Meta-analysis of arterial oxygen saturation monitoring by pulse oximetry in adults

Jensen L A, Onyskiw J E, Prasad N G

Authors' objectives
To describe the aggregated strength of the relationship of arterial oxygen saturation as measured by pulse oximetry with the standard of arterial blood gas analysis as measured by co-oximetry; to examine how various factors affect the relationship; and to describe an aggregate estimate of the bias and precision between oxygen saturation as measured by pulse oximetry and the standard in-vitro measures.

Searching
Computerised and citation indexes of MEDLINE, EMBASE, HealthSTAR, and CINAHL were searched from 1970 to 1995 for articles published in the English language. Bibliographies were reviewed. Published abstracts were retrieved when a published report was not found.

Study selection
Study designs of evaluations included in the review
Studies on pulse oximetry in adults were included if they fulfilled the following criteria: quantitative analysis of empirical data, bivariate correlations, and bias and precision estimates between pulse oximeter and co-oximeter estimates. The repeated measures design was the most frequently used.

Specific interventions included in the review
No specific inclusion criteria were reported. Arterial oxygen saturation was measured using 41 different oximeter models from 25 different manufacturers. Oximeters used ear, finger, forehead and multiple probes. Arterial blood gas was analysed by 17 different models of co-oximeter. Studies were conducted in laboratories, intensive care units, various hospital units and operating rooms.

Reference standard test against which the new test was compared
No inclusion criteria relating to the reference standard were specified. The review compared arterial oxygen saturation measured by pulse oximetry with that measured by co-oximetry.

Participants included in the review
Studies of adult populations were eligible for inclusion. The participants included healthy adult volunteers and a variety of hospital in-patients, such as respiratory patients, thoracic and cardiac surgery patients, critically ill patients, patients with more than one medical condition, individuals with sleep disorders and athletes. The participants' lowest level of oxygen concentration ranged from 36 to 96% (mean 70.29%, standard deviation, SD=15.59), while their highest oxygen concentration ranged from 70 to 100% (mean 98.53%, SD 4.29).

Outcomes assessed in the review
No inclusion criteria relating to the outcome measures were reported. The outcome presented in the review was the correlation between arterial oxygen saturation as measured by pulse oximetry and co-oximetry (considered the reference standard). Bias and precision estimates were also presented.

How were decisions on the relevance of primary studies made?
Two investigators independently assessed each retrieved study, with 100% agreement being required for inclusion.

Assessment of study quality
Validity was assessed using the following criteria: definition of predictor variable; definition of criterion variable; reliability of criterion variable; predictor and criterion variables measured reliably; measurements unbiased; sample size (data points); unbiased sample (homogeneity, confounding characteristics, stability); cross-validation studies; validity estimates (method, technique, accounts for confounding variable); and bias and precision estimates. Two investigators...
Independently rated the studies according to the 12 validity assessment criteria on a 3-point scale (acceptable, unacceptable, unable to assess). Criteria rated as acceptable were assigned a score of one. The scores were summed to determine a total quality score, giving a maximum score of 12. Any disagreements were resolved by consensus.

Data extraction
Methodological and substantive features of each study were coded and entered on a data collection form. Methodological features included year of publication, type of study, quality rating, sample size and number of data points. Substantive features included: type of study cohort; study setting; pulse oximeter model and probe location; co-oximeter and/or arterial blood analyser; range and mean of arterial oxygen saturation; bivariate correlations and/or bias and precision estimates between the pulse oximeter and co-oximeter values; and specific skin conditions that affect pulse oximetry accuracy (skin pigmentation, hypoxia, temperature, perfusion, dyshaemoglobinemia and hyperbilirubinemia).

Methods of synthesis
How were the studies combined?
The Hunter and Schmidt method was used to combine the studies and estimate a weighted mean correlation and a weighted mean bias (see Other Publications of Related Interest).

How were differences between studies investigated?
Heterogeneity of the correlation across studies was assessed by considering the amount of overall variance accounted for by the sampling error, as recommended by Hunter and Schmidt (see Other Publications of Related Interest). Subanalyses were conducted to investigate the effect of six factors on pulse oximetry accuracy: hypoxia, perfusion, dyshaemoglobinemia, temperature, skin pigment and hyperbilirubinemia. The mean correlation was estimated for the following subgroups: studies scoring 9 on the quality rating; decade of publication (1970s, 1980s, 1990s); type of participant (healthy adults, critically ill patients); and probe site (ear probe, finger probe).

Results of the review
Unweighted and weighted mean correlation were estimated using 39 studies (62 oximeter trials). Bias and precision estimates were calculated using 23 studies (82 oximeter trials).

The quality rating ranged from 3 to 11 (mean 8.0, SD 1.75). Lower ratings were due to a failure to report the reliability of criterion variable and a failure to conduct cross validation or include reliability estimates. Based on 39 studies (62 trials), 69% of oximeter trials tested finger probes, 23.7% ear lobes, 4.1% multiple probes and 0.6% forehead probes. Most of the participants were healthy adult volunteers (25.7%).

The mean correlation was 0.910 (variance 0.011) unweighted and 0.895 (variance 0.014) weighted. The absolute mean bias was 1.99% (SD=0.23). The variance caused by sampling error failed to account for most (75%) of the overall observed variance, and was considered to indicate that the correlations were not constant across the studies.

