Future of clinical trials in the delivery room: time for pragmatism
O'Donnell, C.P.F.; Dekker, J.; Rudiger, M.; Pas, A.B. te

Citation

Version: Publisher's Version
License: Licensed under Article 25fa Copyright Act/Law (Amendment Taverne)
Downloaded from: https://hdl.handle.net/1887/3575961

Note: To cite this publication please use the final published version (if applicable).
Future of clinical trials in the delivery room: time for pragmatism

Colm Patrick Finbarr O’Donnell 1,2, Janneke Dekker,3 Mario Rüdiger,4,5 Arjan B te Pas6

ABSTRACT

Despite increased amounts of research, most of the evidence that supports treatment of newborns in the delivery room is rated ‘low’ rather than ‘high’ quality. This assessment stems largely from a lack of evidence from clinical trials. When trials have been performed, the evidence has often been downgraded due to enrolment of small or poorly representative samples, and for lack of blinding of caregivers and outcome assessors. Delivery room trials present particular challenges when obtaining consent, enrolling participants, taking measures to limit bias and identifying appropriate outcome measures. We hope our suggestions as to how future delivery room trials could be more pragmatic will inform the design of large studies that are necessary to allow clinical practice to evolve.

Up to the 1950s, compromised newborns received ad hoc treatment with some very dubious interventions (eg, slapping, electrocution, tobacco smoke enemas) in delivery rooms (DRs) where fatalism largely prevailed.1 In 1953, Virginia Apgar devised a method to systematically assess every baby at 1 min of life.2 A pattern of treatment evolved based on studies of acutely and profoundly asphyxiated term Rhesus monkeys.3 Structured teaching courses appeared in the 1980s.4 5 The International Liaison Committee on Resuscitation published their first recommendations on neonatal resuscitation in 1999.6 These were largely based on consensus opinion and identified a lack of supporting high-quality data. The first quasi-randomised7 and blinded randomised8 DR trials concerned the use of oxygen. Since then, increasing amounts of DR research has been published, and a more elaborate method of evidence evaluation is now used to update recommendations.9

Proof that medical interventions are effective and safe can only be obtained by studying their effects on the target human population. The effects of interventions are best determined in prospective clinical trials that compare the outcomes of exposed and unexposed participants. For the results to be reliable and precise, an adequately large sample of the relevant population must be studied. It is also important that: subjects consent to participate, measures are taken to limit the effects of bias, and relevant and important outcomes are measured in all participants. The larger the study and the lower the risk of bias, the higher the quality of evidence generated.10 Disappointingly, much of the evidence that supports fundamental aspects of DR care—for example, warming adjuncts; heart rate monitoring; whether sustained inflations, positive end-expiratory pressure, continuous positive airway pressure (CPAP) or intermittent positive airway pressure should be used; whether a T-piece or a self-inflating bag should be used; the oxygen concentration that should be used for preterm infants; chest compression technique and frequency; the use of epinephrine—is considered ‘low’ rather than ‘high’ quality.9 This is due to a lack of trials; and where trials have been performed, to enrolment of small or poorly representative samples and absence of blinding of group assignment. Therefore, if we want DR care to evolve, we must consider our current approach to DR trials.

Clinical trials are challenging to conduct in all settings; however, particular considerations apply to obtaining consent, enrolling participants, limiting bias and measuring outcomes for DR studies.

CONSENT

The ethics of doing research in babies are often debated. Less often discussed are the ethics of not doing research in babies. Limited research on interventions means limited proof that they work. More worryingly, if research is limited, not only might we be unaware that our interventions do not work, we might also be unaware that they are harmful. Protecting ‘vulnerable’ populations is a laudable aim. However, excluding people from research does not protect them. Disasters that befell babies due to drugs that they or their pregnant mothers received prompted the demand for their careful study in humans.11 12 It is a cruel irony that babies’ ‘vulnerability’ is a central reason why they are still excluded from drug trials and so remain vulnerable.

In the typical trial, competent adults decide whether they will participate. They should have adequate time to digest clear information about a study before deciding whether they will take part. As babies cannot consent for themselves and depend on their parents to decide, the consent process is more complex. Ideally, parents should have ample time to consider information before deciding whether their baby should participate. However, that is often not possible for DR studies. The majority of participants in DR studies are born prematurely. A proportion deliver spontaneously with little warning. The remainder are delivered early because of rapidly evolving threats to maternal and/or fetal health. Preterm birth is unpredictable, rarely long planned and cannot be deferred. In all

© Author(s) (or their employer(s)) 2023. No commercial re-use. See rights and permissions. Published by BMJ.


