Success of blinding a procedural intervention in a randomised controlled trial in preterm infants receiving respiratory support

Running head: Blinding of a procedural RCT

Number of words: approx. 3265

Elizabeth Reid[1], Omar F Kamlin[2], Francesca Orsini[3], Antonio G De Paoli[4], Howard W Clark[5], Roger F Soll[6], John B Carlin[3, 7], Peter G Davis[2] and Peter A Dargaville[1, 4]

[1] Menzies Institute for Medical Research, University of Tasmania, Hobart, TAS, Australia
[2] Newborn Research, Royal Women's Hospital, Melbourne, VIC, Australia
[3] Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute, Parkville, VIC, Australia
[4] Department of Paediatrics, Royal Hobart Hospital, Hobart, TAS, Australia
[5] Department of Neonatal Research, University College London, London, United Kingdom
[6] Pediatrics, University of Vermont, Burlington, VT, United States
[7] Department of Paediatrics, University of Melbourne, Parkville, VIC, Australia

Corresponding author:
Peter Dargaville, Department of Paediatrics, Royal Hobart Hospital, Hobart, TAS, Australia.
Email: peter.dargaville@ths.tas.gov.au

ABSTRACT

Background: Blinding of treatment allocation from treating clinicians in neonatal randomised controlled trials (RCTs) can minimise performance bias, but its effectiveness is rarely assessed.

Methods: To examine the effectiveness of blinding a procedural intervention from treating clinicians in a multicentre RCT of minimally invasive surfactant therapy (MIST) versus sham treatment in preterm infants of gestation 25-28 weeks with respiratory distress syndrome. The intervention (MIST or sham) was performed behind a screen within the first 6 hours of life by a 'study team' uninvolved in clinical care including decision making. Procedure duration and the study team's words and actions during the sham treatment mimicked those of the MIST procedure. Post-intervention, three clinicians completed a questionnaire regarding perceived group allocation, with the responses matched against actual intervention and categorized as correct, incorrect, or unsure. Success of blinding was calculated using validated blinding indices applied to the data overall (James index, successful blinding defined as >0.50), or to the two treatment allocation groups (Bang index, successful blinding: -0.30 to +0.30). Blinding success was measured within staff role, and the associations between blinding success and procedural duration and oxygenation improvement post-procedure were estimated.

Results: From 1345 questionnaires in relation to a procedural intervention in 485 participants, responses were categorized as correct in 441 (33%), incorrect in 142 (11%) and unsure in 762 (57%), with similar proportions for each of the response categories in the two treatment arms. The James index indicated successful blinding overall 0.67 (95% confidence interval, CI, 0.65–0.70). The Bang index was 0.28 (95% CI 0.23–0.32) in the MIST group and 0.17 (95% CI 0.12–0.21) in the sham arm. Neonatologists more frequently guessed the correct intervention (47%) than bedside nurses (36%), neonatal trainees (31%), and other nurses (24%). For the MIST intervention, the Bang index was linearly related to procedural duration and oxygenation improvement post-procedure. No evidence of such relationships was seen in the sham arm.

Conclusion: Blinding of a procedural intervention from clinicians is both achievable and measurable in neonatal RCTs.

KEYWORDS

Blinding success, procedural intervention, blinding assessment, preterm infants, randomized controlled trials

INTRODUCTION

Blinding is a key strength of randomised clinical trial (RCT) methodology, preventing ascertainment bias and preserving the integrity of estimates of treatment effect. [1,2] However, there is a widespread lack of information provided in trial reports detailing how and whether blinding was achieved. [3,4] This may limit critical appraisal of RCTs, undermine confidence in the results and recommendations, and may ultimately impede their translation to clinical practice. A review of the quality of over 800 neonatal RCTs in newborn infants from the 1950s to 2016 concluded that blinding was 'unsatisfactory' in the majority with no clear improvement evident over time. [5]

Blinding of RCTs in intensive care environments, including in the neonatal intensive care unit, is rendered more challenging by virtue of the intervention in many cases having a physical form, e.g. trials comparing intubation with use of continuous positive airway pressure, [6] or involving a labour-intensive procedure that may be difficult to blind from treating clinicians. An example of the latter is the administration of surfactant to preterm infants with respiratory distress syndrome, where the surfactant delivery procedure requires the engagement of several clinicians for a period of up to 30 minutes. [7] The majority of placebo controlled RCTs of exogenous surfactant therapy conducted before 1990 were unblinded for this reason. More recently, trials comparing different methods for surfactant delivery (via endotracheal tube or thin catheter) have been noted to have a high risk of performance bias due to lack of blinding of clinical personnel. [8] In procedural intervention trials such as these, treatment concealment from treating clinicians is less straightforward than in drug trials. Blinding may require greater ingenuity, be harder to implement and maintain, but is nonetheless highly desirable in reducing the risk of bias. [9,10] Furthermore, clear descriptions of both the strategies used for blinding, and methods of evaluating its success, are needed to improve methodological rigor in future studies.

