PREVALENCE OF POTENTIALLY CLINICALLY RELEVANT COMPLEX EPISODES OF EXTREME SpO2 DURING MANUAL AND AUTOMATIC CONTROL OF INSPIRED OXYGEN

Thomas Bachman, Karel Roubik

Department of Biomedical Technology, Faculty of Biomedical Engineering, Czech Technical University in Prague, Kladno, Czech Republic

Abstract
Continuous monitoring with pulse oximetry is the standard of care for titrating inspired oxygen in the neonatal ICU. However, titrating supplemental oxygen to address frequent desaturations is a challenging task for caregivers. Increasing exposure to SpO2 extremes is associated with increasingly poorer long-term outcomes. More recently, the prevalence of prolonged episodes at extremes and clusters of short episodes have been reported to be also associated with bad outcomes. We speculated that more complex episodes might also have an impact on outcomes. We defined two sets of these: clusters and swings. Automatic control of inspired oxygen based on continuous pulse oximetry is available on many neonatal ventilators. Some have expressed concern that continuous adjustment of inspired oxygen, without observing the infant, might cause instability and thus increased prevalence of clusters and oscillations. The aim of this study was to determine the prevalence of these complex events and determine if they were more common during automated control. To accomplish this, we analyzed data of 58 extremely preterm newborns that were ventilated at least 24 hours with manual inspiratory oxygen control and at least 24 hours with automated FiO2 control, in random order. We found that clusters and swings were quite prevalent, that is similar to the prevalence of prolonged episodes that have been shown to be associated with bad outcomes. We also found that these complex events were reduced during automated control, rather than increased. Finally, we suggest that additional research in this area is warranted.

Keywords
Oximeter, oxygen saturation, neonatal ICU, hyperoxemia, hypoxemia

Introduction
Titrating inspired oxygen (FiO2) to maintain good control of SpO2 is a frustrating task. Infants desaturate many times an hour and a transient increase in FiO2 is sometimes appropriate to offset the potential hypoxemia. Unfortunately, infants are often exposed to excess oxygen as a result of failure to reduce the FiO2 back to baseline levels [1]. Further, increased exposure to hypoxemia and to hyperoxemia have been shown to clearly affect mortality and morbidity [2, 3].

One important report demonstrated that the increasing frequency of longer episodes of hypoxemia is associated with both excessive total time in hypoxemia and with poor long-term outcomes [2]. It has also been shown that an increased rate and clustering of desaturations are also associated with poor outcomes [4–6]. The impact of hyperoxic episodes, is perhaps similar, but has not been studied. Nevertheless, it is well understood that the damage from the hypoxia is exacerbated when followed by hyperoxia [7].

New automated FiO2 control options, now available in many infant ventilators, have been shown to reliably improve SpO2 control, with dramatic reductions in the need for manual adjustments [8, 9]. All automated control systems, if not properly damped, can become unstable, and this is a concern for neonatal FiO2 control systems [10, 11]. While the basic control approaches of the array of automated FiO2 systems is well described, details of differences in dampening are not available and bench simulation tests are not generally available [8]. We previously reported on the dampening of one automated FiO2 system in clinical use, concluding that it was adequately damped with regard to its response to significant desaturation episodes [12]. However, that analysis lacked a comparison to routine manual control. As noted above, clusters of these episodes or re-
peated swings between hyperoxemia and hypoxemia might also be clinically relevant, and control system response could potentially even exaggerate them.

The aim of this current analysis was to identify the frequency of short swings of and between hypoxemic and hyperoxemic during both automated and manual FiO₂ control, and to determine if this potential problem was exacerbated or mitigated during automated control.

**Methods**

This was a prospectively defined analysis of existing SpO₂ data from randomized cross-over studies.

**Subjects**

Our database is populated with continuous 5-second SpO₂ data gathered from clinical trials evaluating the effectiveness of automated SpO₂ control systems. For this analysis we selected only subjects who were born extremely preterm, managed with an SpO₂ target with a midpoint of about 90% with a target range 4 wide. Further only subjects with one day each of auto and manual control were included. Finally subjects that did not spend most of their time requiring supplemental oxygen (at least 75% of time) were excluded to permit better characterization of hyperoxemic episodes. These criteria narrowed the selection to 58 cases from two published studies [13, 14]. The automated FiO₂ control systems in these studies were all AVEA-CLiO₂ (Vyaire, Mettawa, IL USA). The clinical trials all received ethics approval, and the patient information in the database is de-identified.

**Main Measures**

We prospectively defined analyses of three potentially clinically relevant patterns of SpO₂ extremes generally characterized as clusters, and swings. Swings were composed of two categories, oscillations and overshoot. We also supposed that clusters of short episodes were potentially comparable to continuous episodes of at least a minute, which are commonly reported in published studies.