© Author(s) (or their employer(s)) 2023. No commercial re-use. See rights and permissions. Published by BMJ.

cases, parents are stressed and, in many, they may be unwell themselves. The standard approach for DR trials has been to obtain consent to enrol the baby before birth. Approaching stressed people under time pressure is not ideal and may have unintended and undesirable consequences. SUPPORT was a large, multicentre, randomised trial performed in the USA. Consent was sought before birth to randomly assign extremely preterm babies to CPAP or intubation in the DR, and to a higher or lower target oxygen saturation range. More than five women were approached on a mean of two occasions for every baby who was born in the eligible gestational age window and enrolled. Many families whose babies were ultimately ineligible were thus approached and may have experienced unnecessary additional stress. Also, a lot of resources were wasted on fruitless activities. Worryingly, using antenatal consent exclusively may have resulted in enrolment of a poorly representative sample. The babies enrolled in SUPPORT had lower rates of death and other adverse outcomes than babies born at participating centres who were ultimately ineligible were thus approached and may have experienced unnecessary additional stress. Also, a lot of resources were wasted on fruitless activities. Worryingly, using antenatal consent exclusively may have resulted in enrolment of a poorly representative sample. The babies enrolled in SUPPORT had lower rates of death and other adverse outcomes than babies born at participating centres during the same period who were eligible but not enrolled. If prospective antenatal consent is always required, a proportion of newborns will be excluded. Many will be babies who deliver with less advance warning, who may be sicker and derive the greatest benefit—or be at the greatest risk of harm—from the intervention being studied. They may well be the most important babies to study; designing trials to exclude them seems unwise.

ENROLMENT
In most trials, research staff enrol people who have consented to participate within working hours. However, many babies deliver ‘out of hours’, so even if consent is obtained within office hours, many can only be enrolled out of hours. Restricting enrolment to office hours may severely limit participation and result in recruitment of a poorly representative sample. If research staff are required to enrol participants, they should be available around the clock. Maintaining an ‘on-call’ roster of research staff may be difficult, impractical and expensive.

MEASURES TO LIMIT THE EFFECT OF BIAS
In clinical trials, participants should be randomly assigned to intervention or control to prevent caregivers from assigning participants they think will ‘do better’ to their favoured group and/or participants they think will ‘do worse’ to the alternative. The group assignment schedule should be concealed to prevent people from deciding whether or not to enrol subjects according to the group to which they will be assigned. Caregivers and outcome assessors should be blinded to group assignment to prevent differential treatment and measurement of outcomes, respectively. Outcomes should be measured and reported for all participants to prevent the investigators excluding participants whose outcome was undesired.

In DR trials, random assignment, concealment of the group assignment schedule and complete outcome assessment have been consistently achieved. However, masking interventions has proven far more difficult (see figure 1). In drug trials, blinding is relatively straightforward, as a placebo that appears identical to the active agent can usually be sourced. When it is not possible to use a placebo, a ‘sham’ procedure should be considered—that is, clinicians not involved in the clinical care of the patient are the only ones aware of group assignment and they perform the real or a sham procedure shielded from view of caregivers and outcome assessors to maintain blinding. DR care mostly comprises physical interventions—assessment of condition, prevention of hypothermia, and provision of breathing support and circulatory support. A sham procedure should therefore be considered for most studies. However, the DR is often crowded. In addition to the mother and baby, a partner, midwives, obstetricians, neonatal nurses and neonatologists are usually present for preterm births. It is practically impossible to credibly mask most interventions with so many people in a confined space. Moreover, if masking is attempted, a separate ‘round-the-clock roster’ of research clinicians is required, because limiting enrolment to working hours will inevitably restrict participation among eligible babies. This is difficult, impractical, expensive and may only be possible in a select few centres. Performing studies in such centres reduces the applicability of the results.

In the POPART trial, preterm babies were randomised to receive oropharyngeal surfactant or no intervention before umbilical cord clamping at nine hospitals in six countries. There is no placebo safe for babies to aspirate, so we considered using a sham procedure (eg, an injection of air). As the intervention was to be performed before cord clamping (ie, in close proximity to the mother and in the presence of multiple caregivers), we doubted that we could effectively mask it. Also, there was little prospect of having a separate roster of study clinicians to perform a sham procedure at any centre. We considered it unacceptable to exclude babies born out of hours and so performed an unmasked study. Ultimately, 60% of participants were enrolled out of hours; a likely ineffective sham treatment would have precluded the participation of many of them.

OUTCOME MEASURES
The outcomes measured in trials should be important—that is, meaningful to the participant, families, caregivers and society in general. However, there must also be a reasonable chance that they might be influenced by the intervention studied. Many important outcomes for newborns are unlikely to be influenced by DR interventions—blindness, for example. As ‘neonatal resuscitation’ has long been thought a synonym for DR care, death might appear a suitable outcome. However, DR death is rare; the causes—asphyxia, sepsis, pulmonary hypoplasia, exsanguination, malformation—usually long precede birth; and they are unlikely to be greatly affected by a single DR intervention such as a sustained inflation. The ‘gold standard’ outcome for trials enrolling preterm infants is survival free of neurosensory impairment at 2 years’ corrected age. Unquestionably, this is important. However, considering the typical course of a preterm...
infant and the influence of confounding factors in the following months in hospital and years after discharge, it seems fanciful that a 15–30 s intervention at birth would have a measurable effect over 2 years later. Long-term neurodevelopmental assessment for the purpose of a research study is time-consuming and expensive. It should only be considered if there is a real possibility that it may be affected by the study intervention. It makes sense to measure outcomes more closely temporally related to events in the DR. However, it is a challenge to identify outcomes sufficiently short term that are likely to be influenced by the intervention, yet sufficiently long term as to be important and clinically meaningful.

