Home diagnosis of sleep apnea: a systematic review of the literature; an evidence review cosponsored by the American Academy of Sleep Medicine, the American College of Chest Physicians, and the American Thoracic Society


CRD summary
This review assessed the accuracy of various types of portable monitors for the diagnosis of sleep apnoea. The authors concluded that the best evidence was for type 3 monitors in sleep laboratory attended settings. The review had a number of methodological limitations, particularly in the literature search, selection criteria and synthesis of the results, making it impossible to draw reliable conclusions.

Authors' objectives
To assess the utility of portable monitors in diagnosing sleep apnoea in adults.

Searching
MEDLINE was searched from June 1997 to December 2001 for articles published since a previous review of the topic (see Other Publications of Related Interest); the search terms were reported. Studies included in the previous review did not appear to have been included in this review. The bibliographies from two American Sleep Disorder Association reviews and the reference lists of included studies were also checked for other relevant studies. Only studies published in the English language were eligible for inclusion. Reviews, meta-analyses, case reports, abstracts, letters and editorials were excluded.

Study selection
Study designs of evaluations included in the review
Studies that included at least 10 patients in each analysis group (the authors did not define what they meant by analysis group) after completion of the study were eligible for inclusion.

Specific interventions included in the review
Studies that used a portable monitor for the diagnosis of sleep apnoea were eligible for inclusion. Although not specifically defined as exclusion criteria, studies that evaluated older monitors for which more recent research had been published, devices known to be no longer commercially available or not widely used or available, or devices that did not use technology involved with monitoring a physiologic signal were excluded. The monitors evaluated in the included studies were classified as type 2 (comprehensive portable polysomnography), type 3 (modified portable sleep apnoea testing) or type 4 (continuous single or dual bioparameter) monitors. The studies were carried out at home (unattended) or in sleep laboratories (attended). The included studies used a variety of definitions of abnormal breathing event, the most common of which were a reduction in airflow and oxygen desaturation.

Reference standard test against which the new test was compared
Studies that used polysomnography or another acceptable objective test (the authors did not define what was considered acceptable) were eligible for inclusion. The review appeared to assess the ability of monitors to detect an abnormal apnoea/hypopnoea index (AHI).

Participants included in the review
Studies of male or female patients aged at least 18 years with any diagnosis of obstructive sleep apnoea were eligible for inclusion. Although not defined as an inclusion criterion, the authors stated that a study that tested patients following surgical intervention was excluded. The included studies were in patients with an AHI ranging from 10 to 51.

Outcomes assessed in the review
No inclusion criteria relating to the outcomes were reported. The outcomes assessed in the review included sensitivity,
specificity and likelihood ratios (LRs). The review also assessed monitor failure rates.

How were decisions on the relevance of primary studies made?
The authors did not state how the studies were selected for the review, or how many reviewers performed the selection.

Assessment of study quality
Study quality was assessed by rating studies according to 10 separate features. Articles were categorised as level I, II, III or IV, based on three of these items: blinded comparison, consecutive patient sampling, whether the reference standard was performed on all patients. They were also categorised for quality as a, b, c or d, based on the remaining items: prospective patient recruitment, random order of testing, low data loss (<10%), high percentage completed study (>90%), polysomnography methodology or definitions fully described, portable monitor methodology or definitions fully described, and portable monitor scoring fully described.

Two reviewers independently rated the studies for quality. A final evidence level and quality rating was determined by consensus of the evidence review committee.

Data extraction
Two reviewers independently extracted the data from each study. Any disagreements were resolved through consensus. Many studies reported data on sensitivity and specificity for multiple thresholds. For the assessment of the ability of monitors to reduce the probability that a patient had sleep apnoea, values for the best reported sensitivity were extracted. If various combinations of sensitivity and specificity were reported and two sensitivities were close in value, then the 'best reported sensitivity' was taken as the value with the higher corresponding specificity. Data relating to an AHI definition of less than or equal to 15 were preferentially selected when reported. For the assessment of monitor ability to increase the probability of sleep apnoea, best reported specificity values were extracted.

Methods of synthesis
How were the studies combined?
The studies were grouped under the following headings: evidence that monitors reduce the probability of the patient having an abnormal AHI; evidence that monitors increase the probability of the patient having an abnormal AHI; and evidence that a single monitor can both reduce and increase the probability of the patient having an abnormal AHI. The studies were then grouped by type of monitor and study setting and combined in a narrative.

How were differences between studies investigated?
Studies with high levels of evidence and high quality were discussed separately.

Differences between the studies were discussed with respect to disease severity of the patients, co-morbid conditions, gender or race, and patient recruitment.

Results of the review
Fifty-one studies (n=5,901) were included.

Evidence that monitors reduce the probability of the patient having an abnormal AHI.

Type 2 monitors.

Only one small study in a sleep laboratory attended setting was rated as evidence level II. This reported a negative LR of 0.22 with wide confidence intervals (CIs). Two other studies were poorer quality.

Type 3 monitors.

