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)	Changing consumption of resources for respiratory support and short-term outcomes in four consecutive geographical cohorts of infants born extremely preterm over 25 years since the early 1990s.
AUTHORS	Cheong, Jeanie; Olsen, Joy; Huang, Li; Dalziel, Kim; Boland, Rosemarie; Burnett, Alice; Haikerwal, Anjali; Spittle, Alicia; Opie, Gillian; Stewart, Alice; Hickey, Leah; Anderson, Peter; Doyle, Lex

VERSION 1 – REVIEW

REVIEWER	Marina Cuttini Pediatric Hospital Bambino Gesù, Roma, Italy
REVIEW RETURNED	27-Feb-2020

GENERAL COMMENTS	The Authors analysed the data from the longitudinal cohort studies of extremely preterm infants carried out , during 4 periods from Jan 1 1991 to Mar 31 2017, in the State of Victoria, Australia, with the aim of assessing 1) changes in respiratory assistance, neonatal morbidities, and survival to discharge, and 2) use of respiratory resources in relation to changes in survival. The paper is well written and interesting. Comments: 1) in Table 1, it wold be interesting to see also number and % of deaths in the delivery room; 2) Statistical methods. Multiple birth within the same family was taken into account using robust estimates of standad errors. Did you consider the potential role of multiplicity as possible confounder, also at the light of increasing temporal trends due to advance maternal age and medically assisted procreation reported in many countries? 3) together with overall survival at discharge, I would also like to see survival free from major morbidities possibly associated with long-term outcomes (i.e. severe IVH/PVL, BPD, ROP, possibly also NEC). These data might broaden the perspective on the cost-effectiveness of the increasing use of respiratory support. 4) I would like to hear more in the discussion about the decreasing births in tertiary centers. Which could be the reasons? Could this factor explain the lack of decrease of IVH rates? Additional minor comments: 5) how were lethal anomalies defined? 6) in Tables, given that 95% CIs are provided, p values could (and possibly should, at the light of the multiple comparisons) be omitted.

REVIEWER	Sven Schulzke

	University Children's Hospital Basel UKBB, Switzerland
REVIEW RETURNED	06-Apr-2020

GENERAL COMMENTS	Cheong et al. present an epidemiological analysis of respiratory resources use and outcomes in extremely preterm infants in Victoria. The authors compare results from four cross-sectional observation periods across a 25-year period. Overall, it seems that increased survival of extremely preterm infants and changes in respiratory management along with higher oxygen saturation targets are associated with longer duration of hospital stay and respiratory support in this population. Aims, methods, results, and conclusions are well matched. Reasons for changes in respiratory morbidity, other co-morbidities and mortality are adequately discussed. The paper has considerable scientific and clinical relevance in the field. The manuscript is very well written and very clean. I don't see any major issues that need to be addressed to improve the paper. The authors might consider the following minor recommendations: - A sentence referring to the presence or absence of standard criteria for installing or ceasing CPAP or HFNC across the four eras may be worthwhile adding to the discussion. If those criteria have changed over time, could this affect the results? - Page 17, Table 3, line 59: Please add percentage of infants with BPD in the year 2005

REVIEWER	Jeremy Marks University of Chicago, USA
REVIEW RETURNED	10-Apr-2020

GENERAL COMMENTS	In this study, Cheong and co-authors report on changes over time in care delivery and neonatal outcomes of extremely preterm infants born in the state of Victoria, Australia, focusing primarily on the kind of respiratory support provided. They analyze data from all children born in the region during four epochs between 1991 and 2017, an approach similar to that recently used by the same group in their report on long term respiratory outcomes of a similar population. The authors find a marked rise in the duration of respiratory support in the modern epoch compared with all previous epochs, due to prolonged used of high flow nasal cannula (HFNC). However, it appears that this increased length of respiratory support is unaccompanied by improved outcomes, although the presentation of the data makes this a little difficult to parse out. This is an important finding, and continues this group's efforts to determine whether the current reliance on prolonged respiratory support that characterizes care of these infants in most of the developed world is in our patients' and health care system's best interests. It is a little difficult to sort through the tables in their current form, and at least one analysis would benefit from significantly more explanation. Specific concerns follow: Major concerns 1. Table 1 is very long and difficult to read. Separating the first 4 rows, which contain no statistics, into their own first table would be helpful.

