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)	Risk assessment of patient factors and medications for drug-related problems from a prospective longitudinal study of newborns admitted to a neonatal intensive care unit
AUTHORS	Leopoldino, Ramon; Santos, Marco; Costa, Tatiana; Martins, Rand; Oliveira, António

VERSION 1 – REVIEW

REVIEWER	Hideyuki Sawada Department of Neurology, Utano National Hospital, Kyoto, Japan
REVIEW RETURNED	02-Jul-2018

GENERAL COMMENTS	The authors investigated drug-related problems (DRPs) in neonatal ICU. They analyzed data of neonates prospectively. The identified patient-side risks of DRPs, and in addition, high risk drugs. The topic is important, and the data were analyzed carefully. I have several comments as follows. #1. The authors should provide the data of polypharmacy of DRPs, because risk of drug-induced adverse events is related with the number of drugs used. #2. Abstract: Please spell out “PRM” (line 26). Please explain “drug-related problems,” shortly. #3. Result: Please provide the summary of DRPs, classified by Pharmaceutical Care Network Europe System. #4. The distribution of the number of DRPs is not Gaussian. The authors should provide patients data according to the number of DRPs., i.e. group without DRP, with 1 DRP, with 2 DRP, and with 3 or more DRPs.

REVIEWER	Akshaya Bhagavathula United Arab Emirates University- College of Medicine and Health Sciences, Al Ain, UAE
REVIEW RETURNED	17-Jul-2018

GENERAL COMMENTS	The authors manuscript identified DRPs in neonates from 2014-2016. With the increase in the number of drugs uses among neonates this research has pivot role in understanding the prevalence and factors contributing to DRPs. However, the manuscript lacks consistency and needed some revision. Abstract: The first sentence: "It is believed that ...." looks vogue need to focus only on study objective. second: the aim is not clear: first identifying the DRPs, factors associated with the occurrence of DRPs in Neonates attending NICU. Study design: The authors studied from January 2014 and the ethical clearance was obtained in 2014. Do they initiated the study

	prior to the Ethical approval?. Furthermore, why the authors collected up to November 2016. Is it predefined? If so, how they defined it in their protocol. PRM? Please look into all the abbreviations. The outcome variables do not match with their study objective: Do they have already studied the prevalence and incidence of DRPs in neonates prior to study initiation? I am not clear with the outcome variables. Results: 2-278 days of admission for neonates? Strange wording. In think, DRPs are identified, not exposed. What is the COR value, how they adjusted? Conclusion: Irrelevant to their study objective. Introduction: The introduction is not relevant to their study objective and looks like a student notes. Unlicensed medicines, immaturity, and rapid growth, interindividual variability does not have any direct relationship with the DRPs in neonates. Please differentiate what authors are trying to say there was not clear. The third paragraph should be the last paragraph before aim and objectives. What is the incidence of DRPs in Brazil neonates?. The aim presented in the abstract is different from the main manuscript. This is an observational study, cohort includes time: how the authors confined the time and how they evaluated each prescription and how many were excluded? The PCNE scale V6.2 was developed in 2010 and accessed in April 2018? when they used it in the study (see the reference 4). What makes difference from Neofax from Micromedex? and uptodate was not clear? Do these pharmacists recruit for the study purpose or its a part of their routine job? Consensus means? what type of consensus were identified and how they resolved? How was the sample size defined? Why the authors did not show the COR and AOR to make it more clear. Further, how adjustments were done and what variables were adjusted? Previous analysis? Can you site them? Result: I am not clear with the results section and cant able to comment on it. Discussion: Please discuss your findings and highlight the important findings. There is a lot of repetition and not relevant to their study. The conclusion was too general and please focus to provide important intervention for future research. In this considerations, I feel the manuscript needed major revision and cannot be suitable for publications in this present form.

REVIEWER	Dr. Richard Mutemwa University of Zambia School of Public Health Zambia.
REVIEW RETURNED	11-Oct-2018

GENERAL COMMENTS	1. First, on your sample size calculation. It would help to also indicate the estimates you used for: relative precision, confidence

	level, assumed prevalence (or incidence) of DRPs among the NICU population in the selected health facility or target facilities. You have indicated only the expected odds ratio. Alternatively, please provide sufficient information on estimates and procedure you used for your sample size estimation. 2. Discussion of methods is a bit rushed. On what basis were the patient characteristic factors selected? They also needed some kind of systematic introduction. There is mention of characteristics on which data were collected, but that's it – no rationality as to why those were picked? Also authors needed to systematically introduce the medicines or drugs and rationalize their selection – We only come to know which drugs were evaluated much later in the RESULTS section, which is not good enough. There is a problem of assuming a very narrow audience for this paper and in itself that becomes problematic, because such justification information may still be valuable even to professionals familiar with the subject of the article. Each outcome investigated and relationships evaluated need to be stated explicitly…NOT just results reported in the RESULTS section. For instance, results presented in Table 4 are from analysis not described nor hinted at in the Methods Section. How did you move from fitting separate models for each drug as an explanatory factor adjusted for patient characteristics TO analysis for 'CAUSES OF DRPs INVOLVING THE NINE MEDICINES'??? What models were fitted what structure did they take? 3. ALSO, in your presentation of results, PLEASE ensure that you present individual-level and group-level statistics/information separately in order to minimise confusion and risk of being misunderstood. 4. Again, it may be obvious to probably many people in the field but please explicitly state how the medicines examined are prescribed in clinical settings: for instance, can they be prescribed only independently one at a time on each neonate, or they can be prescribed concurrently. If more than one drug may be prescribed at any point, please be explicit on why that was not taken into account in your statistical modelling. In fact, in your RESULTS SECTION, you report that on average each patient received 8 drugs (as opposed to dose) over the treatment period. How come this multiple drug scenario was not reflected in the statistical modelling? Why then were the drugs modelled separately – as you explain in your statistical analysis section???? What about the effect of drug cross-over and interaction over the NICU period and beyond? All these need to be explained in your METHODS Section, detailing how you treated the data. It is rather late to be sharing this detail much later in your RESULTS Section. 5. You report that MORE than 50% of the patients suffered MORE than ONE DRP. How then was this factor not accounted for in the statistical modelling? For that reason, it is important for the authors to explain, in practical terms, what suffering MORE THAN ONE DRP exactly means…is there concurrency or only happened consecutively? Any of these presentations needs explicit accounting for in the statistical modelling…. What about potential association and/or causality between any two DRPs within a patient?? In fact, I feel that this group (with more than ONE DRP) needed separate modelling as a sub-population, for that very

