PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Non-Invasive Positive airway Pressure thErapy to Reduce Postoperative Lung complications following Upper abdominal Surgery (NIPPER PLUS): Protocol for a single centre pilot randomised control trial
AUTHORS	Lockstone, Jane; Boden, Ianthe; Robertson, Iain; Story, David; Denehy, Linda; Parry, Selina M.

VERSION 1 – REVIEW

REVIEWER	Victor, Suresh King's College London, UK
REVIEW RETURNED	07-Aug-2018

GENERAL COMMENTS	The authors have submitted the protocol of NIPPER PLUS Trial for review. The paper describes in good detail the trial protocol and could be shortened. This is a pilot trial and the sample size is small. As such useful interpretation of results is limited with this sample size. Furthermore, I have a few concerns with trial methodology: 1. The use of sealed envelops in boxes should be avoided especially when so many web based randomization tools/ telephone randomization services are readily available. 2. I am surprised there is no stratification other than treatment location. Isn't age of the patient a significant factor? Are the researchers confident of equal distribution of different age groups between the two arms within a sample size of 90 or 40 patients. 3. How will the research team deal with differences in gender, bmi, pre-existing lung disease, smoking etc between groups? Will adaptive randomization be better suited given small numbers? 4. I think blinding of treatment allocation cannot be reliably performed when so many clinical staff are unblinded to the treatment allocation. MGSV includes 'Physician diagnosis', who are not blinded to treatment allocation. Points 1, 2 and 8 of the MGSV score are subjective. Isn't there a more objective score or can this score be modified/ enhanced with stricter criteria to make it more objective. Example: use of color chart for item 2? I am not sure whether the study is ongoing and any protocol changes can be performed at this stage. I am concerned that there are flaws in the study that would prevent a sound interpretation of data. I hope the team will be able to address them.

REVIEWER	Jaber, S. University hospital of Montpellier - France
REVIEW RETURNED	10-Aug-2018

GENERAL COMMENTS	I read with interest the study planned by Lockstone et al. The protocol is clear and well written. The study aims to detect whether there is a possible signal towards postoperative pulmonary complications (PPC) reduction with the use of additional intermittent non-invasive ventilation (NIV) compared to continuous high-flow nasal oxygen therapy alone following high-risk elective upper abdominal surgery in patients at risk. However, I have 2 potential concerns: - First, if the authors consider high-flow nasal oxygen therapy alone as the reference treatment, the SHAM group is not exactly the standard. Indeed, in the OPERA study (Futier el al. ICM 2016), performed in postoperative patients at moderate to high risk of postoperative pulmonary complications who had undergone major abdominal surgery, early preventive application of high-flow nasal oxygen after extubation did not result in improved pulmonary outcomes compared with standard oxygen therapy. In the NIVAS study (Jaber et al. JAMA 2016), the major result was the superiority of NIV over standard oxygen in post-operative ARF. Therefore, the need to perform a study comparing HFNC to NIV following high-risk elective upper abdominal surgery may be questioned. - Second, the number of patients needed to include is based to detect a 75% difference in the PPC rate. I understand the issue of funding, however I wonder if it is ethical to perform such a study. The results will likely show no difference between groups, without having the power to draw a real conclusion. The external validity will also be reduced given the single-center design. A retrospective analysis could be done to justify the conduct of a large multicenter blinded study with adequate power to conclude on the superiority of NIV + high-flow oxygen over high-flow oxygen alone. Finally, the considered sample size seems not correct to detect a significant difference in my point of view. The authors should absolutely consider a new analysis and probably increase the number of included patients.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1
Reviewer Name: S VICTOR
Institution and Country: King's College London, UK

The authors have submitted the protocol of NIPPER PLUS Trial for review. The paper describes in good detail the trial protocol and could be shortened. This is a pilot trial and the sample size is small. As such useful interpretation of results is limited with this sample size.