	2. In (the revised) Table 1, variables which do not differ across the four epochs can be removed from the table, and can be summarized for the whole population in the text prior to presentation of the table. Thus, gestational age, birthweight, and male sex can all be treated this way. Non-significant p values can be replaced with p=NS. In this (revised) table, the N (which appears to be infants offered intensive care) should be clearly stated. The text preceding this table could be more descriptive. For example, not only was the rate of antenatal corticosteroid treatment lower in 1991-91 compared with the other 3 cohorts, but the rate of this treatment did not differ significantly form 1997 onwards. 3. Table 2 is difficult to read. For clarity, the row reporting non-invasive ventilator days seems duplicative, as it is simply the sum of CPAP and HFNC days, and could be removed without compromising understanding. The unadjusted analyses do not need to be presented in the table, as there are no real differences between the unadjusted and the adjusted analyses. A simple statement in the text could say that. Finally, it is unclear what the authors mean when they state that HFNC "did not lead to a decrease in the number of CPAP days". Do they mean "accompanied by"? Is it the practice in this state that infants are always placed on CPAP first and the moved to HFNC? 4. It is unclear what the reader is to take away from presentation of the respiratory utilization data in survivors versus all comers. If the differences in analyses are important, more explanation and discussion of these differences would be helpful. Otherwise, one of these two tables (all comers or babies who survive to discharge) could be removed from the manuscript without compromising its impact. 5. Table 3 is difficult to read. The table begins with survival to discharge, but the common morbidities of prematurity (NEC, surgery, IVH/PVL, treated ROP) are expressed as a percentage of all infants. It is not clear that these data on co-morbidity incidence across time are central to the thrust of the paper. If these data are relevant, they should be in their own table. It seems more central to the paper, and a much clearer narrative, for the authors to present in one table only the incidences of treatments and outcomes that are relevant to respiratory disease among survivors – administration of surfactant, postnatal corticosteroids, and the incidences of BPD. 6. The concept of added duration of respiratory support per additional survivor in the most recent epoch is not well explained, and is confusing. How is an additional survivor defined? What is being simulated? How are these simulations used? The statistics here are more complex than those usually used. and could be explained much more clearly. In like fashion, the reader is given no help in how Figure 1 is constructed or how to interpret it. What is involved in determining, from the current dataset, what is more/less costly and more/less effective? Many more details are required to make this analysis accessible to the reader. Similarly, the paragraph in the discussion about the cost-effectiveness of respiratory support, which appears to depend on this analysis, introduces concepts of efficiency, and the calculation of extra cot allocation for additional survivors. This line of discussion is not well informed by the results as they are currently presented. The

	conclusion that “The relative inefficiency in resource consumption could not be explained by increased consumption of resources by infants who died” is not supported by the text in the results section, or in the legend of Figure 1. If the authors wish to use this analysis, much more explanation in the methods. results and their interpretation should be devoted to it.. 7. The discussion of the factors leading to the prolonged usage of HFNC in the modern era is important, and deserves to be much higher up in the discussion. The authors suggest earlier in the manuscript that much of HFNC treatment may be “wasted”. Elaboration of this idea in this paragraph would tie the argument together. 8. To what do the authors attribute a significant decreased incidence of surfactant use in the most recent epoch compared with 2005, especially since the use of corticosteroids was unchanged? Minor concerns 1. In the introduction, the authors point to the abundance of randomized studies that have failed to show decreased incidence of BPD with non-invasive ventilation compared with invasive ventilation. This is an important point that is underappreciated. If they wish to cite the 2017 JAMA meta-analysis as a counterargument, it would be helpful to interpret that study’s findings in the context of the weaknesses inherent in any meta-analysis. The author’s desire to express cost effectiveness in resource units of days of respiratory support is well taken. The measure could therefore be more directly reported and discussed throughout the manuscript as “duration of respiratory support.” Thus, Table 2 would be more titled “Duration of respiratory support, oxygen therapy, and hospitalization…: 2. Did the infants in the 2016/17 epoch require additional days of respiratory support, or simply receive additional days? 3. Expressions of personal reaction to findings, such as “worrying” or “encouraging” do not really belong in a scientific report. These could be replaced with more descriptive terms.

REVIEWER	Dr Sanja Zivanovic University of Oxford, UK
REVIEW RETURNED	20-Apr-2020

GENERAL COMMENTS	This is a paper with important data that took a lot of effort to collect. However, putting it into certain paradigm and linking it to conclude that there is a causative relationship between certain ventilation methodology and adverse outcomes over 25 years is not appropriate. Furthermore, it is highly inappropriate to suggest how one type of methodology has influenced another (ie HFNC has not shortened the time of CPAP). That depends on the local policies of use and it’s users. Every institution can decide their own non-invasive ventilation support policy, choice of first line treatment and the speed of weaning it. I would suggest to present this data in a form of Epicure data, only for the cohort in question. It is valuable to express the changes through time chronologically for what they are, but not to speculate causation in any form. Please also explain more on the methodology. I am not sure this is a prospective study , unless the protocol for it was written in 1991?

VERSION 1 – AUTHOR RESPONSE

Reviewers' comments:

Reviewer #1: The Authors analysed the data from the longitudinal cohort studies of extremely preterm infants carried out , during 4 periods from Jan 1 1991 to Mar 31 2017, in the State of Victoria, Australia, with the aim of assessing 1) changes in respiratory assistance, neonatal morbidities, and survival to discharge, and 2) use of respiratory resources in relation to changes in survival. The paper is well written and interesting.

Comments:

1. *In Table 1, it would be interesting to see also number and % of deaths in the delivery room.*

   We have data on the number of infants who died on the day of birth, but not the location, and hence we are unable to add the specific data requested to Table 1.

2. *Statistical methods. Multiple birth within the same family was taken into account using robust estimates of standard errors. Did you consider the potential role of multiplicity as possible confounder, also at the light of increasing temporal trends due to advance maternal age and medically assisted procreation reported in many countries?*

   Although maternal age increased over time, a similar time trend was not observed in the rates of multiple births (Table 1). As suggested by the reviewer we repeated the analyses including multiple births as a confounder, but no conclusions were altered. As the issue of the confounding effects of multiple births is not central to the aims of the study, we have not changed the manuscript in response to this question.

3. *Together with overall survival at discharge, I would also like to see survival free from major morbidities possibly associated with long-term outcomes (i.e. severe IVH/PVL, BPD, ROP, possibly also NEC). These data might broaden the perspective on the cost-effectiveness of the increasing use of respiratory support.*

   We have added these data to Table 3 and also an extra sentence to the results (Page 15. Results: (c) In-hospital survival, treatments and morbidities. Paragraph 1, sentence 2): "A similar trend was observed with survival free of major morbidities, as defined by having one or more of IVH grade 3-4, cystic periventricular leukomalacia, necrotising enterocolitis, ROP requiring treatment, or BPD."