	reason of suffering more than ONE DRP. Isn't there some clinical significance for such a complex event of more than one DRP? I also feel that in some of your modelling the infants that died should have been excluded from analysis due to the fact that death is an extreme and unique outcome different from all the other DRPs experienced by neonates that eventually survive – and risk factors for fatality ought to be hypothesized uniquely and separately. If that was not the case then this element in analysis needs to come out explicitly and not glossed over. 6. Moreover, to minimize confusion, please state the actual algebraical expression of the key models that were fitted – even in their generalized forms. It just makes things much easier to understand implementation of the fitted models and the structure of the calculus behind the tabulated results. Thus, as you describe the statistical steps in your METHODS Section, please explicitly describe the factors loaded in each key model and how each factor was calibrated (continuous or categorical/binary; for SCORES, it would be helpful to provide even a brief description of how the Score is/was constructed and the nature of its distribution in this study, presenting key parameters of that distribution. ALSO, since you were fitting predictive models, it is essential that information on model diagnostics is also presented for the reader to evaluate how reliable each fitted model is and hence the weight carried by the confirmed risk factors. 7. It would have been even more informative if the two groups of factors (characteristics and medicines) were evaluated against each other using statistical methods such as counterfactual analysis (PLEASE SEE: "King G, Tomz M, Wittenberg J. 2000. Making the most of statistical analyses: improving interpretation and presentation. American Journal of Political Science 44: 341–55."). That allows a practitioner to weigh their concerns when the two groups of risk factors co-present in a neonate – which, from your findings, appears to be fairly common clinical experience. For instance, what if a neonate of particular gestation period and APGAR Score is put on a particular DRP-risky drug course (versus a less risky drug course, etc.) Your statistical models seem too simplistically neat, as if these two groups of conditions or factors never interact (?). 8. Finally, I'm curious to know why, despite the patient-level follow-up nature of the observational data, a statistic for DRP-incidence is not computed and neither is DRP-prevalence given for the studied NICU population. In addition, I would have preferred use of survival analysis over logistic regression that was applied – in which case Hazard Rates and Hazard Ratios would have been better assessment measures for risk of DRP and risk-comparison over the observation period – with special consideration given to time (time-to-event) spent by the neonate in the health facility. I suspect that using survival analysis would have helped better conceptualization of the problem of DRPs in NICUs and anticipate the potential risk factors and their different potential action-dynamics over the longitudinal observation period. Survival analysis would also much more easily permit estimation of the risk of suffering a second or subsequent DRP under certain conditions,

	given occurrence of the first DRP at any point in time; Or indeed the risk of suffering from more than ONE DRP at any point in time while the neonate, of particular description, is admitted to NICU for more than 24-hours under particular treatment conditions. This is not at all to denigrate the regression methods chosen for the study by the authors, but to just point out relative analytical advantages. In fact, compared to Odds Ratios, Hazard Ratios also tend to suffer less from selection bias – especially given that this was a non-randomized observation study (if it was randomized please explicitly say so in the article, including how such randomization was performed).

REVIEWER	Alan Girling Reader in Medical Statistics, University of Birmingham, UK
REVIEW RETURNED	22-Oct-2018

GENERAL COMMENTS	1. The methods of analysis in this paper are not sufficiently clear to make a judgment of its value. The sample size is declared to be 600 neonates, each of which may experience several (or no) drug-related problems. So the declared outcome is not binary. Perhaps the analysed outcome for the first analysis (i.e. that leading to Table 2) should be declared as “at least one drug-related problem”? The analysis of association with drug-type needs much more explanation. Presumably any particular drug-related problem (i.e. Dose selection, Drug use, Prescription logistics from table 4) can be traced to a specific drug, so DRPs associated with some other drug are not relevant. Does this mean that the outcome in this analysis is restricted to problems that could reasonably be connected with the drug under consideration? And for this analysis, why is it relevant to include subjects that have not received the drug under consideration? Or perhaps they have not been included? Or perhaps the unit of analysis has been taken to be the “prescription event”. None of this is clear, at least to me, so I am unable to interpret the odds-ratios in Table 3. 2. P-values throughout should be given to 3 decimal places, particularly if special significance is attached to the 1% and 5% thresholds. On this point, a 5% p-value achieved after variable selection has been applied would not normally be interpreted as significant (Table 2, and Results.). 3. Abstract. The declared outcome measures do not seem to be correct – they seem to describe exposure variables rather than outcomes. If I have understood, the outcomes are drug-related problems. Given this, it is confusing to say that “neonates were 'exposed' to DRPs” Also what does “risk medicine for PRM” mean? 4. Abstract: The APGAR confidence interval does not contain the estimate. 5. Table 2 It might be more informative to use a finer scale for birthweight given that both confidence limits are the same. Is there a reason not to consider length of stay as an explanatory variable? 6. Table 4 The row-totals do not add up to the numbers in Table 3 in all cases. If there are other causes, an extra column should be included to make this clear.

VERSION 1 – AUTHOR RESPONSE

Hideyuki Sawada (Reviewer 1):

1. The authors should provide the data of polypharmacy of DRPs, because risk of drug-induced adverse events is related with the number of drugs used.

R: Our intention in the analysis of risk factors was to study only patient factors that could be predictors of DRPs. Therefore, we considered only patient variables that could be collected at admission in the NICU. Polypharmacy is a variable that is known only at discharge and thus is not predictive of DRPs. Therefore, in order to comply with the Reviewer's request, we divided the analysis of risk factors of DRPs into risk factors that may predict the occurrence of one or more DRP, and variables associated with DRPs (polypharmacy, number of clinical conditions and length of stay). The analysis of the association of those variables with DRPs was included in Table 2 and the odds-ratios adjusted by the identified risk factors at admission were reported in the Results section.

