# Assessment of predictive ability of Epworth scoring in screening of patients with sleep apnoea

A SIL, G BARR

*Department of ENT, Monklands Hospital, Airdrie, UK*

## Abstract

*Method*: Numerous studies have considered the benefits, and the disadvantages, of the Epworth Sleepiness Scale. Following an extensive literature review, we found that the evidence was inconclusive as regards the diagnostic efficacy of Epworth scoring for obstructive sleep apnoea syndrome. We undertook a retrospective study of 343 patients who underwent a sleep assessment over a 10-year period at the Monklands Hospital.

*Analysis and results*: A total of 238 patients did not have sleep apnoea whereas 105 patients did. The mean Epworth score in patients with obstructive sleep apnoea syndrome was 10.94 (95 per cent confidence interval 9.46–11.42), and in the non-apnoeic group it was 7.73 (95 per cent confidence interval 7.04–8.41). Logistic regression and receiver operating characteristic curves were used to assess the predictive ability of Epworth scoring. The scores only explained 7–10 per cent of the variation in the probability of occurrence of obstructive sleep apnoea syndrome. The odds ratio for Epworth scoring was 1.118, and only 69 per cent of cases were correctly classified by the Scale.

*Conclusions*: The literature review suggested that the Epworth Sleepiness Scale is associated with a low effect size and/or low predictive value when correlated or regressed on the Apnoea–Hypopnoea Index or Respiratory Disturbance Index, thus limiting its value as a screening test. Our study concluded that the Epworth Scale is only marginally useful in predicting the occurrence of obstructive sleep apnoea syndrome. We believe that every patient with a direct or witnessed history of sleep apnoea with obstructive symptoms have some form of sleep assessment.

**Key words:** Obstructive Sleep Apnoea Syndrome; Snoring; Epworth Score; Diagnosis

## Introduction

The Epworth Sleepiness Scale was introduced by Johns in 1991. It assesses eight situations with a likelihood of falling asleep, and assigns scores from zero to three, giving a total scale of 0 to 24.[1] It has gained huge popularity and has been translated into seven languages, and is used all over the world for assessment of daytime sleepiness and sleep apnoea.

There have been many studies of the association between Epworth scoring and sleep apnoea. Several studies have assessed the association between the Epworth Scale and various snoring-related parameters, such as the Apnoea–Hypopnoea Index (AHI), Mean Sleep Latency Test and Mean Wakefulness Test. However, no study to our knowledge has analysed the cause–effect association between Epworth scoring and sleep apnoea on a one-to-one basis, without introducing other confounding factors.

Introducing other parameters into the association has the potential to introduce bias and problems related to intercorrelations between the independent variables, unless the confounding factors have been controlled for.[2] Problems associated with such covariates have rarely been addressed in the literature. It is therefore not surprising that the evidence in the literature is inconclusive with regard to the diagnostic value of the Epworth Scale in patients with sleep apnoea.

It is imperative that the association between the Epworth Scale and obstructive sleep apnoea syndrome (OSAS) should be assessed on a one-to-one basis, in order to assess the efficacy of the Scale in the screening of patients presenting with sleep apnoea. With this objective in mind, we carried out a retrospective analysis using data from the sleep study database at Monklands Hospital.

## Materials and methods

The first part of the project involved a systematic literature review. We searched the Cochrane, DARE (Database of Abstracts and Reviews of Effects),

EMBASE (Excerpta Medical Database), CINAHL (Cumulative Index to Nursing and Allied Health Literature) and Medline (1996 onwards) databases. In addition, we searched for established guidelines, systematic reviews and evidence-based summaries. The search terms used were 'Epworth', 'snoring' and 'apnoea'.

A total of 107 English language studies were identified, including eight Cochrane reviews and three systematic reviews. Each of these papers was studied for any demonstration of an association between the Epworth Sleepiness Scale and sleep apnoea. Only 16 studies were suitable for inclusion in our review. The results of these studies were categorised as type one (a positive association between Epworth scoring and sleep apnoea) and type two (no significant association between Epworth scoring and sleep apnoea). These studies are detailed in Table I.

Each study was reviewed and, wherever possible, the effect size of the Epworth Scale in relation to the AHI or Respiratory Disturbance Index was calculated using Cohen's definition.[18] In studies that used a correlation coefficient between the Epworth scores and the AHI or the Respiratory Disturbance Index, the effect size (unweighted) was the magnitude of the correlation coefficient itself.

Cohen's definition of small, medium and large effect sizes was used to identify effect sizes in these studies. In studies in which comparison of mean Epworth scores was undertaken, the effect size was defined as a standardised mean difference between the groups.[18] RevMan version 5 software (Cochrane Incorporated) was used to compute effect sizes in the various studies and to perform a partial meta-analysis.