4. *I would like to hear more in the discussion about the decreasing births in tertiary centers. Which could be the reasons? Could this factor explain the lack of decrease of IVH rates?*

   There are several reasons why the rates of births outside tertiary centres have increased over eras; which include a changing philosophy in decision-making around resuscitation of extremely preterm infants born <25 weeks' gestation, and delay in/missed opportunities for in-utero transfer. We have added a sentence in the discussion (Page 21. Discussion. Paragraph 6, sentence 2): “The reasons for this are likely to be multifactorial and include decision-making around resuscitation of EP infants born <25 weeks’ gestation, geographical factors (distance to the nearest tertiary centre, precipitous births and births before arrival) and delay in/missed opportunities for in-utero transfer.”

The decrease in the rates of births in a tertiary perinatal centre may be a contributing factor for the trends observed in rates of IVH grades 3 or 4, as suggested by the reviewer. In the State of Victoria, we have previously published rates of IVH in extremely preterm infants born in 2010-2011, comparing those born in a tertiary centre vs not (Boland RA, Davis PG, Dawson JA, Doyle LW. Outcomes of infants born at 22-27 weeks' gestation in Victoria according to outborn/inborn birth status. Arch Dis Child Fetal Neonatal Ed. 2017;102(2):F153-F161). The rates of IVH grade 3 or 4 were 20% [10/49] in the outborn group vs 11% [42/396] in the inborn group. We have added this sentence to the text (Page 21. Discussion. Paragraph 6, sentence 5): “We have previously reported that rates of IVH grades 3-4 were higher in EP infants who were born outside a tertiary neonatal centre compared with those who were (20% vs 11%). The increase in rates of EP born outside a tertiary centre over eras may have contributed to the lack of decrease in rates of IVH grades 3-4 over time.”

5. *How were lethal anomalies defined?*

We have added the following statement to define “lethal anomalies” in the manuscript (Page 8. Patients and methods: (b) Perinatal and neonatal data collection. Paragraph 1, sentence 4): “Lethal anomalies were defined as major malformations coded according to the International Classification of Diseases[13] for which survival was not possible even with provision of neonatal intensive care.” (Ref 13: World Health Organisation. International statistical classification of diseases and related health problems - 10th revision (ICD-PM), 2016.)

6. *In Tables, given that 95% CIs are provided, p values could (and possibly should, at the light of the multiple comparisons) be omitted.*

We acknowledge that there are multiple comparisons and have added a sentence to the main text to highlight this point (Page 10. Statistical methods. Paragraph 3, sentence 2): “Given the multiple comparisons, we have interpreted our findings by focusing on overall patterns and magnitude of differences, rather than on individual p-values.” However, we have left the p-values in the tables as this is often requested by readers. We would be happy to remove the p-values at the request of the Editor and/or Reviewer.

<u>Reviewer #2</u>*: Cheong et al. present an epidemiological analysis of respiratory resources use and outcomes in extremely preterm infants in Victoria. The authors compare results from four cross-sectional observation periods across a 25-year period. Overall, it seems that increased survival of extremely preterm infants and changes in respiratory management along with higher oxygen saturation targets are associated with longer duration of hospital stay and respiratory support in this*

population. Aims, methods, results, and conclusions are well matched. Reasons for changes in respiratory morbidity, other co-morbidities and mortality are adequately discussed. The paper has considerable scientific and clinical relevance in the field. The manuscript is very well written and very clean.

I don't see any major issues that need to be addressed to improve the paper.

1. *The authors might consider the following minor recommendations:*
   - A sentence referring to the presence or absence of standard criteria for installing or ceasing CPAP or HFNC across the four eras may be worthwhile adding to the discussion. If those criteria have changed over time, could this affect the results?

   The commencement and cessation of CPAP or HFNC was at the discretion of the treating clinician. We have no data to suggest that there was a formal change in criteria for management of respiratory support for newborns in the State of Victoria. We have added a sentence to the main text (Page 19. Discussion. Paragraph 3, sentence 6): "There were no standard criteria for commencement or cessation of nCPAP or HFNC, which were at the discretion of the treating clinicians."

2. *Page 17, Table 3, line 59: Please add percentage of infants with BPD in the year 2005*

   Thank you. This is a formatting issue where the percentage of BPD in 2005 (58.2%) was located on the next page.

Reviewer #3: *In this study, Cheong and co-authors report on changes over time in care delivery and neonatal outcomes of extremely preterm infants born in the state of Victoria, Australia, focusing primarily on the kind of respiratory support provided. They analyze data from all children born in the region during four epochs between 1991 and 2017, an approach similar to that recently used by the same group in their report on long term respiratory outcomes of a similar population. The authors find a marked rise in the duration of respiratory support in the modern epoch compared with all previous epochs, due to prolonged used of high flow nasal cannula (HFNC). However, it appears that this increased length of respiratory support is unaccompanied by improved outcomes, although the presentation of the data makes this a little difficult to parse out. This is an important finding, and continues this group's efforts to determine whether the current reliance on prolonged respiratory support that characterizes care of these infants in most of the developed world is in our patients' and health care system's best interests.*

