UvA-DARE (Digital Academic Repository)

Long-term follow-up of obstetric studies
Teune, M.J.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 7

Different follow-up strategies after obstetric studies

Margreet J Teune, Janneke van ’t Hooft, Aleid G van Wassenaer,
Ben Willem J Mol, Brent C Opmeer
Abstract

A main aim of obstetric interventions is to improve the long-term outcome of the child. Unfortunately, the large majority of studies on obstetric interventions does not report such outcomes, as it is expensive, time consuming and a logistic challenge to measure them.

We systematically discuss pros and cons of three different follow-up strategies, and propose a rule of thumb that can facilitate perinatal trialist in their decisions on follow-up. We distinguish three different follow-up strategies: 1. Long-term follow-up of all children participating in the original RCT; 2. Only long-term follow-up of children with abnormal short-term outcomes; and 3. Only long-term follow-up of a random sample of children participating in the original RCT.

We discussed strengths and weaknesses of these strategies and alternative options will be suggested. It is pivotal for trialists to consider follow-up at the design of the original trial. To facilitate future trialist we summarized pros and cons of different strategies and provided a brief guideline to be of help in their decision on long-term follow-up.
Introduction

Obstetric randomized controlled trials (RCTs) are performed to evaluate the effectiveness of perinatal interventions. At the start of such trials, trialists have to determine the primary outcome. Although many perinatal interventions are performed with the ultimate aim to improve the long-term health and development of the child, the large majority of studies on obstetric interventions does not report such outcomes, as it is expensive, time consuming and a logistic challenge. In a recently published review, we showed that only a small minority of large obstetric RCTs (> 350 women) followed the children after discharge from the hospital to evaluate the long-term effectiveness of a perinatal intervention. This review also showed that there is no consistency in the strategy used for the follow-up of these children. Not every study approached all live born children included in the RCT for follow-up and there was also no consistency in the methods used for the follow-up of these children.

Most apparently, a prespecified and early defined strategy for long-term follow-up was lacking, even in those studies that reported on follow-up. If studies did report on follow-up, the plan for follow-up was rarely announced in the initial study protocol, and apparently only conceived during or after the primary study. This hampers the use of adequate arrangements for informed consent, and results often in incomplete follow-up. In a reaction, recent studies have now planned a complete follow-up of all studies, sometimes even leading to designs where short-term outcomes are only reported together with long-term outcomes, as the allocation remain undisclosed until the primary endpoint is known, sometimes taking two year follow-up or even longer.

In this paper, we aim to formulate different concepts for follow-up strategies. We based ourselves on strategies that already have been used in large obstetric RCTs that followed the children long-term after discharge of the hospital. Strengths and weaknesses of these strategies will be discussed and alternative options will be suggested. Also the different follow-up methods used (questionnaires and/or assessments) will shortly be evaluated. In addition of promoting awareness and emphasizing the importance of appropriate consideration and planning of long-term outcomes assessment, the current study could facilitate future trialists in their decision making on follow-up.

Potential follow-up strategies

We based ourselves on our review on long-term outcomes. Only a few studies systematically incorporated long-term child outcomes in their evaluation. Among these studies, we identified three clearly different follow-up strategies, and selected the studies most prominently reporting their follow-up strategy.
These different strategies can be characterised as:
I. Long-term follow-up of all children participating in the original RCT.
II. Only long-term follow-up of children with abnormal short-term outcomes.
III. Only long-term follow-up of random sample of children participating in the original RCT.

Each of these strategies will be illustrated using a clinical study as an example. Subsequently, strength and weaknesses of the strategy will be discussed and alternative options will be suggested. Finally, a brief guideline is presented that can be used to select an appropriate strategy conditional on the characteristics of the study at hand, and recommendations are provided on how to use the guideline in the design phase of a trial.

