Detection of hyperoxaemia in neonates: data from three new pulse oximeters

B Bohnhorst, C S Peter, C F Poets

Aim: To determine the sensitivity and specificity of three newly developed pulse oximeters in the detection of hyperoxaemia, defined as an arterial partial pressure of oxygen (\(\text{PaO}_2\)) of more than 80 mmHg.

Methods: \(\text{SpO}_2\) readings from three oximeters (Agilent Virdia (AgV), Masimo SET (MaS), Nellcor Oxismart (NeO)) were documented in 56 infants (median gestational age at birth 35.5 weeks, range 24–41) whenever an arterial blood gas was taken for clinical purposes. Blood samples were analysed within one minute in a Radiometer ABL 505 blood gas analyser and OSM3 co-oximeter.

Results: Between 280 and 291 blood gases were analysed for each instrument; 105–112 showed a \(\text{PaO}_2\) > 80 mmHg. At an upper alarm limit of 95%, the three instruments detected hyperoxaemia with 93–95% sensitivity. Specificity at this alarm level ranged from 26 to 45%. The mean (SD) difference between arterial oxygen saturation and \(\text{SpO}_2\) (bias) was -0.25 (2.5)% for AgV, 0.06 (2.5)% for MaS, and -0.91 (2.6)% for NeO (p < 0.01; NeO v AgV and MaS).

Conclusion: These instruments detected hyperoxaemia with sufficient sensitivity at an upper alarm limit of 95%, but showed differences in their specificity, which was probably related to differences in measurement bias.
Denmark), and functional SaO, with a co-oximeter (Radiometer OSM 3). Co-oximeter measurements were corrected for fetal haemoglobin. SaO values were plotted against the difference between arterial and pulse oximeter saturation measurements (SaO − SpO), and the mean (bias) and standard deviations (precision) of these differences calculated for each instrument.

Sensitivity was calculated as the proportion of PaO readings ≥ 80 mm Hg associated with an SpO value below the threshold, divided by the number of all instances with a PaO ≤ 80 mm Hg. Statistical analysis was performed using the two sided t test. The study protocol was approved by the ethics committee of Hannover Medical School.

RESULTS

A total of 280 SpO/SaO/PaO determinations were performed for AgV, and 291 each for MaS and NeO; 105 (112 for AgV) in 27 (24) patients showed a PaO > 80 mm Hg (fig 1). A median of 10 measurements (range 2–15) was documented in each infant, with a median of three measurements (range 1–13) showing a PaO > 80 mm Hg. Table 1 shows sensitivity and specificity at various potential upper alarm limits. At an upper alarm limit of 95%, all three instruments detected 93–95% of hyperoxaemic episodes—that is, sensitivity to hyperoxaemia was comparable between instruments. Specificity at this threshold value, however, was more variable, ranging from 26% (NeO) to 45% (MaS).

The two highest PaO values that were > 80 mm Hg, but associated with an SpO < 95% (false negative for hyperoxaemia) were 144 and 92 mm Hg for AgV, 169 and 98 mm Hg for MaS, and 141 and 95 mm Hg for NeO. The lowest PaO with SpO > 95% (false positive for hyperoxaemia) was 46 mm Hg for AgV, 56 mm Hg for MaS, and 50 mm Hg for NeO (fig 1).

Measurement precision, defined as the standard deviation of the difference between SaO and SpO, was similar between instruments (2.5% for both AgV and MaS, 2.6% for NeO), while bias was smaller with the MaS (mean, −0.06%) and AgV (−0.25%) than with the NeO instrument (−0.91%, p < 0.01 for NeO v AgV and MaS).

DISCUSSION

The three pulse oximeters investigated in this study were chosen because they were shown to or purportedly have a relatively low false alarm rate, making them potentially interesting for use in neonates, in whom false alarms are a major problem. Avoidance of hyperoxaemia is particularly important in this age group. As previous studies have shown major differences in the upper alarm limit that has to be used if priority is given to the avoidance of hyperoxaemia, we wanted to know whether such differences still exist with current instruments. We found that all three instruments detected hyperoxaemia, defined as a PaO > 80 mm Hg, with sufficient sensitivity (93–95%) if the upper alarm limit was set at 95%.