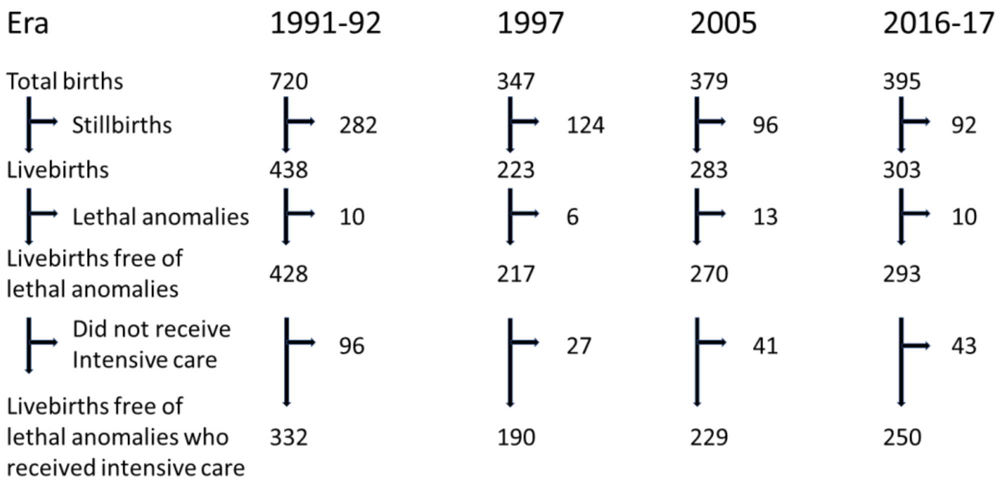
It is a little difficult to sort through the tables in their current form, and at least one analysis would benefit from significantly more explanation. Specific concerns follow:

Major concerns

1. *Table 1 is very long and difficult to read. Separating the first 4 rows, which contain no statistics, into their own first table would be helpful.*

Rather than creating a new table we have created a figure to show the same data that were in the first few rows (Figure 1). We have also added explanatory sentences to the main text (Page 11. Results: (a) Participant characteristics. Paragraph 1, sentence 1-3): "The numbers of total births, stillbirths, livebirths, including those with lethal anomalies and those who did not receive intensive care in each era are shown in Figure 1. The proportion of stillbirths in relation to all births fell over time, from 39.2% in 1991-92 to 23.3% in 2016-17. The proportion of infants who received intensive care was lower in 1991-92 compared with 2016-17 (odds ratio 0.60, 95% confidence interval 0.39, 0.91, p=0.02), particularly for infants born 23 and 24 completed weeks (Supplementary Figure 1a)."

Figure 1. Numbers of births, stillbirths, livebirths with and without lethal anomalies, and livebirths free of lethal anomalies offered intensive care in each era.



2. *In (the revised) Table 1, variables which do not differ across the four epochs can be removed from the table, and can be summarized for the whole population in the text prior to presentation of the table. Thus, gestational age, birthweight, and male sex can all be treated this way. Non-significant p values can be replaced with p=NS. In this (revised) table, the N (which appears to be infants offered intensive care) should be clearly stated. The text preceding this table could be more descriptive. For example, not only was the rate of antenatal corticosteroid treatment lower in 1991-91 compared with the other 3 cohorts, but the rate of this treatment did not differ significantly from 1997 onwards.*

We believe it is important to show all the data in the table, regardless of whether or not they are statistically significant between the eras. The purpose of showing rates of gestational age, birthweight, and male sex, for example, is to enable readers to determine how generalisable our groups are relative to their own. We have also not replaced p values <0.05 with NS. A significance level is arbitrarily determined and we believe it is important to show the p-value to enable the reader to interpret the findings as they see fit.

We have expanded the text prior to Table 1 to be more descriptive, as suggested by the reviewer. (Page 11. Results: (a) Participant characteristics. Paragraph 2): “Among infants who received intensive care, the proportion of infants born in a tertiary perinatal centre was higher in both 1991-92 and 1997 than in 2016-17 (Table 1). Antenatal corticosteroid treatment and rates of caesarean birth were lowest in 1991-92, but were similar to that of 2016-17 from 1997 onwards. Rates of intubation at birth were higher in all earlier eras compared with 2016-17. There were no differences between eras in multiple births, male sex, gestational age, birthweight, or birthweight z-score.”

3. *Table 2 is difficult to read. For clarity, the row reporting non-invasive ventilator days seems duplicative, as it is simply the sum of CPAP and HFNC days, and could be removed without compromising understanding. The unadjusted analyses do not need to be presented in the table, as there are no real differences between the unadjusted and the adjusted analyses. A simple*

statement in the text could say that. Finally, it is unclear what the authors mean when they state that HFNC "did not lead to a decrease in the number of CPAP days". Do they mean "accompanied by"? Is it the practice in this state that infants are always placed on CPAP first and then moved to HFNC?

We acknowledge that non-invasive ventilator days is a sum of CPAP and HFNC. However, we have deliberately presented this variable to show that non-invasive ventilator days had increased over the eras, and that the availability of HFNC was additive rather reducing CPAP days in the 2016-17 era.

Table 2 has been revised and only adjusted analyses included. The same comment applies to the supplementary tables.

As per your suggestion, we have amended the sentence "HFNC did not lead to a decrease in the number of CPAP days" to "HFNC, however, was not accompanied by any decrease in the duration of nCPAP in 2016-17 compared with 2005." (Page 13. Results: (b) Durations of resources for respiratory support, oxygen therapy and hospitalisation. Paragraph 1, sentence 4).

