PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Clonidine as analgesia during retinopathy of prematurity screening in preterm infants (cloROP): Protocol for a randomized controlled trial
AUTHORS	Carlsen Misic, Martina; Eriksson, Mats; Normann, Erik; Pettersson, Miriam; Blomqvist, Y; Olsson, Emma

VERSION 1 – REVIEW

REVIEWER	Carlo Bellieni university of siena, pediatrics
REVIEW RETURNED	12-May-2022

GENERAL COMMENTS	The main concerns about this project are the following: There is no current evidence for an effective analgesic effect of clonidine in newborns (https://pubmed.ncbi.nlm.nih.gov/32270873/), so, why did authors choose this drug for their trial? It is not clear how pain assessment using the PIPP scale would be done: to be careful, the procedure should be video recorded (and this is written), and later revised and point-by-point scored (this is not declared in the text); moreover, one researcher should copy from the monitor the initial values of HR and O2 saturation, to later calculate the % difference as the PIPP scale requires (this is not declared in the text). The control group should receive the best available treatment, for the study be ethically acceptable: are authors sure that pacifier plus facilitated tucking are the most analgesic treatment now available? The muticentric study will be performed by several ophtalmologists: isnt this an important bias for the level of pain/discomfort they can provoke?

REVIEWER	Aslı Vural Istanbul Bakirkoy Dr Sadi Konuk Training and Research Hospital
REVIEW RETURNED	24-May-2022

GENERAL COMMENTS	Dear Authors, In the study named “Clonidine as analgesia during retinopathy of prematurity screening in preterm infants (cloROP), a protocol for a randomized controlled trial” you are investigating an issue that really needs to be resolved. The study is well designed. I think this will yield useful results There are some spelling and gramer mistakes that needs to be reviewed. Eg: - Page 4 line 10 “att” should be at. - “without finding a/an effective” “the aim of this study will be to investigate”

REVIEWER	Rebeccah Slater Oxford University
REVIEW RETURNED	30-Jun-2022

GENERAL COMMENTS	This clinical trial is timely and important. It is asking a critical question as to whether clonidine is an effective analgesic during retinopathy of prematurity (ROP) screening in preterm infants. I fully agree that practical and effective pain-relieving methods for ROP screening need to be identified and that studies investigating the efficacy of pharmacological interventions need to be conducted. I have a few suggested amendments in relation to the trial protocol and study design. In particular, more details are needed throughout the protocol in order to undertake a proper evaluation fo the study design. Ethics and Regulatory Approval It appears that this trial may be a Clinical Trial of an Investigational Medicinal Product (CTIMP). If this is the case, does it require approval from the Swedish Medical Products Agency? This should be clarified in the Protocol. Power, Sample Size Calcualtions and randomisation The sample size (n=18, inflated to 25 to allow for drop-outs – at each site) seems extremely small and sample size calculations are not presented in enough detail to evaluate whether the sample size is appropriate. For the primary outcome measure what is the expected mean PIPP-R value (and standard deviation) in the placebo group for each study site? Why is a 2-point reduction in the PIPP-R score appropriate – this needs a reference or explanation. Can a detailed power calculation be provided? The protocol states that 25 babies will be studied at each centre, but it is not clear how the data will be combined in the analysis. It seems that the intention data may be to present the date for each site separately. Are the two datasets recorded at each site being treated completely independently or together? This needs further clarification. If they are being treated independently do you therefore have co-primary outcome measures requiring the statistical thresholds to be altered? What conclusions will you draw about the efficacy of the intervention if the results differ across the sites? How will balance between trial arms for study characteristics such as (age at birth, time of study, gender be considered)? Will a minimisation algorithm be used to balance the trial arms? Has the impact of RetCam or direct ophthalmoscopy on infants' pain been considered in the design and the power calculations? Is one of these techniques considered to be more painful than the other? Is a power of 0.8 suitable? A higher power for studies of this type is often recommended. Recording techniques Further details of the data acquisition and data analysis (especially in relation to the GSR recording) would be helpful. When exactly will the GSR recording be started and stopped? What specific outcome measures will be recorded and analysed? Do you have an appropriate sample size calculation for this measure?

