Use of two oximeters to investigate a method of movement artefact rejection using photoplethysmographic signals

A. R. VISRAM, R. D. M. JONES, M. G. IRWIN AND J. BACON-SHONE

SUMMARY

Oxygen haemoglobin saturations and plethysmograph signal amplitudes were recorded from two oximeters placed on the fingers and toes of 10 patients undergoing oesophagectomy, to assess a method of removing motion artefact from saturation recordings. By examining changes in the plethysmograph amplitude that preceded changes in saturation, episodes of desaturation caused by movement artefacts were removed from the data. The reliability of the method was then determined by scrutinizing two concurrent oximetric profiles from each patient. A total of 1600 h of data were evaluated. Desaturations occurring contemporaneously in both oxygen saturation profiles were presumed genuine, whereas a desaturation occurring in only one of the profiles was classified as artefactual. Our method had a sensitivity of 96%, a positive predictive power of 98% and a specificity of 60%. We modified the method to increase specificity and re-evaluated our data. We found that a useful increase in specificity was associated with a considerable decline in sensitivity. (Br. J. Anaesth. 1994; 72: 388-392).

KEY WORDS


The computerized collection, display and analysis of oximetry data have facilitated interpretation of hypoxaemia in the postoperative period [1-3]. Retrospective analysis of oximetry data has been used to investigate the respiratory effects of various analgesic regimens [4], but in the awake patient, data are confounded by movement artefact which may account for 90% of recorded desaturation episodes [5]. Coupling the electrocardiograph and oximeter signals has been used to reject movement artefact but synchronization of the oximeter with the electrocardiograph artefact generated by movement has limited its success [6]. Averaging oxygen saturation data over longer intervals has also been used, but as a consequence, short periods of desaturation are missed [7]. We chose the change in pattern of the signal amplitude preceding an oximetric desaturation to identify and retrospectively reject movement artefact from the final data compilation [2].

The aim of this study was to determine the reliability of the signal artefact rejection template in distinguishing artefactual from genuine desaturation, using two pulse oximeters with recording devices.

PATIENTS AND METHODS

We studied 10 male Chinese patients, aged 54-87 yr, undergoing either a Lewis Tanner (n = 8) or a three-phase oesophagectomy (n = 2). The study was approved by the Faculty of Medicine Ethics Committee (University of Hong Kong) and written informed consent was obtained from all patients. The patients were premedicated with pethidine 1 mg kg⁻¹ and atropine 0.6 mg i.m. 1 h before operation. Before induction of anaesthesia, a lumbar extradural catheter was inserted at the L1-2 interspace. Anaesthesia was induced with fentanyl 1-2 μg kg⁻¹ and thiopentone 3-5 mg kg⁻¹ i.v. and maintained with 1-2% isoflurane and 60% nitrous oxide in oxygen. The patient was given atracurium 0.5 mg kg⁻¹ and a left-sided Robertshaw endobronchial tube was passed to facilitate one lung ventilation. Neuromuscular block was maintained with an infusion of atracurium 0.3 mg kg⁻¹ h⁻¹.

FIG. 1. The template used to identify artefactual desaturation data due to patient or probe movement. Reductions in SpO₂ were considered false if: (1) there were changes in signal amplitude strength of ± 2 SD (A) from the previous 60-s mean within 10 s before the reduction in SpO₂ (B); (2) the maximum rate of change of signal amplitude occurred within 2 s of commencement of signal artefact (C); and (3) the saturation signal stabilized to pre-artefact SpO₂ levels within 10 s of cessation of signal artefact (D).

A. R. VISRAM, B.SC., F.R.C.A., R. D. M. JONES, F.A.N.Z.C.A., F.R.C.A., Department of Anaesthesia and Intensive Care, The Royal Brisbane Hospital, Herston Road, Brisbane Q 4029, Australia. M. G. IRWIN, F.R.C.A. (Department of Anaesthesiology); J. BACON-SHONE, PH.D., M.A.C.Q.C, I.S. (Department of Statistics); University of Hong Kong, Hong Kong. Accepted for Publication: October 19, 1993.

