



Published in final edited form as:

Chest. 2007 September ; 132(3): 852–859. doi:10.1378/chest.07-0312.

Variations in Physician Interpretation of Overnight Pulse Oximetry Monitoring

Rory Ramsey, MB, ChB, Reena Mehra, MD, MS, FCCP, and Kingman P. Strohl, MD, FCCP

[†]Department of Medicine (Drs. Ramsey and Strohl), Division of Pulmonary, Critical Care and Sleep Medicine, and Department of Pediatrics (Dr. Mehra), Division of Clinical Epidemiology, Case School of Medicine, Cleveland, OH

Abstract

Background—Overnight pulse oximetry is commonly used for hypoxemia evaluation in patients with COPD and sleep-disordered breathing. There is little information regarding its impact on physician decision making, and therefore an important measure of its clinical utility is untested and unknown. The aim of this study was to describe physician interpretation, use, and opinions regarding overnight pulse oximetry.

Methods—Forty-one pulmonary physicians and fellows participated in structured interviews consisting of three oximetry record interpretations, oral responses to a standard question set, and a questionnaire. Qualitative data were analyzed using an open coding process. Quantitative data were assessed for distributions.

Results—Four measures were consistently used by the majority of physicians in record interpretation: background information, arterial oxygen saturation measured by pulse oximetry (SpO₂) waveform and pattern, and time spent with SpO₂ < 90%. An additional 10 measures were consistently used by 5 to 46% of physicians. No interpretation generated a recommendation with > 60% consensus. There was a wide range of opinions on important matters related to this test, including test utility, indications, variables considered most important for interpretation, and criteria for nocturnal oxygen prescription. Forty-one physicians provided 35 different opinions on when nocturnal supplemental oxygen should be initiated.

Conclusions—The variation in physician interpretation, use, and opinions regarding overnight pulse oximetry calls into question its clinical utility and underscores a need for standardization of presentation, training, and interpretation.

Keywords

COPD; oximetry; sleep-disordered breathing

Overnight pulse oximetry is commonly used for the evaluation of potential hypoxemia in COPD^{1,2} and sleep-disordered breathing.^{3–5} It is also employed in a variety of other medical

Correspondence to: Rory Ramsey, MB, ChB, Case Medical Center, University Hospitals, 11100 Euclid Ave, Cleveland, OH 44106-6003; e-mail: Rory.Ramsey@uhhs.com.

The authors have no conflicts of interest to disclose.

Reproduction of this article is prohibited without written permission from the American College of Chest Physicians (www.chestjournal.org/misc/reprints.shtml).

Publisher's Disclaimer: CHEST is the official journal of the American College of Chest Physicians. It has been published monthly since 1935. Copyright 2007 by the American College of Chest Physicians, 3300 Dundee Road, Northbrook IL 60062. All rights reserved. No part of this article or PDF may be reproduced or distributed without the prior written permission of the copyright holder (<http://www.chestjournal.org/misc/reprints.shtml>). ISSN: 0012-3692.

conditions and hospital settings.⁴ Despite this broad range of use, there is no standardized approach to data presentation or interpretation and discussion of its relevant indications. Additionally, little is known about how physicians view the information from overnight oximetry or how it affects physician decision making. As a result, the clinical utility of overnight oximetry is difficult to ascertain.

Test utility is measured by the ability to diagnose or grade disease, direct physician decision making, and participate in management pathways that lead to improved outcomes. Each of these successive measures of test utility offers a more robust measure of the efficacy of a test.⁶ Diagnostic accuracy is the most basic measure. Pulse oximetry provides an accurate assessment of arterial oxygen saturation in general, and many of the newer oximeters are capable of accurately capturing the periodic hypoxia characteristic of sleep-disordered breathing.⁷ Overnight oximetry has been proposed as a potential diagnostic and screening device for sleep-disordered breathing, a use that has mixed reviews.^{3,5,8–11}

Physician decision making in overnight pulse oximetry has only been investigated in a very preliminary manner.¹² In general, physicians order diagnostic tests to challenge perceived disease probability, and results direct management decisions by confirming or altering disease probability. Decisions are changed when perceived probability crosses critical testing and treatment thresholds.¹³ The lack of test standardization or descriptions of how physicians use this test raises concern that it is not being consistently interpreted or used. Substantial variability in its interpretation or use would imply that physicians are deriving different impressions of disease probability and are not being consistently directed in management decisions.