4. *It is unclear what the reader is to take away from presentation of the respiratory utilization data in survivors versus all comers. If the differences in analyses are important, more explanation and discussion of these differences would be helpful. Otherwise, one of these two tables (all comers or babies who survive to discharge) could be removed from the manuscript without compromising its impact.*

Since survivors consume more resources than do deaths, we wanted to determine that changes in respiratory resource consumption were not merely as a result of the increase in survival rates of EP infants across the eras. We have added a sentence in the methods (Page 9. Statistical methods. Paragraph 1, last sentence): "Since survivors consume more resources than do infants who die, we repeated the analyses including only survivors to discharge to ensure that the increased consumption of resources over time was not merely due to increased survival rates", and a sentence in the discussion (Page 20. Discussion. Paragraph 3, last sentence): "In addition, the increase in survival rates alone across the eras was not responsible for the increased overall consumption of respiratory support in EP infants."

5. *Table 3 is difficult to read. The table begins with survival to discharge, but the common morbidities of prematurity (NEC, surgery, IVH/PVL, treated ROP) are expressed as a percentage of all infants. It is not clear that these data on co-morbidity incidence across time are central to the thrust of the paper. If these data are relevant, they should be in their own table. It seems more central to the paper, and a much clearer narrative, for the authors to present in one table only the incidences of treatments and outcomes that are relevant to respiratory disease among survivors – administration of surfactant, postnatal corticosteroids, and the incidences of BPD.*

The data in Table 3 relate to the denominator of infants offered intensive care, as indicated in the caption to the table; the denominator is not all infants, as suggested by the reviewer.

In regard to the presentation of other morbidities in this paper, it addresses the aims of the study as stated at the end of the Introduction; “the aims of the study were to determine over the 25-year period: (i) the changes in respiratory support, survival to discharge home, and neonatal morbidities of EP infants offered intensive care, and (ii) the consumption of respiratory resources in relation to changes in survival.” Our intent is to report all common morbidities, including IVH, PVL, NEC, ROP, etc., as well as respiratory morbidities such as BPD; which is important and of interest to general readership. To clarify that we are reporting more than just BPD, we have added the words “major respiratory and non-respiratory” before “neonatal morbidities” (Page 7. Introduction. Paragraph 3, sentence 3).

We respectfully decline the suggestion to present respiratory treatments and BPD in the previous table (Table 2). Reviewer 1 has requested that we include data for the combined outcome of survival free of major morbidity (survival free of IVH, PVL, NEC, ROP and BPD). We have added this variable to Table 3, and thus, Table 3 needs to retain the individual variables that make up the combined outcome so the reader can see the contributions of each of the individual components to the combined outcome, and also how they relate to each other. Reporting BPD in a separate table would not be helpful to readers in this regard. Treatments with surfactant and postnatal corticosteroids are linked to the outcome of BPD and logically we feel they should remain in Table 3 alongside BPD, as well as the other treatment of surgery.

6. *The concept of added duration of respiratory support per additional survivor in the most recent epoch is not well explained, and is confusing. How is an additional survivor defined? What is being simulated? How are these simulations used? The statistics here are more complex than those usually used. and could be explained much more clearly. In like fashion, the reader is given no help in how Figure 1 is constructed or how to interpret it. What is involved in determining, from the current dataset, what is more/less costly and more/less effective? Many more details are required to make this analysis accessible to the reader. Similarly, the paragraph in the discussion about the cost-effectiveness of respiratory support, which appears to depend on this analysis, introduces concepts of efficiency, and the calculation of extra cot allocation for additional survivors. This line of discussion is not well informed by the results as they are currently presented. The conclusion that “The relative inefficiency in resource consumption could not be explained by increased consumption of resources by infants who died” is not supported by the text in the results section, or in the legend of Figure 1. If the authors wish to use this analysis, much more explanation in the methods. results and their interpretation should be devoted to it.*

A cost-effectiveness ratio is calculated as the difference in costs divided by the difference in outcomes (effects) between two groups. In our case the costs are expressed as resources for assisted ventilation and the outcome is survival. Hence, we calculate cost-effectiveness ratios between two eras as the difference in the consumption of resources for assisted ventilation per livebirth between the eras (numerator) divided by the difference in survival per livebirth between the eras (denominator), and the results expressed as additional days of resources for assisted ventilation consumed for each additional survivor. We have modified the explanation in the abstract (Page 3. Outcome measures. Sentence 2): “Cost-effectiveness ratios describing the average additional days of respiratory support associated per additional survivor were calculated”, and methods (Page 10. Statistical methods. Paragraph 2, sentence 1-2): “The added duration of respiratory support per additional survivor in 2016-17 was calculated relative to each of the other eras. These are known as incremental cost-effectiveness ratios, and were calculated as the difference in *cost* (mean days of respiratory support at the patient level in 2016-17 compared with other eras) divided by the difference in *effectiveness* (survival outcome at the patient level in 2016-17 compared with other eras), and are interpreted as the average additional cost associated with one additional infant surviving." We have also reworded our results to make this clearer; in the abstract (Page 3. Results. Last sentence): "The average additional costs associated with one additional infant surviving in 2016-17 were 200 (95% confidence interval (CI) 150, 297), 326 (183, 1127), and 130 (70, 267) days compared with 1991-92, 1997, and 2005 respectively", and results (Page 17. Results: (d) Cost-effectiveness ratios - increments in durations of respiratory support associated with one extra infant surviving. Paragraph 1): "The average additional cost associated with one additional infant surviving was estimated to be 200 (95% CI 150, 297) days of respiratory support in 2016-17 compared with 1991-92 (Figure 1). Similarly, in 2016-17 relative to 1997, there were 326 (95% CI 183, 1127) additional days of respiratory support for one additional survival achieved. Noting that for this comparison there were 2.6% of infant simulations in Figure 2 falling in the north west quadrant with greater cost and poorer outcomes, or in other words 'dominated'. The 'dominated' results were excluded from the estimated incremental cost effectiveness ratio to obtain a meaningful 95% CI for the remaining 97.2% of the simulations. The additional days of respiratory support associated with an additional infant surviving in 2016-17 relative to 2005 was estimated to be 130 (95% CI 70, 267) days."

