*Paediatric and Perinatal Drug Therapy, 2005; 6 (4)*

# Analgesic trials in neonates: observations, pitfalls and recommendations

**Monique van Dijk[1], Dick Tibboel[1], John van den Anker[2,3], Sinno Simons[1]**

[1]*Department of Paediatric Surgery, Erasmus MC – Sophia Children's Hospital, Rotterdam, The Netherlands*

[2]*Department of Paediatrics, Erasmus MC – Sophia Children's Hospital, Rotterdam, The Netherlands*

[3]*Division of Paediatric Clinical Pharmacology, Children's National Medical Center, Washington, DC, USA*

Corresponding author

*Monique van Dijk, Department of Paediatric Surgery, Erasmus MC – Sophia Children's Hospital, Rotterdam, The Netherlands. Email: m.vandijk.3@erasmusmc.nl*

**More clinical trials are needed to study the effects of analgesics in neonates and to improve evidence based pain treatment in this group of vulnerable patients. It is not easy, however, to conduct pharmacological studies in neonates. This article explores some of the issues that may complicate clinical trials of analgesics in neonates and gives recommendations for randomised controlled trials (RCTs) in neonates in general.**

Paed Perinatal Drug Ther 2005; 6: 203–210

*Keywords:* analgesics – neonates – clinical trials

## Introduction

It is generally agreed that pain is harmful to neonates and should be adequately treated. Relatively few neonatal pharmacological trials have been conducted. Therefore, neonates are often given drugs without any license or in an "off-label" fashion[1]. Since the Best Pharmaceuticals for Children Act (BPCA) was signed into law in the United Sates of America in 2002, the National Institute of Health and the Food and Drug Administration have undertaken initiatives to delineate this problem and to develop a research agenda to study drugs in newborn infants. A recent WHO report on priority medicines in Europe and the World devotes a chapter to children, again emphasising the need for more drug research in children[2]. Several barriers have to be overcome before a well-designed neonatal trial can be conducted. Hospital ethics committees may be reluctant to agree with drug trials in young infants. In addition, it may be difficult to obtain parental informed consent. Finally, it may be difficult to get funding for trials researching off-patent drugs.

There are, however, pressing arguments to study the effects of analgesics in neonates. Neonates receiving intensive care treatment undergo, on average, 14 painful procedures per day[3]. Surgery in (premature) neonates, e.g. for patent ductus arteriosus, necrotising enterocolitis or major congenital anomalies, requires adequate analgesic treatment. We cannot just base dosing regimens on schedules derived from adults, as the pharmacokinetics and pharmacodynamics of drugs in premature neonates are quite different[4]. Therefore, analgesic trials are essential to determine optimal drug dosing regimens for neonates[5]. This applies both to drugs with a long history of use (morphine, bupivacaine, paracetamol, fentanyl) and to novel drugs (propacetamol).

Trials in neonates are usually phase III trials, with full-scale evaluation, comparing current standard treatment (using a control group) with the new drug. To prevent biased evaluation of the new treatment, each patient is randomly assigned to either new or standard treatment. Trials are then randomised controlled trials (RCTs), which are considered best for research[6]. A recent systematic review concerning opioids in ventilated neonates analysed 13 RCTs[7]. Table 1 details five analgesic RCTs in neonates from 1999 to 2004[8-12]. The opioid trials clearly show lack of consensus on doses. Thus there is an apparent need for studies that will define the right dose for different gestational and postnatal ages. These dosing regimens need to take into account general developmental changes in metabolism and renal clearance, pharmacodynamic changes and the individual's pharmacogenetic background.

## Team

Meeting the quality requirements, from designing the study protocol up to publication of the results, requires multidisciplinary team involvement. Early statistical advice is essential in order to ensure a satisfactory design. The determination of optimal patient numbers alongside an adequate power to detect a difference between the two analgesic regimes is crucial. In addition to the principal investigator, such a team would typically consist of nurses, neonatologists, a develo-pmental pharmacologist, a psychologist, a metho-dologist, a biostatistician, a paediatric intensivist, a pharmacist and a clinical chemist. The role of the pharmacist is important. The test drugs and placebo must have a similar appearance. Furthermore the pharmacy may perform the randomisation and ensure blinding. It is useful to carefully think through design of documentation, data collection, data management and data analysis.