Author Response
Thank you for your comments and thoughtful review. We are pleased that you consider the trial protocol has been described in good detail. Whilst we agree the sample size is small, this study is a pilot. We do not expect to be powered for the primary outcome as stated in the manuscript and supported in the CONSORT statement for pilot and feasibility studies (Eldridge et al. *BMJ* 2016). Rather in this pilot trial, we aim to assess trial feasibility including: safety of modalities, recruitment

and retention, and treatment fidelity. The sample size will provide an estimate of effect only between groups for postoperative pulmonary complications. Our results will inform the protocol amendments required prior to designing and implementing a fully powered, randomised control trial.

Furthermore, I have a few concerns with trial methodology:
1.The use of sealed envelopes in boxes should be avoided especially when so many web based randomization tools/ telephone randomization services are readily available.

Author Response
1.Thank you for your comment. We acknowledge that telephone or web-based randomisation schedules are now readily available and reduces the risk of subversion of the concealed allocation process compared to sealed envelopes (however can never eliminate this risk entirely). We reassure the reviewers that potential threats to internal validity, including selection bias, were considered very carefully prior to commencing NIPPER PLUS and a number of features were implemented to minimise this potential risk. These features included wrapping allocation cards in aluminium foil; locking of the allocation sequence at the research site institute and limiting its availability (only the lead investigator had access via written application to the research committee) and the recording of participant details directly on the corresponding numbered randomisation envelope to facilitate independent audit of possible fraudulent subversion to the randomisation process. Audit will be undertaken at trial completion.

Opaque sealed envelopes were considered the most feasible, low-tech, and cost-effective option of allocation concealment given our investigator clinician-initiated trial had limited funding at trial commencement. It is possible (though in our opinion not likely) that individual physiotherapists involved in the trial could have steamed opened envelopes prior to recruiting patients sabotaging a true randomisation process. We ensure throughout the trial that treating physiotherapists are not responsible for accessing nor opening the randomisation envelope, only the lead investigator or a site investigator can open the envelope once participant eligibility is confirmed.

We have edited the manuscript to make it clearer that envelopes were pre-prepared by non-trial personnel (highlighted on page 7, line 237) and that only site or lead investigators could open the envelopes (highlighted on page 8, line 263).

Methods:
A research assistant independent to the trial pre-prepared 130 sequentially numbered (1-130) opaque envelopes each containing an allocation card wrapped in aluminium foil. Allocation sequence is generated by a web-based computer program (http://www.randomizer.org/).

Eligible consenting patients are then randomised into the trial by the lead or a site investigator only by opening the next sequentially numbered sealed opaque envelope according to the patient's planned postsurgical destination (ward or ICU/HDU). Once opened, participant's details are written on the envelope to ensure that patients were randomised in presenting order and these are filed securely along with the signed consent form.

2.I am surprised there is no stratification other than treatment location. Isn't age of the patient a significant factor? Are the researchers confident of equal distribution of different age groups between the two arms within a sample size of 90 or 40 patients.

Author Response
2.Thank you for your comment. As NIPPER PLUS is a pilot study, the majority of our focus lies in the feasibility of the study. We felt that the ease of providing the modalities, both NIV and high-flow nasal prong oxygen therapy could be biased towards ICU therefore we wanted to ensure we stratified primarily to location. Whilst age was shown to be a significant factor in the increased risk of PPC in surgical patients (Schultz et al. *Eur J Anaesthesiol* 2017), our previous work in this field has not highlighted age as significant predictor of PPC in upper abdominal surgery patients using the same PPC criteria as our pilot study (Scholes et al. *Aust J Physiother* 2009). Due to the uncertainty of known versus possible covariates which may increase risk of PPC, we will be analysing our results and aim to ensure our trial manages this as previously recommended (Roberts and Torgerson. *BMJ*

1999,) and outlined below (highlighted). Whilst we cannot guarantee equal distribution of age between the two arms, we have stated a priori in the manuscript that age is a specific prognostic covariate that will be adjusted in our analysis if found to be imbalanced between groups (highlighted on page 16, line 516).