2. Abstract: Please spell out "PRM" (line 26). Please explain "drug-related problems," shortly. PRM was a typo. R: The following paragraph in the Abstract: "Primary outcome measures: Risk factors and risk medicines for PRM." was replaced by "Primary outcome measures: "DRP (conditions interfering in the patient's pharmacotherapy with potential undesired clinical outcomes)."

3. Result: Please provide the summary of DRPs, classified by Pharmaceutical Care Network Europe System.

R: We attached Supplementary File 1 with the PCNE classification, definition and examples.

4. The distribution of the number of DRPs is not Gaussian. The authors should provide patients data according to the number of DRPs., i.e. group without DRP, with 1 DRP, with 2 DRP, and with 3 or more DRPs.

R: The following paragraph "More than half of the newborns had one or more DRPs (59.8%, 359)." was changed to "There were 237 (39.5%) patients with no DRPs, 132 (22.0%) with one DRP, 71 (11.8%) with two DRPs, and 160 (26.8%) with three or more DRPs."

Akshaya Bhagavathula (Reviewer 2)

1. Abstract: The first sentence: "It is believed that ...." looks vague need to focus only on study objective. second: the aim is not clear: first identifying the DRPs, factors associated with the occurrence of DRPs in Neonates attending NICU.

R: The Objectives in the Abstract were changed from "It is believed that drug-related problems (DRP) are potentially serious in neonates, however, information for neonatal intensive care units (NICU) is scarce. This study aims to identify patient factors and medications associated with the occurrence of DRPs in NICUs." to "To identify patient factors and medications associated with the occurrence of Drug Related Problems (DRP) in neonates admitted to Neonatal Intensive Care Units."

2. Study design: The authors studied from January 2014 and the ethical clearance was obtained in 2014. Do they initiated the study prior to the Ethical approval?. Furthermore, why the authors collected up to November 2016. Is it predefined? If so, how they defined it in their protocol.

R: There was a mistyping of study dates, for which we apologize. Patient enrolment in the study was began in the same month of ethical approval by the IRB in April 2014 and continued until the sample size of 600 patients was obtained. The following text: "This study was an observational, prospective cohort study conducted from January 2014 to November 2016 in the 20-bed NICU of a teaching maternity hospital specialized in high-risk pregnancy." was changed to "This was an observational, prospective longitudinal study conducted from April 2014 to January 2017 in the 20-bed NICU of a teaching maternity hospital specialized in high-risk pregnancy." The sample size was justified in the Statistical analysis section.

3. PRM? Please look into all the abbreviations.
R: PRM was a typo. R: The following paragraph in the Abstract: “Primary outcome measures: Risk factors and risk medicines for PRM.” was changed to “Primary outcome measures: The occurrence of one or more DRP (conditions interfering in the patient’s pharmacotherapy with potential undesired clinical outcomes).”

4. The outcome variables do not match with their study objective: Do they have already studied the prevalence and incidence of DRPs in neonates prior to study initiation?
R: The following paragraph in the Abstract: “Primary outcome measures: Risk factors and risk medicines for PRM.” was changed to “Primary outcome measures: The occurrence of one or more DRP (conditions interfering in the patient’s pharmacotherapy with potential undesired clinical outcomes)”

5. I am not clear with the outcome variables.
R: The following paragraph in the Abstract: “Primary outcome measures: Risk factors and risk medicines for PRM.” was changed to “Primary outcome measures: The occurrence of one or more DRP (conditions interfering in the patient’s pharmacotherapy with potential undesired clinical outcomes)”

6. Results: 2-278 days of admission for neonates? Strange wording.
R: The following paragraph: “The study observed 600 neonates who spent a median of 14 NICU days (range 2 – 278 days).” was changed to “The study observed 600 neonates who had a median length of stay in the NICU of 13 days (range 2 to 278 days).”

7. In think, DRPs are identified, not exposed.
R: The following paragraph in the Abstract: “Most neonates (59.8%) were exposed to DRPs.” was changed to “DRPs were identified in most neonates (60.5%). In the Introduction, the following paragraph: “In paediatric wards, about half of the patients are exposed to DRPs” was changed to “In paediatric wards, DRPs occur in about half of the patients“.

8. What is the COR value, how they adjusted?
R: The main results in our study were obtained from a multivariate model based on logistic regression. Therefore, crude odds-ratios are not relevant once we present adjusted odds-ratios. In order to clarify what was meant by adjusted odds-ratios, the following paragraph in the Abstract: “The factors independently associated with DRPs were” was changed to “In a multivariate logistic regression model, the factors independently associated with DRPs were”.

9. Conclusion: Irrelevant to their study objective.
R: We believe that with the present corrections the conclusions are now directly relevant to the study objectives: the main risk factors found in the study are enumerated in the Conclusion.

10. The aim presented in the abstract is different from the main manuscript. This is an observational study, cohort includes time: how the authors confined the time and how they evaluated each prescription and how many were excluded?
R: Cohort study has different meaning for different people. For some people a cohorts is a patient group set up in the beginning of the study and followed over a given time period (a closed cohort), for other people a cohort is a kind of panel study where the same data is collected repeatedly over time from the same study population, for others is a time-to-event or a event-count study design with an open cohort. The latter corresponds to our study design. However, in order to avoid any ambiguity, we replaced the term cohort by longitudinal. Therefore, the paragraph in the Abstract: “Design: Prospective cohort study.” was changed to “Design: Prospective longitudinal study.” and in the

Methods section, the paragraph: “This study was an observational, prospective cohort study” was changed to “This study was an observational, prospective longitudinal study

11. The PCNE scale V6.2 was developed in 2010 and accessed in April 2018? when they used it in the study (see the reference 4).
R: It was the link to the document that was accessed in April 2018, at the time of writing of the manuscript. This is to inform that the link is alive at the time of writing.

