Studies of New Generation Pulse Oximetry Technologies

Medical Systems
Philips FAST SpO\textsubscript{2} (Fourier Artifact Suppression Technology) is compatible with the widest array of sensors in the industry. It uses a patented frequency analysis algorithm to filter out noise in the sensor signal, thereby overcoming many of the issues associated with traditional pulse oximetry, such as sensitivity to patient movement or intense ambient light.

Many studies have been conducted comparing the various SpO\textsubscript{2} technologies. This guide includes summaries of some of those studies and is meant to help you better understand how Philips FAST SpO\textsubscript{2} performs against other technologies in the marketplace and why Philips FAST SpO\textsubscript{2} is still a smart choice.
Table of Contents

Integrated Review of SpO2 Studies

New Generation Pulse Oximetry in the Care of Critically Ill Patients; 7
*American Journal of Critical Care*, January 2005, Volume 14, No. 1

References 17

Study Summaries

ADULT

New-Generation Pulse Oximetry in the Care of Critically Ill Patients 5

How Reliable are the New Generation Pulse Oximeters in the Detection of Hypoxemia in the Intensive Care Unit? 19

The Clinical Use of Two New Generation SpO2 Devices for Critically Ill Patients with Sepsis 20

A Comparison of the Accuracy of the Philips FAST and Masimo SET Pulse Oximeter Following Cardiac Surgery 21

A Comparison of Three New Generation SpO2 Devices During Ambulation After Open Heart Surgery 22

“Motion-Resistant” Pulse Oximetry: A Comparison of New and Old Models 23

NEONATE

The Clinical Use of Two New Generation SpO2 Devices for Patients in the Neonatal Intensive Care Unit (NICU) 24
The objective of this article was to review the published, peer-reviewed studies that have been done on the new-generation pulse oximeters, which are manufactured with algorithms to filter out patients’ body motions. In addition, this article also describes the application of new-generation pulse oximetry in clinical practice in critical care.

A total of 17 articles were found in MEDLINE from 1995 to 2003. Each article was examined for scientific merit, content, and applicability to clinical practice.

A consistent finding among the studies that were reviewed was the superior performance with regard to reduction in frequency of false alarms and overall accuracy of all the new motion-tolerant devices compared with the performance of the various conventional pulse oximeters. However, more clinical evidence on the performance of new-generation pulse oximetry devices is clearly needed, particularly in the setting of critical care. Although some conclusions can be made, a lack of consistency among study variables and techniques and discrepancies between the results of the studies reviewed, make it difficult to make any overall recommendations about which of the new-generation motion-tolerant devices are best suited for use in critical care in any of the populations of patients studied. Clearly, more clinical trials are needed before the most accurate and reliable motion-tolerant pulse oximetry device can be determined for use in populations of critically ill patients. These trials should address the performance of the algorithms used for obtaining SpO₂ values in critically ill patients of all ages in situations in which patients’ movements are a factor. In addition, each trial should clearly specify which software revisions are being used with each device and should include a comparison with a known gold standard (preferably co-oximetry) for measurement of oxygen saturation. Additionally, the variables of peripheral temperature, finger thickness, hemoglobin concentration, and skin color should be part of every future study on measurement of oxygen saturation by pulse oximetry devices to further assess the significance of these variables on clinical measurement of SpO₂.

Complete article follows.
New-Generation Pulse Oximetry in the Care of Critically Ill Patients

By Karen K. Giuliano, RN, PhD, and Thomas L. Higgins, MD, MBA. From the Boston College William F. Connell School of Nursing, Chestnut Hill, Mass (KKG), Philips Medical Systems, Andover, Mass (KKG), the Adult Critical Care Division, Baystate Medical Center, Springfield, Mass (TLH), and Tufts University School of Medicine, Boston, Mass (TLH).

Objective: To review the published, peer-reviewed studies to date on use of the new-generation pulse oximeters, which are manufactured with algorithms to filter out patients' body motions, and describe the application of new-generation pulse oximetry in clinical practice in critical care.

Methods: MEDLINE was used to locate appropriate articles on pulse oximetry for the years 1995 to 2003. Each article was examined for scientific merit, content, and applicability to clinical practice.

Results: A total of 17 relevant articles on the clinical performance of the new-generation pulse oximeters were reviewed, and the data were organized into a table.

Conclusions: The combination of studies done in both clinical and laboratory settings did not provide any strong and convincing evidence that the performance of any single new-generation device was superior to that of any other new-generation device. However, the clinical performance of all the new generation pulse oximetry devices was better than that of earlier devices. (American Journal of Critical Care, 2005; 14:26-39)
Motion artifact occurs either when a patient’s movements cause the pulse oximeter to incorrectly interpret the movements as a pulse signal or when the motion artifact prevents accurate detection of the patient’s true pulse signal. The resultant increase in false alarms and erroneous measurements can desensitize clinicians to the alarms and increase the chance of missing a clinically important true alarm.6-8

Patients’ motion decreases accuracy of conventional pulse oximeters and has led to the development of “motion tolerant” oximeters.

Manufacturers of pulse oximeters have sought to reduce motion artifact in the pulse signal through improvements in the algorithms. “Motion tolerant” pulse oximetry devices are now commercially available from several different manufacturers, and as a group are referred to as “new generation” pulse oximeters. To date, peer-reviewed scientific publications on evaluations of these new-generation, motion-tolerant devices are few and are primarily limited to 3 of the devices currently on the market: FAST SpO2 (Fourier artifact suppression technology SpO2; Philips Medical Systems, Andover, Mass; first marketed in 1999), SET (signal extraction technology; Masimo Corp, Irvine, California; first marketed in 1998), and Oxismart (Nellcor; Pleasanton, California; first marketed in 1994).

Review of the Literature

MEDLINE was used to identify research studies on the new-generation motion-tolerant pulse oximetry products published in English during the period 1995 through 2004. The search terms “oxygen saturation,” “pulse oximetry,” “SpO2,” and “oxygen measurement” were used. Even though many articles on pulse oximetry in general have been published, the focus of most of the articles was conventional pulse oximetry. In this state-of-the-science review, we focus solely on data-based studies on the performance of the new-generation pulse oximetry devices during patients’ movements. Although we also reviewed some data on the performance of the devices in patients with low perfusion, that aspect was not our focus. We found 17 research studies in which investigators specifically examined the performance of new-generation devices; these studies are listed and described in the Table.