An obvious approach to blinding a procedural intervention from treating clinicians in an RCT is to have the trial intervention performed by a study team consisting of clinical experts not directly involved in patient care. Even with such an approach, there are still barriers to successful blinding. The duration of the intervention may be a factor, particularly in trials comparing a procedural intervention with a sham procedure. Additionally, blinding may be compromised if there is an immediate clinical effect associated with allocation to one of the arms of the RCT (e.g. an oxygenation improvement after surfactant therapy). The influence of such factors has scarcely been examined in any formal way in RCTs of a procedural intervention.

We recently reported the results of multicentre blinded RCT of surfactant administration via thin catheter versus sham treatment in preterm infants with respiratory distress syndrome managed with continuous positive airway pressure – the OPTIMIST-A trial. [7] Herein we describe an analysis of the effectiveness of the approach to blinding in the OPTIMIST-A trial, and of factors influencing blinding success. We hypothesised that blinding of a procedural intervention could be incorporated into the design of an RCT, and that the success of blinding could be measured using validated blinding indices.

METHODS

Study setting and design

The OPTIMIST-A trial was an RCT examining the effect of surfactant administration via thin catheter (minimally invasive surfactant therapy, MIST) in preterm infants at 25 to 28 weeks gestation with respiratory distress syndrome, who were not intubated in early life but required respiratory support with continuous positive airway pressure. The study was conducted in tertiary level neonatal intensive care units in 33 sites, with recruitment occurring between 16th December 2011 and 26th March 2020. [7] The study was approved by ethics committees at all participating sites and was prospectively registered (Australian New Zealand Clinical Trials Registry, ACTRN12611000916943).

Infants were eligible for enrolment in the trial if continuous positive airway pressure level was 5-8 cm H_2O and $FiO_2 \geq 0.30$ in the first six hours of life. Infants in imminent need of intubation, or with an alternative cause of respiratory distress or a congenital anomaly, were excluded. Informed parental consent was obtained prior to study entry. Recruited infants were randomised to receive MIST, in which exogenous surfactant (poractant alfa, 200mg/kg) was delivered via a thin catheter, or a sham treatment. In both cases continuous positive airway pressure was continued unless prespecified intubation criteria were met. The primary outcome was the composite of death or physiological bronchopulmonary dysplasia (BPD), assessed at 36 weeks' postmenstrual age (PMA). Further details of the design and outcomes of the OPTMIST-A trial are described in detail elsewhere. [7,11]

Blinding

Blinding of the intervention in the OPTIMIST-A trial emulated the design of early trials of surfactant therapy [12, 13] and followed the methodology reported in an RCT of brief intubation for surfactant administration versus continuous positive airway pressure alone in preterm infants with respiratory distress syndrome. [14] For each recruited infant in each centre, once eligibility had been confirmed and parental consent gained, a 'study team' was assembled to carry out the randomisation and administer the assigned treatment out of sight of the treating clinicians. The study team consisted of one or more proceduralists skilled in neonatal intubation and trained in the MIST procedure (neonatologist, senior neonatal trainee, or neonatal nurse practitioner), and a neonatal nurse. It was a requirement that all team members were not immediately involved in the clinical care of the infant, with the acknowledgement that they might be involved in clinical care at some point during the hospital stay. The team members were asked to give written assurance that they would carry out the study intervention according to the randomisation and in a blinded manner and would not reveal what intervention had been given to treating clinicians. At all sites, treating clinicians were

aware that the study team would either be delivering surfactant via thin catheter, or performing a sham treatment mimicking the MIST procedure in all respects.

Once assembled, the study team conducted the randomisation, out of sight from the treating clinicians. A web-based portal was used to generate a unique code, and, upon opening of a correspondingly coded opaque sealed envelope, treatment allocation was revealed. Thereafter the activities of the study team in preparation for the intervention were identical regardless of group of assignment, including the assembly of all usual equipment for the MIST procedure. Boxes containing vials of surfactant, or empty boxes made available for the study, were taken from refrigeration to the bedside.

At the infant's bedspace, the study team assumed care of the infant, and screened the infant and activities of the team from view, using curtains and/or portable partitions. Central physiological monitors were disconnected where possible, with a separate study oximeter used for heart rate and oxygen saturation monitoring during the intervention. The MIST intervention comprised placement of the thin catheter into the trachea under laryngoscopic guidance, followed by administration of surfactant in at least 3 aliquots, with a 10 second pause between each. For the sham intervention, the infant was briefly repositioned as for laryngoscopy, but no other physical intervention was performed. The study team were instructed that their actions and verbal interactions were to mimic what would be expected during a MIST intervention, in such a way that the nature of the intervention could not be discerned externally. [10] They were also instructed that the duration of the intervention was to be similar to MIST.