In some studies, categorisation of patients by AHI or Respiratory Disturbance Index values (for comparison with Epworth scores) resulted in more than two groups. In these studies, the group with the highest mean and standard deviation values was compared with the group with the lowest mean and standard deviation values, in order to estimate the effect size as described by Cohen.[18]

A Forest plot was generated using RevMan software, and this plotted effect sizes and 95 per cent confidence intervals (CI) of the effect size (Figure 1).

The sleep study database of Monklands Hospital was used to extrapolate the details of the patients who underwent a sleep study at our centre during a 10-year period from 1998 to 2008. All patients referred with loud, constant snoring, including those with possible sleep apnoea, underwent a sleep study. The sample size was 343 patients.

All patients had a basic sleep assessment in the form of pulse oximetry and periodic observation. Patients having more than two desaturations per hour underwent five-channel home polysomnography. Evidence in the current literature uses a more stringent form of reference, which directs that only patients with more than five desaturations per hour should undergo full polysomnography.[19] In Woodhead and colleagues' study, the presence of OSAS was taken to be suggested by more than 10 falls in saturation of more than 3 per cent from the baseline.[20] Our low threshold for referral ensured the least likelihood of a false negative result as regards identification of sleep apnoea. We used a validated definition[21] of the AHI, involving more than 15 episodes per hour, as the cut-off point for defining sleep apnoea in a patient.

The Epworth score and presence or absence of sleep apnoea were recorded for each patient, using the definitions mentioned in the above paragraph. Various statistical measures were used to analyse the association between the two results, using no other variable apart from the Epworth score to analyse the predictability of sleep apnoea.

The first part of the analysis compared the mean Epworth score in the group with sleep apnoea with that in the group without sleep apnoea. Student's t-test was used to compare the mean values. A logistic regression analysis was used to assess the predictability of the Epworth score with regard to the probability of occurrence of sleep apnoea measured objectively using odds ratios. For the purposes of validation and illustration of results, a receiver operating curve analysis was used to determine any association between the Epworth score and the probability of OSAS occurrence.

## Results and analysis

### Literature review

The first part of the study involved a literature review. This included a large number of studies using Epworth scoring. We found 16 studies directly investigating the association between the Epworth score and OSAS. Studies that involved evaluation of continuous positive airway pressure treatment were not included in our analysis. Five of the 16 studies found a significant association between Epworth score and OSAS (Table I) (i.e. a type one result). Eleven of the 16 studies showed no significant association between Epworth score and OSAS (i.e. a type two result) (Table I).

Significance was determined by the statistical inference drawn by the study authors. The differentiation between type one and type two results was primarily based on the study authors' results and /or conclusions. Post hoc effect sizes were determined wherever possible using the criteria mentioned above. Two studies[3,5] used regression analysis with an outcome measure of OSAS as a dependent variable, and these were the studies in which the cause–effect relationship could be established. The regression coefficient in these studies was small (Table I).

In studies that compared the mean Epworth score among groups defined by the AHI and the Respiratory Disturbance Index, a uniform cut-off threshold for sleep apnoea was not present. Therefore,