Several problems in the published studies made interpretation and/or generalization of the findings difficult. First, various software revisions of each manufacturer’s algorithms for measuring oxygen saturation were used, often making the comparisons somewhat of an apples-and-oranges approach. Second, even when pulse oximetry devices with the same brand name were used, most authors did not fully disclose the actual software revision or signal averaging times used during the testing, making meaningful comparisons between studies even more difficult. Third, studies were characterized by different evaluation criteria, limited evaluation criteria, inadequate description of how device performance was evaluated, or poor study design. Last, Dumas et al.26 evaluated a prototype that was never released commercially; that study, therefore, was excluded from this review. Comparisons, findings, strengths, and weaknesses of the reviewed studies are outlined in the Table.

Laboratory Studies

The performance of new-generation pulse oximetry devices was evaluated in a laboratory setting in 5 studies.9,14,16,17,19 The advantages of doing any kind of clinical study in a laboratory rather than a real clinical setting are generally lower cost and more control over the variables. The major problem, however, is the difficulty of replicating the characteristics of a clinical setting, especially when the clinical setting is as complex as critical care. Additionally, in 4 of the 5 studies, the subjects were healthy volunteers, whose clinical conditions and oxygen saturation measured by pulse oximetry (SpO2) often bear little resemblance to those of critically ill patients.9,14,16,19

One approach usually referred to as the “steady-state reference hand” was used in 2 of the studies.9,17 With this approach, the subject keeps one hand still and the other hand is strapped to a motion arm that is programmed to move in various patterns. The SpO2 readings of the moving hand are then compared with the readings of the nonmoving reference hand. The disadvantage of this approach is that passive motion is used, a kind of motion that has little relevance to clinicians because these simulated passive motions do not mimic the active motion most common in patients in actual clinical settings.27 Therefore, the use of this type of testing is a significant flaw when used in studies to compare the performance of pulse oximetry devices, a flaw that has been recognized by other researchers.16
### Summary of research findings, strengths, weaknesses, and limitations of studies on motion-tolerant pulse oximetry devices

<table>
<thead>
<tr>
<th>STUDY</th>
<th>DEVICES</th>
<th>PATIENTS</th>
</tr>
</thead>
</table>
| Barker and Shah              | • Masimo SET (MT): experimental prototype  
• Nellcor N-3000 (MT) with variable signal averaging time  
• Nellcor N-200 (CPO)                                                            | 10 healthy adult volunteers in a laboratory setting                     |
| Bohnhorst et al             | • Masimo SET (MT) in 8-second signal averaging mode  
• Nellcor N-3000 (MT) with variable signal averaging time  
• Nellcor N-200 (CPO) in 6- to 7-second signal-averaging mode                  | 17 patients in a NICU                                                    |
| Brouillette et al           | Part I  
• Masimo SET (MT) in 4-second signal averaging mode  
• Nellcor N-200 (CPO) in 2- to 3-second signal-averaging mode  
• Transcutaneous oxygen probe (no model given)  
Part II  
• Masimo Radical in 2-second signal averaging mode (MT)  
• Nellcor N-200 (CPO)  
• Nellcor N-395 in 2- to 3-second signal-averaging mode  
• Transcutaneous oxygen probe (no model given)  
• Nellcor N-3000 (MT, but older than N-395)  
• Nellcor N-395 (MT)  
• Datex-Ohmeda 3900P (MT)  
• Nellcor N-3000 (MT)  
• Philips FAST SpO2 (Virida 24C; MT) | 24 children referred to a sleep apnea laboratory because of sleep disordered breathing  
22 children referred to a sleep apnea laboratory because of sleep disordered breathing |
| Durbin and Rostov           | Part I  
• Masimo SET (MT)  
• Arterial blood gas sample  
• Nellcor N-200 (CPO)  
• Nellcor N-395 (MT)  
• Datex-Ohmeda 3900P (MT)  
• Nellcor N-3000 (MT)  
• Philips FAST (MT) | 13 patients with thoracic and cardiovascular conditions in an ICU |
| Durbin and Rostow           | Part I  
• Masimo SET (MT)  
• Nellcor N-200 (CPO)  
• Nellcor N-395 (MT)  
• Datex-Ohmeda 3900P (MT)  
• Nellcor N-3000 (MT)  
• Philips FAST (MT) | 13 patients with thoracic and cardiovascular conditions in an ICU |
| Gehring et al               | • Philips FAST SpO2 (revision B.0; MT)  
• Masimo SET (MT)  
• Nellcor N-395 (MT)  
• Datex-Ohmeda 3900P (MT)  
• Nellcor N-3000 (MT, but older than N-395)  
• Masimo SET (MT)  
• Datex-Ohmeda 3900P (MT)  
• Nellcor N-3000 (MT)  
• Datex-Ohmeda 3900P (MT)  
• Philips FAST SpO2 (Virida 24C; MT) | 10 healthy volunteers in a laboratory setting |
| Hay et al                    | • Masimo SET (MT)  
• Nellcor N-395 (MT)  
• MKPspO2 (MT)  
• Philips FAST SpO2 (Virida 24C; MT)  
• Nellcor N-200 (CPO)  
• Datex-Ohmeda 3900P (MT)  
• Philips CMS (revised A.0; CPO) | 26 infants in a NICU |
| Jopling et al                | • Nellcor N-395 (MT)  
• Masimo SET (MT)  
• NPB-290 | 8 subjects (the characteristics of subjects were not described) |
| Kastle and Konecn            | • Philips M3 (revision B) FAST SpO2 (MT)  
• Ivy 2000 (revision 2.2) with Masimo SET (MT)  
• Nellcor N-3000 (revised 3.03; MT)  
• Philips CMS (revised A.0; CPO) | Output signals derived from patients were used in a laboratory benchmark study |
| Kopotic and Linder           | Part I  
• Ohmeda 2000 with Masimo SET (MT)  
• Nellcor N-295 with Oxismart (MT)  
• Philips FAST (MT)  
• Nellcor N-200 (CPO)  
• Datex-Ohmeda 3900P (MT)  
• Philips CMS (revised A.0; CPO)  
Part II  
• Radical with Masimo SET (MT) | 50 newborns at high risk for respiratory failure  
15 neonates <30 weeks’ gestation |
| Lie et al                    | • Nellcor Symphony N-3000 (MT)  
• Nellcor N-3000 (MT)  
• Philips FAST (MT) | 26 healthy adult volunteers in a simulated PACU |
| Lutter et al                 | • Masimo SET (MT)  
• Nellcor N-3000 (MT)  
• Philips FAST (MT) | 108 patients in a PACU  
118 patients in an ICU  
21 patients treated with IABP |
| Malviya et al                | • Masimo SET (MT)  
• Nellcor N0020 (CPO) | 75 healthy patients in a pediatric PACU |
| Poets et al                  | • Nellcor N-100 and N- 200 (CPO)  
• Ohmeda 3700 (CPO)  
• Nellcor N-3000 (MT)  
• Philips (CPO)  
• Masimo SET (MT) | A total of 267 patients in an NICU |
| Rheineck-Leyssius and Kalkman | • Nellcor Symphony N-3000 with Oxismart (MT)  
• Criticare 504 (CPO) with signal-averaging time at 21 seconds  
• Criticare 504 (CPO) with signal-averaging time at 3 seconds | 53 operating room patients |
| Rheineck-Leyssius and Kalkman | • Nellcor Symphony N-3000 with Oxismart (MT)  
• Criticare 504 (CPO) with signal-averaging time at 21 seconds  
• Criticare 504 (CPO) with signal-averaging time at 3 seconds | 603 patients in a PACU |
| Wouters et al                | • Philips FAST on CMS  
• Nellcor N-3000 (MT)  
• Arterial blood gas analysis | 1483 perioperative patients at 4 sites |