Following the intervention, the study team returned the infant to their position in the cot prior to the screens being drawn, restored pre-intervention continuous positive airway pressure levels (as far as possible), and cleared the bedspace of any identifying items.

Survey of blinding success

Immediately following the completion of each OPTIMIST-A intervention, paper-based questionnaires were administered by the study team to three members of the clinical staff. Where practicable, the questionnaires were given to the bedside nurse, a member of the medical team, and another staff member involved in clinical care. Staff were asked whether they were able to make an informed guess regarding treatment allocation (MIST or sham), and if so, which one they thought had been given. The three possible responses were 'MIST', 'sham', or 'don't know'. Staff could write a comment in justification of their choice and were asked to indicate their professional role: i) bedside nurse; ii) other nurse (float nurse, nurse in charge of shift, nurse in the room but caring for other infant(s)); iii) neonatal trainee (resident, registrar, nurse practitioner); iv) neonatologist; v) respiratory therapist. Questionnaire responses were anonymous, and the actual treatment allocation was not revealed to respondents.

Statistical methods

Staff questionnaires were gathered by the OPTIMIST-A research team at each site and sent electronically to the Data Management Centre. The staff responses, along with the treatment allocation, were entered into the Trial Database, and the responses categorized as correct, incorrect, or unsure ('don't know'). Justifications for the choices made were extracted from the questionnaires and provided additional context to the quantitative data.

Two validated blinding indices were used to evaluate the data gathered in the post-intervention surveys: i) the James index, [15] which provides a simple measure of overall blinding success, returning a value between 0 and 1, with 1 equal to all 'don't know' responses (complete blinding) 0 equal to all correct answers (complete unblinding) and 0.5 where respondents' guesses appear random (50%

correct, 50% incorrect); and ii) the Bang index [3] in which separate measurements of success are calculated for active and sham treatment groups, returning a value between -1 to 1 for each, 1 = all guesses correct (complete unblinding), -1 = all guesses incorrect(may indicate complete unblinding) or0 = perfect blinding. The Bang index can be interpreted as the proportion of correct guesses beyond random chance.

For the James index, successful blinding was defined as a value of ≥ 0.50.[3] For the Bang index, successful blinding was deemed to be indicated by a score with an absolute value of ≤ 0.3 (i.e. - 0.3 to 0.3). [3] Blinding success was determined overall (James index), and separately for the two treatment groups (Bang index), with calculation of the index and 95% confidence intervals (CI) following the approach of Williamson. [16] These analyses were performed using R Statistical Software. [17]

The proportion of correct, incorrect, and 'don't know' responses was compared between clinical roles using a chi square test; the two blinding indices were also calculated within each clinical role. Additionally, the association between blinding success of two procedural factors was examined. These were: i) duration of the interventional procedure (time in minutes, taken from commencement of treatment to restoring the infant to the original position and continuous positive airway pressure settings), and ii) the change in oxygenation post-procedure (FiO_2 5 minutes post-intervention minus FiO_2 immediately prior to intervention). For the duration of the intervention all responses were distributed in bins at 1-minute intervals and the Bang index calculated for each bin. A weighted regression (weighted by the number of responses in each bin) was used to examine the relationship between duration and blinding success. For the change in oxygenation, responses were distributed in sequential bins of 0.05 change in FiO_2 and the same procedure followed as for duration of the intervention.

RESULTS

Baseline characteristics of the 485 randomised infants included in the primary analysis were similar overall between the MIST (n= 241) and sham (n=244) arms, as previously described. [7] Data were analysed from 1345 post-intervention questionnaires completed by treating clinicians (average 2.8 per recruited infant). Of these, 441 (33%) were correct guesses, 142 (11%) were incorrect guesses, and in 762 instances (57%) 'don't know' was recorded. These proportions were similar between the MIST and sham groups (Table 1).

The James index was 0.67 (95% CI 0.65 – 0.70), indicating successful blinding for the trial overall. The Bang index was 0.17 (95% CI 0.12 – 0.21) for the sham arm, and 0.28 (95% CI 0.23 – 0.32) for the MIST arm, indicating successful blinding in the former case and marginal blinding success in the latter (Table 2).