## TABLE I
### STUDIES OF ESS *VS* SNORING PARAMETERS

| Study | Statistics used for association | Findings | ESS effect size (ESS with OSAS paremeters) | Significance of effect size with relation to ESS *vs* AHI |
|---|---|---|---|---|
| *Type 1 associations* | | | | |
| Herzog et al.[3] | Multiple linear regression | Regression coefficient of ESS score = 1.177 $R^2$ (full model) = 0.312 | Spearman's rho (ESS *vs* AHI) = 0.34 | Yes, causal effect established, weak regression coefficient of ESS |
| Kumar et al.[4] | Comparison of means | OSAS patients had higher ESS score | SMD = 1.85 | Yes, but causal effect not established |
| Montoya et al.[5] | Multiple linear regression & logistic regression | Regression coefficient of ESS = 0.1761 $R^2$ (full model) = 0.33 | ESS OR = 1.073 ESS $R^2$ = 0.065 | Yes |
| Gottlieb et al.[2] | Controlled comparison of means (ESS *vs* Respiratory Disturbance Index) + correlations | Significant increase in mean ESS score associated with degrees of OSAS ($p < 0.001$) | SMD = 0.40 | Yes, but causal relationship limited to snoring, & uncertain for OSAS |
| Ibrahim et al.[6] | Comparison of means & multivariate analysis for predictors of OSAS | Significant mean difference ($p < 0.001$) ESS was a significant predictor in multivariate analysis | SMD = 0.98 | Yes, but only patients with ESS >11 were included in multivariate analysis |
| *Type 2 associations* | | | | |
| Olson et al.[7] | Non-parametric correlations (ESS *vs* AHI) | Spearman's rho = 0.0002 (non-OSAS pts) Spearman's rho = 0.06 (OSAS pts) | $R = 0.0002$ (ESS/AHI) $R = 0.15$ (ESS/AHI), if MSLT <5 min | No |
| Chervin & Aldrich[8] | Regression with ESS as outcome variable | ESS did not correlate with AHI | 0.01 (ESS was dependent outcome variable) | No |
| Rosenthal & Dolan[9] | Comparison of sensitivities & specificities (ESS *vs* AHI) | Comparative study of sensitivities ROC curve used for predictive potential | Could not be ascertained AUC not given | Increase in sensitivity for ESS cut-off of 8 (76%), from ESS cut–off of 10 (66%) |
| Smith et al.[10] | Confirmatory factor analysis (2 model comparison) correlations between ESS & RDI | ESS not a good predictor of OSAS Weak correlation between 2 modified ESS scale models & PSG parameters | $R = 0.12$ | Yes, weak correlation between full model and AHI score Causal relationship uncertain |
| Fong et al.[11] | Comparison of means by ANOVA Grouping of pts by severity of OSA, guided by RDI Analysis followed by correlations | No significant difference in ESS scores MSLT varied significantly | $R = 0.046$ (ESS/RDI) | Yes (MSLT), but causal relationship not established between ESS and OSAS |
| Kingshott et al.[12] | Mokken scaling to assess reliability of ESS | 4 out of 8 items in ESS contributed to hierarchical scale | Not given | Yes, but ESS reliability assessed with regards to EDS and not OSAS Causal effect on OSAS uncertain |
| Kingshott et al.[13] | Correlation between ESS & AHI | No significant correlation | $R = 0.02$ | No, causal effect uncertain |
| Osman et al.[14] | Simple correlation between ESS & AHI Comparison of mean AHI events between 2 ESS-defined groups | No significant correlation | $R = 0.12$ | No, causal effect uncertain |
| Chung[15] | Non-parametric correlation & multiple regression using ESS as outcome variable | No significant association between ESS & AHI | Spearman's rho = −0.06 (ESS/AHI) | Causal effect established with MSLT, not with AHI |
| Thong & Pang[16] | Correlation between ESS & AHI | Weak correlation between ESS & AHI | $R = 0.1$ | No, causal relationship uncertain |
| Sauter et al.[17] | Comparison of various EDS measurement scales | No simple correlation between PSG parameters & EDS extent | Could not be ascertained | No |

ESS = Epworth Sleepiness Scale; AHI = Apnoea–Hypoponea Index; OSAS = obstructive sleep apnoea syndrome; SMD = standardised mean difference; OR = odds ratio; MSLT = Mean Sleep Latency Test; min = minute; ROC = receiver operating characteristic; AUC = area under the curve; RDI = Respiratory Disturbance Index; PSG = polysomnogram; ANOVA = analysis of variance; EDS = Excessive Daytime Sleepiness

this differentiation was made based on the authors' recommendations in each individual study.