Abbreviations: CPO, conventional pulse oximeter; IABP, intra-aortic balloon pumping; ICU, intensive care unit; MT, motion tolerant; NICU, neonatal intensive care unit; PACU, postanesthesia care unit; SaO\textsubscript{2}, arterial oxygen saturation; SpO\textsubscript{2}, oxygen saturation measured by pulse oximetry.
Approaches in which active motion is used to test pulse oximetry devices have also been described.14,16,17 Both Jopling et al16 and Gehring et al14 used healthy volunteers but devised active-motion protocols of tapping and scratching motions with a wide variety of amplitudes and velocities. A noteworthy weakness of these studies is that both had small sample sizes (Jopling et al, N=8; Gehring et al, N=10). In addition to motion, the devices in these studies were tested during conditions of hypoxemia and low perfusion. Jopling et al induced hypoxemia in the healthy subjects and thus obtained a range of oxygen saturations of 70% to 100% during the active-motion testing, which provided a much closer approximation to the range of SpO2 values that might occur in a real clinical setting. Gehring et al simulated low perfusion by compressing the brachial artery. The overall findings for the detection of hypoxemia of both of these studies indicated that the performance of the new-generation devices was better than that of the conventional devices. When the new-generation devices were compared with one another, differences in the detection of hypoxemia were not significant.

### FINDINGS

<table>
<thead>
<tr>
<th>Device</th>
<th>Performance</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masimo SET</td>
<td>Best overall performance, 99% accuracy, 100% positive predictive value</td>
<td>Only passive motion tested. The hand with the Masimo SET device was used as the reference hand. The non-motion hand was sometimes moving. Signal-averaging time for the N-200 was not specified.</td>
</tr>
<tr>
<td>N-200</td>
<td>Greatest number of errors in SpO2, 76% accuracy, 73% positive predictive value</td>
<td>No data were reported on overall frequencies of false alarms. No gold standard reference was used.</td>
</tr>
<tr>
<td>N-3000</td>
<td>Lowest dropout rates, 87% accuracy, 81% positive predictive value</td>
<td>Parts I and II. Signal-averaging settings given in the abstract do not agree with those in the text. No gold standard reference was used to verify oxygen saturations. Data on the accuracy of transcutaneous oxygen monitoring were limited but theoretically are of concern because of increased skin thickness in patients other than neonates.</td>
</tr>
</tbody>
</table>

### COMMENTS

- **Part I**
  - Simultaneous readings from an MT device and a CPO were compared during 20 randomly selected desaturation events of ≥4%.
  - With the N-200, 88% of 220 desaturations were detected during wakefulness, and 38% of 194 desaturation events during sleep were classified as motion artifact.
  - With the Masimo Q 400 and the transcutaneous oxygen probe, no desaturation events were detected.
  - A total of 119 desaturation episodes during sleep were detected by all 3 devices.

- **Part II**
  - Newer version MT was compared with CPO for detection of true desaturations.
  - Masimo Radical device detected significantly more nonartifactual desaturation events during sleep than the N-200 did.

- **Performance**
  - Performance of an MT device was evaluated in patients whose CPO did not acquire a reliable signal.
  - Masimo SET was able to obtain a reliable signal in 12 of 13 patients; the patient in whom the signal could not be detected had hypotension, poor peripheral perfusion, and shivering and was being treated with intra-aortic balloon pumping.
  - The difference between direct measurement of SaO2 and the Masimo SET value was 1.1% (±1%).
  - Two thirds of saturation values were >94%.

- **Researchers**
  - Researchers measured amount of data dropout, artifact, or when SpO2 readings differed from arterial blood gas results by more than 10%.
  - With patients randomly assigned to groups, the Masimo SET was significantly more reliable than was the CPO.
  - Patients with the Masimo SET were weaned to a fraction of inspired oxygen of 0.40 more quickly and had fewer arterial blood gas analyses than did the patients with the CPO.
  - The 2 groups did not differ in time to extubation.

**COMMENTS**

- The results may not be applicable to other populations of patients. Only 59 of 86 patients had enough complete data for analysis. Clinicians were not blinded to which oximeter was being used. No gold standard was included for comparison of oximeter readings. Pulse Oximetry with new generation devices was compared with other technologies. Signal-averaging times were not specified.
In a lab setting, all new-generation pulse oximeters outperformed conventional devices in their ability to tolerate motion.

