Author’s response to reviews
Title: Evaluation of two SpO2 alarm strategies during automated FiO2 control in the NICU: a randomized crossover study

Authors:
Malgorzata Warakomska (g.warakomska@gmail.com)
Thomas E. Bachman (TBachman@me.com)
Maria Wilinska (wilinska.maria@gmail.com)

Version: 1 Date: 06 Mar 2019

Author’s response to reviews:

REVIEWER 1
Martin Keszler, MDCM (Reviewer 1):

This is an unblinded randomized crossover study of 21 infants conducted in a single small NICU that routinely uses automated FiO2 control. During the first 6 days of respiratory support a tight and a loose alarm strategy were alternated each 24 hours while the automated FiO2 control was in use. The tight strategy set the alarms "just outside the prescribed SpO2 target range" with a 30-second delay. The loose strategy set the alarms 2% wider, with a 90-second delay. There is little information regarding the appropriate alarm settings during automated FiO2 adjustment in neonates. Alarm fatigue is an important risk in NICU and may have serious adverse consequences - thus this is an important question. As anticipated, the authors found fewer alarms during the loose strategy. Importantly, the time spent within the target range was not different and there was no increase in the time spent at extremes of SPO2.

Specific comments:

What was the patient : nurse ratio during the study? It is well known that this is an important determinant of how well SPO2 targets are maintained. In our unit there is usually one nurse assigned to each two infants receiving respiratory support. We did not include collection of this data for this study. Our years of clinical experience using Auto-FiO2 control has led us to believe that the infant:nurse ratio does not have a significant impact on the maintenance of SpO2 targets. We decided to add a note about staffing to the manuscript along with some other information about care. (RESULTS, 1st paragraph).
The proportion of time at normoxemia appears remarkably high and above what is seen in the literature. Some of this may be a function of the generous definition of normoxemia and to the fact that, as the authors acknowledge, these were relatively stable intubated infants. However, including high SPO2 in normoxemia definition while in 21% oxygen is problematic. Given the duration of the study (6 days), it is quite possible that many infants had largely resolved their lung disease and were in 21% oxygen, which would overstate how well normoxemia was maintained. Thus it would be important to present the data as normoxemia while receiving supplemental oxygen (as you have done in the Figures). Thank you for bringing this potential problem to our attention. In our unit infants are transferred to the Infant Flow after their short stay on the AVEA-CLiO2. As a result of your comment we realized the protocol might have influenced this practice, and looked closer. We found that in about 15% of the days the median FiO2 was less than 0.22. These infants averaged 99% time in normoxemia. However the 10% of days with the highest FiO2 (0.48) the infants spent 90% of the time in normoxemia. We did calculated the time with SpO2 between 86-96 associated with supplemental oxygenation and found them to be 89.7% (6.8) and 91.5% (8.9). We have added that to the manuscript to avoid others reaching the same speculation as yours. We added this to Table 3 and also amended the Methods and Results (METHODS, 5th paragraph, line 6, RESULTS, 4th paragraph, last sentence)

Relevant to the above point, is sedation/analgesia routinely employed in your NICU for ventilated infants? If so, this could be another explanation for the stability of oxygenation. Please clarify. We do not routinely use sedation or analgesia during respiratory support. These newborns are subject to the standard care known as minimal handling, according to their clinical condition. 20% oral glucose is routinely used before the uncomfortable procedures. We decided to add sedation/analgesia information to the manuscript along with some other information about care. (RESULTS: 1st paragraph)

A substantial amount of data was excluded because of limit settings inconsistent with the study. Such high rate of protocol violation is concerning in an unblinded study. Please include this as a limitation.

We add this to the limitations. (DISCUSSIONS: 6th paragraph, line 15)

Though relatively infrequent, were the prolonged episodes of hyperoxemia and hypoxemia > 3 min evenly distributed between the two strategies?

These did tend to be more common during the loose strategy, especially for hyperoxemia. The differences were not statistically significant even as a proportion of days. There were 12 &16 for hypoxemic and 3 &11 for hyperoxemic. Nevertheless in light of the trend associated with the
longer time delay, we have added the summary data to the table and refined the comment in the text. (RESULTS: Table 4, and 6th paragraph, last sentence).

Without actual data, the statement that: "Since the 90-second delay in the loose alarm strategy should have eliminated essentially all the high and low SpO2 alarms, the residue level of a couple per hour is clearly persistent signal quality alarms" is a speculation that may or may not be accurate. This statement should be eliminated or at least made explicitly speculative.