There was a difference in the proportion of correct and incorrect guesses and 'don't know' responses between staff in different roles (Figure 1), $\chi^2_8 = 31.824$, P = <0.01. Neonatologists were the only group having a greater proportion of correct guesses (47%) than 'don't know' responses (41%), and as a group had a higher frequency of correct guesses than other staff: bedside nurses (36%), neonatal trainees (31%), other nurses (24%) and respiratory therapists (19%). The James index lower confidence limit was >0.50 for all clinicians apart from neonatologists (Table 2). The Bang index showed that blinding was more successful in the sham arm across all staff groups (Table 2).

In the MIST arm, there was a negative relationship between the duration of the interventional procedure and blinding success measured by the Bang index (Figure 2). This relationship was not seen in the sham arm. The degree of improvement in oxygenation also was associated with a decrease in blinding success in the MIST arm, but this relationship was not seen in the sham arm (Figure 3).

The majority of the questionnaires were returned with comments. Common reasons provided for choosing MIST included: 'decrease in work of breathing', 'FiO_2 weaned from 30% to air', 'better oxygenation' and 'FiO_2 decreased'. Justifications for choosing sham included: 'no change in oxygen requirement', 'minimal change in FiO_2' and 'no clinical change'. Examples of comments accompanying don't know responses were: 'FiO_2 has gone down but work of breathing remains moderate', 'possibly had surfactant but only small improvement in FiO_2', 'well camouflaged' and 'did not know what was happening behind closed curtains'.

DISCUSSION

Optimal design of parallel group RCTs should include blinding of clinicians to treatment allocation, but this is more difficult to achieve with procedural interventions, and its success is rarely measured. In this neonatal RCT comparing a procedural intervention (MIST) with a sham treatment, we found that blinding of clinicians was achievable, and the degree of success measurable, with the confidence intervals of the James blinding index being >0.50 overall, and >0.50 for clinicians in all roles apart from neonatologists. Where analysed by intervention group, blinding was more successful in the sham treatment group, and both prolonged procedural duration and greater improvement in oxygenation post-intervention were identified as barriers to successful blinding.

Two complementary blinding indices providing different insights into blinding success were used to evaluate the staff responses from the post-intervention questionnaires. Our intent was to draw on the combined strength of the two indices, as suggested by several authors[3,18]. The James blinding index, which is recommended for use in hypothesis testing, returned values indicating successful blinding for the trial overall and for staff in most roles. This index places emphasis on the 'don't know' responses, arguably the ideal feedback to be received if a blinding strategy is successful, unless indicative of 'a socially desirable response'.[3] The protection of anonymity in the OPTIMIST-A

questionnaire did allow clinicians the freedom to choose 'don't know'. Comments giving justification where 'don't know' was selected certainly appeared to indicate that the chosen response was appropriate.

The Bang blinding index can be directly interpreted as the proportion of respondents guessing the treatment allocation correctly beyond chance, and allows insight into different guessing behaviours in the treatment and control arms of an RCT. In our study the Bang Index indicated a rate of guessing correctly beyond chance (i.e. unblinding) of 17% in the sham group, and 28% in the MIST group. The determination of how much unblinding is acceptable in a particular RCT is not straightforward; [19] we followed the suggestion of Bang et al. in nominating a threshold of 30%, to indicate an undesirable level of unblinding.

The assessment of blinding amongst staff in different roles followed the same pattern of blinding success with both the James and Bang index with neonatologists showing the highest degree of unblinding. Neonatologists also had the highest frequency of correct responses and lowest frequency of 'don't know' answers. This may indicate an effect of experience and awareness of the clinical course of infants enrolled in the study and may also indicate a reluctance to respond indecisively.

The higher frequency of correct responses amongst neonatologists need not be considered in a negative light. It is clear from the comments on the questionnaires that while many of their correct guesses of MIST as the allocated treatment were supported by such phrases as 'Lower FiO_2 needs', incorrect responses from neonatologists indicating MIST in relation to infants receiving a sham procedure were on occasions also supported by similar commentary, e.g., 'Baby stabilised and improved after procedure'. This suggests that while the neonatologists may have had a strong suspicion regarding the treatment allocation, they could not be certain and there was no evidence of

systematic unblinding. We would contend that this lack of certainty allowed neonatologists and other treating clinicians the freedom to make unbiased decisions about management in the crucial hours following a study intervention. This was a particularly important consideration in the OPTIMIST-A trial, where a choice regarding whether to intubate an infant with respiratory distress syndrome was often shortly after the trial intervention.