Kastle and Konecny\textsuperscript{17} used an approach referred to as noise-mix-composition. In this approach, 136 different combinations of raw patient-generated signals from the operating room, the intensive care unit, and the PACU were used. The range of oxygen saturation in the signals was 39\% to 100\% (mean 89.9\%; SD 10.7\%), and the range of pulse rates was 49/min to 201/min (mean 108.2/min; SD 45.9/min). These signals were then fed digitally into each of the pulse oximeters being studied. The advantage of this approach is that real, not simulated, signals were used. In addition, a wide range of both oxygen saturations and pulse rates were used to test the pulse oximeters. This range of values more closely approximated the types of clinical phenomena encountered in the critical care setting.

In general, the overall results of the 5 laboratory studies were consistent. In studies in which passive motion was used, the Masimo SET device tended to have the best performance, whereas in studies in which active motion was used, the Nellcor and Philips devices tended to have the best performance. For example, in the study by Kastle and Konecny,\textsuperscript{17} the Masimo device had the best performance (2.6-fold improvement over conventional pulse oximeters compared with 1.6-fold improvement for both Philips and Nellcor) during the steady-state reference hand, passive-motion part of the testing. In contrast, during the noise-mix-composition, active-motion part of the study, the Philips device had the best performance (1.6 fold improvement compared with 1.5-fold for the Masimo device and 1.3-fold for the Nellcor device). However, none of these differences were clinically significant.

In 4 of the 5 laboratory studies, conventional devices were used as one means of comparison.\textsuperscript{9,14,17,19} A consistent and important finding of these 4 studies was that regardless of the differences in measurement criteria, the performance of all the new-generation pulse oximetry devices was better than the performance any of the conventional devices to which they were compared.
<table>
<thead>
<tr>
<th>FINDINGS</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researchers used the N-3000 as a reference oximeter and looked at errors in both pulse rate and SpO2&lt;br&gt;During desaturation events, all the devices were within 3% of the reference reading 95% of the time&lt;br&gt;With either low perfusion or motion, all the devices were within 3% of the reference reading 75%-85% of the time&lt;br&gt;During both low perfusion and motion, all the devices were within 3% of the reference reading less than 62% of the time&lt;br&gt;The Masimo SET, Philips FAST SpO2, and N-395 devices did not differ significantly in overall performance&lt;br&gt;Large errors (significant) occurred only between the N-3000 and the N-395 during both motion and low perfusion</td>
<td>Compression of the brachial artery was used to simulate low perfusion, which is probably not representative of low perfusion in the real clinical setting&lt;br&gt;The sample size was small&lt;br&gt;The N-3000 was used as the reference oximeter for all the devices</td>
</tr>
<tr>
<td>Part I&lt;br&gt;The Masimo SET was compared with the N-200 CPO&lt;br&gt;The Masimo was better than the N-200 for false hypoxemias, false bradycardias, total false alarms, data dropout, and true bradycardias</td>
<td>The software revision for the Philips device was not given, so it is uncertain whether the device is really an MT device or a CPO&lt;br&gt;No gold standard was used to assess the accuracy of the SpO2 values for any of the devices</td>
</tr>
<tr>
<td>Part II&lt;br&gt;The 3 new-generation devices were compared with one another&lt;br&gt;The Masimo SET was better than the other new-generation devices for false hypoxemias, desaturations, and data dropout&lt;br&gt;The performances of the Masimo SET, the N-395, and the Philips FAST SpO2 were better than the performance of the MARSpO2 for false bradycardias</td>
<td>The methods used and results lacked clarity and sufficient detail for adequate evaluation of the study&lt;br&gt;The laboratory situation may not adequately reflect motion in clinical practice&lt;br&gt;The sample size was small</td>
</tr>
<tr>
<td>Different testing methods (active and passive [machine generated] motion) were compared in subjects with various levels of hypoxemia-induced states (SaO2, 70%-100%) during a 2- to 3-minute period&lt;br&gt;The performances of the N-395 and the Masimo SET were better than the performance of the N-290 in both active- and passive-movement tests</td>
<td>Steady-state reference hand part&lt;br&gt;The performance of each of the 3 new-generation devices was better than that of the CPO: 1.6-fold improvement over the CPO for the Masimo SET and 1.5-fold for the Masimo SET, and 1.3-fold for the Nellcor</td>
</tr>
<tr>
<td>Noise-mix-composition part&lt;br&gt;The performance of each of the 3 new-generation devices was better than that of the CPO: 1.6-fold improvement over the CPO for the Philips FAST SpO2</td>
<td>The charts and graphs in the article supply a level of detail that is sometimes difficult to follow</td>
</tr>
<tr>
<td>Steady-state reference hand part&lt;br&gt;The performance of each of the 3 new-generation devices was better than that of the CPO: 1.6-fold improvement over the CPO for the Philips FAST SpO2, 1.5-fold for the Masimo SET, and 1.3-fold for the Nellcor</td>
<td></td>
</tr>
</tbody>
</table>
Clinical Studies in Adults

Real clinical environments and adult patients were used to assess the performance of new-generation pulse oximeters in 6 studies.1,2,12,20,23,25 The new generation devices included the Masimo SET, Nellcor N-3000, and Philips FAST SpO₂. The clinical areas included the operating room,1,12 the PACU,20,24,25 and the critical care unit.12,13,20 Because patients’ movements should not be a factor in the operating room, we considered studies done in the operating room separately from those done in the PACU or critical care unit.

The new-generation devices used in the operating room included the Nellcor N-3000 and the Philips FAST SpO₂. In the first study, Rheineck-Leyssius and of the N-3000 with a 5- to 7-second signal averaging time and the performance of a conventional pulse oximetry device with a 21-second signal-averaging time.

None of the new-generation oximeters differed from the other in clinical performance.

Signal-averaging time for pulse oximeters represents the amount of time (in seconds) used by the device to calculate the displayed SpO₂ value. The shorter the signal-averaging time, the more sensitive the device will be to changes in the SpO₂ value. However, the device will also be more prone to false alarms due to artifact or error. Conversely, as the signal-averaging time is increased, the number of alarms due to error or artifact is decreased, but the amount of time needed to detect a true hypoxemia is concomitantly increased, because the SpO₂ values are averaged over a longer period.