Clusters were defined as three-minute periods in which 1 or more minutes were at an extreme SpO₂ (<80%, >98%) for at least 5 seconds, but excluding episodes of 1 minute in length or longer. We reported the hypoxemic and hyperoxemic clusters separately. Oscillations were defined as any 5-minute period during which, regardless of episode length, there was at least 1 minute <87% SpO₂ and also 1 minute >96% SpO₂, regardless of episode length. Lastly, we defined hyperoxemic overshoot as a 2-minute period, following an episode of <80% SpO₂ (of at least 5 seconds), in which 1 minute or more of the 2 minutes was >96% SpO₂ (see Figure 2 in Appendix).

**Analysis**

Extraction of the endpoints from the 5-second database was accomplished with purpose-built software. Because the resolution of the data was 5 seconds, one data point was defined as <5 seconds (i.e., it could be just instant, or nearly 5 seconds), two consecutive data points as 5 seconds, seven as 30 seconds, thirteen as 60 seconds, twenty-five as 2 minutes, thirty-seven as 3 minutes and sixty-one as 5 minutes. Differences among the episode length categories and modes of control were determined with the Wilcoxon signed-rank test for paired samples. Statistical tests were conducted with XLSTAT v19.03 software (Addinsoft, Paris, France).

**Results**

Two days of SpO₂ control of 58 preterm infants receiving mechanical ventilation and supplemental oxygen were evaluated. Ten were intubated. The cases came from seven neonatal ICU’s. The inspired oxygen was randomized with manual control (M-FiO₂) on one day and automatic control (A-FiO₂) on the other. The subject’s median age was 20 days (IQR 15–29) with a postmenstrual age of 28 weeks (IQR 21–30). The baseline oxygen needs of the subjects were relatively low (median FiO₂ 0.28, IQR 0.25–0.32). Most subjects spent no time on room air (39%), and the balance a median 4% of the time (IQR 1–9%). Figure 1 is a histogram of their SpO₂ of exposures.

![Fig. 1: SpO₂ Histogram with confidence limits. The points mark the mean proportion in each SpO₂ bin, the lines reflect the 95% bounds (CI of the proportion). The dotted lines are M-FiO₂, and solid lines A-FiO₂.](image)
methods was nearly identical [A-FiO2 91% (88–93%) and M-FiO2 91% (88–94%)].

There were frequent episodes of at least 5 seconds below normoxemia (<80% <87% SpO2) and above normoxemia (SpO2>95, >98%). Details are shown in Table 1. Episodes of extreme hypoxemia (SpO2<80%) were much more prevalent than episodes of extreme hyperoxemia (SpO2>98%). The frequency of hyperoxemic episodes was markedly reduced during automated control (p<0.001). In contrast episodes for SpO2 <87% were increased during automated control but decrease for episodes <80% SpO2 (p<0.001). Not shown in the Table is the frequency of episodes of 1 minute of longer. These make up about 10% of the total episodes during manual control and 1% of the episodes during automated control.

Table 1: Frequency of Hypoxic and Hyperoxic Episodes.

<table>
<thead>
<tr>
<th>Episodes (/day)</th>
<th>Auto</th>
<th>Manual</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpO2 &lt;80%</td>
<td>(115–175)</td>
<td>(175–352)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SpO2 &lt;87%</td>
<td>(307–577)</td>
<td>(210–423)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SpO2 &gt;96%</td>
<td>(105–404)</td>
<td>(318–758)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SpO2 &gt;98%</td>
<td>(8–63)</td>
<td>(34–147)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Episodes presented as median and (IQR).

The analysis of patterns of episodes of instability, our primary endpoint, is shown in Table 2. They are all markedly less frequent during A-FiO2 (p<0.001).

Table 2: Frequency of Episodic Instability.

<table>
<thead>
<tr>
<th>Episodes (/day)</th>
<th>Auto</th>
<th>Manual</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster (&lt;80%)</td>
<td>(1–7)</td>
<td>(18–29)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cluster (&gt;98%)</td>
<td>(0–4)</td>
<td>(8–29)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Oscillation</td>
<td>(3–24)</td>
<td>(11–47)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Overshoot</td>
<td>(0–8)</td>
<td>(2–20)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Endpoints defined as, Cluster: spending at least 1 minute of 3 minutes with SpO2 <80% or >98% excluding episodes of 1 minute or longer. Oscillation: spending at least 1 minute <87% and also 1 minute >96% in a 5-minute period. Overshoot: response to SpO2<80% resulting in SpO2>96% in more than 50% of the subsequent 2 minutes. Presented as median and (IQR).