Six of the 16 studies showed a weak correlation between the Epworth score and the AHI; two studies used regression analysis with the AHI as the outcome variable, and the Epworth Scale was found to have small regression coefficients. Two studies found significant correlations between the Epworth score and
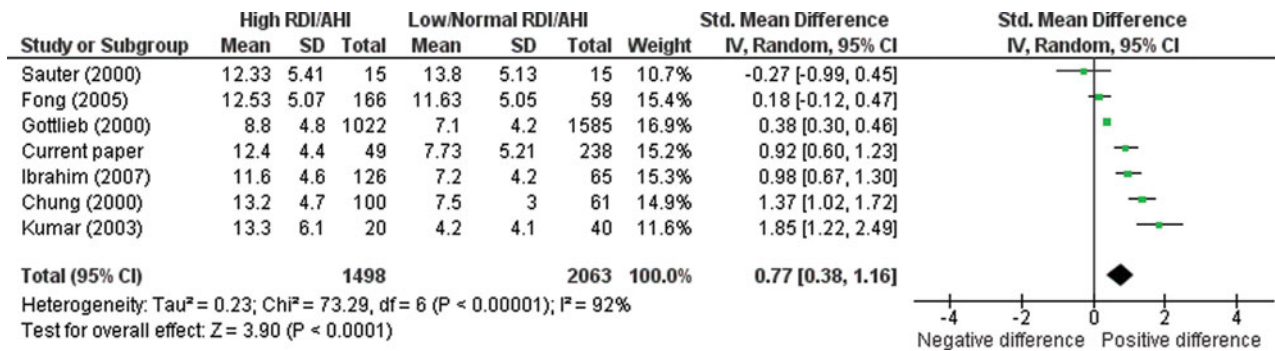
| Study or Subgroup | High RDI/AHI Mean | SD | Total | Low/Normal RDI/AHI Mean | SD | Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|---|---|
| Sauter (2000) | 12.33 | 5.41 | 15 | 13.8 | 5.13 | 15 | 10.7% | -0.27 [-0.99, 0.45] |
| Fong (2005) | 12.53 | 5.07 | 166 | 11.63 | 5.05 | 59 | 15.4% | 0.18 [-0.12, 0.47] |
| Gottlieb (2000) | 8.8 | 4.8 | 1022 | 7.1 | 4.2 | 1585 | 16.9% | 0.38 [0.30, 0.46] |
| Current paper | 12.4 | 4.4 | 49 | 7.73 | 5.21 | 238 | 15.2% | 0.92 [0.60, 1.23] |
| Ibrahim (2007) | 11.6 | 4.6 | 126 | 7.2 | 4.2 | 65 | 15.3% | 0.98 [0.67, 1.30] |
| Chung (2000) | 13.2 | 4.7 | 100 | 7.5 | 3 | 61 | 14.9% | 1.37 [1.02, 1.72] |
| Kumar (2003) | 13.3 | 6.1 | 20 | 4.2 | 4.1 | 40 | 11.6% | 1.85 [1.22, 2.49] |
| **Total (95% CI)** | | | **1498** | | | **2063** | **100.0%** | **0.77 [0.38, 1.16]** |

Heterogeneity: Tau² = 0.23; Chi² = 73.29, df = 6 (P < 0.00001); I² = 92%
Test for overall effect: Z = 3.90 (P < 0.0001)

FIG. 1

Forest plot showing effect sizes as standard mean difference and 95 per cent confidence intervals, for studies using comparison of means. RDI = Respiratory Disturbance Index; AHI = Apnoea–Hypopnoea Index; IV = Inverse Variance method; SD = standard deviation; CI = confidence interval

the Mean Sleep Latency Test, but concluded that the latter was more useful in predicting OSAS.

A formal, complete meta-analysis was not performed as the data and outcome measures in all the studies could not be pooled. This was primarily because of the use of a combination of parametric and non-parametric measures to investigate the correlations between the Epworth score and the AHI.

Effect sizes were calculated wherever possible, and are shown in Figures 1 and 2. It was possible to perform a partial meta-analysis, in order to determine the effect size, in those studies that used comparison of mean Epworth scores among patient groups. Standardised mean difference was used to calculate the effect size.

Weighting of effect sizes was done by the inverse variance method, and a random-effect model was applied for meta-analysis, which assumes that the effect size was not uniform across the variable populations studied. This was statistically justified as the chi-square test for heterogeneity was significant ($p < 0.01$). However, this may also be clinically explained by the lack of objectivity in the Epworth Scale scores across the population samples. This may partially be due to the lack of a uniform AHI cut off threshold to define sleep apnoea.
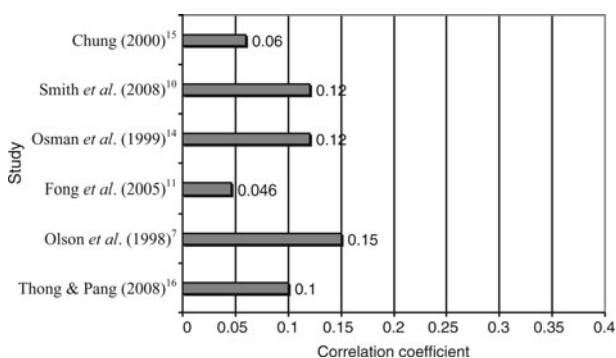
The total effect size was 0.77, which is deemed to be large by Cohen[18] (Figure 1). In the case of correlations, Cohen described a correlation coefficient of up to 0.2 as a low effect, up to 0.4 as moderate, and greater than 0.4 as a large effect.[18]

*Study*

In our study, 105 patients had OSAS whereas 238 patients did not (Figure 3), and the mean Epworth score was 8.75 (95 per cent CI 8.30–9.19). The sample of 343 patients had a representative age distribution (median age 45 years).