Overnight oximetry-directed pathways for sleep-disordered breathing have not been examined for cost or patient outcome. The value of identifying nocturnal hypoxia that is not sleep apnea is also unclear. Several trials^{2,14,15} have evaluated the use of nocturnal oxygen therapy in patients with COPD and isolated nocturnal hypoxia, but results are conflicting and no clear consensus on its utility exists.

We chose to assess the clinical utility of overnight pulse oximetry by investigating its impact on physician decision making. Specifically, we evaluated the information physicians use to interpret overnight oximetry records and described the diagnoses and recommendations physicians generated from reading the same records. We also surveyed physician opinions with regards to important preselected components of overnight oximetry. Substantial variability in these assessments would imply that overnight oximetry is not consistently directing physician assessment of disease state, probability, and severity, and is therefore not helpfully directing clinical management decisions. We hypothesize that there is substantial variability in these parameters and that this finding would adversely impact its clinical utility by creating inconsistent clinical decision-making paradigms.

Materials and Methods

Participants

A total of 24 board-certified pulmonary physicians and 17 pulmonary fellows in three academic medical centers in Cleveland participated in the study. The division chief from each participating institution granted permission to approach their physicians for enrollment into this study. Requests were made to keep the subject matter of the interview confidential to avoid priming subsequent subjects. Interviews were conducted in physician offices or at a conveniently located conference room. The University Hospitals of Cleveland Institutional Review Board Committee approved the study protocol, and all participants provided full written informed consent.

Procedures

The first author attempted to approach in person and/or telephone all members of the three divisions. The first author conducted all interviews. All interviews were recorded.

Measures

Oximetry Record Interpretation—From a sample of 30 overnight oximetry records, the authors selected 3 records by consensus. These records were selected for their representation of findings commonly encountered in evaluation for “rule out sleep-disordered breathing” and assessments for home oxygen or dyspnea. Two different oximetry record formats were used. Both formats contained numeric summaries of the data and graphs of arterial oxygen saturation measured by pulse oximetry (SpO₂) and heart rate waveforms in a compressed and expanded format. Background information that accompanied these reports was also included: record 1, age 57 years, male gender, body mass index [BMI] of 24 kg/m², and room air; record 2, female gender, room air; and record 3, age 81 years, male gender, BMI of 31 kg/m², and room air. Summary data in these studies included the following: recording time, mean SpO₂, low SpO₂, percentage of time with SpO₂ < 90%, number and duration of desaturation events, and pulse rate information. Copies of the compressed SpO₂ waveforms for records 1, 2, and 3 are presented in Figures 1–3, respectively. The records were presented to study subjects separately and in random order. The subjects were instructed to interpret the overnight oximetry record by identifying all the information used in the order and manner they were accustomed. Subjects were also asked to leave an impression and recommendation. The interviewer did not interact with or assist the subject during the study interpretation but was available only for clarification of instructions.

Semistructured Discussion—The interviewer led a discussion directed around six predetermined questions designed to elicit opinions and attitudes in the following categories: test interpretation, utility, indications, and treatment thresholds.

Quantitative Questionnaire—At the conclusion of the structured discussion, participants completed a written 21-item questionnaire. These questions were designed to assess physician opinions on matters related to the utility of clinical and test information and measures of physician experience. Questions 1 to 11 used a 5-point Likert scale to assess opinions on the importance of specific patient history data and test information on test interpretation and outcome generation. Questions 12 to 14 focused on physician opinions of the relevance of technical aspects of pulse oximeters. Questions 15 and 16 collected information related to physician use of the oxygen desaturation index (ODI). Questions 17 to 20 measured experience with pulmonary medicine and specifically with personal exposure to using overnight oximetry. Question 21 dealt with a global impression of the usefulness of overnight oximetry.

Data Analysis

Qualitative analyses were performed to assess the oximetry record interpretation and semistructured discussion transcripts by using an open coding process. Investigators (R.R., R.M., K.P.S.) separately analyzed eight transcripts and developed codes and classification by consensus. As codes evolved additional transcripts were reviewed to challenge, expand, and refine the categories. The final coding structure was applied to all transcripts (by R.R.) using a spreadsheet (Excel; Microsoft; Redmond, WA).

Results

A total of 24 board-certified pulmonary physicians and 17 pulmonary and critical care fellows participated in the study. This represented 65% of all potential subjects in the three academic

centers. Reasons for not participating included the following: co-author, busy schedule, failure to return contact, and absence (*eg*, vacation). All participants completed the interview process.

Record Interpretation

There was substantial variation in the information physicians used to interpret the records (Table 1). Only four variables were consistently used (defined as used in two of three record interpretations) by > 50% of physicians. These were background information, SpO₂ waveform and pattern, and percentage of time with SpO₂ < 90%. An additional 10 measures were consistently used by 5 to 46% of physicians.