The influence on blinding success of two procedural factors, the duration of the study intervention and the change in FiO_2 post-procedure, was investigated, these being two obvious and measurable factors that could have a bearing on blinding of surfactant delivery. Both factors had a significant relationship with blinding success for the MIST group only, such that blinding was less effective with longer procedural duration and greater improvement in FiO_2, in particular beyond a duration of 10 minutes, or FiO_2 improvement of 0.10. Unblinding may be inevitable if the administered treatment is immediately effective, as was apparent in some of the comments where MIST was correctly chosen as the group of allocation (e.g. 'FiO_2 has fallen from 42 to 21%'). An immediately apparent clinical improvement after a procedural intervention in an unavoidable factor that may limit the effectiveness of blinding in an RCT. It would be rare, however, for the difference in effect to be so stark between active and sham treatments as to negate that value of blinding; the findings of the present study support this. Duration of the intervention should be tightly controlled, with a pre-specified timeframe for the intervention being more rigidly imposed regardless of group of allocation. Importantly, knowledge of the specific factors that compromised blinding is highly valuable in planning future studies.

It is acknowledged there may be negative impacts of undertaking a blinding strategy on trial recruitment. [2] In the OPTIMIST-A trial a separate study team carried out the intervention. This team took time to mobilise, and in some instances added to the trial costs. Analysis of the reasons given

for non-enrolment of the 1192 eligible infants[7] showed that only 69 (6%) were not randomised due to lack of availability of a study team.

There is a paucity of studies of blinding of neonatal RCTs with which to make direct comparisons. Indeed, there are very few studies that provide a detailed account of blinding procedures and their success. However, we note two unrelated studies that provide helpful descriptions of how the success of blinding was assessed using the James and Bang indices. [20, 21]

The act of blinding aims to provide protection from bias and the assessment of the success of a blinding strategy provides useful insights into the protection of allocation concealment in future studies. To improve methodological quality in RCTs and deliver more robust results, it should become commonplace for studies, where blinding is possible, to provide detailed accounts and assessment of the efficacy of blinding procedures.

CONCLUSION

This analysis of the success of blinding in the OPTIMIST-A trial suggests that blinding of a procedural intervention from clinicians can be achieved, and is measurable, in neonatal RCTs.

ACKNOWLEDGEMENTS

We thank clinical and research staff in the 33 participating centres, and parents and infants involved in the study.

Trial Registration: Australian New Zealand Clinical Trial Registry (ACTRN12611000916943).

Disclaimer

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Royal Hobart Hospital Research Foundation [grant number 11-382]; and the Australian National Health and Medical Research Council [grant number 1049114]. Chiesi Farmaceutici provided the exogenous surfactant at reduced cost.

REFERENCES

1. Hrobjartsson A, Forfang E, Haahr MT, et al. Blinded trials taken to the test: an analysis of randomized clinical trials that report tests for the success of blinding. *Int J Epidemiol* 2007; 36: 654–663.

2. Monaghan TF, Agudelo CW, Rahman SN, et al. Blinding in clinical trials: Seeing the big picture. *Medicina* 2021; 57: 647.

3. Bang H, Flaherty SP, Kolahi J, et al. Blinding assessment in clinical trials: A review of statistical methods and a proposal of blinding assessment protocol. *Clin Research and Regulatory affairs* 2010; 25: 143–156.

4. Kolahi J, Bang H and Park J. Towards a proposal for assessment of blinding success in clinical trials: up-to-date review. *Community Dent Oral Epidemiol* 2009; 37:477–484.

5. Nai Ming Lai, Ong JMJ, Chen KH, et al. Are neonatal trials better conducted and reported over the last 6 decades? An analysis on their risk-of-bias status in Cochrane reviews. *Neonatology* 2019; 116:123-131.

6. Enhorning G, Shennan A, Possmayer F, et al. Prevention of neonatal respiratory distress syndrome by tracheal installation of surfactant: A randomised controlled trial. *Pediatrics* 1985; 76:145.

7. Dargaville PA, Kamlin COF, Orsini F, et al. Effect of minimally invasive surfactant therapy vs sham treatment on death and bronchopulmonary dysplasia in preterm infants with respiratory distress syndrome. *JAMA* 2021; 326(24):2478-2487.

8. Abdel-Latif ME, Davis PG, Wheeler KI et al. Surfactant therapy vis thin catheter in preterm infants with or at risk of respiratory distress syndrome (Review). *Cochrane Database of Systematic Reviews* 2021; 5: CD011672.

9. Boutron I, Guittet L, Estellat C, et al. Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. *PLoS Med* 2007; 4: e61.

10. Wartolowska K, Beard D and Carr A. Blinding in trials of interventional procedures is possible and worthwhile. *F1000research,* 2017; 6: 1663.

11. Dargaville PA, Kamlin COF, De Paoli AG, et al. The OPTIMIST-A trial: evaluation of minimally invasive surfactant therapy in preterm infants 25-28 weeks gestation. *BMC Pediatr*. 2014; 14:213.

12. Bose C, Corbet A, Bose G, et al. Improved outcome at 28 days of age for very low birth weight infants treated with a single dose of a synthetic surfactant. *J Pediatr*. 1990; 117:947-953.