RECOMMENDATIONS

Clinical research is crucial to the evolution of clinical care. Much uncertainty lingers in the DR, even after trials have been performed. We therefore need to study larger and more representative samples. This requires international collaboration between caregivers at many hospitals, including sites where babies have not previously participated in DR trials. If we are to enrol appropriate numbers in a timely fashion, the studies must be pragmatic. We once feared that, applied to research studies, ‘pragmatic’ meant lazy, ill-considered or badly organised. It need not. To be pragmatic is to deal with things sensibly and realistically based on practical rather than theoretical considerations. Few centres can maintain a ‘round-the-clock’ roster of research staff. Thus, it must be possible for clinicians to perform necessary elements of a study without it adversely affecting their ability to care for the baby in the DR, even in the middle of the night. If clinicians see a study as a competing interest, they would not participate. Elements of study design that may sabotage enrolment should be reconsidered. Because an imperfect study with lots of participants is more valuable than a scientifically pure study with few participants. Imperfect information is better than no information.

We need to consider alternative approaches to consent. Deferred consent—where subjects are enrolled and as soon as is reasonably possible, the patient’s family are informed of their inclusion, advised of the option to withdraw and their consent for continued participation is sought—has been used successfully in DR trials.22–25 Trials should meet certain criteria to be eligible for a deferred consent process. The condition studied should arise in emergent circumstances that make it difficult or impossible to seek consent prospectively. The research should be based on a valid scientific hypothesis that supports a reasonable prospect of benefit over standard care. Alternatively, the research should study interventions that are both considered as standard of care at different institutions. This approach is appropriate and is used in circumstances where serious consequences may occur with or without the treatment (eg, adults with out-of-hospital cardiac arrest).26 Many institutional review boards have been hesitant to approve the use of deferred consent, even though it would be appropriate for most DR studies. However, the majority of parents find this approach acceptable, and twice as many find it preferable to prospective consent.27 It should be used more widely.

We need to accept that it may be difficult, if not impossible, to credibly mask some DR interventions, and that trying to do so may severely limit enrolment and undermine the study. We need to acknowledge that unmasked studies are more prone to bias, and that the best way to offset this weakness is to study more babies. We need to choose outcome measures carefully. Ideally, they should be selected from an agreed set of important core outcomes.20 Many studies will be unblinded, so the measures should be objective, precise and repeatable. Asking clinicians to perform additional complex tasks, use new equipment or collect extra pieces of information that they would not ordinarily is unlikely to be successful. We need to be judicious about the number of outcomes we measure. The more information clinicians are asked to collect, the greater the burden it becomes, and the greater the risk that data will be wrong, missing or fabricated. Moreover, the more outcomes we measure, the greater the chance that one will appear different between the groups, when in truth it is not. If an outcome is not absolutely necessary, do not collect it.

An example of a DR intervention that warrants study is tactile stimulation. Though it has been used since ancient times and recommended in guidelines for more than 30 years, information is limited. In an unblinded, single-centre study, preterm infants breathed more effectively when repetitively stimulated.23 However, the number of participants was small (44) and there was a lot of ‘contamination’—infants in the control group were stimulated more frequently than was the case before the study was performed. We will evaluate the effect of repetitive stimulation in a stepped-wedge cluster randomised trial in babies born before 32 weeks’ gestation. This design is used to study interventions that are implemented stepwise, where individual randomisation is difficult because of ethical, logistical or financial reasons.28 Participating hospitals, rather than individual babies, will be the unit of randomisation. Each hospital will begin in the ‘control’ group, with clinicians selectively stimulating babies as is their usual practice. Hospitals will be randomly assigned to cross over at a specified interval to the ‘intervention’ group, whereupon clinicians will repetitively stimulate babies for 5 min after birth, until each hospital has been in both arms. It is not possible to credibly blind the intervention; this design should limit the ‘contamination’ seen in the earlier study.23 We will seek parental consent to use their babies’ data, before birth if time permits or after enrolment if it does not. We will measure the effect on oxygen saturation at 5 min of life; the demographic and other outcome data that we collect will fit on one A4 page. To participate, hospitals need to deliver babies <32 weeks and use pulse oximetry in the DR. We plan to study >3000 infants at around 40 hospitals in Europe. This study will demonstrate that it is possible to perform DR studies at scale and in a timely fashion. Tactile stimulation is a long-used and relatively benign intervention that does not require new equipment or extensive training to implement. Subsequent studies that test new or more complex interventions in preterm and sick term infants will be more difficult to perform. They will require support from professional bodies and funding agencies to enable us to refine the care we give to newborns in the DR.
Data availability statement  Data sharing not applicable as no datasets generated and/or analysed for this study.

ORCID iD
Colin Patrick Finbarr O’Donnell http://orcid.org/0000-0001-7453-5452

REFERENCES