SpO₂ waveform pattern interpretations were given by 51%, 54%, and 56% of respondents for studies 1, 2, and 3, respectively. In records 1 and 3, there was almost unanimous agreement that the waveform constituted a cyclic pattern (95% and 100%, respectively). In record 2, 91% identified the pattern as noncyclic and 9% as cyclic. Eighteen percent of respondents reported a mixed pattern.

Heart rate waveform pattern interpretations were given by 51%, 37%, and 44% for studies 1, 2, and 3, respectively. Heart rate waveform patterns were interpreted as consistent with sleep-disordered breathing by 71%, 85%, and 78% of these respondents in the three respective studies.

Potential diagnoses were offered by 41%, 46%, and 54% of participants for studies 1, 2, and 3, respectively. In study 1, sleep-disordered breathing was suggested by 94%, lung disease by 24%, heart disease by 24%, rapid eye movement (REM) related by 35%, and neuromuscular disease by 6%. In study 2, sleep-disordered breathing was suggested by 58%, lung disease by 74%, heart disease by 26%, REM related by 21%, and neuromuscular disease by 4%. In study 3, sleep-disordered breathing was suggested by 82%, lung disease by 9%, heart disease by 18%, REM related by 18%, and neuromuscular disease by 5%.

Recommendations generated from study interpretations are represented in Figure 4. Polysomnography, nocturnal oxygen prescription, and the need for more medical information were the most commonly selected recommendations. In no study did a single recommendation achieve a consensus > 60%.

Structured Discussion

Study participants expressed a range of opinions on the following themes: test interpretation, test indications, test utility, and criteria for ordering supplemental oxygen. There was substantial variation in what physicians believed to be the “most important” information for interpretation (Fig 5). Many expressed uncertainty with regards to the best way to interpret this test. One reported, “I don't have any strong basis for interpreting them other than trying to fumble through it with a little common sense.” Another noted, “Often, I see someone else's interpretation but wonder if I would read it the same way”.

Physicians expressed a range of opinions on the proper indications for this test (Fig 6). Screening for nocturnal hypoxia was the most agreed-on indication, but even within this category there was substantial variation in opinion. Many used the test for anyone suspected of nocturnal hypoxia, while others confined their use of this test to narrow indications (*eg*, pulmonary hypertension, neuromuscular disease).

Participants expressed negative opinions on the utility of overnight oximetry on a number of levels. First, there was concern that test interpretation was unduly subjective and not properly validated or standardized. “The problem with this test is that it's available but I don't think we know what to do with it. Each one of us has their own thing about it (the test). So its utility is

very poor.” Second, many participants expressed doubt that the test added any value to a history and physical examination or affected decision making. “It almost never obviates the need to do a polysomnography”. “If I was really worried that a patient had obstructive sleep apnea I wouldn’t go fishing with this test. I would go fishing with my history and physical examination.”

Many participants had positive comments about the utility of overnight oximetry. Much of this was in reference to its use as a screening device for nocturnal hypoxia. Other favorable comments were mostly general in nature and did not refer to test performance characteristics. “I think it’s one of the most important tests we have available to us since we can’t enter the patient’s room environment other than through the use of this test. It’s an important part of any pulmonologist’s practice, particularly in patients that have periods of sleep that require oxygen therapy”.

Lastly, there was tremendous variation in what physicians constituted criteria for nocturnal supplemental oxygen prescription. Forty-one physicians provided 35 different opinions on when nocturnal supplemental oxygen should be initiated.

Questionnaire

There was general agreement on the importance of several demographic, clinical, and test-related information by a majority of participants (Table 2). On questions related to the technical aspects of pulse oximetry: 30% expected a physical calibration of the oximeter to be recorded, 46% believed the oximeter averaging time was important, and 7% believed that other matters related to the device had a meaningful impact on test results.

Two participants calculated their own ODI. There was a range in what participants considered an abnormal ODI: > 5 by 46%, > 10 by 37%, and > 15 by 10%. Seventeen percent of subjects had “formal” exposure to overnight oximetry interpretation. There was a wide range in the number of studies interpreted per year (0 to 160). Fifty-four percent of subjects interpreted $< 10/\text{yr}$, while 46% interpreted $\geq 10/\text{yr}$. Only 7% participants agreed with the statement that “overnight oximetry is worthless.”

Discussion

We found a wide variation in physician interpretation and use of overnight pulse oximetry. When presented with three separate oximetry records, physicians weighed different information and generated substantially different recommendations. They also had varying opinions regarding test indications, test utility, the “most important” information, and criteria for nocturnal supplemental oxygen prescription.