12. What makes difference from Neofax from Micromedex? and uptodate was not clear?
R: The consultation of several databases is necessary because neither of them has complete drug information and the two databases may be seen as complementary, and sometimes the information is conflicting between databases. The following paragraph: “supported by the Neofax® textbook (Thomson Reuters, New York, USA), as well as the Micromedex® (Truven Health Analytics, Michigan, USA) and Uptodate® (WoltersKluwer, Alphenaanden Rijn, NL) databases.” was changed to “supported by the Neofax® textbook (Thomson Reuters, New York, USA), as well as the Micromedex® (Truven Health Analytics, Michigan, USA) and Uptodate® (WoltersKluwer, Alphenaanden Rijn, NL) databases that provided authoritative information on adverse drug reaction and drug-drug interactions.”

13. Do these pharmacists recruit for the study purpose or its a part of their routine job?
R: The following paragraph was added at the end of the section: “The pharmacists involved in this research were permanent members of the clinical pharmacy team allocated to the NICU of our institution. The identification of DRPs and their notification to the medical team is an important part of their routine work and all were experienced in the detection of DRPs.”

14. Consensus means? what type of consensus were identified and how they resolved?
R: The following paragraph: “A third pharmacist (TXC) was consulted when there was lack of consensus between the two evaluators.” was changed to “Whenever the two evaluators disagreed upon the classification of the cause of a PRM, a third pharmacist (TXC) was called in to break the tie.”

15. How was the sample size defined?
R: The definition of the sample size was presented in the statistical analysis section. The text was amended in order to complete the assumptions of the calculations. The following text: “The defined sample size of 600 subjects would afford 70% power to identify associations with an odds-ratio of 1.30 or more.” was changed to “The target sample size was set at 600 patients, a number that would afford 70% power to identify associations with an odds-ratio of 1.30 or greater for factors with a prevalence over 30%.”

16. Why the authors did not show the COR and AOR to make it more clear. Further, how adjustments were done and what variables were adjusted? Previous analysis? Can you site them?
R: The main results in our study were obtained from a multivariate model based on logistic regression. Therefore, the odds-ratios adjusted by the other variables in the multivariate model are the relevant statistics and thus the only reported in the main text.. Nevertheless, Table 2 presents the crude and the adjusted odds-ratios for each study variable.

17. Result: I am not clear with the results section and cant able to comment on it.
R: We hope that with the corrections and improvements suggested by the Reviewers, the results section is clearer.

18. Discussion: Please discuss your findings and highlight the important findings. There is a lot of repetition and not relevant to their study.

R: We structured the Discussion section to first present the main results of our study, then to present the review of the literature, then we discussed each find separately and proposed a theoretical explanation for each one, then we identified some limitation of the study and concluded with suggestions for further research in this topic. We really do not see repetition and redundancy in the discussion. Perhaps if the Reviewer is more specific we may attend to the request.

19. The conclusion was too general and please focus to provide important intervention for future research.
R: In the conclusion we specifically mentioned all the findings of this study. Suggestions for future research were presented just before the conclusions.

Richard Mutemwa (Reviewer 3)
1. First, on your sample size calculation. It would help to also indicate the estimates you used for: relative precision, confidence level, assumed prevalence (or incidence) of DRPs among the NICU population in the selected health facility or target facilities. You have indicated only the expected odds ratio. Alternatively, please provide sufficient information on estimates and procedure you used for your sample size estimation.
R: The definition of the sample size was presented in the statistical analysis section. The aim of this study was to identify risk factors for DRP in neonates and, accordingly, sample size was calculated in order to be able to detect factors with an odds-ratio of DRP above 1.3, and the value of 1.3 was chosen because it was considered that associations with an odds-ratio below that value are not clinically relevant. The sample size was also calculated limited to patient factors that were present in at least 30% of patients, because risk factors that occur uncommonly are not clinical helpful. The text was amended in order to complete the assumptions of the calculations. The following text: “The defined sample size of 600 subjects would afford 70% power to identify associations with an odds-ratio of 1.30 or more.” was changed to “The target sample size was set at 600 patients, a number that would afford 70% power to identify associations with an odds-ratio of 1.30 or greater for patient factors with a prevalence over 30% 17”. The following reference, describing the method of calculation of sample size, was added: 17. Chow S, Shao J, Wang H. Sample Size Calculations in Clinical Research. 2nd Ed. Boca Raton, FL: Chapman & Hall/CRC Biostatistics Series 2008.

2. Discussion of methods is a bit rushed. On what basis were the patient characteristic factors selected? They also needed some kind of systematic introduction. There is mention of characteristics on which data were collected, but that’s it – no rationality as to why those were picked? Also authors needed to systematically introduce the drugs and rationalize their selection – We only come to know which drugs were evaluated much later in the RESULTS section, which is not good enough. There is a problem of assuming a very narrow audience for this paper and in itself that becomes problematic, because such justification information may still be valuable even to professionals familiar with the subject of the article.
R: The following paragraph was added before the description of the patient variables: “In the absence of information in the literature on risk factors for PRM in neonates, the patient variables selected as candidates for assessment in a multivariate risk model were those that could be collected at NICU admission on every neonate and that reflect serious conditions that are usually associated with enhanced pharmacotherapy.” In addition, the following paragraph: “Throughout the hospitalization period, every neonate was evaluated for the number of clinical problems, number of prescribed drugs and occurrence of DRPs.” was changed to “In addition to patient variables, the study wanted to identify medications that were associated with increased risk of DRPs in neonates and, therefore, all the medications prescribed to each neonate during the NICU stay were recorded.”

3. Each outcome investigated and relationships evaluated need to be stated explicitly…NOT just results reported in the RESULTS section. For instance, results presented in Table 4 are from analysis

not described nor hinted at in the Methods Section. How did you move from fitting separate models for each drug as an explanatory factor adjusted for patient characteristics TO analysis for 'CAUSES OF DRPs INVOLVING THE NINE MEDICINES'??? What models were fitted what structure did they take?
R: Further detail was provided in the Statistical analysis section on the analysis plan and models used. The following text was added in the Statistical analysis section: "In the drugs identified in the previous analysis as high risk medications, the respective causes of DRPs are presented descriptively."

4. ALSO, in your presentation of results, PLEASE ensure that you present individual-level and group-level statistics/information separately in order to minimise confusion and risk of being misunderstood.
R: We believe that with the increased detail we added to the results section after the Reviewer's comments, that issue is minimized or resolved.

