, there was no difference in the rate of hypoxic/ischemic encephalopathy (RR 1.03; 95% CI = 0.86 to 1.23) among the survivors at the short-term evaluation, suggesting that the two treatment modalities could be equivalent and there would be no ethical contraindications to a large multicentre trial. However, subtle neurological abnormalities are not always easy to detect in the first few months of life. A trend towards a one-third reduction in the incidence of neurological abnormalities (cerebral palsy/developmental delay) at 20 months in the Resair 2 study raises serious concerns on the use of room air.12 The large number of patients lost to follow-up in this study, however, makes these results difficult to interpret. The percentage of patients lost to follow-up was similar across the two treatment groups but twice as high as the percentage of patients suffering from a neurological abnormality. Patients lost to follow-up in the room air group had a lower heart rate at 90 sec than those lost to follow-up in the oxygen group. Therefore, patients lost to follow up in the room air group might have been more severely asphyxiated than those lost to follow-up in the oxygen group. Hence those results may even underestimate the difference between the two groups in the incidence of neurological abnormalities at 20 months.

In North America, the incidence of cerebral palsy is estimated at 1.4/1,000 deliveries.17 Thus, the sample size of a study trying to determine a difference between room air and oxygen in the overall incidence of cerebral palsy would be much higher than the one needed to look for a difference in the death rate. In their study, Wang et al. reported that the incidence of grade III-IV intraventricular hemorrhage was 11% (2/18) in the room air group and 0% (0/23) in the 100% oxygen group, with a similar mortality in both groups (1/18 vs 1/23 for room air and 100% oxygen, respectively).11 Based on this single small study, we can infer that 570 preterm neonates would be required to attempt to eliminate a difference of 5% (alpha value of
0.05, beta value of 0.2, and two-sided test) in the incidence of severe intraventricular hemorrhage. However, considering that neither Wang et al. nor Escrig et al. were able to stabilize premature newborns with room air, we are of the view that room air should not be the initial gas mixture used for resuscitation of premature newborns.\textsuperscript{11,18}

The trend towards a reduction in the number of patients who require tracheal intubation (RR 0.85; 95\% CI = 0.69 to 1.05) if oxygen is used should be interpreted cautiously because the expertise/training of the resuscitators was not mentioned in half of the studies, and the resuscitators were blinded to the gas mixture for only two studies.

First reported in 1952, the APGAR score has gone through various periods of favour and disfavour; it has been criticized as having a poor inter-rater reliability and a poor calibration when used as an isolated criterion to predict mortality and long-term morbidity, particularly in preterm newborns.\textsuperscript{19} The relationship between the five minute APGAR score and long-term neurological outcome is too fragile to base any evaluation of a resuscitation protocol on it, and future studies should not consider the five minute APGAR score as a final outcome. Finally, even if air were chosen as the first gas mixture, due to the high number of neonates who were switched to oxygen during these studies (approximately 25\%), supplementary oxygen should continue to be made available.\textsuperscript{2,3}

In conclusion, the literature is insufficient to make any statement regarding the superiority of oxygen or room air as the initial gas mixture for neonatal resuscitation.

Disclosure Access to some of the references was provided by University of Montreal. We have no other relevant disclosure.

Competing interests None declared.

References