These findings have important implications for the clinical utility of overnight oximetry. The utility of a test can be measured by its ability to modify physicians' perception of disease probability and guide physician decision making. The substantial variability in physician recommendations described in this study suggests the test is not consistently informing physician perception of disease state, probability, or severity. This implies overnight oximetry is not being used as an objective marker of disease and therefore cannot be consistently informing physician decision making and contributing to improving patient outcomes.

The results of this study provide at least three reasons for this variable effect on decisions. First, physicians have different opinions on what information is important. They considered different variables and weighed them differently. This probably reflects a lack of applicable clinical outcome data and also an uncertainty on what is currently known. Second, physicians had varying opinions of what constituted normal and what constituted treatment thresholds.

For example, there was a substantial spread on what constituted a normal ODI, and there were 35 different opinions on the criteria for nocturnal supplemental oxygen. Third, physicians use this test for different reasons. Some of the indications offered included sleep-disordered breathing screening, risk stratification, and screening for nocturnal hypoxia.

It is not unusual to have variation in physician use of a test. Physicians weigh several factors when determining a management strategy: the pretest probability of disease, test performance characteristics, the cost/harm of testing and treating, and estimation of patient preference for cost and clinical improvement. A management decision is a composite of these factors, and variation in estimation of any of them can lead to different decisions. The major problem with the variability seen in the current study was not only the degree of interpretation inconsistency, but also that physicians used different information to generate their recommendations. There was no consistent approach to interpretation and no apparent links between test output and recommendations. Spirometry is a good counterexample. It provides an objective measure of lung function for which variable management approaches may be taken based on interpretation of test results or other patient factors. The main difference though, is that there is a well-standardized approach to performing, presenting, and interpreting spirometry. These standards provide important safeguards that limit unwarranted use. Spirometry has been extensively studied, and its value and limitations are well known and publicized.^{16–20} Whether physicians use these recommendations is a matter of choice rather than lack of evidence.

This study has limitations. Physicians were presented with records of unknown patients and with minimal clinical information. In practice, physicians usually have the opportunity to interact with patients and acquire some additional information that informs test indication, interpretation, and decision making. We chose not to provide additional clinical information because we were primarily interested in the physician's approach to the oximetry information, especially to identify that information that physicians believed to be important and relevant to any subsequent diagnoses or recommendations. Adding additional clinical information might have narrowed physician consideration of test information, diagnoses and/or decisions, and decreased our ability to discern differences in physician approaches to the information. Our approach of not providing additional clinical information could have increased the variability observed in this study compared to what one might expect in practice. Nonetheless, we believe the degree of variability in approach and interpretation described in this study is impressive. Also, physicians may have been asked to interpret records with an unfamiliar format. All records, however, were from the most commonly used outpatient portable devices used in 2006, and as such represent the variability a physician would be expected to encounter in a Cleveland practice at the time of study. Finally, it is also conceivable that physician behavior in front of an interviewer and microphone might differ from what would happen in practice.

The variability in physician interpretation and use of overnight oximetry described in this study suggests this test is not effective in general practice. Outcome studies are needed to better determine the meaning of overnight oximetry data before their role can be better defined. Until then, two simple interventions may help promote more effective use of overnight oximetry. In the first, standardizing test presentation would provide a more uniform format for data and waveform presentation. For instance, every printout should have the make, model, and averaging characteristics used for the test, along with a selected group of other relevant information. In the second, an expert-based review of overnight oximetry could summarize the evidence, grade its strength, and provide recommendations for test use and improvement. These measures would not make the test more efficacious at diagnosing or measuring disease but can reduce variability and focus on the issue of effectiveness. These are simple quality control initiatives that could have a significant and positive impact on practice.

Acknowledgments

Funding was provided by the American Heart Association National Scientist Development Award (0530188N), CHEST Foundation of the American College of Chest Physicians T. Franklin Williams Geriatric Development Research Award, and National Institutes of Health/National Heart, Lung, and Blood Institute grant K23 HL079114-01A2.