## CONSORT statements and registration of trials

Evidently, all team members should be familiar with the CONSORT guidelines[13] and ultimately the study protocol should comply with these guidelines. Using the CONSORT guidelines as a checklist helps to consider issues which otherwise could be overlooked, such as follow-up of refusals, adequate description of randomisation and blinding. Furthermore, many peer reviewed journals require trials to conform to the CONSORT guidelines. This may stimulate uniformity in reporting RCTs.

The International Committee of Medical Journal Editors (ICMJE) recently ruled that for trials to be published, subscription into a public trials registry at or before start of patient enrolment is required as from July 2005[14]. This measure is expected to reduce publication bias, i.e. the tendency that non-significant or 'less interesting' results are less likely to be published than are significant results. Paediatric trials may be entered into the Drug Evaluation in Children register (www.dec-net.org). The DEC-net Register is supported under the European Union's Fifth Framework programme 'Quality of Life', and was activated on 1st July 2004.

## Pilot study

Once a study protocol has been created, its practical execution may present unforeseen problems. For instance, it may be far too optimistic with regards to number of patients included and speed of recruitment. This may be highly relevant as studies are often sponsored by grants of limited duration. A pilot study may give an estimation of required sample size and facilitates a realistic power analysis. In addition, a pilot study may reveal logistic or practical

**Table 1** Neonatal analgesic trials (1999–2004)

| Clinical trial | *n* | Gestation (weeks) | Morphine | | Alternative analgesic | | Outcome measures | Reference |
|---|---|---|---|---|---|---|---|---|
| | | | Loading dose (mcg/kg) | Infusion (mcg/kg/h) | Loading dose (mcg/kg/h) | Infusion (mcg/kg/h) | | |
| Morphine v Fentanyl | 163 | 24–40 | 140 | 20 | 10.5 | 1.5 | BPS, plasma adrenaline and noradrenaline, adverse events | 8 |
| Morphine v Placebo v Midazolam | 67 | 24–32 | 100 | 10-30 | 200 | 20-60 | Neurological outcomes | 9 |
| Morphine v Morphine | 68 | 35–40 | 100 | 10 | 100 | 30* | BPS, stress hormone levels, BP, HR | 10 |
| Morphine v Placebo | 150 | 25–42 | 100 | 10 | 0 | 0 | BPS, neonatal death, IVH, PVL | 11 |
| Morphine v Placebo | 898 | 23–32 | 100 | 10-30 | 0 | 0 | BPS, neonatal death, IVH, PVL | 12 |

*Given 3 hourly
BPS = Behavioural pain scale

problems and provides for instrument testing and improvement of task allocation. However, the pilot study may result in overoptimistic expectations about patient inclusion. In one of our studies in premature babies[11], we were confronted with a refusal rate as high as 28% in the RCT (Figure 1). We feel this may have been due to having to seek consent within 8 hours after start of mechanical ventilation, when parents are often still in shock from the premature birth of their child. For some parents, morphine was an emotionally charged drug, associated with addiction and terminal care.

## Communication

Continuous communication between team members, parents and neonatal staff throughout the study is crucial. On completion of the study, parents appreciate receiving a report stating the most relevant results of the study. Close collaboration between medical staff and nursing staff is important. It should be clear who has final responsibility for trial participants. Special meetings and newsletters help to improve communication and collaboration. Finally, in multicentre studies, research meetings with all collaborating centres are crucial to guarantee study protocol adherence.

## Consent

Communication with parents is challenging at the start of the study when informed consent is sought, especially if the study protocol dictates enrolment soon after birth. Parents who are still in shock because their baby was born prematurely or presented with severe conditions such as meconium aspiration syndrome or congenital anomalies, may be inclined to refuse consent. This hurdle may be overcome in part by having the attending neonatologist explain the study antenatally, when parents are not exhausted and upset[15, 16]. However, this remedy is of little help in unexpected premature or difficult deliveries. Another option is to have the clinical researchers seek consent independently and not to involve the care giving physicians. This may even be preferable to avoid (suspicion of) conflict of interest.