We did not feel comfortable deleting this thought, as the residual alarm rate was much higher than the rate of episodes greater than 1 minute in duration. We have reworded this thought and made the comments explicitly speculative. (DISCUSSION: 4th paragraph, last sentence.)

Minor comments:

The definition of tight control as settings "just outside the target range" could be interpreted variously. Why not be specific and say 1% above and 1% below (unless the setting was variable, which would be problematic). The units for the alarm range is missing "2 wider" - I assume 2% wider? But is it 2% both above and below, or is it 1% above and 1% below the 'tight' for a total of 2? From Table 3 it appears to be the latter, but this should be stated more clearly in the text.

Thank you for bringing this to our attention, we have changed the wording to make it clearer. (METHODS: paragraph 4, starting line 3)

In fact, it appears that the target ranges were somewhat variable, though not widely divergent. Still, for a study like this, it would be preferable to pick a target range and alarm limits that are uniform. If the clinical team wished the settings to be different, then the patient is not eligible…

When designing the study we assumed there would be more variation in the set target range, with a significant portion narrower (i.e., 90-95%). We now see that the study design would have been better if we had fixed the target range, and as you suggested excluded subjects whose attending physician felt a different range was appropriate.

Although generally very well written, there are a few grammatical and syntax errors (which should be correctable in copy editing).
We are sorry these crept in. Our co-author, Tom Bachman, a native speaking Californian has carefully reviewed this aspect of the revised manuscript.

Page 7, line 51, I believe that by "projectability", the authors mean generalizability.

We have made this change. (DISCUSSION: paragraph 6, 1st line)

REVIEWER 2

Wissam Shalish, M.D. (Reviewer 2): The following manuscript evaluated, in a randomized crossover fashion, the impact of two SpO2 alarm strategies during automated FiO2 control on frequencies of alarms as well as SpO2 extremes. The study overall provides interesting and novel findings regarding the consequences of setting a tight or loose alarm strategy, and setting a broad or relatively narrower alarm delay, when using automated FiO2 control systems. However, a few suggestions and important clarifications are needed:

Overall comments

- A more in depth review of the evidence regarding alarm fatigue may strengthen the rationale for this study: What constitutes alarm fatigue (i.e. what is the threshold for which alarm fatigue may pose a problem, based on the available evidence)? Are there data reporting frequency of alarms and/or perceptions of alarm fatigue in patients exposed to automated FiO2 control as opposed to manual control? In other words, does a median alarm frequency of 5 per hour, as shown in the tight SpO2 target strategy, correlate with alarm fatigue based on the available literature?

- A similar review on SpO2 alarm delay is needed: what is currently reported in the literature? On what basis were alarm delays of 30-sec and 90-sec selected?

We have expanded this section of the Background, addressing poor compliance with alarm settings, as well as description and impact of alarm fatigue. The kind of information that you outlined is not available. As we already reported, the AAP and European guidance don’t agree on alarm settings, and their recommendations are not evidence driven. We do hope our work will stimulate more studies. (BACKGROUND: paragraph 1)

The system permits delays as long as 120 seconds. The 30-second alarm delay used in the tight strategy was the typical delay most often used in the unit. It was also the time delay used in a
multicenter trial in which we participated. The time delay in the loose strategy, 90-seconds, was selected following a discussion among the investigators.

- In the introduction, the authors rightfully indicate that the aim of this study was to determine whether a loose alarm strategy could significantly reduce alarm frequency "without increasing over reliance on automation". The fact that the loose alarm strategy did not create an increased exposure to SpO2 extremes is reassuring. However, it is also very important to look at the automated fluctuations in FiO2 for each patient as a marker of safety. The results section seems to indicate a wide IQR for FiO2 changes, but it is not clear how much oscillation occurred per patient? This is critical, because in the first days of life, small FiO2 changes may indicate changes in lung state that could require more prompt intervention (Example: lung derecruitment from improper ETT placement, pneumothorax, pulmonary hemorrhage, need for surfactant etc). Thus, it may be warranted that the clinical team be notified (using alarms) of wide fluctuations in FiO2 even in the absence of SpO2 extremes.