References

1. Lewis CA, Eaton TE, Fergusson W, et al. Home overnight pulse oximetry in patients with COPD: more than one recording may be needed. *Chest* 2003;123:1127–1133. [PubMed: 12684303]
2. Chaouat A, Weitzenblum E, Kessler R, et al. A randomized trial of nocturnal oxygen therapy in chronic obstructive pulmonary disease patients. *Eur Respir J* 1999;14:1002–1008. [PubMed: 10596681]
3. Whitelaw WA, Brant RF, Flemons WW. Clinical usefulness of home oximetry compared with polysomnography for assessment of sleep apnea. *Am J Respir Crit Care Med* 2005;171:188–193. [PubMed: 15486338]
4. Netzer N, Eliasson AH, Netzer C, et al. Overnight pulse oximetry for sleep-disordered breathing in adults: a review. *Chest* 2001;120:625–633. [PubMed: 11502669]
5. Chiner E, Signes-Costa J, Arriero JM, et al. Nocturnal oximetry for the diagnosis of the sleep apnoea hypopnoea syndrome: a method to reduce the number of polysomnographies? *Thorax* 1999;54:968–971. [PubMed: 10525553]
6. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991;11:88–94. [PubMed: 1907710]
7. Davila DG, Richards KC, Marshall BL, et al. Oximeter's acquisition parameter influences the profile of respiratory disturbances. *Sleep* 2003;26:91–95. [PubMed: 12627739]
8. Kirk VG, Bohn SG, Flemons WW, et al. Comparison of home oximetry monitoring with laboratory polysomnography in children. *Chest* 2003;124:1702–1708. [PubMed: 14605038]
9. Magalang UJ, Dmochowski J, Veeramachaneni S, et al. Prediction of the apnea-hypopnea index from overnight pulse oximetry. *Chest* 2003;124:1694–1701. [PubMed: 14605037]
10. Vazquez JC, Tsai WH, Flemons WW, et al. Automated analysis of digital oximetry in the diagnosis of obstructive sleep apnoea. *Thorax* 2000;55:302–307. [PubMed: 10722770]
11. Levy P, Pepin JL, Deschaux-Blanc C, et al. Accuracy of oximetry for detection of respiratory disturbances in sleep apnea syndrome. *Chest* 1996;109:395–399. [PubMed: 8620711]
12. Martinez MW, Rodysill KJ, Morgenthaler TI. Use of ambulatory overnight oximetry to investigate sleep apnea in a general internal medicine practice. *Mayo Clin Proc* 2005;80:455–462. [PubMed: 15819281]
13. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med* 1980;302:1109–1117. [PubMed: 7366635]
14. Fletcher EC, Donner CF, Midgren B, et al. Survival in COPD patients with a daytime Pao₂ greater than 60 mm Hg with and without nocturnal oxyhemoglobin desaturation. *Chest* 1992;101:649–655. [PubMed: 1541127]
15. Connaughton JJ, Catterall JR, Elton RA, et al. Do sleep studies contribute to the management of patients with severe chronic obstructive pulmonary disease? *Am Rev Respir Dis* 1988;138:341–344. [PubMed: 3195833]
16. Pellegrino R, Viegi G, Brusasco V, et al. Interpretative strategies for lung function tests. *Eur Respir J* 2005;26:948–968. [PubMed: 16264058]
17. Macintyre N, Crapo RO, Viegi G, et al. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. *Eur Respir J* 2005;26:720–735. [PubMed: 16204605]
18. Wanger J, Clausen JL, Coates A, et al. Standardisation of the measurement of lung volumes. *Eur Respir J* 2005;26:511–522. [PubMed: 16135736]
19. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J* 2005;26:319–338. [PubMed: 16055882]
20. Miller MR, Crapo R, Hankinson J, et al. General considerations for lung function testing. *Eur Respir J* 2005;26:153–161. [PubMed: 15994402]

Abbreviations

BMI	body mass index
ODI	oxygen desaturation index
REM	rapid eye movement
Sp_o2	arterial oxygen saturation measured by pulse oximetry