Valid informed consent suggests that parents understand the purpose of the study as well as its benefits and risks. In addition they should be aware of the voluntary nature of their participation. Mason et al reported the lack of **valid** informed consent from 71.5% of 200 parents who had been asked for consent for inclusion of their babies in a trial[17]. Problems were related to lack of understanding, coercion, incompetence and information problems[17]. Ballard et al. recently found that only 5% of 64 parents understood any potential risk from participating in the NEOPAIN study in which neonates received 10 to 30 mcg morphine/kg/h or placebo[18]. These examples show that devising the best possible strategy of asking parental consent is quite essential. Table 2 summarises some of the issues and solutions in the consent process.

## Intention to treat

Intention to treat (ITT) is a strategy for the design, conduct and analysis of RCTs that compare patients in the groups to which they were originally randomly assigned. Or in brief: once assigned, remains assigned. ITT is especially recommended in pragmatic trials that focus on measuring the benefit of treatment in routine

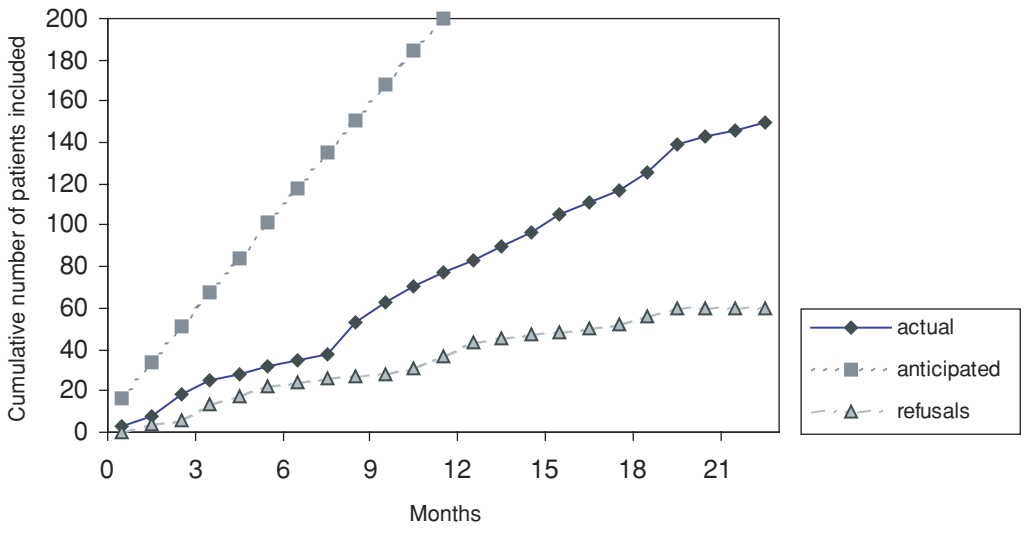


**Figure 1** Comparison of actual and planned inclusion of patients in NICU pain study.

**Table 2** Some issues and possible solutions around informed consent

| Issues | Possible solutions |
|---|---|
| Time pressure: emergency research*: less than 24 hours reflection time for parents. | Give antenatal information and ask presumed consent (with possibility of opting out) or adapt research question and protocol including more time before inclusion. |
| Parents are in shock and cannot properly decide. | Antenatal information with presumed consent and possibility of opting out. |
| No full understanding of procedure, benefits and risks associated with study, right to withdraw from study at any time. | Both oral and written information and check if information was understood. Point out that an independent neonatologist is available for consultation. |
| Parents feel pressure/involuntariness | Researcher, and not attending physician, should ask consent. |

*Intervention must take place within 24 hours

clinical practice[19]. The design should describe the inclusion criteria which when violated, would justify exclusion from ITT analysis. During data collection it should be attempted to minimise numbers of missing responses on the primary outcome. When inevitable, the effects of the missing responses and the different strategies to cope with missing data in the analysis should be explored. Even false inclusions should remain in the groups to which they were allocated.

## Internal validity

The concept of internal validity of a study is crucial because it questions if the observed outcome truly depends on the explanatory variable(s) and not on biases. For instance, it questions whether it is indeed the treatment that causes pain scores to go down. Internal validity may be improved by random allocation of subjects to the treatment conditions. Confounding factors, on the other hand, are major threats to internal validity. For example, critically ill neonates may not be able to exhibit pain behaviours. These are factors influencing both the treatment and the outcome so that they may alter the measure of association. Proper randomisation may reduce the threat of confounding.