5. Again, it may be obvious to probably many people in the field but please explicitly state how the medicines examined are prescribed in clinical settings: for instance, can they be prescribed only independently one at a time on each neonate, or they can be prescribed concurrently. If more than one drug may be prescribed at any point, please be explicit on why that was not taken into account in your statistical modelling. In fact, in your RESULTS SECTION, you report that on average each patient received 8 drugs (as opposed to dose) over the treatment period. How come this multiple drug scenario was not reflected in the statistical modelling? Why then were the drugs modelled separately – as you explain in your statistical analysis section???? What about the effect of drug cross-over and interaction over the NICU period and beyond? All these need to be explained in your METHODS Section, detailing how you treated the data. It is rather late to be sharing this detail much later in your RESULTS Section.
R: The details of the statistical analysis were added in the Statistical Analysis section. The following text: "For the estimation of the risk of DRP associated with medications commonly used in a NICU a multiple logistic regression model adjusted by the risk factors identified in the previous analysis." was changed to: "It was hypothesized that some medications could be singled out because they are associated with a significantly higher risk of DRP, through a combination of complex dosing and/or administration, and of frequency of use. Those drugs would be high-risk medications requiring close monitoring from the clinical pharmacy team. In the NICU setting, very often several medications are prescribed concurrently, sometimes by simultaneously through the same intravenous line, and accounting for the interplay of all medications administered to a patient at a given in a statistical model would be unmanageable. Therefore, the estimation of the risk of DRP associated with each medication was based on a simpler model, where the risk of DRP observed with a given medication was compared to the average risk observed with all other medications prescribed to this patient population, controlling for the risk factors for DRPs identified in the previous analysis. For this analysis, a set of multiple logistic regressions with each drug as independent variable and adjusted by the risk factors identified in the previous analysis was evaluated and, for those medications where a statistically significant association with the occurrence of one or more DRPs was found at the 5% significance level, results are presented as adjusted odds-ratios of DRP with that medication to the average risk of all the other medications prescribed. In the drugs identified in the previous analysis as high risk medications, the respective causes of DRPs are presented descriptively. The interaction of each of those high risk medications with each risk factor previously identified was tested with multiple logistic regression, with significant interactions assumed at the $p<0.10$ level."

6. You report that MORE than 50% of the patients suffered MORE than ONE DRP. How then was this factor not accounted for in the statistical modelling? For that reason, it is important for the authors to explain, in practical terms, what suffering MORE THAN ONE DRP exactly means…is there concurrency or only happened consecutively? Any of these presentations needs explicit accounting

for in the statistical modelling…. What about potential association and/or causality between any two DRPs within a patient?? In fact, I feel that this group (with more than ONE DRP) needed separate modelling as a sub-population, for that very reason of suffering more than ONE DRP. Isn't there some clinical significance for such a complex event of more than one DRP?
R: Our aim was not to identify predictors of increased risk of DRPs, but to identify patient factors at NICU admission that may signal neonates at increased risk of suffering a DRP. Therefore, it was not relevant to our aim whether a patient had one or more DRPs, even so because there is yet no information on whether DRPs are associated with worse outcomes, neither on whether having more than one DRP, either concurrently or sequentially, is associated with worse outcomes. We did add an explanation on the meaning of having more than one DRP, by adding the following text after the reporting of the frequency distribution of DRPs: "Multiple DRPs in the same patient could occur concurrently or simultaneously."

7. I also feel that in some of you modelling the infants that died should have been excluded from analysis due to the fact that death is an extreme and unique outcome different from all the other DRPs experienced by neonates that eventually survive – and risk factors for fatality ought to be hypothesized uniquely and separately. If that was not the case then this element in analysis needs to come out explicitly and not glossed over.
R: We gave information on the number of deaths as part of the description of the study population, so that readers can evaluate the clinical seriousness of the population in our NICU, and we never had the intention to identify risk factors for NICU mortality. Excluding the fatalities from the analysis set would, in our opinion, bias the results and make the risk estimates no longer applicable in practice because at NICU admission one cannot tell which patients will survive.

8. Moreover, to minimize confusion, please state the actual algebraic expression of the key models that were fitted – even in their generalized forms. It just makes things much easier to understand implementation of the fitted models and the structure of the calculus behind the tabulated results.
R:The following text was added: after the first model: "The model was ln[p(DRP=1)/p(DRP=0)] = □0 + □ixi, where □0 is the regression constant, □i the partial regression coefficients and xi the independent variables."; after the second model: "The model was ln[p(DRP=1)/p(DRP=0)] = □0 + □1x1 + □ixi, where □0 is the regression constant,□□ are the partial regression coefficients and xi is a binary variable coding for the medication, □ and xi the co-variables."; after the third model: "The model was ln[p(DRP=1)/p(DRP=0)] = □0 + □1x1 + □ixi + □jx1xi, where □0 is the regression constant,□□ are the partial regression coefficients, x1 a binary variable coding for the medication, xi the co-variables and □jx1xi the interaction of the medication with each co-variable."

9. Thus, as you describe the statistical steps in your METHODS Section, please explicitly describe the factors loaded in each key model and how each factor was calibrated (continuous or categorical/binary; for SCORES, it would be helpful to provide even a brief description of how the Score is/was constructed and the nature of its distribution in this study, presenting key parameters of that distribution.
R: The following text was added in the Statistical Analysis section after "For the identification of risk factors of DRP…": ", … an initial selection of patient variables at NICU admission were tested for association with the occurrence of one or more DRPs with logistic regression. All variables were binary, except gestational age and birth weight that were continuous."

10. ALSO, since you were fitting predictive models, it is essential that information on model diagnostics is also presented for the reader to evaluate how reliable each fitted model is and hence the weight carried by the confirmed risk factors.

R: Actually, we were not trying to fit a predictive model. Rather, our aim was just to identify some important clinical variables collected at NICU admission that are independently associated with DRPs. We believe that model fit statistics like the Hosmer-Lemeshow test or deviance will be of no interest to the average reader, so we added the area under de ROC curve of the multivariate model, which is a measure of model accuracy that most readers are probably familiar with. The following text was added in the Methods section: “The c-statistic for the multivariate model with five variables was 0.72.”