Hence, the results of the study23 are not surprising because patients’ movements should not be a factor in the operating room. Furthermore, a 21-second signal averaging time on the conventional comparison device would obscure most of the signal artifact because the signal is averaged over such a long period. In contrast, in most of the new-generation devices, signal-averaging times are less than 10 seconds. This decrease in time enhances detection of hypoxemia, and, at the same time, the improved algorithms are better able to filter out motion artifacts due to patients’ movements, thus helping to control for the increase in false alarms that accompany decreased signal-averaging times with conventional pulse oximetry.

Factors that influenced oximeters’ degree of error were peripheral temperature, finger thickness, hemoglobin concentration, and skin color.

In the second study done in the operating room, Wouters et al25 used a total of 2694 SpO₂ recordings in 1483 patients and compared the measurement accuracy of the Philips FAST SpO₂ and the Nellcor N-3000. Values obtained via co-oximetry were used as the reference SpO₂ measurements. The overall results indicated accurate, comparable, and clinically acceptable performance for both devices, with neither device outperforming the other. In addition, this study was the only one in which multivariate analyses were used to determine other patient-related variables that might affect the performance of pulse oximetry devices. Four variables had a significant effect on the degree of error of both pulse oximeters: peripheral temperature, finger thickness, hemoglobin concentration, and skin color. Although finger thickness was addressed in a few of the studies by rotating the placement of the sensor on the patient’s finger (see Table), none of the other variables was addressed. These variables should be included and analyzed in all future studies on pulse oximeters. The new-generation Masimo SET, Philips FAST SpO₂, and the Nellcor N-3000 devices were used in the studies done in the PACU and the critical care unit.12,13,20,24,25 For these studies as a group, two overall conclusions can be made. First, the performance of the Masimo SET and the Nellcor N-3000 devices was significantly better than that of the comparison conventional devices.12,13,20,24 Second, when compared with one another, none of the new-generation devices differed significantly in clinical performance. Lutter et al20 compared the performance of Masimo SET, Nellcor N-3000, and Philips FAST SpO₂ devices during intra-aortic balloon pumping in the intensive care unit. They found that all 3 devices had degradations in clinical performance during pumping, suggesting that SpO₂ measurements in general should be used with caution during pumping. This finding seems reasonable, because intra-aortic balloon pumping creates a rhythmic disturbance in the arterial waveform that is in sync with the cardiac cycle, thereby making it difficult for pulse oximetry devices to distinguish between a true cardiac signal and the “artifact” signals generated by the intra-aortic balloon pump.
## Findings

### Part I

Admission rates to a special care nursery or a regular nursery were compared for matched infants who did (experimental, N=25) or did not (control, N=25) have $\text{SpO}_2$ monitoring in the delivery suite.

The rate of admission to the special care nursery was significantly higher for infants in the control group. The performance of the Masimo SET was better than that of the Oxismart; the percentage of minutes with continuous data output was significantly higher for the Masimo SET.

### Part II

The effects of pulse oximetry monitoring on immature infants were evaluated. The signal adequacy in the delivery suite of the MT device and the CPO was compared.

The time from delivery to adequate signal for the MT device was compared. Mean time from delivery to adequate signal was 2.3 minutes (range 1-5.3 minutes). All of the 15 infants had an initial reliable pulse oximetry signal.

The N-3000 had the best performance, with fewer false alarms and dropouts than the N-200.

The alarm rates for the 3 devices studied in each of the populations of patients were evaluated. No statistical or clinical differences were detected between any of the devices studied for true-positives, false-positives, and false-negatives for the PACU and ICU parts of the study.

In the PACU and ICU, although some individual differences between devices for types of alarm (true-positive, false-positive, false-negative, true-negative, inoperative) were detected, the overall performance of the devices was comparable. During IABP, no devices performed well (although the errors were different); therefore, pulse oximetry is probably not very accurate during IABP.

The Masimo SET detected 100% of true alarms (27), accounted for 4 of 31 false alarms, and had a positive predictive value of 87%.

The N-200 detected 59% of true alarms, accounted for 10 of 26 false alarms, and had a positive predictive value of 61%.

Data dropout was similar for both devices.

Five published studies on detection of hyperoxemia in neonates ($>95\%$ oxygen saturation) were reviewed. The performance of all the devices studied varied widely and was imprecise.

Overall findings were that the performance of new-generation pulse oximeters was much better than that of the CPOs for detection of hyperoxemia; differences between the new-generation devices were not significant. The new-generation devices generated 93% fewer false alarms than did the CPOs. Even the new devices did not reliably detect hyperoxemia during periods of patients’ movements.

The N-3000 and the Criticare 504 with 21-second signal-averaging time had equal performance, with 1 false alarm each. The Criticare 504 with 3-second signal-averaging time had the highest number of false alarms, with a total of 20.

The total number of alarm events in the study was low.

## Comments

### Part I

Sample selection was biased because of a lack of randomization of subjects into groups. The poorly described methods, particularly for dependent variables and matching of the control group, limit evaluation of the study. Difference in “continuous data” variable could be due to different manufacturers’ approaches to data display during potential artifact conditions. The lack of simultaneous $\text{SaO}_2$ data makes interpretation of the study results difficult.

### Part II

The poorly described methods, particularly for dependent variables, limits the study findings. A weakness was that they did not do anything with the experimental and control groups in Part II.

No gold standard was used. The 2 new-generation devices were not compared with each other.

No gold standard was used to assess accuracy of $\text{SpO}_2$ readings. Very limited information was provided on how clinicians determined the true and false alarms. No alarm thresholds were specified. The software revision for the Philips CMS was not specified, so it is difficult to be sure that it was an MT device.

No gold standard was used. The devices were not compared with each other.

Many different devices were used in the studies reviewed. The study protocols varied widely, making comparisons difficult.