An additional survivor is obtained from the difference in survival rates. For example, in Table 3 the survival rate to discharge home is 86.8% in 2016-17 and 74.2% in 2005. The difference is 12.6%, meaning there are 12.6 additional survivors per 100 livebirths in 2016-17 than there are in 2005. In the cost-effectiveness calculations, the difference in survival rates is the denominator, and it is expressed per livebirth (in the example comparing 2016-17 with 2005 it is 0.126 additional survivors per livebirth). Consequently the denominator is always a value less than one; hence a smaller increase in survival rates will translate into a larger cost-effectiveness ratio, if the difference in consumption of resources in the numerator is constant.

To explain the bootstrapping, and having introduced the concepts of costs (duration of respiratory support at patient level in 2016-17 compared with other eras) and effectiveness (survival outcome at patient level in 2016-17 compared with other eras), we have added the following to the sentence to the methods (Page 10. Statistical methods. Paragraph 2, sentence 4-5): "To account for sampling uncertainty, probabilistic sensitivity analysis was undertaken, using bootstrapping with 1,000 replications drawing from costs and effectiveness at the patient level. The bootstrapped results were presented using cost-effectiveness planes, with each of the 1,000 dots representing one mean cost and effectiveness outcome from random sampling, with replacement of the original sample."

Regarding efficiency, we have added a sentence to the methods. (Page 10. Statistical methods. Paragraph 2, sentence 3): "A higher incremental cost effectiveness ratio is less efficient than a lower ratio."

To present results supporting the relative efficiencies between eras, we have removed the numerical data from the Discussion into the end of the Results where its relationship with the cost-effectiveness ratios can be seen more readily, as follows (Page 18. Results: (d) Cost-effectiveness ratios - increments in durations of respiratory support associated with one extra infant surviving. Paragraph 2): "Using the above cost-effectiveness ratios, each additional cot in 2016-17 would produce just under two additional survivors compared with 1991-92 (365/200 days), just over one additional survivor per cot compared with 1997 (365/326 days), and just under three additional survivors per cot compared with 2005 (365/130 days) all of which compare unfavourably with approximately five extra survivors per cot if each additional survivor was to consume only the average days of assisted ventilation for survivors in 2016-17, which was 71 days (365/71 days)."

To explain the sentence in the Discussion (Page 20. Paragraph 4, sentence 3): "The relative inefficiency in resource consumption could not be explained by increased consumption of resources by infants who died," we have added the following to the end of that sentence "because the days consumed by infants who died were much lower than those consumed by those who survived, and were not substantially different between eras." The results for infants who died are already in the Results (Page 15. Results: (c) In-hospital survival, treatments and morbidities. Paragraph 1, last sentence): "The median durations of hospital stay for those who died were similar across eras (median days [interquartile range; 25^{th}-75^{th} centiles] 1991-92: 4 [1, 13]; 1997: 3 [1, 17]; 2005: 14 [2, 69]; 2016-17: 6 [1, 20])" to which we have added ", and are much lower than for infants who survived (Table 2)", from which readers can quickly see that the medians for infants who died are much lower than for those who survived.

7. *The discussion of the factors leading to the prolonged usage of HFNC in the modern era is important, and deserves to be much higher up in the discussion. The authors suggest earlier in the manuscript that much of HFNC treatment may be "wasted". Elaboration of this idea in this paragraph would tie the argument together.*

The discussion of the factors leading to the prolonged use of HFNC is in Paragraph 3 of the Discussion, after the introductory paragraph which summarises the main findings of the whole study, followed by a paragraph discussing survival rates. It would seem logical to discuss survival first followed by respiratory support, and hence we have not moved the paragraph higher. As we already discuss in this paragraph, we think that our findings may be explained by the lower thresholds for starting and stopping HFNC, than for nCPAP. The interpretation of whether or not HFNC may be "wasted" rests with the reader.

8. *To what do the authors attribute a significant decreased incidence of surfactant use in the most recent epoch compared with 2005, especially since the use of corticosteroids was unchanged?*

The decrease in surfactant use is likely a consequence of less intubation at birth in 2016-17 and a move towards more non-invasive respiratory support to treat respiratory distress syndrome rather than intubation and mechanical ventilation. Minimally invasive surfactant administration, whereby surfactant is administered to non-intubated infants by inserting a small catheter into the trachea, injecting the surfactant, and then removing the catheter, was not widely practised in these cohorts, including the most recent 2016-17 cohort.