**Strategy I: Long-term follow-up of all children participating in the original RCT.**

*Clinical example: The ORACLE I study* 4,5

**Objective of the study**
The ORACLE I study (n=4826 women) is an example of a large RCT in which all children were followed long-term regardless the results of the short-term outcomes. This study evaluated broad spectrum antibiotics (erythromycin and/or co-amoxiclav or both or placebo) in women with preterm, prelabour rupture of fetal membranes (pPROM). The composite primary outcome was a composite of death before discharge from hospital; chronic lung disease; or a major cerebral abnormality on ultrasonography before discharge.

**Short-term results**
The results of this study indicated a range of short-term health benefits with the prescription of erythromycin for pPROM. Co-amoxiclav was not associated with any change in primary outcome. However, although women who received co-amoxiclav had prolonged pregnancy compared to those who did not receive any co-amoxiclav, it was also associated with a significantly higher incidence of necrotising enterocolitis and was therefore not recommended.

**Follow-up strategy**
Based on the short-term outcomes the trialist expected a probable reduction in childhood disability. In the original trial, women were already informed of the intention to do a subsequent follow-up assessment when they gave written informed consent. After the original trial was finished and short-term outcomes were known, the trialist sought follow-up information for children at 7 years of age who were born to the 4809 women with pPROM who completed ORACLE I. The follow-up rate was 75%.

**Follow-up methods used**
Data were collected with a structured parental questionnaire to assess the child’s health status. Furthermore, results from national curriculum tests were used to assess the children’s educational attainment.


**Long-term results**

The long-term results showed that the prescription of antibiotics for women with pPROM seems to have no effect on the health of children at 7 years of age and therefore should not be given to the mother in the case of premature rupture of membranes. This study showed, that the overall conclusion about the effectiveness of a perinatal intervention can change, after long-term outcomes have become known.

**Strengths, weaknesses & suggestions**

_**Follow-up strategy: strengths, weaknesses & suggestions**_

The strength of following all children on the long-term, like in the ORACLE I study, is that the effects of a specific perinatal intervention can be determined without losing much information. On the other hand, this strategy is time consuming and consequently very expensive. Another strength is that the trialist already informed the participants at the start of the original study of the intention to do a subsequent follow-up. Nonetheless, the researchers had to trace all children with the help of the UK Office of National Statistics and by contact with their family doctors. Maybe it would have saved time if the trialist maintained contact with the parents immediately from the inclusion and perhaps this would have resulted in a follow-up rate higher than 75%.

_**Follow-up methods used: strengths, weaknesses & suggestions**_

In the ORACLE study follow-up data were collected with a structured parental questionnaire to assess the child’s health status. Questionnaires have the advantage that they are relatively inexpensive and easy to organise. A disadvantage of questionnaires in the follow-up of children is that they are often designed as developmental screening tools and therefore are less useful to detect mild problems, that tend to be common in preterm born children. Other disadvantages are that parents frequently over- and under report the health status of their child and that response rates vary considerably between studies. Instead of using questionnaires as sole outcome measure, an alternative strategy could be a step up approach. An example of this approach is to use the questionnaire to identify the children with a high risk for a developmental problem. The next step could be an invitation of this group of children for further physical and neurological assessments, like the Bayley scales of infant development. Furthermore, the ORACLE I study used the results from national curriculum tests to assess the children’s educational attainment. This information was already collected and no extra money had to be spent on IQ assessments. We think that trialist should more often consider the use of regularly collected information. For example, in The Netherlands the CITO test is administered in every child at the end of primary school. Based on this school performance test, recommendations are formulated for an appropriate-secondary school level. The result of this test could provide valid information to evaluate the effect of perinatal interventions on school performance.
**Strategy II: Only long-term follow-up of children with abnormal short-term outcomes.**

*Clinical example: the DUBLIN study* ⁹, ¹⁰

**Objective of the study**

The Dublin study is an example of a large RCT, where only children with abnormal short-term outcomes were followed long-term (n=12,964 women, n=13,079 children). This study was about intrapartum fetal heart rate monitoring in which a strategy of continuous electronic intrapartum fetal heart monitoring was compared with an alternative strategy of intermittent auscultation. During both strategies it was allowed to measure fetal scalp blood pH if there was doubt about the condition of the baby. Primary outcomes were intrapartum and neonatal deaths (within 28 days of delivery) and neurological abnormalities (seizures, simultaneous abnormalities of both tone and reflexes, other neurological abnormalities that persisted one week after birth).