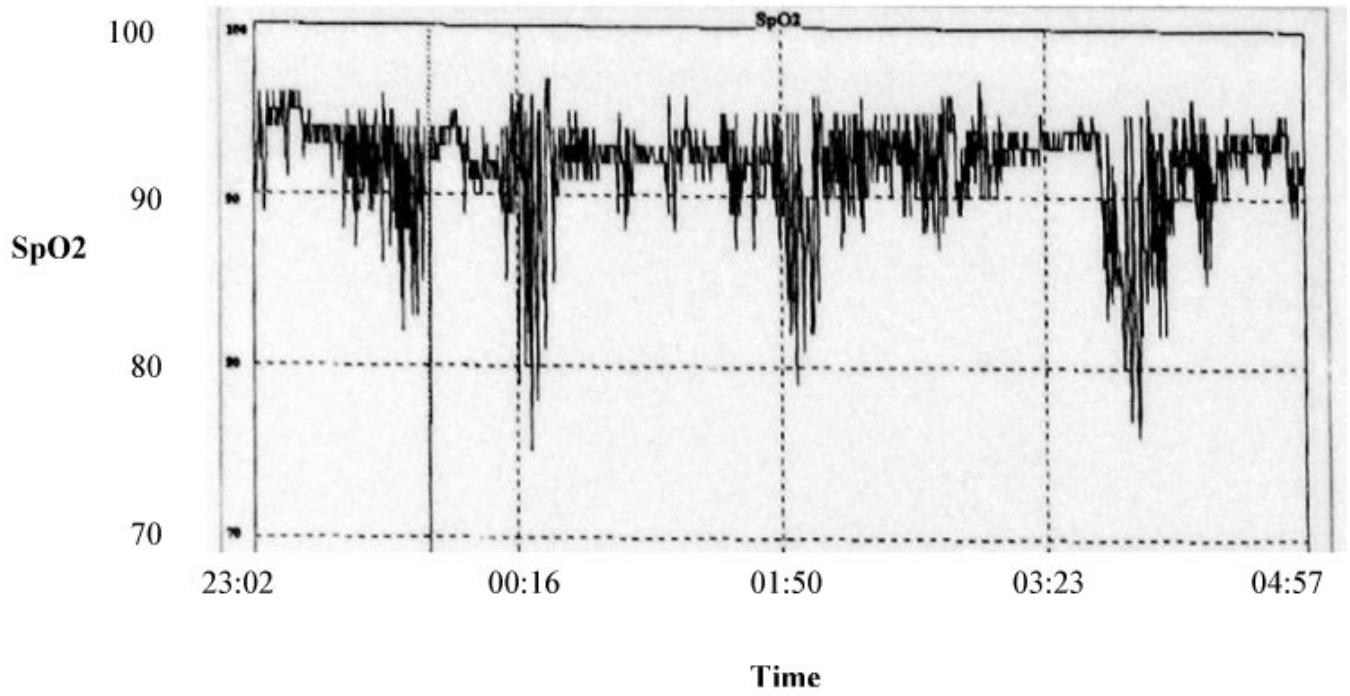


Figure 1.
Compressed SpO₂ waveforms from record 1.

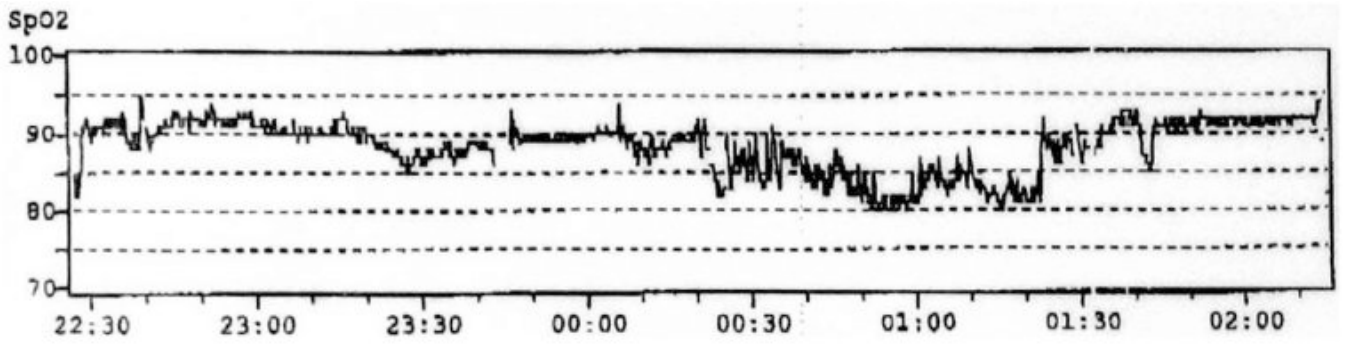


Figure 2.
Compressed SpO₂ waveforms from record 2.