Because we aimed to have a substantial proportion of hyperoxaemic values, we did not restrict our study to preterm infants, as did Bucher et al, but also included term infants. As SaO measurements were corrected for the proportion of fetal haemoglobin, and pulse oximeters are not influenced by fetal haemoglobin, we are confident that the decision to include with PaO ≥ 80 mm Hg. Specificity was calculated as the proportion of PaO readings < 80 mm Hg associated with an SpO value below the threshold, divided by the number of all instances with a PaO ≤ 80 mm Hg. Statistical analysis was performed using the two sided t test. The study protocol was approved by the ethics committee of Hannover Medical School.

RESULTS

A total of 280 SpO/SaO/PaO determinations were performed for AgV, and 291 each for MaS and NeO; 105 (112 for AgV) in 27 (24) patients showed a PaO > 80 mm Hg (fig 1). A median of 10 measurements (range 2–15) was documented in each infant, with a median of three measurements (range 1–13) showing a PaO > 80 mm Hg. Table 1 shows sensitivity and specificity at various potential upper alarm limits. At an upper alarm limit of 95%, all three instruments detected 93–95% of hyperoxaemic episodes—that is, sensitivity to hyperoxaemia was comparable between instruments. Specificity at this threshold value, however, was more variable, ranging from 26% (NeO) to 45% (MaS).

The two highest PaO values that were > 80 mm Hg, but associated with an SpO < 95% (false negative for hyperoxaemia) were 144 and 92 mm Hg for AgV, 169 and 98 mm Hg for MaS, and 141 and 95 mm Hg for NeO. The lowest PaO with SpO > 95% (false positive for hyperoxaemia) was 46 mm Hg for AgV, 56 mm Hg for MaS, and 50 mm Hg for NeO (fig 1).

Measurement precision, defined as the standard deviation of the difference between SaO and SpO, was similar between instruments (2.5% for both AgV and MaS, 2.6% for NeO), while bias was smaller with the MaS (mean, −0.06%) and AgV (−0.25%) than with the NeO instrument (−0.91%, p < 0.01 for NeO v AgV and MaS).

DISCUSSION

The three pulse oximeters investigated in this study were chosen because they were shown to or purportedly have a relatively low false alarm rate, making them potentially interesting for use in neonates, in whom false alarms are a major problem. Avoidance of hyperoxaemia is particularly important in this age group. As previous studies have shown major differences in the upper alarm limit that has to be used if priority is given to the avoidance of hyperoxaemia, we wanted to know whether such differences still exist with current instruments. We found that all three instruments detected hyperoxaemia, defined as a PaO > 80 mm Hg, with sufficient sensitivity (93–95%) if the upper alarm limit was set at 95%.

Because we aimed to have a substantial proportion of hyperoxaemic values, we did not restrict our study to preterm infants, as did Bucher et al, but also included term infants. As SaO measurements were corrected for the proportion of fetal haemoglobin, and pulse oximeters are not influenced by fetal haemoglobin, we are confident that the decision to include

<table>
<thead>
<tr>
<th>Upper alarm limit (%)</th>
<th>Agilent Viridia</th>
<th>Masimo SET</th>
<th>Nellcor Oxismart</th>
<th>Radiometer OSM3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity</td>
<td>Specificity</td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>92</td>
<td>0.99</td>
<td>0.12</td>
<td>1.00</td>
<td>0.16</td>
</tr>
<tr>
<td>93</td>
<td>0.99</td>
<td>0.18</td>
<td>0.99</td>
<td>0.21</td>
</tr>
<tr>
<td>94</td>
<td>0.96</td>
<td>0.21</td>
<td>0.98</td>
<td>0.31</td>
</tr>
<tr>
<td>95</td>
<td>0.93</td>
<td>0.30</td>
<td>0.94</td>
<td>0.45</td>
</tr>
<tr>
<td>96</td>
<td>0.86</td>
<td>0.43</td>
<td>0.83</td>
<td>0.54</td>
</tr>
<tr>
<td>97</td>
<td>0.72</td>
<td>0.55</td>
<td>0.68</td>
<td>0.72</td>
</tr>
<tr>
<td>98</td>
<td>0.48</td>
<td>0.69</td>
<td>0.36</td>
<td>0.82</td>
</tr>
<tr>
<td>99</td>
<td>0.30</td>
<td>0.81</td>
<td>0.23</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Figure 1 PaO plotted against SpO for Agilent Viridia (A), Masimo SET (B), and Nellcor Oxismart (B). The vertical lines indicate the threshold used to define hyperoxaemia, the horizontal ones the alarm limit at which the instruments detected at least 93% of hyperoxaemic episodes.