13. Stevenson D, Walther F, Long W, et al. Controlled trial of a single dose of synthetic surfactant at birth in premature infants weighing 500 to 699 grams. *J Pediatr*. 1992; 120: S3-S12.

14. Reininger A, Khalak R, Kendig JW, et al. Surfactant administration by transient intubation in infants 29 to 35 weeks' gestation with respiratory distress syndrome decreases the likelihood of later mechanical ventilation: A randomized controlled trial. *Journal of Perinatology* 2005;25: 703-708.

15. James KE, Bloch DA, Lee KK, et al. An index for assessing blindness in a multi-centre clinical trial: Disulfiram for alcohol cessation-A VA cooperative study. *Statistics in Medicine* 1996; 15:1421-1434.

16. Williamson, CF. On the Measurement of Blinding in Randomized, Controlled Trials. PhD Thesis, 2014, Baylor University.

17. R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

18. Webster RK, Bishop F, Collins GS, et al. Measuring the success of blinding in placebo-controlled trials: Should we be so quick to dismiss it? *Journal of Clinical Epidemiology*, 2021; 135: 176 – 181.

19. Freed B, Williams B, Situ X, et al. Blinding, sham, and treatment effects in randomized controlled trials for back pain in 2000-2019: A review and meta-analytical approach; *Clinical trials* 2021; 18 (3) 361-370.

20. Houweling AH, Shapiro S, Cohen JM, et al. Blinding strategies in the conduct and reporting of a randomized placebo-controlled device trial; *Clinical Trials* 2014; 11 (5) 547-552.

21. Hertzberg V, Chimowitz M, Lynn M, et al. Use of dose modification schedules is effective for blinding trials of warfarin: evidence for the WASID study. *Clinical Trials* 2008; 5: 23-30.

Table 1. Clinician responses by actual treatment allocation.

Actual group allocation	Guess, *n*			
	MIST	Sham	Unsure	Total
MIST	**246 (36%)**	59 (9%)	371 (55%)	676
Sham	83 (12%)	**195 (29%)**	391 (58%)	669
Total	329 (24%)	254 (19%)	762 (57%)	1345

Correct responses highlighted in bold. Responses were categorized in the data overall as correct in 441 (33%), incorrect in 142 (11%) or unsure ('don't know') in 762 (57%).

Table 2. Blinding indices by clinical role.

Clinician role	n	James index [a]	Bang index [b]	Bang index [b]
			Sham arm	MIST arm
All	1345 [c]	0.67 (0.65-0.70)	0.17 (0.12-0.21)	0.28 (0.23-0.32)
Bedside nurse	594	0.64 (0.60-0.68)	0.20 (0.13-0.27)	0.30 (0.23-0.38)
Other nurse	301	0.76 (0.71-0.81)	0.08 (-0.01-0.16)	0.22 (0.13-0.31)
Neonatal trainee	331	0.69 (0.64-0.74)	0.16 (0.07-0.25)	0.26 (0.17-0.35)
Neonatologist	93	0.53 (0.43-0.63)	0.28 (0.09-0.47)	0.44 (0.25-0.64)
Respiratory therapist	21	0.81 (0.62-1.0)	0.00 (-0.39-0.39)	0.00 (-0.36-0.36)

[a] James index (95% confidence interval), values ≥0.50 defined as satisfactory blinding.

[b] Bang index (95% confidence interval), values -0.30 to 0.30 defined as satisfactory blinding.

[c] Includes 5 staff members who did not specify their clinical role.