11. It would have been even more informative if the two groups of factors (characteristics and medicines) were evaluated against each other using statistical methods such as counterfactual analysis (PLEASE SEE: “King G, Tomz M, Wittenberg J. 2000. Making the most of statistical analyses: improving interpretation and presentation. American Journal of Political Science 44: 341–55.”). That allows a practitioner to weigh their concerns when the two groups of risk factors co-present in a neonate – which, from your findings, appears to be fairly common clinical experience. For instance, what if a neonate of particular gestation period and APGAR Score is put on a particular DRP-risky drug course (versus a less risky drug course, etc.). Your statistical models seem too simplistically neat, as if these two groups of conditions or factors never interact (?).
R: We agree with the Reviewer in that patient factors and medications may interact and this may provide a better risk stratification. Our answer to that hypothesis was to test the interaction of each medication with the patient risk factors in multiple logistic regression models separately for each drug. In the impossibility of doing does tests for all the drugs prescribed, we performed the analysis only for the set of 9 medications that had been identified in previous analyses as high risk medications. Therefore, we added the following text in the Statistical analysis section: “The interaction of each of those high risk medications with each risk factor previously identified was tested with multiple logistic regression, with significant interactions assumed at the p<0.10 level. The model was ln[p(DRP=1)/p(DRP=0)] = □0 + □1x1 + □ixi + □jx1xi, where □0 is the regression constant,□□ are the partial regression coefficients, x1 a binary variable coding for the medication, xi the co-variables and □jx1xi the interaction of the medication with each co-variable.” In the Results section, we added the following text: “There were statistically significant interactions between renal disease and the prescription of amphotericin (p=0.084) and of meropenem (p=0.054), and between a 5-minute APGAR score < 7 and prescription of Vancomycin (p=0.038).”

12. Finally, I’m curious to know why, despite the patient-level follow-up nature of the observational data, a statistic for DRP-incidence is not computed and neither is DRP-prevalence given for the studied NICU population. In addition, I would have preferred use of survival analysis over logistic regression that was applied – in which case Hazard Rates and Hazard Ratios would have been better assessment measures for risk of DRP and risk-comparison over the observation period – with special consideration given to time (time-to-event) spent by the neonate in the health facility. I suspect that using survival analysis would have helped better conceptualization of the problem of DRPs in NICUs and anticipate the potential risk factors and their different potential action-dynamics over the longitudinal observation period. Survival analysis would also much more easily permit estimation of the risk of suffering a second or subsequent DRP under certain conditions, given occurrence of the first DRP at any point in time; Or indeed the risk of suffering from more than ONE DRP at any point in time while the neonate, of particular description, is admitted to NICU for more than 24-hours under particular treatment conditions. This is not at all to denigrate the regression methods chosen for the study by the authors, but to just point out relative analytical advantages. In fact, compared to Odds Ratios, Hazard Ratios also tend to suffer less from selection bias – especially given that this was a non-randomized observation study (if it was randomized please explicitly say so in the article, including how such randomization was performed).
R: Obtaining estimates of prevalence and incidence of DRP was not the objective of our study, which focus strictly on risk factors for DRPs. Time-to-event analysis with the Cox proportional hazards model has the limitation of the proportionality assumption, and since we had no censored data we believe that logistic regression is more robust in this problem. We did consider using event-count analytical

methods, namely Poisson regression and negative binomial regression, but the results were very poor with those models. Those were the reasons for our preference for logistic regression.

Alan Girling (Reviewer 4)

1. 1. The methods of analysis in this paper are not sufficiently clear to make a judgment of its value. The sample size is declared to be 600 neonates, each of which may experience several (or no) drug-related problems. So the declared outcome is not binary. Perhaps the analysed outcome for the first analysis (i.e. that leading to Table 2) should be declared as “at least one drug-related problem”? The analysis of association with drug-type needs much more explanation. Presumably any particular drug-related problem (i.e. Dose selection, Drug use, Prescription logistics from table 4) can be traced to a specific drug, so DRPs associated with some other drug are not relevant. Does this mean that the outcome in this analysis is restricted to problems that could reasonably be connected with the drug under consideration? And for this analysis, why is it relevant to include subjects that have not received the drug under consideration? Or perhaps they have not been included? Or perhaps the unit of analysis has been taken to be the “prescription event”. None of this is clear, at least to me, so I am unable to interpret the odds-ratios in Table 3.

R: The outcome actually is binary and we clarified that in the Statistical analysis section by adding the text: “For the identification of risk factors of DRP, an initial selection of patient variables at NICU admission were tested for association with the occurrence of one or more DRPs with logistic regression.” We gave many more details of the statistical analysis methods in that section: The text: “For the identification of risk factors of DRP, the set of patient variables whose association with DRPs was statistically significant at the 0.10 significance level in univariate logistic regression was analyzed by backward stepwise multiple logistic regression, and those variables significant at the 0.05 level were retained in the final model. Results are presented as adjusted odds-ratios (AOR) and 95% confidence intervals (CI). For the estimation of the risk of DRP associated with medications commonly used in a NICU, each prescribed drug was analyzed by a multiple logistic regression model adjusted by the risk factors identified in the previous analysis.” was changed to: “All variables were binary, except gestational age and birth weight that were continuous. The set of patient variables whose association with DRPs was statistically significant at the 0.10 significance level in univariate logistic regression was analyzed by stepwise backward multiple logistic regression, and those variables significant at the 0.05 level were retained in the final model. Variables collected only at discharge from the NICU were analysed in a separate model. Results of this analysis are presented as odds-ratios adjusted by the other variables in the model (AOR) and 95% confidence intervals (CI). The model was ln[p(DRP=1)/p(DRP=0)] = □0 + □ixi, where □0 is the regression constant, □i the partial regression coefficients and xi the independent variables.