Statistical methods

Adjustment covariates will be selected by backward stepwise regression from covariates that may have the potential for clinically significant alterations in effect sizes. These include: smoking history, age, length in time of operation, operation category (upper gastrointestinal, colorectal, urological, other), incision type and location[59], intraoperative ventilation strategies[3 60], fluid delivery[61], blood transfusions[62], and mode of post-operative analgesia[63].

3. How will the research team deal with differences in gender, bmi, pre-existing lung disease, smoking etc between groups? Will adaptive randomization be better suited given small numbers?

Author Response

3.Thank you for your comment. We recognise differences between baseline groups can influence and bias the outcome. We carefully considered this and discussed this with our statistician before implementing the trial. We aim to ensure our trial manages this as previously recommended (Roberts and Torgerson. *BMJ* 1999,) and outlined above in response to comment 2. We have recently undertaken this same approach in our RCT in abdominal surgery published in the BMJ this year (Boden et al. *BMJ* 2018). Considering that in at least 1 in 20 demographic characteristics baseline imbalances between randomised groups will occur by chance (Lydersen S. *Ann Rheum Dis* 2015), we have stated that findings would be adjusted if imbalances were found in any of the specified prognostic covariates that could potentially confound the primary outcome. These include BMI, prior respiratory co-morbidity, smoking etc. Both adjusted and unadjusted results will be reported in the final manuscript. In response to comment 2 and 3 we have also amended the manuscript (highlighted on page 16 lines 511 - 514) to state more clearly there may be baseline differences.

Statistical methods

As our study is stratified to postoperative location (ICU/WARD) only, there is a possibility of significant baseline differences between groups. This will be managed according to the prognostic strength and size of imbalances due to potential confounding baseline variables between groups being assessed[58]. Adjustment covariates will be selected by backward stepwise regression from covariates that may have the potential for clinically significant alterations in effect sizes. These include: smoking history, age, length in time of operation, operation category (upper gastrointestinal, colorectal, urological, other), incision type and location[59], intraoperative ventilation strategies[3 60], fluid delivery[61], blood transfusions[62], and mode of post-operative analgesia[63].

4.I think blinding of treatment allocation cannot be reliably performed when so many clinical staff are unblinded to the treatment allocation. MGSV includes 'Physician diagnosis', who are not blinded to treatment allocation. Points 1, 2 and 8 of the MGSV score are subjective. Isn't there a more objective score or can this score be modified/ enhanced with stricter criteria to make it more objective. Example: use of color chart for item 2?

Author Response

4.Thank you for your comment. We agree that many of the postoperative clinical staff will be unblinded to the treatment allocation (such are the issue of undertaking pragmatic clinical trials) and some aspects of the MGS PPC diagnostic tool may be considered as subjective. However, the PPC assessments are undertaken by a research assistant independent to the trial and who has no clinical involvement with the participants. The four clinical factors (auscultation, sputum, room air oxygenation, and temp>38) are assessed by direct patient assessment and from the medical record,

and the four diagnostic factors (CXR, WCC, sputum microbiology, and physician documentation of a diagnosis) are extracted from the medical record by the blinded assessor. Whilst the individual physician may not be blinded to group allocation, this is only one of eight criteria and to be positive for a PPC a minimum of 4 are necessary.

Additionally, the PPC score is applied equally to both intervention and control groups, minimising the effects of differences in diagnostic methods. The PPC tool used in this pilot study is shown to be sensitive to therapeutic interventions designed to ameliorate postoperative atelectasis and alveolar de-recruitment (Boden et al. *BMJ* 2018). There are indeed many different methods published in the literature used to define a PPC. We have previously utilised and published this same method in upper abdominal surgery trials (Boden et al. *BMJ* 2018, Parry et al. *Physiotherapy* 2014, Haines et al. *Physiotherapy* 2013 and Scholes et al. *Aust J Physiother* 2009). We have amended the manuscript to highlight the assessor has no clinical involvement (page 11, lines 359 - 360) and have included below a participant's data collection sheet example completed by the blinded assessor.