## Multiple testing

Multiple testing is when too many statistical tests are performed to 'trawl' for results of a sample. It is advisable to determine the primary outcome beforehand and not use multiple outcome variables. Multiple testing may present a problem in terms of obtaining false positive results. In case of multiple testing, a correction such as the Bonferroni correction on the *P*-value should be applied. Alternative solutions are multivariate analysis, for instance principal component analysis or multiple regression analysis.

## Outcome measures

Outcome measures in analgesic trials generally relate to efficacy and safety. Efficacy is tested by pain assessment. As self-reporting cannot be used in neonates, pain assessment tools have been devised. These are discussed in detail later.

Stress-related outcome variables such as plasma levels of noradrenaline and cortisol are useful to determine relationships between opioids and stress responses[20]. We have previously found that continuous morphine infusion significantly reduced plasma noradrenaline levels in ventilated newborns compared with placebo treatment[20]. Though relevant for research purposes, the determination of hormonal plasma levels has limited value in clinical practice, for instance for adjustment of the individual patient's drug dosing. Finally, time will tell if neuroimaging techniques will yield a 'gold' standard of neonatal pain assessment[21].

## Use of placebo

Although neonates will not experience a placebo effect of drug therapy, it could bias those who measure the effect of pharmacological interventions. As placebo is required to ensure blinding during the trial, inclusion of a placebo group into the study design is warranted even in neonatal trials. It may be appropriate to use a placebo, if the efficacy of an analgesic drug has not been proven in a specific clinical situation. However, placebo is only an option for the control group when there is no existing standard drug therapy against which the new drug should be compared. The World Medical Association has developed the Declaration of Helsinki as a statement of ethical principles to provide guidance to physicians and other participants in medical research involving human subjects.

The use of placebo in trials was first mentioned in the 1996 version. The latest version (2000) includes Clause 29 stating: 'The benefits, risks, burdens and effectiveness of a new method should be tested against those of the best current prophylactic, diagnostic, and therapeutic methods. This does not exclude the use of placebo, or

no treatment, in studies where no proven prophylactic, diagnostic or therapeutic method exists'. This seemed to rule out the use of placebo wherever proven treatment existed. In reaction to protests from the medical world, in 2001 a Note of Clarification was added to Clause 29, describing the circumstances when a placebo-controlled trial is acceptable even if proven therapy is available[22].

## Dose regimen

One of the most important issues in designing an analgesic trial is probably the determination of the dose regimen of the investigational drug. Effective doses (per kg bodyweight) vary with changing gestational and postnatal ages. Optimal dosing regimens should be based on extensive literature reviews. In general it is a matter of balancing between high doses that are effective and low doses that cause less toxicity. Another consideration, aimed at avoiding staff uneasiness, is using doses that approximate those that were standard of care before the trial started.

To find the appropriate dose to be used in RCTs a pharmacokinetic study needs to be designed that will use the least amount of blood samples to arrive at the correct dose recommendation. For scientific investigations in neonates, ethical review boards will typically allow the use of maximal 3 ml/kg of blood. In other words, for a preterm infant of 800 grams a total of 2.4 ml may be used for both the pharmacokinetic analysis and any biochemical tests performed to evaluate safety. Evidently we need precise microanalytical assays that will allow us to do these studies. The use of HPLC-assays and more recently HPLC-MS/MS techniques have revolutionised the opportunities for even the smallest prematurely born neonates.

A recent development is the application of sparse pharmacokinetic sampling in larger neonatal populations, i.e. population pharmacokinetics. In this analytical approach, some precision in the pharmacokinetic parameters of individual participants is sacrificed by taking fewer samples to allow inclusion of a wider spectrum of participants likely to receive the drug clinically. Reduction in frequency and number of samples has obvious appeal in neonatal populations. Additionally, the ability of population methods to analyse unbalanced data collected at various time points is attractive in these populations. This method allows pooling of data across studies to provide a uniform, robust, single pharma-cokinetic analysis rather than attempting to compare results of separate, smaller studies that may be complicated by significant analytical methodology differences.