The mean correlation for studies scoring a quality rating of 9 (22 studies, 43 oximeter trials) was 0.908 (variance 0.011) unweighted and 0.883 (variance 0.016) weighted.

The correlation was 0.957 for healthy adult volunteers (13 studies, 318 participants, 32 oximeter trials), 0.950 for anaesthetised patients (1 study, 34 patients, 1 oximeter trial), 0.948 for athletes (2 studies, 21 athletes, 2 oximeter trials), 0.930 for thoracic surgical patients (2 studies, 15 patients, 2 oximeter trials), 0.904 for cardiac surgical patients (2 studies, 72 patients, 2 oximeter trials), 0.880 for respiratory patients (8 studies, 558 patients, 11 oximeter trials), and 0.760 for critically ill or intensive care unit patients (8 studies, 329 patients, 8 oximeter trials).

The mean correlation by probe location was: for the ear (3 studies, 3 oximeters), 0.938 (variance 0.002) unweighted and 0.934 (variance 0.001) weighted; and for the finger (3 studies, 3 oximeters), 0.963 (variance 0.001) unweighted and 0.967 (variance 0.001; P<0.0001) weighted.

The mean correlations for factors affecting pulse oximeter accuracy were as follows.
Hypoxia (5 studies, 15 oximeter trials; oxygen saturation: 67.6 to 87.8%): 0.924 (variance 0.008) unweighted and 0.938 (variance 0.006) weighted. The data showed substantial differences in bias and precision estimates between pulse oximeters at low saturation, the most common being an underestimation of saturation and falling precision.

Perfusion (3 studies, 3 oximeter trials): 0.717 (variance 0.049) unweighted and 0.582 (variance 0.004) weighted.

Dysaemoglobinaemia (5 studies, 6 oximeter trials; carboxyhaemoglobin: 5.87 to 9.10%): 0.817 (variance 0.028) unweighted and 0.717 (variance 0.035) weighted. There was a tendency to overestimate oxygen saturation with increasing carboxyhaemoglobin.

Temperature (3 studies, 3 oximeter trials; mean temperature: 28.6 to 34.8 degrees C): 0.760 (variance 0.043) unweighted and 0.665 (variance 0.024) weighted.

Skin pigment (1 study, 2 oximeter trials): 0.800 (variance 0.0002) unweighted and 0.800 (variance 0.0002) weighted.

Hyperbilirubinaemia (1 study, 1 oximeter trial; serum bilirubin: 2.7 to 35 mg/100 mL): 0.850 unweighted. Oxygen saturation was significantly underestimated.

The 21 pulse oximeters used in the studies were ranked by correlation with oxygen saturation. The correlations ranged from 0.986 for the Datascope Accusat finger/flex probe (2 oximeter trials, 25 participants; 245 data points) to 0.591 for the Ohmeda BIOX 111 finger/multiple probe (2 oximeter trials, 76 participants; 464 data points).

**Authors' conclusions**
Pulse oximeters were found to be accurate within 2% (1 SD) or 5% (2 SD) of in-vitro oximetry in the range of 70 to 100% oxygen saturation. When comparing ear and finger probes, readings from finger probes were more accurate. Pulse oximeters may fail to record accurately the true oxygen saturation during severe or rapid desaturation, hypotension, hypothermia, dysaemoglobinaemia, or low perfusion states.

**CRD commentary**
The objective of the review was clearly stated and the research question was partially defined by a limited set of inclusion criteria. The literature search involved several databases, but the search terms were not provided. By limiting the primary studies to those published in the English language some relevant studies might have been omitted. The methods used to select the studies for inclusion and assess validity were described. Validity was evaluated using 12 defined criteria. The results were clearly presented and heterogeneity was assessed statistically. A considerable number of variables were used to determine the effect of factors on the correlation of results from pulse oximetry with the standard measure.

The discussion considered the accuracy limits of pulse oximeters and potential sources of bias in the review: the uncertainty in measurements from pulse oximetry and the ‘gold’ standard of arterial blood samples; variation in the co-oximeters used in the primary studies; a lack of reporting the validity of the criterion measure; the technological and physiological limitations affecting the accuracy of pulse oximetry; instrument error; missing data and the lack of reporting the sample size and number of data points; and no studies aggregated bias and precision estimates.

The authors' conclusions are supported by the evidence presented, but should be considered with reference to the limitations described.

**Implications of the review for practice and research**
Practice: The authors think that clinicians must clearly understand that pulse oximeters do have accuracy limitations and that they may fail to record accurately the true oxygen saturation during physiological extremes.

Research: The authors consider that further study, to determine the impact on pulse oximetry accuracy of condition such as circulatory compromise from hypotension and vasoactive drugs, is needed.
Bibliographic details

PubMedID
9835670

Other publications of related interest

Indexing Status
Subject indexing assigned by NLM

MeSH
Adult; Clinical Trials as Topic; Female; Humans; Male; Monitoring, Physiologic /methods; Oximetry /methods /standards; Oxygen /blood; Oxygen Consumption /physiology; Reference Values; Reproducibility of Results; Sensitivity and Specificity

AccessionNumber
11998002122

Date bibliographic record published
31/03/2005

Date abstract record published
31/03/2005

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.