Outcome measures:
An assessor blinded to group allocation, who has no clinical involvement with the study, assesses participants prospectively and daily for a PPC until the seventh postoperative day. Thereafter, additional PPC assessments are only performed if clinically indicated when there are signs of respiratory deterioration reported in the medical record until postoperative day 14 or hospital discharge, whichever occurs first.

WEEK 1							
POD	1	2	3	4	5	6	7
Date	31/03/2017	01/04/2017	02/04/2017	03/04/2017	04/04/2017	05/04/2017	06/04/2017
Weekday	Fri	Sat	Sun	Mon	Tues	Wed	Thurs
Location	ICU	Surg ward	Surg ward	Surg ward	Surg ward	Surg ward	Surg ward
Oxygen/ventilation	NP	NP	nil	nil	nil	nil	nil
Analgesia (1-7)	3	3	3				
Analgesia (1-7)			5	5	5	5	5
Analgesia (1-7)							
Assess patient for a PPC EVERY DAY for 7 days or until d/c from hospital							
1: Does the patient have a PPC? If no, progress to D/C Ax. If yes, treat as necessary.							
Observations: Mark presence of sign or symptom for each day that it is present (y/n/not avail)							
Ausc changes	y	n	n	n	n	n	n
Sputum changes	n	n	n	n	n	n	n
SpO2 < 90% on RA*	n	n	n	n	n	n	n
Temp > 38*	n	n	n	y	y	y	n
sputum culture +ve	not avail	not avail	not avail	not avail	not avail	not avail	not avail
WCC >11 or AB's	n	n	n	n	n	n	n
CXR changes	not avail	not avail	not avail	not avail	n	not avail	not avail
Dr dx of a PPC	n	n	n	n	n	n	n
PPC diagnosis	n	n	n	n	n	n	n

5. I am not sure whether the study is ongoing and any protocol changes can be performed at this stage. I am concerned that there are flaws in the study that would prevent a sound interpretation of data. I hope the team will be able to address them.

Author Response
5. Thank you for your comments. Unfortunately, protocol changes cannot be performed at this stage as we are in the recruitment phase. However, I do hope that our responses have helped to address

your concerns. Your comments are an excellent summary of the known and recognised limitations of this pilot study (thank you) and we believe that these methodological flaws have been openly articulated in this revised manuscript.

Reviewer: 2
Reviewer Name: S. JABER

Institution and Country: University hospital of Montpellier - France

Please state any competing interests or state 'None declared': None

Please leave your comments for the authors below I read with interest the study planned by Lockstone et al. The protocol is clear and well written. The study aims to detect whether there is a possible signal towards postoperative pulmonary complications (PPC) reduction with the use of additional intermittent non-invasive ventilation (NIV) compared to continuous high-flow nasal oxygen therapy alone following high-risk elective upper abdominal surgery in patients at risk.

However, I have 2 potential concerns:

- First, if the authors consider high-flow nasal oxygen therapy alone as the reference treatment, the SHAM group is not exactly the standard. Indeed, in the OPERA study (Futier el al. ICM 2016), performed in postoperative patients at moderate to high risk of postoperative pulmonary complications who had undergone major abdominal surgery, early preventive application of high-flow nasal oxygen after extubation did not result in improved pulmonary outcomes compared with standard oxygen therapy. In the NIVAS study (Jaber et al. JAMA 2016), the major result was the superiority of NIV over standard oxygen in post-operative ARF. Therefore, the need to perform a study comparing HFNC to NIV following high-risk elective upper abdominal surgery may be questioned.