## Toxicity

As well as positive effects, such as analgesia or stress relief, drugs may cause toxicity. Drug toxicity may be due to the drug itself or secondary to a chemical agent that the drug has been dissolved in. Drug toxicity may be idiosyncratic or dose dependent. As only a few patients may show adverse effects, trials which focus on side effects should have a large sample size for sufficient power. Adverse effects known to result from certain drugs should be monitored at predetermined time-points. For instance, hypotension due to opioids may be assessed by continuous blood pressure measurement or by document-ation of the need for inotropic drugs. Respiratory depression, is a significant side effect of opioids and needs to be considered even in neonates receiving ventilatory support[23].

## Effect size

Effect size (ES) is the magnitude of a treatment effect, independent of sample size. For instance, in studies comparing treatment with placebo, the effect size is the difference between the means of the two groups divided by the pooled standard deviations. This effect size is also referred to as Cohen's *d*; Cohen provided broad rules of thumb for characterising effect sizes: with *d* 0.2 small, *d* 0.5 moderate and *d* 0.8 large[24].

When designing trials aiming to determine the analgesic effect of a certain drug, a statistical difference in pain scores (compared to a placebo group or compared to another drug) should detect the quantified effect of the studied agent, if any. Nevertheless, a statistically significant difference between pain scores does not necessarily imply clinical relevance[25]. The effect size aimed at should be feasible and clinically significant. This may be difficult to estimate when there is limited information on expected means and standard deviations. The required sample size can be determined by a power analysis. The power is commonly set at 80 to 90% and represents the degree of certainty that a difference between groups, if present, would be detected. The type I error ($\alpha$) is usually set at 0.05 and is the probability of detecting a significant difference when there is really no difference. When power calculations show that a large sample is required to obtain a significant result, it is advisable to consider a multicentre trial, especially when data collection otherwise would take more than two to three years.

## Issues in neonatal pain assessment

With a gold standard for neonatal pain assessment still lacking, behavioural observation and physiological parameters are used to estimate pain intensity in neonates. Some focus on facial expression. Others combine scoring of behavioural and physiological items to obtain a total pain score, e.g. the CRIES[26] (acronym for Crying, requires Increased oxygen, Increased vital signs, Expression, Sleeplessness), Premature Infant Pain Profile (PIPP)[27] and Neonatal Infant Pain Scale (NIPS)[28].

Although these instruments have been recommended by experts on pain in newborns[29], we feel that current pain assessment is far from ideal[30]. Any pain instrument must be age-appropriate and be both valid and reliable. Reliability or consistency for observational tools is primarily checked by testing inter-rater reliability to assure that different observers/raters assess in a similar vein[31]. Other reliability issues are intra-rater reliability (stability of the rater) and internal consistency of the items. Validity is often defined by asking the question: are we measuring what we think we are measuring? This should be tested by comparing the pain instrument to an expert opinion or an existing instrument. Sensitivity to change or responsiveness (for some scientists part of validity) is established when scores decrease substantially after analgesic treatment[32].

Most instruments unfortunately have been validated for procedural pain only, such as from heel lances. Another drawback is the fact that neonates, especially in extreme prematurity, may show blunted behavioural responses due to exhaustion, severity of illness or prolonged pain. Physiological parameters such as heart rate, blood pressure and oxygen saturation are of limited value for pain assessment because they are also influenced by severity of illness and medical interventions (e.g. inotropic drugs). Sick preterm neonates often respond to minor stimuli with bradycardia or a fall in oxygen saturation, even during nursing care.

For various reasons, financial or logistic, researchers may not be able to assess pain continuously. Yet in postoperative trials the most relevant assessment period may be at night, as patients usually return from surgery in the afternoon or evening. Scoring may be done either retrospectively using videotapes or in real time at the bedside. Videotaping has the advantage that the observer can be more easily blinded to treatment and is not distracted by information from the monitor or medication pumps. Bedside observation provides a better perspective, especially when observing neonates in damp incubators. Finally, patients should be assessed on a regular basis, taking into account analgesic interventions for both efficacy and safety.