The investigators attempted to assess motion tolerance in patients in the operating room, where motion is not a clinical factor. No gold standard was used.
### Findings

<table>
<thead>
<tr>
<th>FINDINGS</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>The N-3000 had the best performance, with only 1 false alarm in 199 alarms and 36 dropouts</td>
<td>No gold standard was used Assessing exactly how true and false alarms were measured and classified is difficult</td>
</tr>
<tr>
<td>The Criticare 504 with 21-second signal-averaging time had 32 false alarms in 172 false alarms and 172 dropouts</td>
<td>Different co-oximeters were used for measuring dependent variable, and quality control measures to ensure the accuracy of the devices were not described Small numbers of abnormal SaO2 values were studied Saturation readings were compared only when the pulse oximetry signal was optimal</td>
</tr>
<tr>
<td>SpO2 values were compared with SaO2 values in 2694 recordings obtained by using 2 MT devices</td>
<td>Little difference was found between the 2 values (bias, 0.19%; precision, 2.2%; limits of agreement, +4.63% to -4.25%) over a 60%-100% range of saturation values</td>
</tr>
<tr>
<td>5% of SaO2 readings were &lt;93%</td>
<td>Peripheral temperature, finger thickness, hemoglobin concentration, and skin color were reasons for error rates</td>
</tr>
</tbody>
</table>

### Clinical Studies in Children

The Masimo SET and the N-295 (Nellcor’s earliest motion-tolerant pulse oximetry device) devices were used in 2 clinical studies done in children. Malviya et al studied a group of 75 healthy children in a PACU and compared the Masimo SET device with a conventional Nellcor N-200 pulse oximeter. No gold standard was used and only a single new-generation pulse oximetry device was used. However, the findings from this study were consistent with the results from all other studies; the clinical performance of the motion-tolerant Masimo SET was better with regard to both true and false alarm conditions than was the comparison conventional device. Brouillette et al compared the Masimo SET with the Nellcor N-200 and N-395 in children in a sleep laboratory. Unfortunately, this study was so methodologically flawed that no meaningful conclusions can be made. Significant weaknesses included discrepancies in the reported signal-averaging times of the devices and lack of a gold standard reference SpO2 value, making it impossible to assess accuracy for either device.

### Clinical Studies in Patients in the Neonatal Intensive Care Unit

A total of 3 studies were done in patients in the neonatal intensive care unit (NICU). The motion-tolerant devices used in these studies included the Philips FAST SpO2, the Masimo SET, and the Nellcor N-3000. One study was so poorly designed and had such ambiguous findings that it was excluded from this review. Bohnhorst et al compared the Masimo SET and the Nellcor N-3000 motion-tolerant devices with each other and with the Nellcor N-200 conventional device. The performance of both of the motion-tolerant devices was better than that of the comparison conventional devices, again consistent with findings from previous studies. In addition, when the 2 motion-tolerant devices were compared with each other, the performance of the Masimo SET was significantly better than that of the Nellcor N-3000 for correct alarms, detection of hypoxia, and detection of bradycardia. These data indicate that the overall alarm rate for the N-3000 may have been achieved at the expense of some missed alarms for hypoxemia. If true, this possibility is clearly something that would be a point of concern for critical care clinicians. However, because the rate of false alarms was not evaluated in this study, fully evaluating the performance of each of the devices in the study is difficult.

More specifically, in a recent study, Hay et al examined the performance of the Philips FAST SpO2, Masimo SET, and the Nellcor N-395 devices and compared them with conventional devices for detection of hyperoxemia. The number of false alarms (specificity) and data dropouts did not differ significantly between the new-generation devices, and the Masimo SET had the best overall clinical performance. However, the software revision of the Philips device was not provided, again making a meaningful comparison difficult.

An important clinical issue specific to patients in the NICU is the performance of new-generation motion-tolerant pulse oximetry devices in detecting hyperoxemia. Hyperoxemia in the NICU is generally defined as a PO2 greater than 80 mm Hg, and only a few investigators attempted to determine the appropriate alarm level of pulse oximetry devices needed for detection of hyperoxemia. Hyperoxemia is associated with a number of complications in neonates, including retinopathy of prematurity, a complication that is not a clinical concern in children and adults. Poets et al reported that at an upper alarm limit of 95%, new-generation pulse oximeters are much better than the older generation devices in detecting hyperoxemia.
Summary
A consistent finding among the studies reviewed was the superior performance with regard to reduction in frequency of false alarms and overall accuracy of all the new motion-tolerant devices compared with the performance of the various conventional pulse oximeters. However, more clinical evidence on the performance of new-generation pulse oximetry devices is clearly needed, particularly in the setting of critical care. Although some conclusions can be made, a lack of consistency among study variables and techniques and discrepancies between the results of the studies reviewed make it difficult to make any overall recommendations about which of the new-generation motion-tolerant devices are best suited for use in critical care in any of the populations of patients studied. Clearly, more clinical trials are needed before the most accurate and reliable motion-tolerant pulse oximetry device can be determined for use in populations of critically ill patients. These trials should address the performance of the algorithms used for obtaining SpO\textsubscript{2} values in critically ill patients of all ages in situations in which patients’ movements are a factor. In addition, each trial should clearly specify which software revisions are being used with each device and should include a comparison with a known gold standard (preferably co-oximetry) for measurement of oxygen saturation. Finally, on the basis of the findings from Wouters et al,\textsuperscript{25} the variables of peripheral temperature, finger thickness, hemoglobin concentration, and skin color should be part of every future study on measurement of oxygen saturation by pulse oximetry devices to further assess the significance of these variables on clinical measurement of SpO\textsubscript{2}.

"New-generation oximeters have fewer false alarms and higher accuracy than do conventional oximeters."

New technology has, without doubt, advanced the science of pulse oximetry. Evidence supports the contention that the performance of the new-generation devices with incorporated enhanced algorithms for SpO\textsubscript{2} monitoring is superior to the performance of conventional pulse oximeters. However, until more evidence is available from clinical studies in which the new-generation oximeters are compared with one another, no valid judgments or decisions can be made about which new generation device has the best clinical performance in different populations of critically ill patients.

Commentary by Mary Jo Grap (see shaded boxes).

To purchase reprints, contact The InnoVision Group, 101 Columbia, Aliso Viejo, CA 92656. Phone, (800) 809-2273 or (949) 362-2050 (ext 532); fax, (949) 362-2049; e-mail, reprints@aacn.org.
REFERENCES

Adult - Accuracy

How Reliable are the New Generation Pulse Oximeters in the Detection of Hypoxemia in the Intensive Care Unit?
Appavu, S., Haley, T., Giuliano, K., and Duller, B.
Cook County Hospital, Chicago, IL

• This was presented at the 2004 Society of Critical Care Medicine National Conference.
• The purpose of this study was to evaluate the reliability of pulse oximeters in detecting true hypoxemia in the critically ill patient using continuous invasive blood gas monitoring as the standard.
• Hypoxemias were defined as mild (PaO$_2$<80mmHg), moderate (PaO$_2$<60mmHg), severe (PaO$_2$<50mmHg). Fourteen critically ill surgical patients with acute respiratory failure requiring mechanical ventilation were enrolled. All patients were monitored with all three brands (Philips, Masimo, and Nellcor) of pulse oximeters as well as a co-oximeter calibrated continuous blood gas monitor.
• The new generation of Philips, Masimo and Nellcor pulse oximeters detected severe and moderate hypoxemia equally well.
• The following abstract was published in Critical Care Medicine, Vol. 31 (12):A 96 (2003).