Correspondence to A. R. V.
Routine monitoring included electrocardiography, direct arterial and central venous pressure monitoring, capnography and pulse oximetry. An extradural bolus of morphine 5 mg was given at induction followed by a 3-8 ml h\(^{-1}\) infusion of morphine 0.05 mg ml\(^{-1}\), which was continued for the next 24-36 h. Thereafter, analgesia was managed with pethidine i.m. on demand. After operation all patients were given supplementary oxygen therapy via a Hudson Multi-vent air entrainment mask (\(FIO_2\) = 0.3-0.4) and managed in the intensive care unit for the entire study period.

The collection of oximetry data commenced on the evening before surgery and was continued for 72 h after operation. A Nellcor N-200E pulse oximeter, adjusted to a 2-3-s averaging time, was attached via D-25 Oxisensor probe (Nellcor) to a finger on the hand contralateral to the arterial cannula and another oximeter was attached via a sensor to the patient's toe. The serial communication ports of the oximeters were connected to two 386DX laptop computers and the internal clocks of the computers were synchronized at the beginning of each study. Oxygen saturation, heart rate and signal amplitude data were recorded 60 times per minute by Satmaster (EMG Scientific, CA, U.S.A.). Data associated with zero amplitude were discarded automatically from compilation by the software. On completion of the study, data were archived for subsequent analysis.

Independently, each oxygen saturation profile from both oximeters was examined by the authors who identified every acute reduction in the oxygen saturation signal during the entire study period. For the purpose of this study an acute desaturation episode was defined as a decrease in oxygen saturation of more than 2% to less than 94% for more than 15-s duration. Each acute reduction in \(\text{SpO}_2\) was viewed with its concomitant signal amplitude recording. If the signal amplitude changed from baseline in a pattern that conformed to the template criteria (fig. 1), the episode was classified as a movement artefact, otherwise it was recorded as a genuine desaturation episode.

Concurrent desaturation episodes in the two oximeters were compared visually on two computer screens. Toe probe data were examined for a 2-min period after a finger probe desaturation episode, to take account of the delay in appearance of any concurrent toe probe desaturation. Paired desaturation episodes were categorized as true negative to describe a genuine desaturation and true positive to describe a genuine desaturation (fig. 2). For the purposes of our study, we defined sensitivity as (true negative/true negative + false negative), specificity as (true positive/true positive + false negative) and positive predictive power as (true negative/true negative + false negative) and these were calculated for the template. The template criteria were then modified to increase the specificity (table I) and the paired data were analysed.

The mean duration and severity of acute desaturations in those pairs classified as false negative and true positive were compared. The unpaired \(t\) test was used to test the null hypothesis with significance set at 5%.
TABLE II. Patient and clinical details (mean (range or SD) or number)

<table>
<thead>
<tr>
<th>n</th>
<th>Age (yr)</th>
<th>Weight (kg)</th>
<th>Preoperative $P_{aO_2} (F_{O_2} = 0.21)$ (kPa)</th>
<th>Duration of surgery (min)</th>
<th>Total dose of extradural morphine per patient (mg)</th>
<th>Total dose of pethidine l.m. per patient (mg)</th>
<th>Smokers (n)</th>
<th>Operation (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>64.7 (54-87)</td>
<td>54.8 (6.7)</td>
<td>11.2 (1.4)</td>
<td>310 (92.2)</td>
<td>20.7 (4.9)</td>
<td>124 (64.6)</td>
<td>4</td>
<td>Lewis Tanner 8</td>
</tr>
</tbody>
</table>

TABLE III. Specificity and sensitivity of the template after modification

<table>
<thead>
<tr>
<th>Modification</th>
<th>Original template (desaturation pairs)</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>True negatives</td>
<td>1221</td>
<td>1063</td>
<td>599</td>
<td>317</td>
</tr>
<tr>
<td>False positives</td>
<td>43</td>
<td>201</td>
<td>665</td>
<td>947</td>
</tr>
<tr>
<td>False negatives</td>
<td>15</td>
<td>40</td>
<td>36</td>
<td>31</td>
</tr>
<tr>
<td>True positives</td>
<td>23</td>
<td>59</td>
<td>75</td>
<td>112</td>
</tr>
<tr>
<td>Discarded data</td>
<td>269</td>
<td>208</td>
<td>196</td>
<td>126</td>
</tr>
<tr>
<td>Positive predictive power (%)</td>
<td>97</td>
<td>93</td>
<td>89</td>
<td>84</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>96</td>
<td>84</td>
<td>47</td>
<td>25</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>60</td>
<td>60</td>
<td>67</td>
<td>78</td>
</tr>
</tbody>
</table>

The incidence of severe desaturation ($S_{PO_2} < 85\%$) decreased by 96% from 1030 to 40 episodes in the finger probe and from 643 to 53 episodes (92%) in the toe probe.