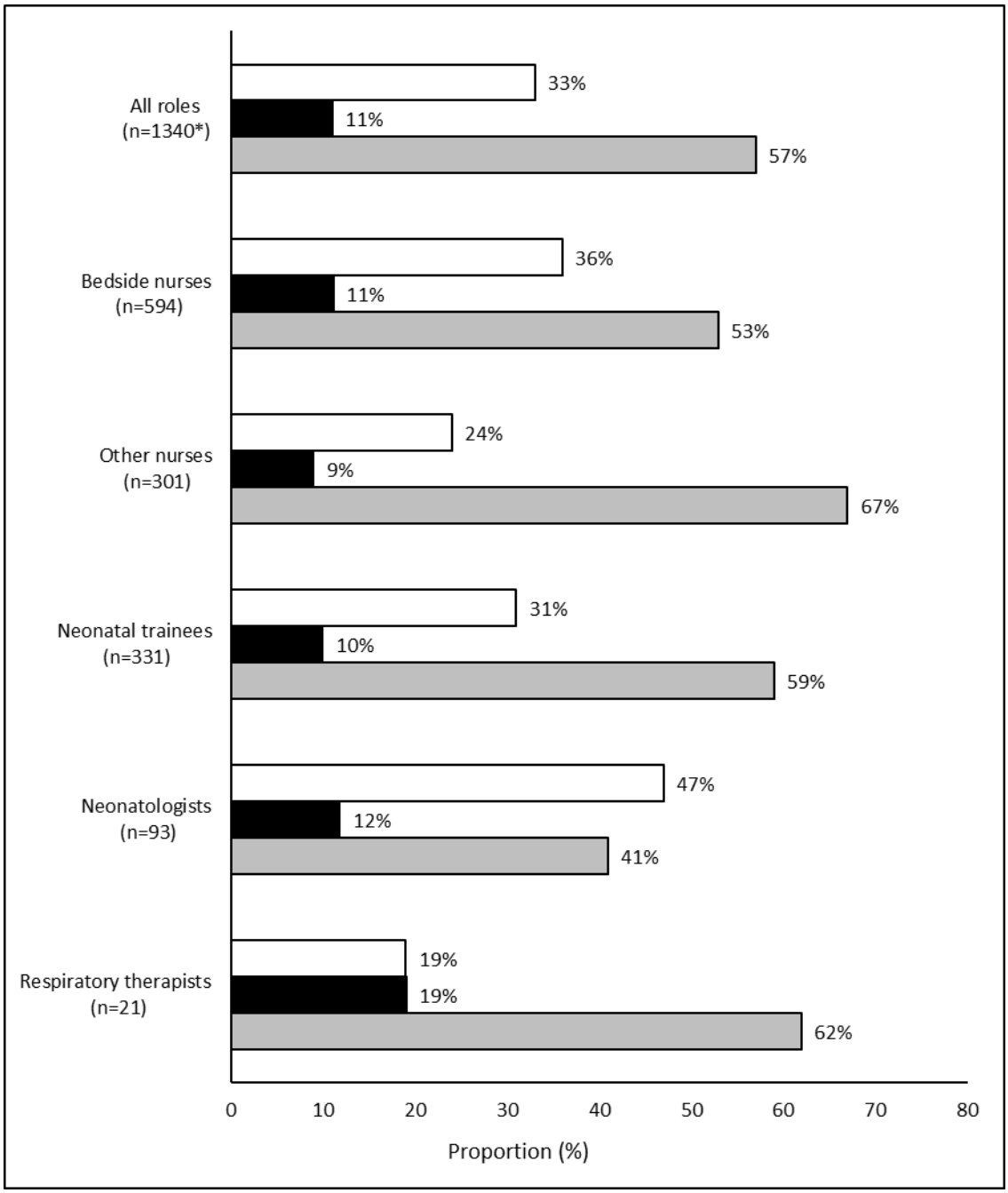

Figure 1. Clinician responses

Proportion of correct (white bar), incorrect (black bar) and "don't know" (grey bar) responses for all clinical staff (at top), and by role.