The most prevalent were oscillations, which occurred about once an hour during M-FiO2 and every two hours during A-FiO2. A cluster occurred every two hours during M-FiO2 and every 5–6 hours during A-FiO2. Overshoot from hypoxemia to hyperoxemia was also less prevalent (A-FiO2 2 per day, M-FiO2 8 per day). It should be noted that these metrics are not additive; specifically, clusters and overshoot might well be included in episodes of oscillation.

Discussion

We characterized the frequency of complex episodes of SpO2 that might be clinically relevant in a population of extremely preterm infants receiving respiratory support and supplemental oxygen. These included clusters of shorter episodes of hypoxemia and hyperoxemia, as well swings between the two. With the exception of overshoot, this report is to our knowledge the first to characterize the prevalence of instability in SpO2 control that results in cluster of or swings between hypoxemia and hyperoxemia. We found that all these events were prevalent but markedly reduced during automated control.

We used two different levels of hypoxemia in our definitions of events (<80% and <87% SpO2). A large study of 972 preterm infants reported on the outcomes infants who spent an average of 3% of the time with SpO2<80%, comparable to our M-FiO2 population [2]. They found that bad long-term outcomes increased with increasing time with SpO2<80% and decreased as the exposure shortened. SpO2 levels less then 80% are often reported as the threshold for severe hypoxemia, and are associated with likely PaO2 levels <40mmHg [15]. SpO2 levels of 87% are often reported as the threshold for low SpO2 alarms for levels and are associated with PaO2 levels likely <50mmHg [16].

We found clusters of hypoxemia to be somewhat prevalent, though markedly less frequent during automated control. The interquartile range was 18–29 per day during manual control and 1–7 during automated control (p<0.001). Several other studies have reported the prevalence of prolonged episodes during manual and automated control [13, 17, 18]. These were 0–11 during automated control and 1–26 during manual control, rates comparable to the prevalence that we found for clusters of shorter hypoxic episodes. Our definition of clusters excluded prolonged episodes and thus these represent an independent metric, that might be appropriate to report when evaluating oxygen control methods.

We used two levels of hyperoxemia in our definitions of events (>96% and >98% SpO2). SpO2 levels of 99–100%, are often used as a marker of severe hyperoxemia when associated with supplemental oxygen. They are associated with PaO2 levels markedly higher than 80 mmHg in extremely preterm infants, with a risk of greater than 50% of being higher than 100 mmHg [15, 16]. A SpO2 of 97% would be considered by most as well above normoxemia, and is three above the
European consensus of 94% as the top of the desired range [19]. A SpO2 of 96-97% represents a risk of 25% that the PaO2 levels are higher than 80 mmHg [16].

Studies have reported prevalence of prolonged episodes of severe hyperoxemia during manual and automated control [13, 17]. The average interquartile ranges in these studies of automated and manual were 0–16 and 0–26 per day for hyperoxic episodes. These rates of these prolonged episodes of high SpO2 are similar to the prevalence we found for clusters of shorter episodes, (0–4 and 8–29). Automated control being significantly lower. Certainly, the prevalence of prolonged episodes of high SpO2 during manual control is a marker of nurse attentiveness, and to the extent that clusters are also clinically relevant they should be reported when comparing differences between and among manual and automated control.

Swings between hypoxemia and hyperoxemia have an additional likely clinical relevance. In fact, one theoretical concern about automated SpO2 control has been that making adjustments to FiO2 without observing the patient might result in excessive prevalence of instability, as compared to best manual practices. Studies of A-FiO2 have consistently shown reductions in hyperoxemia [8, 9], but that does not rule out swings. We evaluated two types, oscillation and overshoot. We found the prevalence of oscillations between hypoxemia and hyperoxemia (IQR 11–47 A-FiO2 and 3–24 M-FiO2 per day) lower for automated control. Oscillations were actually more prevalent that the rate of prolonged extreme episodes that are noted above. The difference is at least partially due to the SpO2 threshold differences (<80% vs >98% vs <87% vs >96%). We also found the prevalence of hyperoxic overshoot from hypoxemia also common, though significantly less frequent during automated control. That is, an interquartile range of 0–8 per day for automated control and 2–20 for manual control. While we are the first to report the prevalence of clusters, there are two reports of overshoot during manual and automated control and they are consistent with our findings, suggesting projectability our findings [12, 20]. Claure et al, reported an average of 12 and 7 per day for manual and automated control respectively. The other report only studied automated control in routine use with an interquartile range of 1–2 per day [12].