*Inter-rater reliability*

If different raters are involved in a trial, we need to estimate their inter-rater reliability. This concept refers to the consistency with which the same information is rated by different raters. Inter-rater reliability is tested by comparing a rater's test scores with those of experienced observers. As a measure for inter-rater reliability we suggest using either the intra-class correlation coefficient[33] or the linearly or quadratic weighted Cohen's kappa[34]. Pearson's product correlation coefficient is not recommended because a perfectly linear relationship represented by a Pearson's product correlation coefficient of one (e.g. one observer consistently rates 2 points lower than the other observer) does not imply perfect agreement between observers (which implies identical scores on all observations)[35].

## Rescue medication

The available tools for neonatal pain assessment often fail to suggest cut-off points guiding pain treatment. Consequently, administration of rescue medication in RCTs is primarily guided by the subjective clinical judgement of staff present at the time. Pre-assigned pain score thresholds should be set for rescue analgesia. The use of an algorithm including decision rules for the administration of rescue medication would enhance objectivity and standardisation of treatment. However, the very administration of rescue medication may diminish the strength of the study design, especially when rescue medication is required frequently. Excessive use of rescue analgesia in one arm should result in early termination of the trial. As rescue medication reflects normal clinical practice, a study design that provides for the use of rescue medication is preferred.

## Multicentre trials

In order to obtain sufficient sample size and to increase the generalisability of the results, multicentre trials may be needed. Multicentre trials, however, have significant disadvantages. All the logistical problems of single centre RCTs are even more apparent in multicentre trials. Furthermore, formalities, such as who is funding the research, whose name is on the published manuscripts, may be required. Then there are statistical considerations in that observations within centres may be correlated[36]. Failure to consider the centre in statistical analysis may

result firstly in incorrect P values and confidence intervals. Secondly, biased estimates may present themselves because of uncontrolled confounding. For instance, if the proportion of very premature neonates varies significantly between centres, and this variation affects treatment as well as the outcome, confounding is apparent. Thirdly, effect modification may occur if the effect of treatment on outcome varies significantly between centres. This effect could be efficiently tested for example by random-coefficient models[36].

## Conclusions

For many years now it has been recognised that newborns do feel pain and that their pain should be treated. Regrettably, neonatal pain is still poorly managed, and effecting a change in clinicians' minds might be more difficult than supposed. Large multicentre RCTs are needed to produce evidence about effects of analgesics. We have described the major issues to be considered in designing such trials. An algorithm (Figure 2) highlights the main points. Our ultimate goal is the improvement of neonatal pain management.

## Acknowledgements

We thank J. Hagoort for his help in preparing the manuscript. We received invaluable suggestions from H.J. Duivenvoorden.