You have brought up an interesting point about Auto-FiO2 control. The system we used has FiO2 alarms that address these issues. There is a high and a low FiO2 alarm. The nurse is alerted when either condition is met. The Auto-FiO2 system will not deliver O2 levels below the set low level, but will deliver FiO2 levels higher than the threshold, if it deems it necessary. These alerts have a 60 second delay so as not to trigger alerts in response to short transient changes. There is also another alarm that warns the nurse if the FiO2 baseline average value has increased significantly in the previous two hours. The operator also selects the desired change threshold. This is intended to alert the staff to the situation where the SpO2 is being well controlled by the AutoFiO2 but the infant is deteriorating.

Studies of this Auto-FiO2 system have consistently reported more time in the set target range and decreases in extreme SpO2 episodes. However the findings on FiO2 exposure have not been consistent, with some showing small increases, decreases and no difference associated with automatic control. These different findings are likely a reflection of differences in manual control.

The characterization of more small swings in FiO2 associated with Auto-FiO2 that result in better SpO2, while obviously present, have not been described. Your question has provoked our interest in evaluating this using data from studies that included both auto and manual control.

Further elaboration on this notion could strengthen the manuscript, as follows:

- Provide histograms of FiO2 variability (similar to Figures 1a and 1b), but making sure to separate the tight and loose strategies for comparison sake.
We originally considered this as we prepared the manuscript. We decided it would not be useful for two reasons. First the basal FiO2 needs changed so much from initiation of treatment to weaning it wouldn’t be of much help describing the changes in FiO2 that you are interested in. To illustrate, at the start of the study the FiO2 (1st hour median) ranged between 33-90% for the 21 infants. The median for day1 ranged between 23% and 58%. Second, the set target range effects the FiO2 delivery, and the differences in SpO2 alarm settings are not even considered in the control algorithm.

In another study in which we participated, infants were randomized to one of two target ranges, and then were managed in a 24-hour crossover with Auto and Manual FiO2 control. (vanKaam JPeds 2015) A histogram of FiO2 was produced but not included in the publication. We have included it below, to give you a better perspective of the system’s performance, compared to manual. But again, we would suggest that much of the variation is a result of the basal O2 need of each infant even in the 2 days of study.

- Provide more data, if possible, on the patient population: did they all receive surfactant? What was the mode of ventilation (volume-targeted ventilation requires less accrued monitoring as it auto-adjusts to patient's lung compliance), were blood gases any different during the tight or loose alarm strategies? Discuss above findings in the Discussion section

We have added information, including the use of surfactant, the initial FiO2 levels, more detail on the modes of ventilation, staffing ratio and the use of sedation and analgesia. (RESULTS, Table 2, paragraph 2 starting line 4, paragraph 3 last two sentences)

We did not perform systematic arterial blood gas assessments, and those taken were not recorded as part of the study. Transcutaneous PtcCO2 was monitored in many infants, but not recorded as part of the study. Our impression is that there were no differences in either associated with the alarm mode, but subtle differences would not have been evident.

Other minor clarifications
- What was the averaging time for SpO2?

The oximeter used with this Auto-FiO2 system is the Masimo Set, and the averaging in the system is always 8 seconds.
- Based on Table 3, it appears that there is some overlap in the high and low SpO2 alarms between the 2 groups. This might be better shown in the form of histograms comparing the distributions of high SpO2 alarms (tight vs. loose strategy) and low SpO2 alarms (tight vs. loose strategy). This could be provided in the supplementary appendix. If the overlap appears important, then could one speculate that the observed results stem more from the 90-second delay rather than the alarm strategy itself?

Thank you for this idea. Unfortunately we do not have data separating the High and Low alarm frequency. As a result of your question produced a histogram that shows the distribution of the differences between the frequency of alarm events for the two alarm strategies. As expected there is some overlap, but we are not led to any speculation. However we do believe the histogram is helpful in describing the primary effect of the two alarm strategies, and have added the figure to the manuscript. (RESULTS: Figure 2)

- In Table 3, should FiO2 be shown as median (IQR) or mean (SD)? It is not clear based on the fact that it says 'median' but only one value is presented between parentheses.

We understand the point of confusion; our presentation was inadequate. The SpO2 and FiO2 values in this table are the daily mean of the hourly median values. To make this clear we have changed Table 3, and added a description in the Methods (paragraph 5, last sentence).

- There are several minor typos and syntactical errors throughout the manuscript. Careful revision is required.

We are sorry these crept in. Our co-author, Tom Bachman, a native speaking Californian has carefully reviewed this aspect of the revised manuscript.