	The heart rate and oxygen saturation values will be recorded by video-recording the vital signs monitors. This does not seem to be the most efficient method to record this data. Did you consider directly downloading the data from the monitors? Objectives and Outcome Measures It would be helpful to have a table that clearly defines the exact objectives and precise outcome measures that will be evaluated for each of the primary and secondary outcomes. It is difficult to measure the PIPP-R score during the eye exam as the view of the face is obscured by the speculum. I suggest that time timing of the PIPP-R scoring is changed such that the 'PIPP-R score is calculated in the 30 s period after the retinopathy of prematurity screening is completed (i.e. after removal of the speculum following examination of the second eye). In addition, it may be more appropriate to record the GSR during the same period as the PIPP-R score and therefore change the timing for this measure too. How are adverse events being reported within 72 hours after the examination going to be recorded? Should this be a specific list of adverse events that are considered to be related to the intervention rather than a set of more general observations? If this is not specified, then all adverse events that occur (including those that are considered to not be directly related to the procedure) will be reported and analysed, which could impact your results Is the examining physician's assessment of how easy the infant was to examine a recognised validated scale? If not, do you have any experience to confirm the suitability of using this outcome measure? Can this outcome measure be referenced? Drug and placebo Who will be preparing and making the clonidine and placebo syringes? More details are required about where this is manufactured and how it is being prepared for trial use. How is blinding going to be maintained? Further explanation in the protocol would be helpful 'Strengths and Limitation of this Study' section The section titled 'Strengths and limitation of this study' does not really highlight the strengths and limitations of there study but rather provides a description of the 'Study Highlights'. The strengths and limitation need to be more clearly articulated. It seems that a major limitation is that 'In unit A RetCam (Clarity Medical Systems, Pleasanton, CA, USA) is used for ROP screening while direct ophthalmoscopy is used almost exclusively at unit B.' This should be discussed. In addition, given that clonidine is a known sedative a discussion about whether your primary outcome measure can discriminate between pain and sedation is needed. There is a typo on Page 4 – line 10

VERSION 1 – AUTHOR RESPONSE

Reviewer 1 comments	Reply
There is no current evidence for an effective analgesic effect of clonidine in newborns (https://pubmed.ncbi.nlm.nih.gov/32270873/), so, why did authors choose this drug for their trial?	There are some studies of the painrelieving/sedating properties of Clonidine despite the empty Cochrane review you are referring to. A few examples are the references we have used (Läkemedelsverket 2014, Pichot et al 2012) as well as Hunseler et al 2014 - doi: 10.1097/PCC.0000000000000151 and Donato et al 2019 https://doi.org/10.1016/j.clp.2019.08.004. We also use Clonidine in the clinical setting in neonatal care for other painful conditions such as ventilator treatment. This together with the fact that there is still no consensus in the optimal treatment of pain during ROP-screening makes Clonidine a medicine worth investigating for the painful and stressful procedure of ROP screening.
It is not clear how pain assessment using the PIPP scale would be done: to be careful, the procedure should be video recorded (and this is written), and later revised and point-by-point scored (this is not declared in the text); moreover, one researcher should copy from the monitor the initial values of HR and O2 saturation, to later calculate the % difference as the PIPP scale requires (this is not declared in the text).	This has been clarified.
The control group should receive the best available treatment, for the study be ethically acceptable: are authors sure that pacifier plus facilitated tucking are the most analgesic treatment now available?	As of now there are no consensus of the “most analgesic treatment” during ROP eye examinations. The units involved in the study both uses the non-pharmacological interventions described as part of the clinical routine.
The muticentric study will be performed by several ophtalmologists: isnt this an important bias for the level of pain/discomfort they can provoke?	You are correct in that the examining ophthalmologist can affect the level of pain/discomfort, however the two units will be assessed as two different “cohorts” and the examinations compared within that unit/cohort as described more in detail below.