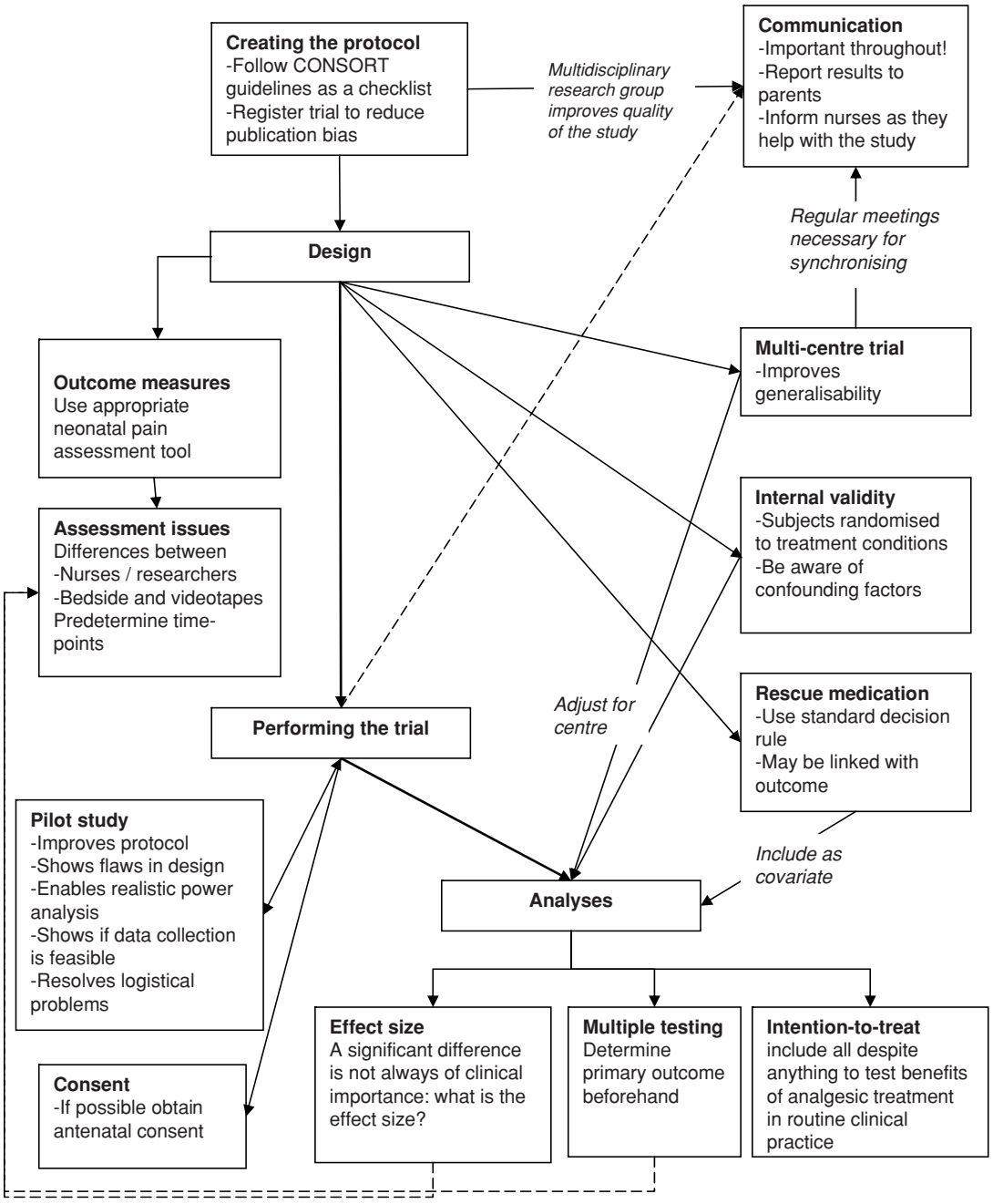


**Figure 2** Algorithm showing steps in well designed trial and potential confounding factors.

## References


1. Choonara I. Unlicensed and off-label drug use in children: implications for safety. Expert Opin Drug Saf 2004; 3: 81-83.
2. Kaplan W, Laing R. Priority medicines for Europe and the World: World Health Organisation, 2004: 91-93.
3. Simons SH, van Dijk M, Anand KS et al. Do we still hurt newborn babies? A prospective study of procedural pain and analgesia in neonates. Arch Pediatr Adolesc Med 2003; 157: 1058-1064.
4. van Lingen RA, Simons SH, Anderson BJ, Tibboel D. The effects of analgesia in the vulnerable infant during the perinatal period. Clin Perinatol 2002; 29: 511-534.
5. Choonara I. Why do babies cry? We still know too little about what will ease babies' pain. BMJ 1999; 319: 1381.
6. Pocock SJ. Clinical trials: a practical approach. Chichester: John Wiley & Sons Ltd, 1983.
7. Bellu R, Waal K, Zanini R. Opioids for neonates receiving mechanical ventilation. Cochrane Database Syst Rev 2005: CD004212.
8. Saarenmaa E, Huttunen P, Leppaluoto J, Meretoja O, Fellman V. Advantages of fentanyl over morphine in analgesia for ventilated newborn infants after birth: a randomized trial [see comments]. J Pediatr 1999; 134: 144-150.
9. Anand KJ, Barton BA, McIntosh N et al. Analgesia and sedation in preterm neonates who require ventilatory support: results from the NOPAIN trial. Neonatal Outcome and Prolonged Analgesia in Neonates [published erratum appears in Arch Pediatr Adolesc Med 1999 Aug; 153: 895]. Arch Pediatr Adolesc Med 1999; 153: 331-338.
10. Bouwmeester NP, Hop WC, van Dijk M et al. Postoperative pain in the neonate: age-related differences in morphine requirements and metabolism. Intensive Care Med 2003; 29: 2009-2015.
11. Simons SH, van Dijk M, van Lingen RA et al. Routine morphine infusion in preterm newborns who received ventilatory support: a randomised controlled trial. JAMA 2003; 290: 2419-2427.
12. Anand KJ, Hall RW, Desai N et al. Effects of morphine analgesia in ventilated preterm neonates: primary outcomes from the NEOPAIN randomised trial. Lancet 2004; 363: 1673-1682.
13. Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomised trials: a comparative before-and-after evaluation. JAMA 2001; 285: 1992-1995.
14. De Angelis CD, Drazen JM, Frizelle FA et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. JAMA 2004; 292: 1363-1364. Epub 2004 Sep 08.
15. Burgess E, Singhal N, Amin H, McMillan DD, Devrome H. Consent for clinical research in the neonatal intensive care unit: a retrospective survey and a prospective study. Arch Dis Child Fetal Neonatal Ed 2003; 88: F280-285; discussion F285-286.
16. Manning DJ. Presumed consent in emergency neonatal research. J Med Ethics 2000; 26: 249-253.