Minor concerns

1. *In the introduction, the authors point to the abundance of randomized studies that have failed to show decreased incidence of BPD with non-invasive ventilation compared with invasive ventilation. This is an important point that is underappreciated. If they wish to cite the 2017 JAMA meta-analysis as a counterargument, it would be helpful to interpret that study's findings in the context of the weaknesses inherent in any meta-analysis. The author's desire to express cost*

effectiveness in resource units of days of respiratory support is well taken. The measure could therefore be more directly reported and discussed throughout the manuscript as "duration of respiratory support." Thus, Table 2 would be more titled "Duration of respiratory support, oxygen therapy, and hospitalization….:

We have added the following sentence to highlight the weaknesses inherent to the JAMA meta-analysis (Page 6. Introduction. Paragraph 2, sentence 5): "However, the meta-analysis was limited by the overall poor quality of evidence, heterogeneity of the non-invasive ventilation strategies, and lack of robust high-quality randomised trials."

We have altered the titles for Table 2 (Page 14) and subsection (b) of the Results (Page 13) as suggested.

2. *Did the infants in the 2016/17 epoch require additional days of respiratory support, or simply receive additional days?*

   We acknowledge that we cannot be certain of the distinction between "requiring" or "receiving" additional days of respiratory support. We have removed the term "required" throughout the manuscript where it pertains to additional days of respiratory support.

3. *Expressions of personal reaction to findings, such as "worrying" or "encouraging" do not really belong in a scientific report. These could be replaced with more descriptive terms.*

   As requested, we have modified sentences and removed these terms.

   - Page 19. Discussion. Paragraph 2, sentence 1: "The increase in EP survival from the current study is in keeping with trends worldwide over a similar time period."
   - Page 21. Discussion. Paragraph 6, sentence 1: "We noted a decrease in births within a tertiary neonatal centre," and sentence 4: "The decrease in cystic periventricular leukomalacia by almost 80% over time is similar to reports from other cohorts, although rates of IVH grades 3-4 remained unchanged."

<u>Reviewer #4</u>: *This is a paper with important data that took a lot of effort to collect.*

1. *However, putting it into certain paradigm and linking it to conclude that there is a causative relationship between certain ventilation methodology and adverse outcomes over 25 years is not appropriate.*

   We have reported observed trends throughout the paper of survival, respiratory support methodologies and short term morbidities. There has been no intention to link any causation between these factors as we recognise that observational cohort studies are only able to establish associations and not causation.

2. *Furthermore, it is highly inappropriate to suggest how one type of methodology has influenced another (ie HFNC has not shortened the time of CPAP). That depends on the local policies of use and it's users. Every institution can decide their own non-invasive ventilation support policy, choice of first line treatment and the speed of weaning it.*

Our comment (Page 19. Discussion. Paragraph 3, sentence 4) on "...the use of HFNC, which was mostly additive rather than an alternative to the use of nCPAP" was based on the interpretation of the data. There is no intention to suggest that one methodology had influenced the other, rather an observation that the availability of a newer mode of non-invasive respiratory support (HFNC) in the 2016-17 cohort has not been associated with a decrease in the duration of use of a more established mode i.e. nCPAP.

3. *I would suggest to present this data in a form of Epicure data, only for the cohort in question. It is valuable to express the changes through time chronologically for what they are, but not to speculate causation in any form.*

   We have presented the results to be in keeping with the aims of the study, which was to determine changes in respiratory support, survival and morbidity in infants born <28 weeks since the early 1990s. The EPICure studies reported data on less mature infants; <26 weeks (EPICure 1) or <27 weeks (EPICure 2). We have reported some outcomes by week of gestation in the supplementary figures for those who wish to compare our outcomes more directly with the EPICure studies. As per our response to your Point 1, we have reported observations of trends and possible explanations to explain these trends over time. We have not speculated causation.

4. *Please also explain more on the methodology. I am not sure this is a prospective study, unless the protocol for it was written in 1991?*

   A prospective study refers to one that "goes ahead in time", rather than retrospective whereby we would have "gone back in time to comprise the cohorts and follow them up to the present" [Ref: Grimes DA, Schulz KF. Cohort Studies: Marching towards outcomes. Lancet 2002;359(9303):341-5.] Our cohorts were recruited at birth for each era, and data prospectively collected from the time of recruitment, which fulfils the definition of a prospective study.

VERSION 2 – REVIEW

REVIEWER	Marina Cuttini Ospedale Pediatrico Bambino Gesù, Italy
REVIEW RETURNED	03-Jun-2020

GENERAL COMMENTS	The Authors have answered all my queries. As for my suggestion to cancel p values from tables and rely on 95% confidence intervals, (possibly with footnotes stating that lack of inclusion of value 1 in the 95% CI indicates a p<0.05), I leave the decision to the Editor.

REVIEWER	Jeremy Marks University of Chicago United States
REVIEW RETURNED	14-May-2020

GENERAL COMMENTS	The authors have addressed most of the concerns raised in the previous reviews, and the manuscript is much improved, with accompanying clearer impact.

	There remain several issues to do with the presentation of the cost analysis, which has already benefitted enormously from the newly added text explaining the analysis, results and interpretation. 1. Please explain how the point estimates of average additional costs (e.g. 200, CI 150, 297; 2016 vs 1991-2) are derived from the data presented in Fig. 2. 2. Please include the unit of measure when reporting these point estimates in the abstract. 3. Regarding the exclusion of the 2.6% of simulations in Fig. 2 that are “dominated”: The results will be easier to read if the statement that these simulations were excluded is moved to the Methods section, along with a rationale for why they were excluded. If the authors wish to use the word “dominated”, this should also be defined. 4. In the survivors per cot analysis, the authors annualize the reported point estimates of the incremental cost effectiveness ratios between 2016 and the prior epochs. While this is straightforward, the rationale for comparing these annualized estimates with the average days of assisted ventilation in 2016 is not. Please provide a rationale for this comparison, and an interpretation of the unfavorable result. 5. When reporting the additional survivors, please present the actual ratios, rather than rounding to the nearest whole survivor (“just under two survivors”). 6. The figure citation in this paragraph is probably intended to be to Figure 2.