Reviewer 2 comments	Reply
In the study named “Clonidine as analgesia during retinopathy of prematurity screening in preterm infants (cloROP), a protocol for a randomized controlled trial” you are investigating an issue that really needs to be resolved. The study is well designed. I think this will yield useful results	Thank you.
There are some spelling and gramer mistakes that needs to be reviewed. Eg: Page 4 line 10 “att” should be at.	This has been corrected.
“without finding a/an effective”	This has been corrected.
the aim of this study will be to investigate	Not sure we understand what you mean with this?

Reviewer 3 comments	Reply
This clinical trial is timely and important. It is asking a critical question as to whether clonidine is an effective analgesic during retinopathy of prematurity (ROP) screening in preterm infants. I fully agree that practical and effective pain-relieving methods for ROP screening need to be identified and that studies investigating the efficacy of pharmacological interventions need to be conducted. I have a few suggested amendments in relation to the trial protocol and study design. In particular, more details are needed throughout the protocol in order to undertake a proper evaluation fo the study design-	Thank you, we have done our best to address your concerns below.
Ethics and Regulatory Approval It appears that this trial may be a Clinical Trial of an Investigational Medicinal Product (CTIMP). If this is the case, does it require approval from the Swedish Medical Products Agency? This should be clarified in the Protocol.	You are absolutely correct and we had only described this with the EudraCT number. We have now clarified this under “Ethics and dissemination”.
Power, Sample Size Calcualtions	Thank you for this remark. We have rewritten this section to

and randomisation The sample size (n=18, inflated to 25 to allow for drop-outs – at each site) seems extremely small and sample size calculations are not presented in enough detail to evaluate whether the sample size is appropriate.	give more details.
For the primary outcome measure what is the expected mean PIPP-R value (and standard deviation) in the placebo group for each study site? Why is a 2-point reduction in the PIPP-R score appropriate – this needs a reference or explanation. Can a detailed power calculation be provided?	We have added information about PIPP-R values in previous studies and argue for using a reduction of 2 points as clinically significant.
The protocol states that 25 babies will be studied at each centre, but it is not clear how the data will be combined in the analysis. It seems that the intention data may be to present the date for each site separately. Are the two datasets recorded at each site being treated completely independently or together? This needs further clarification. If they are being treated independently do you therefore have co-primary outcome measures requiring the statistical thresholds to be altered? What conclusions will you draw about the efficacy of the intervention if the results differ across the sites?	The data from the respective unit will not be combined but analysed independently for each unit. Since the examinations are done with different methods on the two units, the study will investigate the pain-relieving effect of clonidine for each of the methods separately. This has been explained in the manuscript.
How will balance between trial arms for study characteristics such as (age at birth, time of study, gender be considered)? Will a minimisation algorithm be used to balance the trial arms?	We will report any statistically significant differences in demographic and possibly confounding variables. Since the data from the two methods (direct ophthalmoscopy, RetCam) are analysed and reported separately, this will not be a confounder. We have re-written the analysis section in the manuscript to be clear about this.
Has the impact of RetCam or direct ophthalmoscopy on infants' pain been considered in the design and the power calculations? Is one of these techniques considered to be more painful than the other?	There are inconsistent previous research findings about pain intensity from the two methods (Mehta M, Adams GG, Bunce C, Xing W, Hill M. Pilot study of the systemic effects of three different screening methods used for retinopathy of prematurity. Early Hum Dev. 2005;81(4):355–60, (Mukherjee AN, Watts P, Al-Madfai H, Manoj B, Roberts D. Impact of retinopathy of prematurity screening examination on cardiorespiratory indices: a comparison of indirect