Introduction: The purpose of this prospective study was to compare the clinical performance of three new generation pulse oximeters (Philips FAST, Nellcor N-395, and Masimo SET [Signal Extraction Technology]) during hypoxemia in critically ill surgical patients.

Hypothesis: No difference in performance.

Methods: Critically ill surgical patients (N=14) with acute respiratory failure requiring mechanical ventilation were enrolled in the study. IRB approval and informed consent were obtained prior to enrollment. Subjects were monitored continuously with all three pulse oximeters, a co-oximeter calibrated continuous blood gas monitor (Trendcare, Diametrics), and intermittent stat laboratory blood gas measurements. The hands and fingers were randomly selected for sensor placement, and the fingers were protected from ambient light. Monitoring fingers were randomly changed at specific intervals. Hypoxemias were defined as mild (PaO$_2$<80mmHg), moderate (PaO$_2$<60mmHg) or severe (PaO$_2$<50mmHg). ROC curves were calculated for each device at each hypoxemia level.

Results: More than 250 hours of continuous data, including motion artifact and poor perfusion signals, were analyzed. For severe hypoxemia at a sensitivity of >=90%, the specificity for each device was: Philips=89.2%; Masimo=86.1% and Nellcor=84.8%. For moderate hypoxemia at a sensitivity of >=85%, the specificity was: Philips=91.5%; Masimo=89.3% and Nellcor=82.5%. For mild hypoxemia at a sensitivity of >=80%, the specificity was: Masimo=71.4, Philips=70.6 and Nellcor=59.1.

Conclusions: These data support that as PaO$_2$ levels get higher, the balance between sensitivity and specificity can only be achieved by accepting increasingly smaller sensitivity and specificity values. All three devices performed well for both moderate and severe hypoxemia detection. In contrast, all three devices showed a significant false alarm rate of >25% during mild hypoxemia, a finding that highlights one of the general limitations of all pulse oximeters. Future research should concentrate on improving the reliability for all three devices by reducing false readings during mild stages of hypoxemia.
The Clinical Use of Two New Generation SpO₂ Devices for Critically Ill Patients with Sepsis

Giuliano, K, Elliot, S., Scott, S.
Baystate Medical Center, Springfield, MA

• This study will be presented at the 2005 American Association of Critical Care Nurses Conference, New Orleans, LA, May 7-15, 2005.

• The purpose of this study was to assess the clinical performance of two new generation SpO₂ devices (Philips FAST and Masimo SET) for critically ill patients with sepsis.

• Fourteen subjects with a signal quality index of 1.3 or less were monitored with two devices simultaneously.

• In this study both SpO₂ devices performed equally well in this septic patient population. No significant statistical or clinical differences in device performance were found and the performance of both devices met the AAMI standard for SpO₂ device performance.

Purpose: The purpose of this study was to assess the clinical performance of 2 new generation SpO₂ devices for critically ill patients with sepsis.

Background/Significance: SpO₂ monitoring is a standard of care for the critically ill. The new generation SpO₂ devices have been designed for improved clinical performance during low perfusion, however little is known about the differences among these new generation devices.

Methods: Two new generation devices were used simultaneously to monitor septic patients (N=14) with a perfusion index (PI) of 1.3 or less during this study. The two devices used were Philips FAST and Masimo SET (V3). Randomization was used for digit selection, data were only recorded when waveforms were of good quality, and data were recorded every 5 minutes for a period of two hours on each patient enrolled.

Results: The mean SAPSII score was 49.2, mean age was 64 years and mean PI of .56. There was no significant mean difference between the two devices across all time measurements (mean SET SpO₂=97.09, mean FAST SpO₂=97.00). Repeated measures ANOVA was not significant for time (p=.627), device (p=.256) or the interaction (p=.954). Finally, there were no individual readings between the 2 devices that differed by more than 3% in any case where both waveforms were adequate, and less than 5% of the total data was missing due to poor quality waveforms.

Conclusions: Physiologic monitoring in critical care requires accurate data. The results of this study support that both SpO₂ devices performed equally well in this septic patient population. No significant statistical or clinical differences in device performance were found. In no case would treatment have differed based on an individual reading from either device. Finally, the performance of both devices met the AAMI standard for SpO₂ device performance (+/-3%).
Introduction: The purpose of this study was to compare the clinical performance and accuracy of 2 new-generation pulse oximetry (SpO₂) devices, Philips FAST (Philips Fourier Artifact Suppression Technology), and Masimo SET (Signal Extraction Technology) in cardiac surgery patients.

Hypothesis: There is no difference in accuracy between Philips FAST and Masimo SET pulse oximeters.

Methods: This study was approved by the IRB of the Medical Center. Informed consent was obtained from the subject or family member. Subjects were a convenience sample admitted to the Cardiovascular ICU (CVICU) following cardiac surgery. Once admitted to the CVICU, a pulse oximetry finger probe from each oximeter was placed on the subject using random finger assignment for probe placement. SpO₂ and signal quality index (PI) were recorded at baseline and every 5 minutes for 2 hours. The arterial blood gases (ABG) used for comparison were drawn as clinically indicated.

Results: The mean age of the sample (N = 50) was 64.9 years (36 men and 14 women). Pulse oximetry data with corresponding ABG values were available at 65 different instances. The overall correlation between the Philips and Masimo SpO₂ for the 65 instances was .765 (p < .01). The correlations were as follows: the Philips SpO₂ to the SaO₂ was .287 (p < .05); the Masimo to the SaO₂ was .765 (p < .01). However, when looking at patients where both devices indicated low signal quality (PI ≤ 1.3 (N = 34)) the correlation between SaO₂ and SpO₂ for the Philips device was .426 (p = .012) and the Masimo device was .468 (p < .01).