Eight of 9 studies in a sleep laboratory attended setting were rated as evidence level I or II; all had low LRs (0 to 0.15) and a small percentage of false-negatives (4 to 8%). Two of 4 studies in a home unattended setting were rated as
evidence level II; both had low LRs and one had a substantial percentage of false-negatives.

Type 4 monitors.

Thirteen of 16 evidence level I or II studies in a sleep laboratory attended setting had low LR s (0.15 or less) and most had a low percentage of false-negatives. Four of 8 studies in a home unattended setting were rated as evidence level I or II. Two level II studies had modestly low LRs, whilst false-negative rates across all 4 studies ranged from 3 to 34%.

Evidence that monitors increase the probability of the patient having an abnormal AHI.

Type 2 monitors.

The only study in a sleep laboratory attended setting that was rated as evidence level IIa reported high specificity (90%) and a high LR for a positive result (LR=8). False-positives ranged from 4 to 11%. Two other studies were of poor quality. The only study in a home unattended setting was of poor quality.

Type 3 monitors.

Eight studies in a sleep laboratory attended setting were rated as evidence level I or II and one was rated level IV. Most had high specificity (>90%), high sensitivity and high LRs for a positive result. Two of 4 studies in a home unattended setting were rated as evidence level II and both had modest LRs (5.1 and 9); the false-positives ranged from 2 to 31% across all studies.

Type 4 monitors.

Sixteen of 25 studies in a sleep laboratory attended setting were rated as evidence level I or II. LRs for positive results were 5 or more in 11 studies, >10 in nine, and >20 in three. Studies with LRs of 10 or more reported small false-positive rates (0 to 12%). Four of 8 studies in a home unattended setting were rated as evidence level I or II; three had LRs >10 with false-positive rates of 0 to 12%.

Evidence that a single monitor can both reduce and increase the probability of the patient having an abnormal AHI.

Type 2 monitors.

Only 2 evidence level IV studies were found.

Type 3 monitors.

All 9 of the sleep laboratory attended studies produced both high positive and low negative LRs. Eight of these studies were rated as evidence level I or II and only one had patients who were not classified as positive or negative; most studies had misclassification rates of about 5%. Two of 4 studies in a home unattended setting produced both high positive and low negative LRs. Both were rated as evidence level II, both used different thresholds to produce low and high LRs, and misclassification rates were 5% and 16%.

Type 4 monitors.

Fifteen of 25 sleep laboratory attended studies produced both high positive and low negative LRs. Nine were evidence level I or II. Ten studies used different thresholds to produce low and high LRs. Misclassification rates were low (<7%). One study in a home unattended setting classified some patients as having or not having sleep apnoea; half of the patients were not classified and the misclassification rate was 4%.

Other results were also reported.

Cost information
A limited number of studies reported on cost. These generally reported that the cost of monitoring with type 2, 3 and 4 monitors was less than that of standard polysomnography.
Authors' conclusions
Most evidence for the use of monitors in classifying patients as having or not having sleep apnoea came from high-quality studies of type 3 monitors in the sleep laboratory attended setting. These studies had low false-positive rates and most studies found a threshold that distinguished patients with sleep apnoea from those without.

CRD commentary
This was a poorly conducted and presented review. The review assessed a fairly broad objective and the inclusion criteria were not well defined. The literature search was limited to one electronic database and only studies published in English were eligible for inclusion, thus raising the possibility of publication and language bias. The search terms were narrow and searches were only carried out from 1997, when a previous review was conducted. However, neither the studies included in the previous review, nor its results, were included in the current review; thus, relevant studies are likely to have been missed. Methods were used to minimise errors and bias in the validity assessment and data extraction, but it was unclear whether similar steps were taken in the study selection process. A formal quality assessment was carried out, but this involved grading studies on the basis of whether they fulfilled a number of items rather than reporting which individual items each study fulfilled, which would have been more informative.

Although some details of the studies were presented in tabular format, the tabulation of additional details on test and population, for example, would have been helpful. The results were summarised in a narrative synthesis, but this was very difficult to follow. In view of these limitations, the authors' conclusion may not be reliable.

Implications of the review for practice and research
Practice: The authors did not state any implications for practice.

Research: The authors made many recommendations for future research. For example, the need to study more diverse populations; to address study methods (including potential bias from funding sources, definition of hypopnea, recruitment, comorbidities, adequate sample size, description of population, details about performance and scoring of polysomnogram and portable monitor, scoring of monitor); and to address study design factors.

Funding
American Academy of Sleep Medicine; American College of Chest Physicians; American Thoracic Society.

Bibliographic details

PubMedID
14555592

DOI
10.1378/chest.124.4.1543

Original Paper URL

Other publications of related interest
Indexing Status
Subject indexing assigned by NLM

MeSH
Biomedical Research; Equipment Design; Humans; Infant; Male; Monitoring, Ambulatory /instrumentation; Polysomnography /instrumentation; Reproducibility of Results; Research Design; Severity of Illness Index; Sleep Apnea Syndromes /complications /diagnosis; Societies, Medical

AccessionNumber
12003006813

Date bibliographic record published
31/01/2007

Date abstract record published
31/01/2007

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.