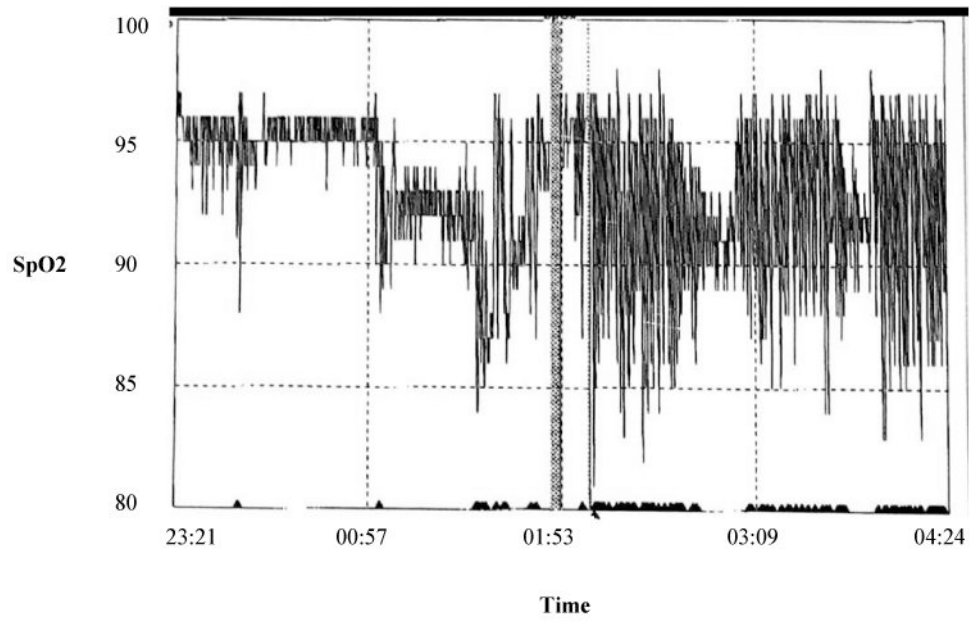


Figure 3.
Compressed SpO₂ waveforms from record 3.

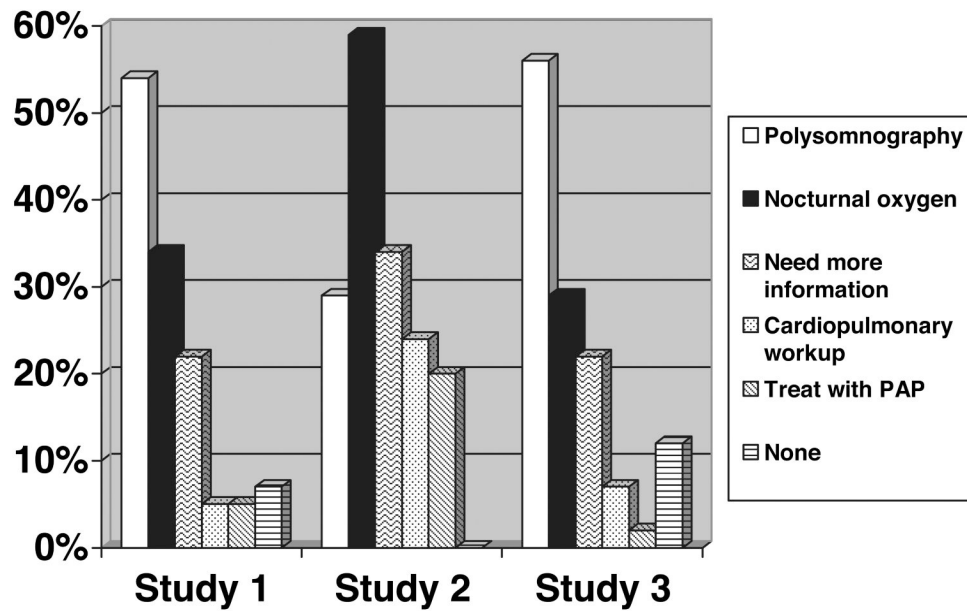


Figure 4.
Physician participant recommendations.