Author Response
Thank you for your comment. Whilst we recognise high-flow nasal prong oxygen therapy may not be considered standard care post major abdominal surgery in other hospitals and this may also affect the generalisability of our results, our pilot work prior to implementing this study demonstrated significant increases in the use of high-flow in this patient population in our hospital from 2013 to 2017, and similarly in Australian hospitals generally. Since the use of high-flow is current practice within our ICU, our Intensivists could not support a non high-flow nasal prong oxygen therapy group. We undertook a prior observational, pre-post cohort, single centre study, consisting of 182 consecutive high-risk elective upper abdominal surgery patients. This manuscript is currently under peer review for publication. The pre-cohort group received standardised pre-operative physiotherapy, standardised early ambulation and no additional respiratory physiotherapy postoperatively (Boden et al, *BMJ* 2018). This historical cohort was then compared to the post-cohort group who received standardised pre-operative physiotherapy, early ambulation (non standardised) and additional prophylactic postoperative intermittent NIV. A significant reduction in PPC, once baseline differences were accounted for was demonstrated in the post-cohort group. However, there were also significant confounders including increased use of high-flow nasal prong oxygen therapy in the post-cohort group. This current pilot study was therefore designed to eliminate the confounders from our observational study including high-flow nasal prong oxygen therapy.

We read with interest the results from the OPERA trial and whilst the authors found no differences in pulmonary outcomes compared to standard care, the duration which high-flow nasal prong oxygen therapy was applied for was a considerably shorter time (median duration of 15 [IQR 12-18] hours following extubation) than our current clinical practice in our setting (48 hours following extubation) and we felt this was worth investigating.

The provision of NIV in response to known acute respiratory deterioration prevents reintubation rates and improves clinical outcomes compared to standard oxygen therapy (Jaber et al. *JAMA* 2016), however the benefits, feasibility and safety of NIV to prevent acute respiratory deterioration compared with longer use of high-flow following major abdominal surgery is uncertain.

- Second, the number of patients needed to include is based to detect a 75% difference in the PPC rate. I understand the issue of funding, however I wonder if it is ethical to perform such a study. The results will likely show no difference between groups, without having the power to draw a real conclusion. The external validity will also be reduced given the single-center design. A retrospective analysis could be done to justify the conduct of a large multicenter blinded study with adequate power to conclude on the superiority of NIV + high-flow oxygen over high-flow oxygen alone. Finally, the considered sample size seems not correct to detect a significant difference in my point of view. The authors should absolutely consider a new analysis and probably increase the number of included patients.

Author Response

These are all limitations we recognise. Our sample size was an estimate only as the CONSORT guidelines for pilot studies (Eldridge et al. *BMJ* 2016) recommend not including a sample size for feasibility studies. Indeed, we have already undertaken a retrospective analysis (submitted and under review for publication). We now are interested in feasibility (safety, fidelity of intervention, recruitment, consent and retention as well as costs involved) and these results will then inform a planned, powered multicentre RCT in the future that will require full funding. We do not expect to be powered for the primary outcome as stated in our manuscript (highlighted on page 2, lines 69-70).

Strengths and limitations of this study

- *This is a pilot, single centre study unlikely to be powered to determine treatment effectiveness.*
- *Results of this pilot study will assist the design and conduct of future definitive multicentre trials.*

We recognise this study will unlikely show significant difference in PPC between groups. Our observational study mentioned in our response to the previous comments prior to this pilot study demonstrated a significant reduction in PPC incidence in the post-cohort group (7% vs 18%) compared to the pre-cohort group once baseline differences were accounted for. Although as mentioned there were many confounding factors which may have biased our results. Our 75% power calculation in this pilot study (highlighted on page 15, lines 490 - 492) was reported to demonstrate to the audience what difference between groups would be required for this pilot study to have any power.

Sample Size:

This sample will only be adequately powered (80%) if there is a large 75% relative risk reduction in PPC with the application of NIV (18% down to 4%).

VERSION 2 – REVIEW

REVIEWER	Suresh Victor King's College London, UK
REVIEW RETURNED	28-Aug-2018

GENERAL COMMENTS	Unfortunately, concerns have not been addressed. The trial suffers from significant methodology issues including sample size and is unlikely to generate any useful data.

REVIEWER	Jaber, S. Montpellier University Hospital - France
REVIEW RETURNED	02-Sep-2018

GENERAL COMMENTS	The authors satisfactorilly respond to the queries.