Desaturation episodes which did not comply with our definition of an acute clinically significant desaturation episode (i.e. desaturation < 15-s duration with a reduction in $S_{PO_2}$ of < 2% from 94%) were discarded and this reduced the total number of desaturation episodes to 1878 (1244 from the finger and 634 from the toe probe): 1264 (67.3%) of these desaturations were not accompanied by any desaturation in the concurrent toe or finger recording and were therefore regarded as artefactual. Of the remaining 614 desaturations (307 pairs) occurring contemporaneously in both oximeters, 269 pairs were preceded by changes in signal amplitude that conformed to the movement artefact criteria and were regarded as movement artefacts. Analysis of template performance required that these data be set aside and the remaining 76 concurrent desaturations (38 pairs) were regarded as genuine. In 15 of these pairs, a desaturation in one probe was accompanied by signal amplitude changes that conformed to the movement artefact template criteria (a false negative), indicating a template specificity of 60%. Of the 1264 desaturations (true negatives and false positives) which were not accompanied by a saturation in the concurrent record, the template accurately identified 1221 as being caused by movement artefact (96% sensitivity). There was less than a 3% chance of a desaturation being identified as a movement artefact when it was actually a genuine desaturation (i.e. 97.6% positive predictive power).

The results of the three modifications of the template on specificity and sensitivity are shown in table III. When the signal amplitude strength change was expanded to 3 SD from the previous 60-s mean (modification I), sensitivity decreased by 12% without an increase in specificity. When the time criterion was changed so that the change in signal amplitude occurred between 5 and 15 s before the reduction in $S_{PO_2}$ (modification II), there was a 7%
increase in specificity but a 50% decrease in sensitivity. Application of modification III to the template resulted in an increase in template specificity to 75% but with a concomitant 71% decrease in sensitivity. Modification of both criteria within the template decreased the number of desaturation pairs that were discarded on the grounds of movement artefact affecting both probes by 56%. There was a five-fold increase in the number of desaturation pairs where both desaturations were regarded as genuine (true positives) while there was only a two-fold increase in the desaturation pairs where one of the desaturations was identified as being caused by movement artefact (false negatives).

Figure 4 shows the severity and duration of all 76 genuine desaturations. There was no difference in severity ($P = 0.86$) or duration ($P = 0.95$) between desaturations misidentified as movement artefact (false negatives) and those which were classified as genuine (true positives).

**DISCUSSION**

Pulse oximetry in a moving postoperative patient has severe limitations [8]. In 123 patients studied in the postoperative recovery area, it was found that 75% of alarms caused by low oxygen saturation readings were "trivial" and attributed to movement artefact or poor signal quality [9]. Wilson showed that 87–90% of desaturation episodes in children sedated for dental day surgery were due to movement artefact [5]. We have confirmed these findings by demonstrating that of 1878 desaturations seen, 1264 (67%) were evident in only one of the two concurrent oximetric recordings and would therefore appear to be artefactual. We have used two oximeters to verify independently artefactual desaturation as other workers have tried with limited success to filter electronically the oxygen saturation signal [10].

Synchronization of the R wave of the ECG with the oximeter waveform has been shown to reduce the number of motion artefacts by 50% (from 4.1 to 2.1%) when used in the neonatal intensive care unit [11]. This method has not been assessed for its reliability; furthermore it has been suggested that the oximeter may synchronize with the ECG artefact generated by movement [6]. Computerized oxygen saturation monitoring systems have used the alarm status of the oximeter and also oxygen saturation to invalidate data when the alarm is activated [3]. However, the oximeter alarm does not distinguish a genuine desaturation from one produced by movement. Unfortunately, the Nellcor N200 oximeter does not incorporate a plethysmographic waveform display which would assist in distinguishing artefactual from genuine data, although in any case this interpretation is beyond the present signal processing capabilities of the software [10]. Interestingly, there were fewer episodic desaturations sensed by the toe probe compared with the finger probe, both for the raw and evaluated data. This suggests that the finger probe is either more vulnerable to artefacts of other types or our template has been unsuccessful in isolating all movement artefact in the hand. However, for long-term oxygen saturation data collection in the absence of peripheral vascular disease, the foot may provide a more reliable location for probe attachment.