The clinical impact of extreme episodes is not well studied. Importantly it has been reported that increased time with SpO2 <80% is associated with an increase in late death or disability at 18 months [2]. Therapies that reduced these prolonged events could improve long-term outcomes, they speculated. One group reported that increased frequency of hypoxic events, regardless of length, was associated with severe retinopathy of prematurity and chronic lung disease at discharge [4, 5]. They also found that clusters of short hypoxic episodes had more impact than when they were evenly dispersed [6]. These findings may reflect a cause-effect relationship or rather be a marker of pathology with a corresponding adverse outcome. The impact of hyperoxemia has not been studied as thoroughly, but a landmark study published two decades ago demonstrated that high levels are associated with severe retinopathy of prematurity and chronic lung disease [21]. We speculate that these swings in hyperoxemia, whether from hypoxemia or normoxemia, would be, as they are in hypoxemia, potentially highly relevant.

The primary limitation of our report is that our metrics for these complex episodes were arbitrary. Certainly, it makes sense that clusters of short episodes, might have the same physiological impact as a single longer episode of the same total time, and that in addition reperfusion injury would likely be associated with swings between high and low SpO2 levels. Nevertheless, we might have selected definitions of duration or SpO2 cut-offs that are not related to outcomes or whose relative frequency or impact might change dramatically with different definitions.

There are other limitations of our study. First, it reflects a small population of infants, and while the endpoints were prospectively defined and while it is a multicenter experience, it is observational. There is another aspect of the study impacting the generalization of our work. Our study reflects experience with a lower SpO2 target range resulting in a median of 91%. A higher target range, if resulting in a higher median SpO2, would likely have a different profile. A study of two target ranges reported a shift with the higher target range that resulted in less hypoxemia and more hyperoxemia during both automated and manual control [13]. It has also been shown that manual control of SpO2 at higher target ranges is better than at lower targets because of reduced desaturations. All these issues of thresholds and target ranges should be further explored.

Finally, while it is encouraging that the automatic FiO2 control system that we evaluated performed better than manual control, these results only suggest that it is possible that all other automated control systems might, as well. Clinicians might watch the response of systems that they use carefully. Manufacturers should share their bench testing results and encourage evaluation of their system to insure they are mitigating, rather than exacerbating complex swings in oxygenation.

Besides the precise definition of swings and clusters, other factors need to be explored. It is also possible that these metrics might not be clinically relevant and are just markers of underlying pathophysiology. On the other hand, if these swings and clusters prove to be clinically relevant it is impractical to mitigated them with refined manual FiO2 control strategies. Thus, the incentive to adopt automated control would then have added benefits beyond improving time in the desired SpO2 range and labor savings [22]. These benefits, if marked, could be reflected in the current vary large randomized outcomes studies of automated control [23, 24]. It is also possible that these, as well as
prolonged episodes are adequately reflected in the current leading metric of exposure, percent time. Regardless, exploring these exacerbations ought to help engineers refine control algorithms.

Conclusions

Our study characterizes the prevalence of complex hypoxemic and hyperoxemic events. We found the clusters and oscillations to be quite prevalent. We demonstrated that the use of automated control results in reducing their prevalence. We believe this information should not only raise awareness of the potential risk from events currently not evaluated but also mitigate concerns that automated control might increase exposure to oxygenation instability.

Acknowledgements

The work was supported by grant SGS17/203/OHK4/3T/17 of the Czech Technical University in Prague.

References


[7] Sola A. Oxygen saturation in the newborn and the importance of avoiding hyperoxia induced damage. Neoreviews. 2015; 16(7):e393–405. DOI: 10.1542/neo.16-7-e393


[22] Poets CF, Franz AR. Automated FiO2 control: nice to have, or an essential addition to neonatal intensive care? Arch Dis Child Fetal Neonatal Ed. 2017 Jan;102(1):F5–6. DOI: 10.1136/archdischild-2016-311647
Fig. 2: Superimposed examples of overshoot, clusters and oscillation in a 5-minute period. Each dot on the chart represents a 5-second data point, the lines are added to reflect what would be seen on the oximeter display. The intent of the format is to contrast the 3 types of complex episodes evaluated.

OVERSHOOT: The gray line with open squares is an example of overshoot from hypoxemia. In the 120 seconds following a short severe desaturation (15 seconds) the SpO2 is greater than 96% for 65 seconds (longest episode 35 seconds).

CLUSTER: The black line with closed triangles is an example of a hypoxemic cluster. Three short episodes with SpO₂ <80% in an 180-second period. The three severe desaturations total 60 seconds (20 seconds each).

OSCILLATION: The black line with closed circles is an example of an oscillation above and below normoxemia (SpO₂ 87-96%). In the 300 seconds there are three episodes greater than 96% (longest 25 seconds) and 2 below 87% (the longest 35 seconds).