	ophthalmoscopy and retcam imaging. Ophthalmology. 2006;113(9):1547–52). This is the reason for the separate data sets being analysed independently. The power calculation is valid for each method separately since we are aiming for the same difference in the primary outcome variable. We have tried to be clear about this in the manuscript.
Is a power of 0.8 suitable? A higher power for studies of this type is often recommended.	There is always a delicate balance between increasing the power and thereby expose study participants for possible risks and discomfort. We chose the 80% level which still is very common in clinical studies and gives a reasonable low risk of getting a type I-error.
Recording techniques Further details of the data acquisition and data analysis (especially in relation to the GSR recording) would be helpful. When exactly will the GSR recording be started and stopped? What specific outcome measures will be recorded and analysed? Do you have an appropriate sample size calculation for this measure?	The GSR electrodes will be attached at the same time as the saturation/heart rate-probe (used for the PIPP-R assessment) which has been clarified in the manuscript. The same time period as the PIPP-R (the first 30 seconds of the ROP examination) will be used for the analysis of the GSR-values (stated under “secondary outcomes” together with the outcome measures). We have made the sample size calculation from our primary outcome, the PIPP-R.
The heart rate and oxygen saturation values will be recorded by video-recording the vital signs monitors. This does not seem to be the most efficient method to record this data. Did you consider directly downloading the data from the monitors?	After several studies with PIPP-R as the outcome in our research group we are accustomed with this way of recording the data. Since we record with a mixer we can see both vital signs and the infant´s face in the same picture making the assessment easier than with two different medias.
Objectives and Outcome Measures It would be helpful to have a table that clearly defines the exact objectives and precise outcome measures that will be evaluated for each of the primary and secondary outcomes.	Thank you, we have added a table with this information according to your suggestion.
It is difficult to measure the PIPP-R score during the eye exam as the view of the face is obscured by the speculum. I suggest that time timing of the PIPP-R scoring is changed such that the ‘PIPP-R score is calculated in the 30 s period after the retinopathy of prematurity screening is completed (i.e. after removal of the speculum following	We agree that it is more challenging to assess PIPP-R during ROP examinations, however we have used videorecordings during ROP examinations in a previous study and was able to assess the PIPP-R values. We are also more interested in the pain relieving effect of Clonidine *during* the examination compared to *after* the examination.

examination of the second eye). In addition, it may be more appropriate to record the GSR during the same period as the PIPP-R score and therefore change the timing for this measure too.	
How are adverse events being reported within 72 hours after the examination going to be recorded? Should this be a specific list of adverse events that are considered to be related to the intervention rather than a set of more general observations? If this is not specified, then all adverse events that occur (including those that are considered to not be directly related to the procedure) will be reported and analysed, which could impact your results.	Potential adverse events are recorded in the patients´ CRF according to instructions from the Swedish Medical Products Agency, this has been clarified in the manuscript. In the patient´s CRF there are also examples of what is considered AE´s (possible side effects of Clonidine such as bradycardia, hypotension and fatigue).
Is the examining physician's assessment of how easy the infant was to examine a recognised validated scale? If not, do you have any experience to confirm the suitability of using this outcome measure? Can this outcome measure be referenced?	No, this is not a recognized validated scale and there are unfortunately no references. It is a way to ask the examining ophthalmologists how easy the infant was to examine in a more standardized way.
Drug and placebo Who will be preparing and making the clonidine and placebo syringes? More details are required about where this is manufactured and how it is being prepared for trial use. How is blinding going to be maintained? Further explanation in the protocol would be helpful	As stated under "procedure" an unblinded nurse or pharmacist will prepare the study solution. He/she is the only person not blinded for the allocation. Since the syringe is only marked with the infants´ study ID blinding will be maintained (the solutions look the same). We have clarified where the study solutions will be accessed (from the units´ ordinary medical supply).
'Strengths and Limitation of this Study' section The section titled 'Strengths and limitation of this study' does not really highlight the strengths and limitations of the study but rather provides a description of the 'Study Highlights'. The strengths and limitation need to be more clearly articulated.	Thank you, this section has been revised.
It seems that a major limitation is	As stated above we are not comparing retcam vs indirect

that 'In unit A RetCam (Clarity Medical Systems, Pleasanton, CA, USA) is used for ROP screening while direct ophthalmoscopy is used almost exclusively at unit B.' This should be discussed.	ophthalmoscopy but we will be analyzing them independently.
In addition, given that clonidine is a known sedative a discussion about whether your primary outcome measure can discriminate between pain and sedation is needed.	Pain and stress are closely linked and newborn infants will benefit from a decrease in any of them. PIPP-R is a validated tool for pain assessment in preterm infants.
There is a typo on Page 4 – line 10	This has been corrected.