Table 1 Sensitivity and specificity at various potential upper alarm limits

www.archdischild.com
term infants did not affect our results. Also, the number of measurements varied between infants. This was because blood gases were only taken for clinical purposes, and arterial lines were removed at the discretion of the attending neonatologist. There was nothing to suggest, however, that this decision in any way biased our results.

The sensitivity of these pulse oximeters was within the same range as that of some conventional instruments, with only minimal differences between instruments. This is probably due to the fact that all instruments measure functional rather than fractional oxygen saturation, which is in contrast with an earlier study, which found that an instrument that measured fractional saturation had to be used with an upper alarm limit of 88% if hyperoxaemia were to be detected reliably. With regard to specificity, however, the MaS seemed to perform better than the other two instruments, which may be related to differences in measurement bias. Although these differences were small (< 1%), they may still be relevant, as small changes in SaO2 may be associated with changes in PaO2, in the hypoxic range.

As shown in table 1, both sensitivity and specificity depend on the alarm limit used. Sensitivity can be increased by decreasing the upper alarm limit, but the specificity, which is already low, will then decrease even further. This carries the risk of keeping infants hypoxaemic if priority is given to the avoidance of hyperoxaemia. This is because a wide range of Pao values may be associated with an SpO2 above the upper alarm limit (fig 1). In this study, the lowest Pao value measured with an SpO2 > 95%, however, was always above 45 mm Hg, which is somewhat reassuring.

Although our findings largely reflect blood physiology, i.e. the shape of the O2 dissociation curve, one might expect that the new generation instruments investigated here would have performed better than their first generation counterparts. Their major strength, however, is their improved handling of conditions with low signal to noise ratios, for which they apply frequency domain analysis (AgV), adaptive filtering (NeO), or a combination of frequency and time domain analysis and adaptive filtering (MaS). Such conditions, however, although not specifically excluded, occur comparatively rarely and are thus unlikely to have significantly affected our results.

A principal problem with a study validating medical instruments is that software algorithms are constantly updated by manufacturers. For example, Nellcor's Oxismart technology has already been replaced by Oxismart/XL, and Masimo had two major software upgrades since initiation of this study. All three manufacturers, however, assured us that, although handling of motion and/or low perfusion has changed with these software upgrades, the “lookup tables” used for deriving the displayed saturation values from the red/infrared ratios that are actually measured have not.

In conclusion, this study has shown that all three pulse oximeters investigated are similarly sensitive in the detection of hyperoxaemia, but differ somewhat in their specificity; probably reflecting differences in measurement bias. Nevertheless, despite these encouraging results, we do not recommend using pulse oximetry as the sole means of monitoring oxygenation in the neonatal intensive care unit. We must also warn against uncritically applying the definition used for hyperoxaemia in this study, although valid for preterm infants receiving additional inspired oxygen, to healthy neonates breathing room air at sea level, in whom both an Pao above 80 mm Hg and an SpO2 above 95% are part of normal infant physiology.

ACKNOWLEDGEMENTS

We are grateful to the medical and nursing staff of the neonatal intensive care unit at Hannover Medical School for their help with this study and to Agilent, Masimo, and Nellcor for providing the instruments investigated in this study.

Authors’ affiliations

B Bohnhorst, C S Peter, Department of Neonatology and Pediatric Pulmonology, Hannover Medical School, Hannover, Germany

C F Poets, Department of Neonatology, University Hospital, Tübingen, Germany

REFERENCES