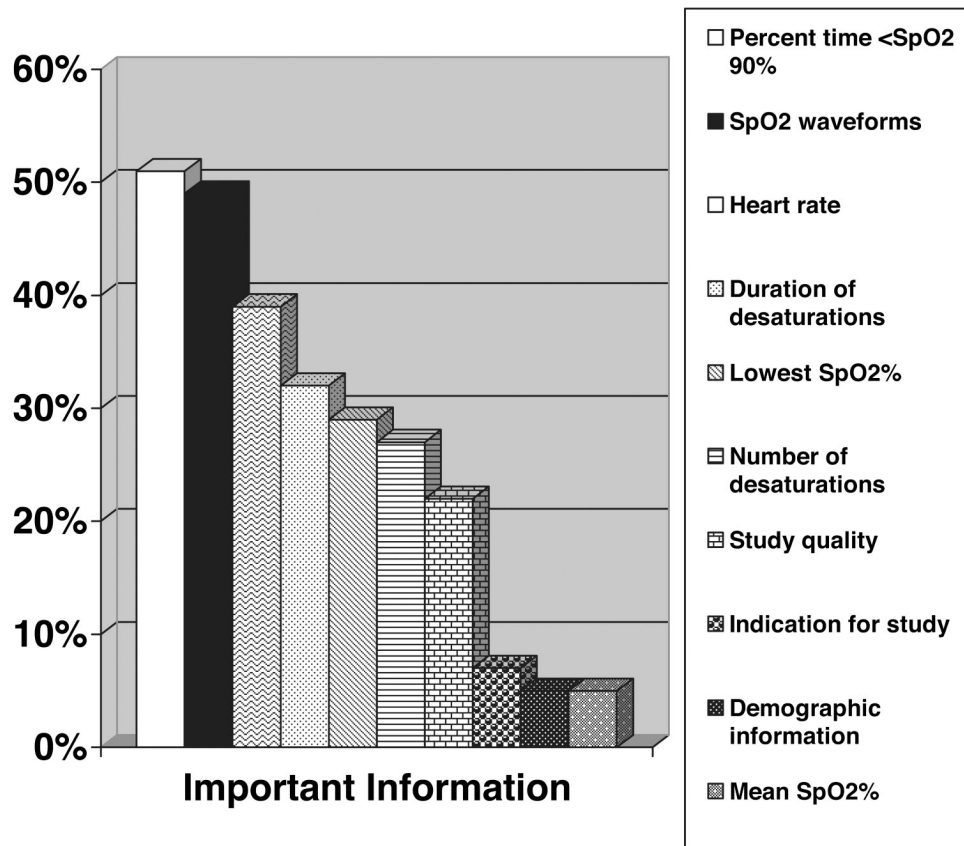


Figure 5. Physician participant opinion regarding "most important" information for overnight oximetry interpretation.

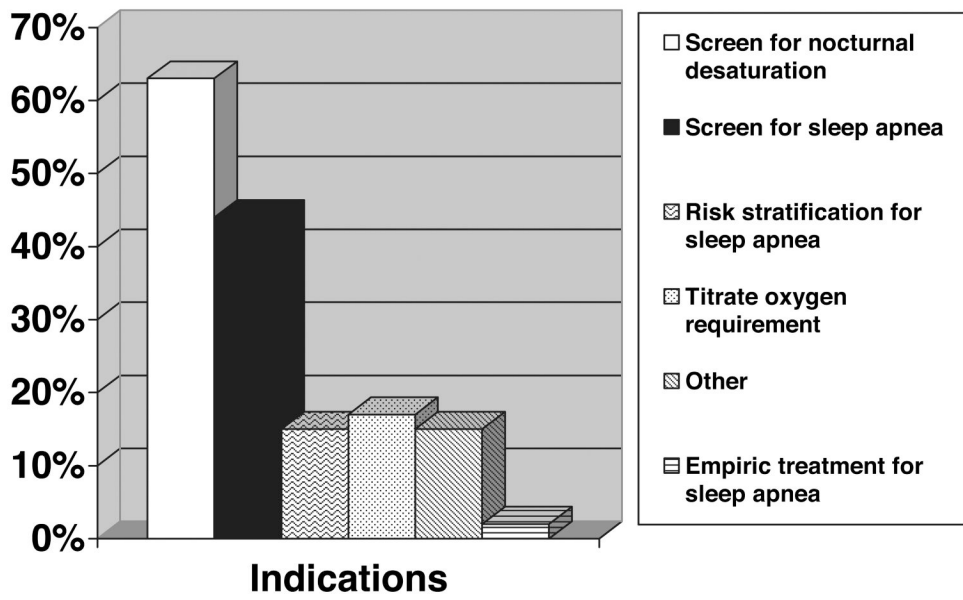


Figure 6.
Physician participant opinions on overnight oximetry indications.

Table 1
Information Considered by Physician Participant for Interpretation in Two of Three Oximetry Records

Variables	Information Used, % of Physicians
Introduction	
Background information	76
Indications	7
Waveforms	
Sp _o ₂ waveform	80
Sp _o ₂ waveform pattern	66
Heart rate waveform	44
Timing of heart rate pattern	27
Heart rate waveform pattern	5
Numeric summary	
Percentage of recording time with Sp _o ₂ < 90%	80
Lowest Sp _o ₂	46
Recording time	44
Number of desaturation events	37
Duration of desaturation events	27
Pulse rate	17
Mean Sp _o ₂	15

Table 2
Variables Perceived To Affect Overnight Oximetry Interpretation or Recommendations

Variables	Agree, %
Witnessed apneas	93
Heart disease	90
Lung disease	90
Study length	90
Sleepiness	88
Heart rate variability	80
BMI	80
Hypertension	63
Smoking	61
Gender	20
Race	2