*1340/1345 returned questionnaires identified a staff role

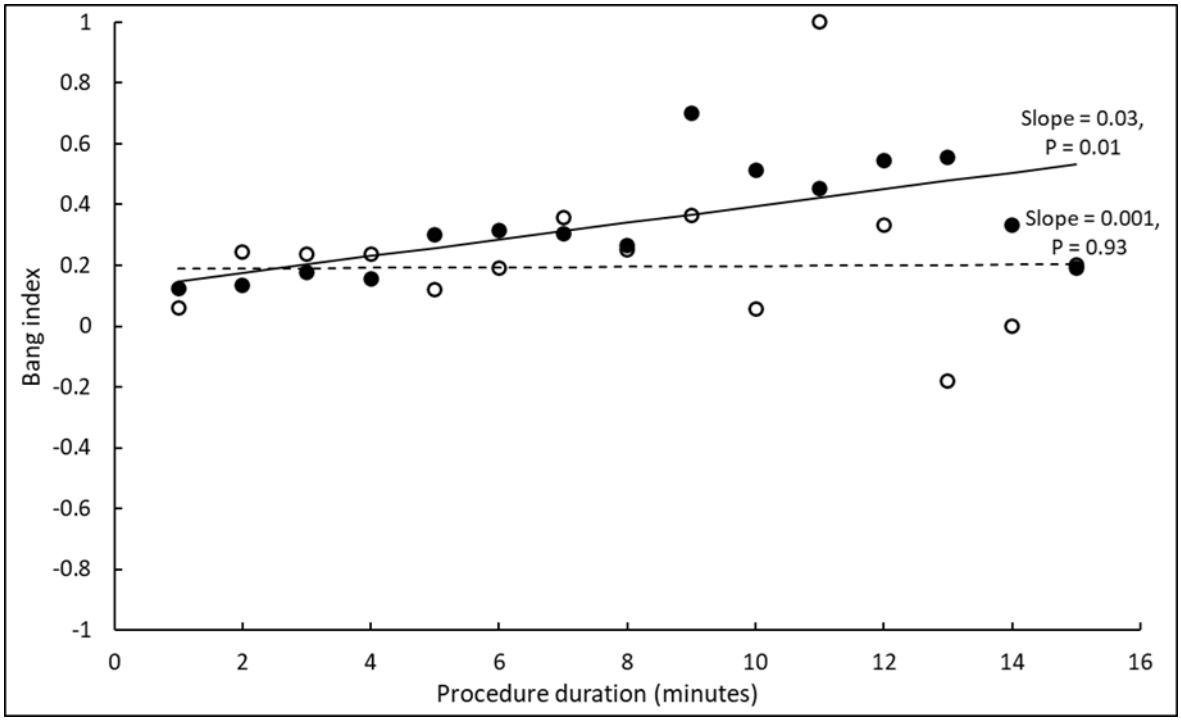

Figure 2. Association between blinding success and procedure duration.

Weighted linear regression of within-bin Bang index on duration of the intervention, where x-axis data are the mid-points of interval bins containing all observations of procedure duration. pre-procedure; negative values represent improvements in oxygenation and are shown on the right side of the plot. Each data point represents the Bang index for the pooled staff response in that bin. Separate plots for MIST arm (black circles, solid regression line) and sham arm (white circles, dashed regression line). Slope of the regression line and associated P value are shown in each case.

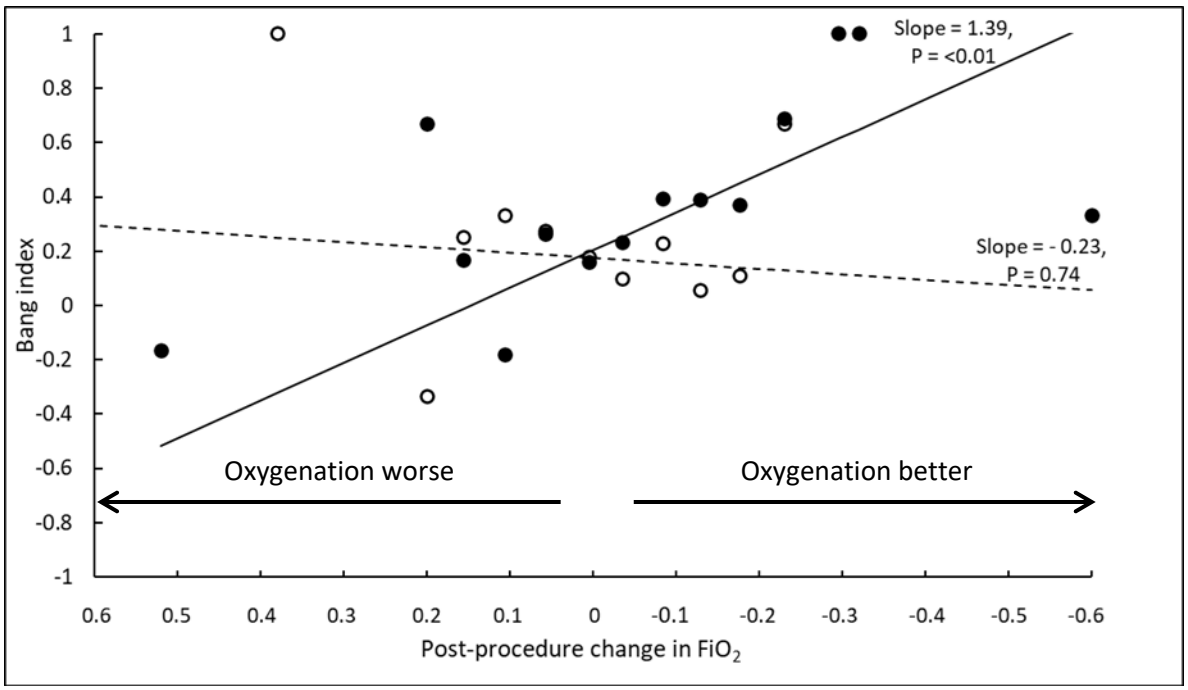

Figure 3. Association between blinding success and oxygenation improvement post-procedure.

Weighted linear regression of within-bin Bang index on post-procedure change in FiO_2, where x-axis data are the mid-points of bins containing all observations of FiO_2 alteration. Change in FiO_2 calculated as FiO_2 5 min post-procedure minus FiO_2 immediately pre-procedure; a negative value represents improvement in oxygenation. Each data point represents the Bang index for the pooled staff response in that bin. Separate plots for MIST arm (black circles, solid regression line) and sham arm (white circles, dashed regression line). Slope of the regression line and associated P value are shown in each case.