VERSION 2 – REVIEW

REVIEWER	Carlo Bellieni university of siena, pediatrics
REVIEW RETURNED	12-Aug-2022

GENERAL COMMENTS	Risultati della traduzione I congratulate the authors on their responses. One point remains unclear. The authors replied that there are studies on the efficacy of cloidine in the newborn as an analgesic, but they cite in their response some studies that do not report this finding (Pichot, who does not speak of newborns and speaks of clonidine as a sedative and not as an analgesic) or which merely report negative conclusions on the analgesic efficacy made by the Cochraine review I quoted (Thorkesson). Another work is not in English (if the authors have important data there they should provide a translation). A last work that they cite (Hunseler) does not demonstrate the analgesic effect on the newborn, but the saving of fentanyl and diazepam, which could be due to an adjuvant effect. Therefore, I urge the authors to provide substantial evidence to their hypothesis that clonidine is effective as an analgesic in the newborn.

REVIEWER	Rebeccah Slater Oxford University
REVIEW RETURNED	29-Aug-2022

GENERAL COMMENTS	The majority of my concerns have been addressed in the revised protocol. I have made a few further comments in your attached 'reviewer response' document.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1	Response
I congratulate the authors on their responses. One point remains unclear. The authors replied that there are studies on the efficacy of cloidine in the newborn as an analgesic, but they cite in their response some studies that do not report this finding (Pichot, who does	Thank you for your positive response. Regarding Clonidine as we replied previously "*there are some studies of the painrelieving/sedating properties of Clonidine*". As you probably know a lot of the pharmacological treatments that are used in neonatal care today do not have the same extensive research foundation as treatments in older children/adult

not speak of newborns and speaks of clonidine as a sedative and not as an analgesic) or which merely report negative conclusions on the analgesic efficacy made by the Cochraine review I quoted (Thorkesson). Another work is not in English (if the authors have important data there they should provide a translation). A last work that they cite (Hunseler) does not demonstrate the analgesic effect on the newborn, but the saving of fentanyl and diazepam, which could be due to an adjuvant effect. Therefore, I urge the authors to provide substantial evidence to their hypothesis that clonidine is effective as an analgesic in the newborn.	population and we clinically use these drugs any way. In our opinion this makes it even more important to perform pharmacological research within this patient population. On page 225 in Pichot it says “the core effect of alpha-2 agonists is to modify the perception of pain”. In Hunseler´s article it´s true that Clonidine seemed to have an opioid sparing effect but the authors also conclude in page 519 that the newborns in the Clonidine groups had a deeper degree of sedation and *better analgesia*” We have tried in our manuscript to only say what is known e.g. Clonidine is used in neonatal care and that is safe to use due to minimal risk of respiratory depression and that it can also reduce the need for other sedation and pain-relieving medications (and that is what we are referring to with references 24-26). This study has been started. We are including patients and we think the study will bring important results and a possibility for us to know more about the pain-relieving effects of Clonidine.

Reviewer 3	Response
The majority of my concerns have been addressed in the revised protocol. I have made a few further comments in your attached 'reviewer response' document.	Thank you. For clarity we have copied your comments and answered them below.
I still consider that the study limitations are not fully described here. For example, the two different sites having two different approaches is not mentioned or the limitations recording the PIPP scoring during the procedure.	We agree that we cannot compare the two sites and have added this to the limitations section. Recording facial expressions during eye examinations can be difficult but still gives us the opportunity to go back and review over again. We have added a sentence about this.
There is a typo ‘findings’	Thank you, this has been revised.
I still consider that it would be helpful if a full statistical review is undertaken by the journal in relation to the primary and	We leave this to the editor to decide.

secondary outcomes.	
A reference would be helpful in the main text (about PIPP-recording)	We have added this.
This should be explained in the protocol (about VAS-scoring of examination difficulties)	We have added this information.