**Short-term results**

More than three times as many fetal blood samples were obtained for fetal heart rate abnormalities in the electronic fetal heart rate monitoring group. The overall risk of intrapartum and neonatal death in children without neurological problems was identical in both groups (2.1/1000). However, children allocated to the electronic fetal monitoring group had a lower risk for neurological abnormalities compared to children allocated to the intermittent auscultation group (OR 0.67, 95% C.I 0.43-1.04), a non-significant difference. When only seizures were considered, the risk of seizures was significantly lower for children in the electronic fetal monitoring group (OR 0.45, 95% C.I. 0.22-0.91).

**Follow-up strategy**

Follow-up at one year of age was planned for surviving infants who had neonatal seizures or other simultaneous abnormalities of tone and reflexes in the neonatal phase. Nevertheless, since the age of one year is too young to exclude a diagnosis of cerebral palsy with confidence, the follow-up was extended to the age of four for all children with neurological abnormalities after birth. Cases of cerebral palsy among babies who did not have abnormal neurological signs in the neonatal period and who therefore did not have been followed systematically, were identified at the time of referral to specialist clinics.

**Follow-up methods used**

At one year of age a general assessment of posture and large movements, vision and fine movements, hearing, speech, and social behaviour was performed. The re-examinations at age four included psychological and neurological assessments as well as test of hearing, sight, and speech when indicated.

**Long-term results**

Follow-up at one year has been successfully accomplished in 42 of the 43 children babies who survived seizures and other simultaneous abnormalities of tone and reflexes in the neonatal period. At this examination, three babies in each trial group, all of whom who
experienced neonatal seizures, were judged to have major neurological abnormalities, including cerebral palsy. Reassessment at the age of four of 9 children in the intensively monitored group and 21 in the control group who survived after neonatal seizures showed that 3 children in each group had cerebral palsy. A fourth child in the intensively monitored group with cerebral palsy showed transient abnormal neurological signs during the neonatal period. Eight other children in the intensively monitored group and 7 in the control group who showed no abnormal neurological signs in the neonatal period also developed cerebral palsy. Sixteen of the total of 22 cases of cerebral palsy had not shown clinical signs suggestive of intrapartum asphyxia. This study concluded that continuous electronic monitoring decreases the risk for seizures, but not for cerebral palsy.

**Strengths, weaknesses & suggestions**

*Follow-up strategy: strengths, weaknesses & suggestions*
A possible strength of this study is that a subset of children was systematically selected for long-term follow-up. Taking the pathophysiological background into account, the trialists decided at the start of the original study to follow only children with neurological abnormalities long-term. This selection was based on the hypothesis that neonatal asphyxia could cause neurological abnormalities in the new-born and that these infants have a higher risk to develop cerebral palsy. Nevertheless, this can also be a weakness. This study started in 1981 and since that time a lot more is known of the relationship between neonatal asphyxia and long-term outcomes. It seems that children with neonatal asphyxia are at increased risk for long-term intellectual, verbal and motor deficits and these children show often no short-term neurological abnormalities \(^{11,12,13}\). Therefore, if the considerations about the pathophysiological background are incorrect, the possibility arises that the wrong follow-up strategy and method will be selected. Because of this, children without short-term neurological abnormalities, but with long-term intellectual, verbal and motor deficits are missed in the DUBLIN study because they were not followed long-term. Another strength is that the authors decided to extend the follow-up to the age of four which is legitimated by the fact that the age of one is too young an age at which to exclude a diagnosis with cerebral palsy with confidence. Nevertheless, they could have considered not to assess the children at the age of one and only at the age of four at the start of the original trial because cerebral palsy was their primary long-term outcome. A strength is also that the trialist were aware of the fact that children without short-term neurological abnormalities could also develop cerebral palsy. They have overcome this problem by searching for these when they were referred to specialist clinics.