Our computerized data compilation registers every desaturation below a set level and time limit and therefore grossly overestimates desaturation incidence; 50% of the computer compiled episodic desaturations lasted less than 15 s and even a transient dip in oxygen saturation below a preset value is counted as a desaturation. Confirming our previous findings, we found a 33–44% reduction in the number of desaturations compiled by computer when our template was applied manually to the data [2]. The most dramatic overestimation of desaturation episodes occurred in the severe desaturation group ($S_{O2} < 85\%$) and reiterates the misleading inferences which may be drawn from unevaluated computerized compilation of oximetric data. Beydon and colleagues used a system where rapidly changing oxygen values were given less weight in determining the mean oxygen saturation and if there was a decrease in oxygen saturation of more than 25% from one 5-s value to another, it was considered artefactual [12]. Increasing the time over which oxygen saturation data are averaged reduces the number of episodic desaturations detected and differential weighting may not distinguish a sudden genuine desaturation from movement artefact in a postoperative patient.

To increase the objectivity of our template application to oxygen saturation data we used the
concurrent recordings of two oximeters attached to the patient. The template was constructed after observations on postoperative patients and volunteers showing that characteristic large variations in signal amplitude occurred consistently during oximeter probe movement and that these heralded false reductions in oxygen saturation [2]. A reduction in oxygen saturation without signal amplitude instability occurring concurrently in the two oximeters is likely to be genuine. We found that the template was sensitive, identifying 96% of all desaturations unaccompanied by a concomitant desaturation in the other probe and the positive predictive power was such that less than 3% of the false desaturations were identified incorrectly; although this may reflect the high prevalence of artefactual desaturations. Concurrent desaturations recorded in the two oximeters may be genuine or may be caused by patient movement affecting both probes simultaneously. Nearly one-third of all evaluated desaturations had a concurrent change in signal amplitude pattern that suggested that both oximeter probes were affected by movement and for the determination of template specificity, these data were discarded. The resulting specificity of 60% implies that 40% of desaturations which were genuine could be misidentified by the template as being caused by movement artefact. Postoperative oxygen therapy resulted in a small genuine desaturation sample size and this may have distorted our determination of specificity. A larger pool of genuine desaturations may have increased our ratio of true positives to false negatives and therefore increased specificity. However, the template specificity was similar, irrespective of whether it was applied to severe or mild desaturation episodes (fig. 4). We felt this was not acceptable clinically. We modified the template so that more vigorous movement was required to produce the required change in the signal amplitude and any coincidental minor movement would be less likely to be interpreted as a false desaturation. In addition, we increased the time separation between changes in signal amplitude and reduction in $SpO_2$ to distinguish between hypoxaemic arousal causing patient movement and artefactual desaturation subsequent to patient movement. Neither modification greatly increased specificity and both reduced sensitivity. When both modifications were tested on our data, specificity increased marginally but sensitivity declined to unacceptable levels. Unfortunately, this method of evaluation does not allow independent assessment of specificity and sensitivity. The modifications introduced to increase specificity result in more of the discarded data being considered genuine and therefore the increase in specificity may only reflect decreasing sensitivity of the template. This also explains the large increase in true positive desaturation pairs with each modification of the template and the relatively small increase in false negatives.

Prolonged oxygen saturation data collection may be an important tool in assessing the benefit or otherwise of various therapeutic regimens but the data must be genuine before inferences are drawn. Although this template lacks some specificity, it does provide high sensitivity and positive predictive power when filtering computerized data. Oxygen saturation data which are evaluated against a plethysmographic waveform template may improve specificity.

ACKNOWLEDGEMENT
We thank Mr Andy Lai, senior technician in the Department of Anaesthesiology, University of Hong Kong for his expert assistance.

REFERENCES