Conclusions: This study demonstrates that the Philips and Masimo devices have a high degree of accuracy compared to the arterial blood SaO₂. In patients with low signal quality (PI < 1.3) the correlations with SaO₂ are very similar and although there is a slight difference in statistical significance, this does not represent a clinical significance in the performance of the devices.
Adult - Motion

A Comparison of Three New Generation SpO₂ Devices During Ambulation After Open Heart Surgery

Richards, N., Giuliano, K., Jones, P., Krentz, K., Basinger, S.
Saint Luke's Hospital, Mid America Heart Institute, Kansas City, MO

This study was presented at the 2004 AACN National Conference, Orlando, FL, May 15-20, 2004.

The purpose of this study was to assess the clinical performance of 3 new generation SpO₂ devices (Philips FAST, Masimo SET and Nellcor N-3000) while patients were walking after their open-heart surgery.

Thirty-six patients were enrolled in this study. All three devices were used continuously while patients were walking. Data on dropouts (DO) false alarms (FA) and correlation with heart rate (HR) were recorded.

Statistical analysis shows that there was significant difference between the three devices for both DO and FA. The N-3000 had the most data dropout, whereas the converse was true for FA; the Masino SET had highest real FA.

Purpose: The purpose of this study was to assess the clinical performance of 3 new generation SpO₂ devices (Philips FAST, Masimo SET and Nellcor N-3000) during ambulation after open-heart surgery (N=36).

Background/Significance: The new generation SpO₂ devices have been designed for improved clinical performance during motion, however little is known about the differences in performance among these new generation devices during patient motion.

Methods: Randomization was used for digit and hand selection, and all three devices were used continuously during ambulation. Data on dropouts (DO), false alarms (FA), and correlation with heart rate (HR) were recorded.

Results: Pairwise comparisons indicated significant differences across all three devices for both DO and FA. The N-3000 had the most data dropout (odds ratio of 31.9 to Masimo SET and 5.6 to Philips, at the 95% CI). However, converse was true for FA, with the Masimo SET being highest (odds ratio of 17.9 to Nellcor and 2.3 to Philips, at the 95% CI). There was also a significantly higher amount of data dropout for all three devices when readings were taken in the hand where a radial graft had been used (p=.004). For HR correlation, the mean absolute difference across all three devices was similar (Philips=4.3, Masimo=5.1, Nellcor=3.0).

Conclusions: Physiologic monitoring in acute care requires accuracy and minimal false alarms. This study shows that there are differences across all three devices with regard to both DO and FA. High amounts of DO are problematic, because no clinical patient information is available during DO. However, false alarms are even more problematic, because they desensitize the clinicians to alarms and call into question the accuracy of displayed data. While these data highlight the statistical differences in the SpO₂ devices that were studied, the clinical implications of these differences warrants further study.
This study tested all commercially available motion resistant pulse oximeters along with numerous conventional pulse oximeters during reduced perfusion and mechanically controlled motion on volunteers breathing room air and hypoxic gas mixtures.

Seventy healthy volunteers participated in this study which used 20 different pulse oximeters including Masimo, Philips, Datex-Ohmeda, and Nellcor.

This study showed that the newer-generation pulse oximeters exhibit improved performance during patient motion, which is important for displaying accurate SpO\textsubscript{2} values with hypoxic patients when they are awake and tend to be agitated and moving. Of the 20 different oximeters tested, the Philips overall SpO\textsubscript{2} performance resulted in a second place ranking.

Address correspondence and reprints to Steven J. Barker, PhD, MD, Department of Anesthesiology, University of Arizona, P.O. Box 245114, Tucson, AZ 85724-5114.

Address e-mail to sjbarker@u.arizona.edu
Neonate - Motion

The Clinical Use of Two New Generation SpO₂ Devices for Patients in the Neonatal Intensive Care Unit (NICU)

Fegegta Workie, K. Rais-Bahrami, Billie L. Short.
Children's National Medical Center & The George Washington University
School of Medicine, Washington, D.C.

- This study was presented at the 2004 American Academy of Pediatrics, San Francisco, CA, October 9-13, 2004.
- The purpose of this study was to assess the clinical performance of two new generation pulse oximetry devices (Philips FAST and Masimo SET (V3)) in NICU patients, especially during the patient's movement.
- In this study two devices were used simultaneously to monitor patients. Thirty-six patients were involved in this study. The data for SpO₂, number of true and false alarms, number of dropouts and the duration of dropouts for each monitor were recorded.
- The two new generation pulse oximeters were equally sensitive and provided accurate measurements, although Masimo SET was less sensitive to motion artifact and had fewer data dropouts over time.

Purpose: The purpose of this study was to assess the clinical performance of two new generation pulse oximeters (SpO₂) devices in NICU patients.

Background/Significance: Continuous monitoring of pulse oximetry (SpO₂) is a common practice among preterm and critically ill neonates. The new generation SpO₂ devices have been designed for improved clinical performance. A substantial reduction in alarm frequency is achieved with newer devices that claim to offer motion resistant measurements. However, little is known about the differences among these new generation devices in the NICU patient population.

Methods: During the study period, two new generation devices were used simultaneously to monitor patients (N=36) in the NICU. The two devices studied were Philips FAST and Masimo SET (V3). Patients were randomized for their digit selection and data were only collected when waveforms were of good quality. The data for SpO₂, number of true and false alarms and number of dropouts as well as the duration of dropouts for each monitor were recorded for a period of two hours on each patient enrolled.

Results: The mean gestational age of the patients was 32.4 weeks, the mean current age was 18.6 days, with 20 males and 16 females. Repeated measures ANOVA indicated no significant difference between the two devices across all time measurements (time: p=0.345, device: p=.057, interaction: p=0.357). In addition, paired t-tests for true alarms and false alarms were non-significant (p-values of 0.468 & 0.299). There was a significant difference in the amount of data dropout, with the Philips device having more data dropout than the Masimo device (p<0.001).

Conclusions: Physiologic monitoring in critical care requires accurate data measurement. The two new generation pulse oximeters (Philips FAST and Masimo SET) were equally sensitive and provided accurate measurements, although Masimo SET was less sensitive to motion artifact and had fewer data dropouts over time.