Supplementary material

1. Supplemental Appendix One – Weaning Methods

Infants who had been off respiratory support for >7days were part of the exclusion criteria for the CHiPS study. Seven days was chosen as a cut off as infants who had to go back on respiratory support after this period were deemed likely to have an intercurrent illness or event, rather than to have failed primary weaning. Infants off respiratory support for <7days who then required respiratory support were, however, eligible for the study, and were placed on nCPAP for at least 48 hours prior to randomisation.

For all infants the study commenced at 12 midday on the day of randomisation with weaning as follows:

1. nCPAP arm: Weaned over 96 hours (all changes at 12 midday). Day 1 and 2 remained on nCPAP 6cm water (48 hours), day 3 reduced to nCPAP 5cm water for a further 48 hours. Day 5 off nCPAP at 12 midday.
2. nHF arm: Weaned over a period of 96 hours (all changes at 12 midday). Day 1 commenced on nHF flow 6L/min, Day 2 4L/min, Day 3 3L/min, Day 4 2L/min. Day 5 off nHF at 12 midday.

Failure criteria included any of the following: 6 desaturations (<80% saturations) and associated bradycardias (<100/min) in 6 hours requiring bedside intervention, one episode requiring intermittent positive pressure via T piece, or respiration rate >70/min consistently for 30 minutes.

2. Supplemental Appendix Two – Statistical Analysis

Observational data obtained over a 10-year period for 310 infants <30 weeks GA was used as pilot data. Review of the pilot data showed that over 80% of cases came off nCPAP in the 4 to 5 weeks between PMA of 31 and completion of 34 weeks PMA. This was regarded as the time of active weaning and was taken into account in the simulation models.

A bootstrap method was used for each set of study parameters ie: sample, size, effect size and non-inferiority threshold. Four hundred bootstrap samples with replacement were drawn from the data. From the 400 samples the proportion which would have concluded in non-inferiority using one-sided (2.5%) confidence intervals based on a heteroscedastic least-squares model was determined. These computations yielded the estimated power of each model and indicated a total of 100 infants (50 per arm) would provide a 92% power to concluded non-inferiority of the weaning time between the two methods of respiratory support. A non-inferiority margin of 15% was chosen. For nHF weaning to be shown to be non-inferior to nCPAP weaning, the mean time in the nHF arm plus the 15% non-inferiority margin would need to lie within the 95% confidence interval of the nCPAP weaning time. For a mean birth gestation of 29 weeks, the model indicated that this 15% margin amounted to a difference of approximately 4 days.

The primary outcome was analysed using a per protocol (pp) analysis. This was chosen as in the context of a non-inferiority trial it is likely to be a more conservative estimate. However, it was also planned to present the intention to treat analysis as a secondary analysis and discuss any differences that might be detected. Data was analysed as hours on respiratory support (rather than days) to avoid the problem of interpreting incomplete days.

The Restricted Mean Survival Time (RMST) comparison was used to analyse the primary outcome. It is a widely used and well described alternative time-to-event analysis method. It accommodates censoring but is not specific to censored data. The basic assumption behind RMST is that hazards in both arms may converge at some point and that to detect the difference between interventions; one should look at differences before such a point of convergence. RMST is conservative in a non-inferiority setting, as it does away with the portion of the follow-up where interventions behave similarly. RMST focuses on duration rather than hazard, which has greater relevance to the study. RMST does not

assume proportionality of hazards. In order to obtain the RMST a survival curve was estimated through standard survival analysis. The cut-point was selected using the modified Sheldon hazard method, (1) which estimated the hazard at each event time and then searched for a point which gave the smallest sums of squared error for a model. This allowed for different hazards between the two groups and the same hazard after the chosen time point.(1) This method was chosen as, per the Sheldon reference above, competing methods underestimate the cut-point, potentially washing out the differences in restricted mean durations. This is why, in a non-inferiority setting we settled on this procedure for the cut-point selection.

Time on the study methods of support was not available for all as clinically at 36 weeks PMA, infants who are still on pressure support such as nCPAP or nHF are changed to low flow oxygen instead. This is because infants on nCPAP cannot establish suck feeds and use of oxygen at this later period after birth does not cause deleterious effects whereas before this all reasonable attempts are made to avoid oxygen if possible. Therefore, at 36 weeks PMA, the 2 study arms effectively cease and oxygen is given.

A subgroup analysis of time on respiratory support for infants born at <27 weeks GA was planned. The data was subject to the same procedure as for the primary outcome above and a cut-point was similarly determined.

The PSS: NICU survey which is one of the secondary outcomes consisted of four domains and a total score. The score ranged on a continuous scale (0-100); this was analysed using linear mixed modelling and empirical sandwich estimators that adjust for the standard errors and test statistics involving the fixed-effects model. A parent factor was regarded as a random effect. Interactions between group and time were investigated across the 4 domains and the total score. A Tukey test was used for multiple comparisons within the PSS:NICU.

3. Supplemental Appendix Three - Results

Twenty infants with more severe evolving CLD did not meet our inclusion criteria, and therefore were not included.

Using intention to treat analysis, the restricted mean hours from randomisation to 72 hours off respiratory support or 36 weeks PMA was 409 hours in the nCPAP group giving an upper bound of 470 hours (mean plus 0.15 non-inferiority margin), and 390 hours in the nHF group with an upper confidence interval of 435 hours. These results supported the findings of non-inferiority reported in the per protocol analysis (Figure 3 in article).

A post hoc analysis indicated that if it was assumed the nCPAP group could have come off support at 24 hours instead of 96 hours, this did not significantly change the time on support. The median difference between groups in this case was -23 hours (95%CI -72 to 83 hours) with the shorter median time in the nHF group.

In our subgroup analysis of our <27 week GA infants we were unable to conclude nHF non-inferior to nCPAP when weaning from nCPAP within a 15% non-inferiority margin. The cut point was selected using the Sheldon hazard method(1) and for this sub-group was 504 hours. The restricted mean time from randomisation to wean off respiratory support in the nCPAP group was 304 hours with an upper boundary of 350 hours (mean plus 0.15 margin) and when weaning using nHF this was 362 hours with an upper confidence interval of 408 hours.

Results from the PSS:NICU(2) showed stress scores were reduced in the nHF group, in regards to the 'relationship' section of the survey (p=0.045) which assesses parents perceptions on separation, contact and interaction, but increased in both groups in the sights and sounds section, over time (p=0.005). Although the total PSS:NICU score did not show significant interaction between group and time, the interaction terms were retained for each

domain as it was felt likely that parents' perceptions would likely change with time in regard to some of the categories of the questionnaire.

REFERENCES:

1. Sheldon EH. Choosing the Cut Point for a Restricted Mean in Survival Analysis, a Data Driven Method. VCU Scholars Compass. Virginia Commonwealth University. https://scholarscompass.vcu.edu/etd/3121. Accessed May 18, 2020.
2. Miles MS, Funk SG, Carlson J. Parental Stressor Scale: neonatal intensive care unit. Nursing research. 1993 May.