REVIEWER	Dr Sanja Zivanovic University of Oxford , UK
REVIEW RETURNED	31-May-2020

GENERAL COMMENTS	Changes in the article have been made according to reviewers comments. I would be happy to accept without further changes

VERSION 2 – AUTHOR RESPONSE

<u>Reviewers' comments:</u>

<u>Reviewer #1</u>:

1. *The Authors have answered all my queries. As for my suggestion to cancel p values from tables and rely on 95% confidence intervals, (possibly with footnotes stating that lack of inclusion of value 1 in the 95% CI indicates a p<0.05), I leave the decision to the Editor.*

   We have not altered the tables, as per our response to Q3 above.

<u>Reviewer #3</u>: The authors have addressed most of the concerns raised in the previous reviews, and the manuscript is much improved, with accompanying clearer impact. There remain several issues to do with the presentation of the cost analysis, which has already benefitted enormously from the newly added text explaining the analysis, results and interpretation.

1. *Please explain how the point estimates of average additional costs (e.g. 200, CI 150, 297; 2016 vs 1991-2) are derived from the data presented in Fig. 2.*

We have estimated the mean and the 95% CI using bootstrapping. The mean is estimated as the average of the 1000 replications, and the 95% CI the 2.5 and 97.5 percentiles. We have revised the wording in the method to make it clearer that bootstrapping was used for the point estimates and have also now included an additional reference (reference 17: Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes: Oxford University Press; 2015) for the method.

Page 10; Statistical methods; paragraph 2; sentence 4-6: "To account for sampling uncertainty, a bootstrapping method with 1,000 replications drawing from costs and effectiveness at the patient level was used to estimate the mean and the 95% CI of the cost-effectiveness ratios, where the 95% CI were estimated using the 2.5 and 97.5 percentiles.[17] Each of the 1,000 replications represents one mean cost and effectiveness result from a random sampling with replacement of the original sample. The bootstrapped results were also presented using cost-effectiveness planes."

2. *Please include the unit of measure when reporting these point estimates in the abstract.*

We have now included the unit 'days' after all point estimates in the abstract (Page 3; Results; last sentence).

3. *Regarding the exclusion of the 2.6% of simulations in Fig. 2 that are "dominated": The results will be easier to read if the statement that these simulations were excluded is moved to the Methods section, along with a rationale for why they were excluded. If the authors wish to use the word "dominated", this should also be defined.*

We added the following to the Methods, which explains that any "dominated" results were to be excluded, and the definition of "dominated".

Page 10; Statistical methods; paragraph 2; sentence 7: "In the event that any of the simulations were 'dominated' (i.e., they represented both greater costs and poorer outcomes), they were excluded from the calculations because the negative incremental cost effectiveness ratios they produce are not interpretable.[17]"

We have also altered the section in the Results, as follows:

Page 18; subsection (d); paragraph 1; sentence 3: "There were 2.6% of the replications in Figure 2 falling in the north west quadrant, which were excluded from the estimated incremental cost effectiveness ratio to obtain a meaningful 95% CI for the remaining 97.2% of the replications (the excluded replications are available in Figure 2)."

4. *In the survivors per cot analysis, the authors annualize the reported point estimates of the incremental cost effectiveness ratios between 2016 and the prior epochs. While this is straightforward, the rationale for comparing these annualized estimates with the average days of*

assisted ventilation in 2016 is not. Please provide a rationale for this comparison, and an interpretation of the unfavorable result.

We agree with the reviewer that the rationale was not clear so we have deleted this last comparison. (Page 19; Results; subsection (d); paragraph 2; last sentence). For clarity, we have added a sentence in the Discussion (Page 22; paragraph 4; sentence 2: "Considering the substantial value of an infant's life, the additional investment in respiratory support in recent years is likely worthwhile.")

5. *When reporting the additional survivors, please present the actual ratios, rather than rounding to the nearest whole survivor ("just under two survivors").*

   We have revised this to present actual ratios. (Page 19; Results; subsection (d); paragraph 2; sentence 1: Using the above cost-effectiveness ratios, each additional cot dedicated for respiratory support in 2016-17 led to 1.8 survivors compared with 1991-92 (365/200 days), 1.1 additional survivors compared with 1997 (365/326 days), and 2.8 additional survivors per cot compared with 2005 (365/130 days).)

6. *The figure citation in this paragraph is probably intended to be to Figure 2.*

   Thank you for pointing out this oversight. We have rectified this in the text. (Page 18; Results; subsection (d); paragraph 1; sentence 1: The average additional cost associated with one additional infant surviving was estimated to be 200 (95% CI 150, 297) days of respiratory support in 2016-17 compared with 1991-92 (Figure 2)).

<u>*Reviewer #4*</u>: *Changes in the article have been made according to reviewers comments. I would be happy to accept without further changes.*

Thank you.

We hope that these amendments will meet the requirements of the reviewers and editorial team.

VERSION 3 – REVIEW

REVIEWER	Jeremy D. Marks University of Chicago USA
REVIEW RETURNED	14-Jul-2020

GENERAL COMMENTS	My remaining concerns have been addressed appropriately. My recommendation is to accept this manuscript in its current form.