It was hypothesized that some medications could be singled out because they are associated with a significantly higher risk of DRP, through a combination of complex dosing and/or administration, and of frequency of use. Those drugs would be high-risk medications requiring close monitoring from the clinical pharmacy team. In the NICU setting, very often several medications are prescribed concurrently, sometimes by simultaneously through the same intravenous line, and accounting for the interplay of all medications administered to a patient at a given in a statistical model would be unmanageable. Therefore, the estimation of the risk of DRP associated with each medication was based on a simpler model, where the risk of DRP observed with a given medication was compared to the average risk observed with all other medications prescribed to this patient population, controlling for the risk factors for DRPs identified in the previous analysis. For this analysis, a set of multiple logistic regressions with each drug as independent variable and adjusted by the risk factors identified in the previous analysis was evaluated and, for those medications where a statistically significant association with the occurrence of one or more DRPs was found at the 5% significance level, results are presented as adjusted odds-ratios of DRP with that medication to the average risk of all the other medications prescribed. The model was ln[p(DRP=1)/p(DRP=0)] = □0 + □1x1 + □ixi, where □0 is the regression constant,□□ are the partial regression coefficients, x1 a binary variable coding for the medication, and xi the co-variables. In the drugs identified in the

previous analysis as high risk medications, the respective causes of DRPs are presented descriptively. The interaction of each of those high risk medications with each risk factor previously identified was tested with multiple logistic regression, with significant interactions assumed at the $p<0.10$ level. The model was $\ln[p(DRP=1)/p(DRP=0)] = \beta 0 + \beta 1x1 + \beta ixi + \beta jx1xi$, where $\beta 0$ is the regression constant, $\beta\beta$ are the partial regression coefficients, x1 a binary variable coding for the medication, xi the co-variables and $\beta jx1xi$ the interaction of the medication with each co-variable."

2. P-values throughout should be given to 3 decimal places, particularly if special significance is attached to the 1% and 5% thresholds. On this point, a 5% p-value achieved after variable selection has been applied would not normally be interpreted as significant (Table 2, and Results.).
R: P-values were changed to 3 decimal places throughout. Regarding the last comment, we believe the Reviewer is referring to the issue of multiple comparisons. However, statistical tests in the variable selection were done only as a rule to eliminate irrelevant variables from the multiple logistic regression analysis, and were not done for hypothesis testing. Therefore, we feel it is not inadequate to interpret p-values < 0.05 as significant in the multiple logistic regression analysis.

3. Abstract. The declared outcome measures do not seem to be correct – they seem to describe exposure variables rather than outcomes. If I have understood, the outcomes are drug-related problems. Given this, it is confusing to say that "neonates were 'exposed' to DRPs" Also what does "risk medicine for PRM" mean?
R: We have corrected the outcome measures in the Abstract. The text: "Risk factors and risk medicines for PRM." was changed to "The occurrence of one or more DRP (conditions interfering in the patient's pharmacotherapy with potential undesired clinical outcomes)". The following paragraph in the Abstract: "Most neonates (59.8%) were exposed to DRPs." was changed to "DRPs were identified in most neonates (60.5%). In the Introduction, the following paragraph: "In pediatric wards, about half of the patients are exposed to DRPs" was changed to "In pediatric wards, DRPs occur in about half of the patients". In the Abstract, the text: "The risk medications for DRP" was changed to "The medications with greater risk for DRP"

4. Abstract: The APGAR confidence interval does not contain the estimate.
R: That was corrected.

5. Table 2 It might be more informative to use a finer scale for birthweight given that both confidence limits are the same. Is there a reason not to consider length of stay as an explanatory variable?
R: Birthweight was reanalysed converted to kg. Length of stay may be an explanatory variable, but cannot be used as a predictive variable because its value can be obtained only at discharge time, and our original aim was to focus on predictor variables. Nevertheless, we recognize that length of stay, as well as the number of unique medications prescribed and the complexity of care, represented by the number of clinical problems, should be evaluated as possible explanatory variables. However, since they cannot be considered predictive variables, their association with the occurrence of DRPs was analysed separately. Therefore, we added the results of the logistic regression analysis in Table 2 and the following text in the Statistical analysis section: "Variables collected only at discharge from the NICU were analysed in a separate model."

6. Table 4 The row-totals do not add up to the numbers in Table 3 in all cases. If there are other causes, an extra column should be included to make this clear.
R: We corrected table 4, modifying the domains of DRP causes according to the PCNE classification system v6.2, and we included a column with causes that had low frequency for most drugs. In addition, we added PCNE system as a supplementary file.

VERSION 2 – REVIEW

REVIEWER	Dr Richard Mutemwa University of Lusaka, UNILUS, Department of Public Health, Zambia London School of Hygiene & Tropical Medicine, LSHTM, London, UK.
REVIEW RETURNED	27-Feb-2019

GENERAL COMMENTS	It is a very important subject in maternal and neonatal clinical care. No major issue. I'm aware of the word-limit and other journal-related restrictions. But there is a missed opportunity, given the importance of the topic, for others in the practice to replicate and learn some of the statistical procedures deployed in this paper, especially in relation to the rounds of modelling of DRP on selected administered drugs. Yet, that does not take away the importance of the subject and integrity of the results.

REVIEWER	Alan Girling University of Birmingham
REVIEW RETURNED	23-Jan-2019

GENERAL COMMENTS	P4 line 21 “interfere in patients’ pharmacotherapy” does not convey the intended meaning. “arising from the patient’s pharmacotherapy that may actually or potentially interfere with health outcomes” is perhaps closer to the sense. P6 last line. Length of stay, number of drugs and number of clinical problems are not binary variables. These “discharge variables” were analysed separately (P7 line 10), but no details are presented (i.e. with or without adjustment for baseline risk-variables?). P6 2nd para. The risk associated with particular medications must surely be affected by length of stay, number of distinct medications and case-complexity. Since none of these is causally related to the medication prescribed it would be logical to adjust for them when assessing medication-specific risk, so as to compare different medications on a level playing field. Otherwise, a medication might be identified as high-risk simply because it is often used in complex cases, or in cases that remain in PICU for a long time. In their response, the authors say their focus is on predictive models – i.e. models that use information at admission only. In so doing they may miss the most important contributing factors. To give one example, their analysis does not eliminate the possibility that excess DRPs occur mostly in patients with longer lengths of stay, rather than being associated with any particular drugs. P6 2nd para. It seems that the “previous analysis” mentioned in lines 45-47 refers to the analysis of admission variables, and does not include any discharge variables. But this is not clear from the statistical analysis section. There have, in fact, been two “previous analyses”. P8 line 7. “significant interactions assumed at the $p < 0.10$ level”. This is not a level usually regarded as significant. P9 The style of reporting the adjusted results is inconsistent between the first two paragraphs. (i.e. p-values) P9 line 35 “could be obtained” is redundant. P9 last line. “DRP involving medications” What does this imply? Do not all DRPs involve medications? Table 3. This table is incomprehensible as it stands, and not explained in the text. The row labelled “Total” is clearly not the