17. Mason SA, Allmark PJ. Obtaining informed consent to neonatal randomised controlled trials: interviews with parents and clinicians in the Euricon study. Lancet 2000; 356: 2045-2051.
18. Ballard HO, Shook LA, Desai NS, Anand KJ. Neonatal research and the validity of informed consent obtained in the perinatal period. J Perinatol 2004; 24: 409-415.
19. Roland M, Torgerson DJ. Understanding controlled trials: what are pragmatic trials? BMJ 1998; 316: 285-289.
20. Simons SH, van Dijk M, van Lingen RA et al. Randomised controlled trial evaluating effects of morphine on plasma adrenaline/noradrenaline concentrations in newborns. Arch Dis Child Fetal Neonatal Ed 2005; 90: F36-F40.
21. Chen AC. New perspectives in EEG/MEG brain mapping and PET/fMRI neuroimaging of human pain. Int J Psychophysiol 2001; 42: 147-159.
22. Carlson RV, Boyd KM, Webb DJ. The revision of the Declaration of Helsinki: past, present and future. Br J Clin Pharmacol 2004; 57: 695-713.
23. Bhandari V, Bergqvist LL, Kronsberg SS, Barton BA, Anand KJ. Morphine administration and short-term pulmonary outcomes among ventilated preterm infants. Pediatrics 2005; 116: 352-359.
24. Cohen J. Statistical power analysis for the behavioral sciences. 2nd edition: Lawrence Erlbaum Ass., 1988.
25. McQuay HJ, Barden J, Moore RA. Clinically important changes-what's important and whose change is it anyway? J Pain Symptom Manage 2003; 25: 395-396.
26. Krechel SW, Bildner J. CRIES: a new neonatal postoperative pain measurement score. Initial testing of validity and reliability. Paed Anaesth 1995; 5: 53-61.
27. Stevens B, Johnston C, Petryshen P, Taddio A. Premature Infant Pain Profile: development and initial validation. Clin J Pain 1996; 12: 13-22.
28. Lawrence J, Alcock D, McGrath PJ et al. The development of a tool to assess neonatal pain. Neonatal Network 1993; 12: 59-66.
29. Anand KJ. Consensus statement for the prevention and management of pain in the newborn. Arch Pediatr Adolesc Med 2001; 155: 173-180.
30. van Dijk M, Simons S, Tibboel D. Pain assessment in neonates. Paed Perinatal Drug Ther 2004; 6: 97-103.
31. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. 2nd edition New York: Oxford University Press, 1995.
32. Johnston CC. Psychometric issues in the measurement of pain. In: Finlay GA, McGrath PJ, eds. Measurement of pain in infants and children. Progress in pain research and management. Seattle: ISAP Press, 1998: 5-20.
33. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin 1979; 86: 420-428.
34. Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. Psychological Bulletin 1968; 70: 213-220.
35. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1: 307-310.
36. Localio AR, Berlin JA, Ten Have TR, Kimmel SE. Adjustments for center in multicenter studies: an overview. Ann Intern Med 2001; 135: 112-123.


PPDT – 0139, *Accepted for publication*: 5 October 2005
*Published Online*: 17 November 2005