*Follow-up methods used: strengths, weaknesses & suggestions*
A strength of this study is that they collected information through psychological and neurological assessments at one and four year of age in babies who survived seizures and other simultaneous abnormalities of tone and reflexes in the neonatal period. As mentioned above, through such assessments also mild neurological problems can be
detected. If the relation between asphyxia and increased risk for long-term intellectual, verbal and motor deficits was already known at the start of the study, the researchers perhaps would have selected other assessments, as for example an IQ test, in their follow-up protocol.

**Strategy III**: Only long-term follow-up of random sample of children participating in the original RCT.

*Clinical example: the HUNGARIAN study* ¹⁴,¹⁵

**Study objective**
The HUNGARIAN study is an example of a large obstetric RCT in which only a random sample of infants is followed long-term (n=4753). This RCT evaluated the use of periconceptional multivitamin supplementation in reducing the incidence of a first occurrence of neural-tube defects. Women planning a - in most cases their first - pregnancy were randomly assigned to receive a single tablet of a vitamin supplement (containing 12 vitamins, including 0.8 mg of folic acid; 4 minerals; and 3 trace elements) or a trace-element supplement (containing copper, manganese, zinc, and a very low dose of vitamin C) daily for at least one month before conception and until the date of the second missed menstrual period or later.

**Short-term results**
Pregnancy was confirmed in 4,753 women. Six cases of neural-tube defects were reported in the group receiving the trace-element supplement, as compared with none in the vitamin-supplement group (*P* = 0.029). Given these results, the authors recommend that all women planning pregnancy should receive a vitamin supplement including folic acid.

**Follow-up strategy**
Because of this effectiveness and the subsequently expected widespread use of periconceptional folic acid/multivitamin supplementation, impacts of this new primary preventive method had to be carefully evaluated and follow-up was planned at 2 and 6 years of age. Four hundred children in the intervention group and 400 control children were randomly selected with equal sex distribution for the age group of 2 and 6 years of age. No power calculation is mentioned in the article.

**Follow-up methods used**
Children were examined by the following methods: evaluation of available medical records, general paediatric examination (including neurological examination), anthropometric examination, audiological and ophthalmological examination, mental development at two years of age (Brunet-Lezine method) and IQ at 6 years of age (Binet test).

**Long-term results**
Of 800 invited children, 625 children participated in the long-term follow-up examinations. The results showed no differences in the rate and distribution of disorders including allergies (except otitis media), audiological and ophthalmological anomalies, anthropometric and
mental development and IQ between children who belonged to the multivitamin or trace element group.

**Strengths, weaknesses & suggestions**

*Follow-up strategy: strengths, weaknesses & suggestions*

A strength of this study is that the trialist decided to perform long-term follow-up because they were aware of the fact that the impacts of this new widespread primary preventive method had to be carefully evaluated. It is not clear whether the trialists decided to perform long-term follow-up at the start of the original trial. The disadvantage of not planning follow-up at the start of the trial, is that it will hampers the use of adequate arrangements for informed consent, and therefore often results in incomplete follow-up. Furthermore, there is an economical advantage of choosing only a random sample for follow-up. With this follow-up strategy money is saved because fewer children have to undergo an assessment. Nevertheless, this strategy can also be a weakness because enough children have to be included in the follow-up to detect a significant difference between the long-term outcomes of the two groups of the study. In the article the authors didn’t mention a power calculation for the long-term outcomes.