	total of the other rows. The discrepancy may be due to other medicines. But the total under "n" for "Cases of DRP" is given as 4,917, yet this number is 1,115 on p8 line 50. Also the "%" column does not give n as a percentage of either of these numbers. Similar inconsistencies are evident in the "frequency of prescriptions" columns. Table 4. It would be better to use the same order of medicines as in Table 3.

VERSION 2 – AUTHOR RESPONSE

Response to Reviewers

The authors are grateful to the reviewers for their comments that helped us to improve our paper. Below are the responses to the reviewers.

Alan Girling (Reviewer 4):

1. P4 line 21 "interfere in patients' pharmacotherapy" does not convey the intended meaning. "arising from the patient's pharmacotherapy that may actually or potentially interfere with health outcomes" is perhaps closer to the sense.

R: The following sentence: "DRP are events or circumstances that, actually or potentially, interfere in the patient's pharmacotherapy and that may lead to undesired clinical outcomes." Was changed to: "DRP are events or circumstances arising from the patient's pharmacotherapy that may actually or potentially interfere with health outcomes.".

2. P6 last line. Length of stay, number of drugs and number of clinical problems are not binary variables. These "discharge variables" were analysed separately (P7 line 10), but no details are presented (i.e. with or without adjustment for baseline risk-variables?).

R: Changed the sentence to: "For the identification of risk factors of DRP, an initial selection of patient variables at NICU admission (sex, gestational age, birth weight, type of delivery, occurrence of PROM, 1-minute and 5-minute APGAR, a diagnosis of neurological, renal or cardiac disorder, and malformations) were tested for association with the occurrence of one or more DRP with logistic regression. All variables were binary, except gestational age and birth weight that were continuous.". The sentence "Variables collected only at discharge from the NICU were analysed in a separate model." was changed to "Variables collected only at discharge from the NICU (number of unique medications, length of stay and number of clinical problems) were analysed in a separate logistic model consisting of those three variables."

3. P6 2nd para. The risk associated with particular medications must surely be affected by length of stay, number of distinct medications and case-complexity. Since none of these is causally related to the medication prescribed it would be logical to adjust for them when assessing medication-specific risk, so as to compare different medications on a level playing field. Otherwise, a medication might be identified as high-risk simply because it is often used in complex cases, or in cases that remain in PICU for a long time. In their response, the authors say their focus is on predictive models – i.e. models that use information at admission only. In so doing they may miss the most important contributing factors. To give one example, their analysis does not eliminate the possibility that excess DRPs occur mostly in patients with longer lengths of stay, rather than being associated with any particular drugs.

R: Length of stay and number of distinct medications are so tightly associated with the occurrence of DRP that the inclusion of those variables in a multivariate model will obscure the influence of any

other variable, including the medication prescribed. This fact, as well as our option to focus on admission variables, led us to not to consider adjustment by those variables.

4. P6 2nd para. It seems that the “previous analysis” mentioned in lines 45-47 refers to the analysis of admission variables, and does not include any discharge variables. But this is not clear from the statistical analysis section. There have, in fact, been two “previous analyses”.

R: The following text: “ ... the risk of DRP observed with a given medication was compared to the average risk observed with all other medications prescribed to this patient population, controlling for the risk factors for DRP identified in the previous analysis. For this analysis, a set of multiple logistic regressions with each drug as independent variable and adjusted by the risk factors identified in the previous analysis was evaluated ... “ was changed to “ ... the risk of DRP observed with a given medication was compared to the average risk observed with all other medications prescribed to this patient population, controlling for the risk factors at NICU admission for DRP that were identified in the previous analysis. For this analysis, a set of multiple logistic regressions with each drug as independent variable and adjusted by the risk factors at NICU admission identified in the previous analysis was evaluated ... “

5. P8 line 7. “significant interactions assumed at the $p < 0.10$ level”. This is not a level usually regarded as significant.

R: Test for interactions between factors typically have low power and, because of this but also because testing for interactions was performed in an exploratory perspective, it is common practice to adopt a greater significance level, often 0.10 or 0.15. In an exploratory perspective the greatest concern is not on increasing the alpha error but rather on increasing the beta error, and this justifies the choice of a greater significance level for interaction tests.

6. P9 The style of reporting the adjusted results is inconsistent between the first two paragraphs. (i.e. p-values).

R: The paragraph was corrected to make all reporting styles consistent.

7. P9 line 35 “could be obtained” is redundant.

R: This term was removed from the phrase.

8. P9 last line. “DRP involving medications” What does this imply? Do not all DRPs involve medications?

The following term “DRP involving medications” was changed to “problems involving medications”

9. Table 3. This table is incomprehensible as it stands, and not explained in the text. The row labelled “Total” is clearly not the total of the other rows. The discrepancy may be due to other medicines. But the total under “n” for “Cases of DRP” is given as 4,917, yet this number is 1,115 on p8 line 50. Also the “%” column does not give n as a percentage of either of these numbers. Similar inconsistencies are evident in the “frequency of prescriptions” columns.

Table 3 was changed with the correction in the number of DRP and frequency of prescription, and the addition of the row referring to other medicines. The NICU had a total of 4,917 prescription drugs in which 1,252 had problems. The number of drugs involved in problems was greater than the number of DRP because a DRP could be caused by more than one medicine, for example, problems of drug interactions.

10. Table 4. It would be better to use the same order of medicines as in Table 3.

R: Table 4 was changed according to the above suggestion.