*Follow-up methods used: strengths, weaknesses & suggestions*

A strength in this study is the follow-up method that is used. Not only parental questionnaires were used, but also developmental and IQ test were assessed.

**Discussion**

We discussed three different follow-up strategies, and illustrated each with an example of a large obstetric RCT followed by child follow-up. In summary, the decision whether follow-up is necessary and what strategy is most appropriate, seems to come down to four considerations:

I. Whether there is a relation between short-term and long-term outcomes.
II. Whether questionnaire based information is sufficient or whether physical assessments are required in all or part of the follow-up sample.
III. What health information is required and the age at which the relevant outcomes can be appropriately measured.
IV. Whether the decision based on short-term results would change current practice.

All three strategies discussed presume that knowledge on the association between short-term and long-term outcomes is lacking, otherwise extrapolation of short-term findings to the long-term would suffice. At present, this is the case for most short-term
perinatal outcomes, so for the time being empirical long-term follow-up measurements should be seriously considered for trials on perinatal interventions.

When more information about long-term child outcomes becomes available, one can explore the association between short-term and long-term child outcomes. These associations can then be used to develop prediction models for different long-term child outcomes \(^\textsuperscript{16,17}\). As far as we know, such prediction models for long-term child outcome do not exist, but these models have the potential to be a good alternative strategy for follow-up, especially when there are time and money restraints. Nevertheless, the development of such models requires large birth cohorts, in which data surrounding pregnancy, delivery and short-term outcomes are available, as well as follow-up data on various health related outcomes. More long-term data is even required for the external validation of these models because prediction models are known to be optimised for the specific dataset in which they have been derived and external validation is required to correct for this overfit bias \(^\textsuperscript{18}\). Acquiring this long-term information is expensive, but can be cost-effective on the long-term. By using the models for extrapolation of short-term outcomes to a long-term horizon, less children have to be followed long-term.

We advise every trialist to consider long-term follow-up at the start of the original randomized controlled trial (table 1). Depending on the type of intervention trialists have to decide if long-term follow-up is necessary for fully evaluating the consequences of the intervention. If so, one has to determine which long-term outcome(s) are relevant, taking the pathophysiological mechanism of the intervention into account. In addition, the choice for particular long-term outcomes may also have implications for the timing of the follow-up, e.g. at what age the follow-up has to take place. For example, IQ can only be measured accurately from the age of 5. Thereafter, the decision has to be made how many children have to be followed long-term to receive an answer to the primary study question. A power calculation for the long-term outcome should be considered.
### Table 1: Different strategies for long-term child follow-up

#### Start of original RCT

1. Is long-term child follow-up relevant?  
   **Considerations:** Type of intervention; does the intervention have the potency to influence long-term outcomes, taking the pathophysiological background of the intervention into account.

   - Yes → question 2
   - No → no planning of follow-up

2. Determine long-term outcome(s)  
   **Considerations:** Again, take the pathophysiological background of the intervention into account. Does the intervention influence growth, hearing or vision, neurological development, IQ etc.

3. Determine the duration of follow-up  
   **Considerations:** Duration of follow-up depends on the long-term outcome. For example, a definitive diagnosis of cerebral palsy can be made at the age of two, IQ can be measured adequate at the age of 5 etc.

4. Determine how many children have to be followed long-term  
   **Considerations:** Three options:
   - A: follow all children long-term
   - B: follow only children long-term with abnormal short-term outcomes
   - C: follow a random sample of children long-term

   - Consider a power calculation for long-term outcome(s)
   - Consider also no long-term follow-up if no differences in short-term outcomes are found in both arms of the study and current clinical practice will not change.

5. Determine which follow-up methods will be used  
   **Considerations:** The optimal method depends on the type of the long-term outcome(s): physical examinations, neurological examination, questionnaires, visual and hearing assessments, neurological assessments (motor development, mental development), IQ test etc.
Chapter 7

References