The Epworth scores were normally distributed, as suggested by the Kolmogorov–Smirnov test ($z = 1.22$, $p = 0.103$). The mean Epworth score was 10.94 (95 per cent CI 9.46–11.42) in the OSAS group and 7.73 (95 per cent CI 7.04–8.41) in the non-OSAS group (Figure 4). The difference in mean values was statistically significant (Student's $t$-test, $p = 0.003$).

Logistic regression was used to assess the predictive potential of the Epworth score as regards the probability of OSAS occurrence. Table II shows the mean Epworth scores in the various groups with sleep apnoea as defined by the AHI. It is however worth noting that, for the purposes of analysis, the 'mild'

FIG. 2

Bar chart showing correlation coefficients (Pearson's *R* or Spearman's rho) for studies correlating Epworth Sleepiness Scale with the Apnoea–Hypopnoea Index or the Respiratory Disturbance Index.
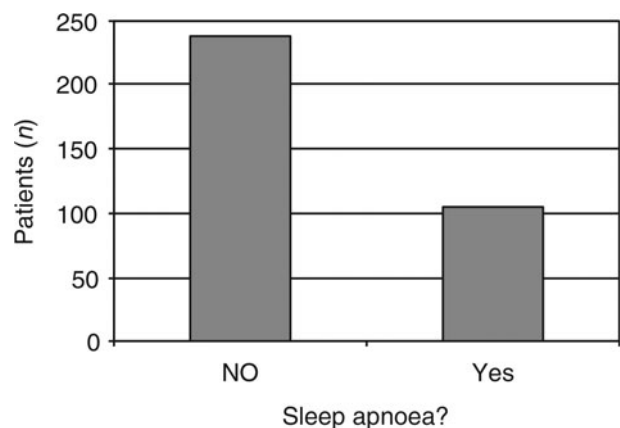
FIG. 3

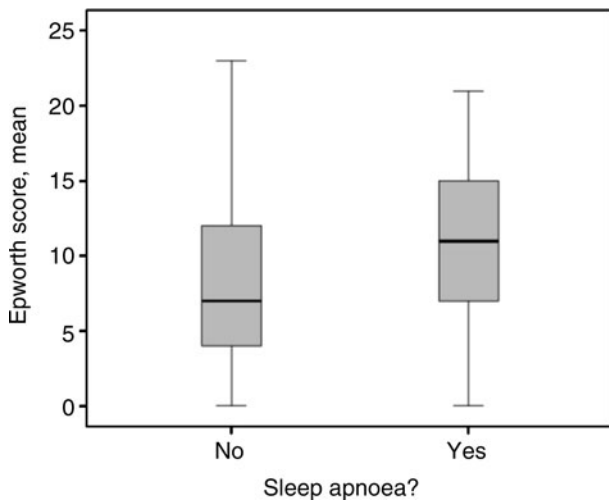Bar chart showing patients with and without sleep apnoea.

FIG. 4

Box plot showing mean Epworth scores for patients with and without obstructive sleep apnoea syndrome. The difference was statistically significant ($p = 0.003$).
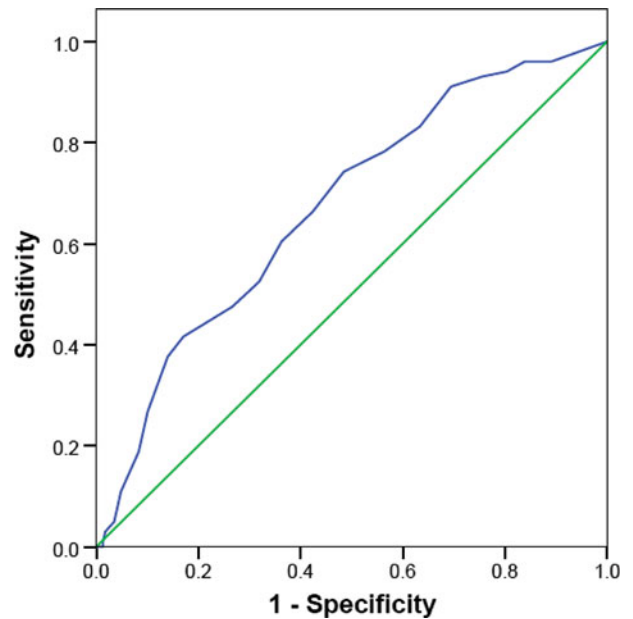


FIG. 5

Receiver operating characteristic curve showing the relationship between the sensitivity and the false positive rate (i.e. 1 − specificity). Blue plot = sensitivity $vs$ (1 − specificity); green line = reference line

group (with an AHI of 15 or less) was not included in the group defined as having sleep apnoea. This approach has been validated in recent literature.[21] The goodness of fit statistic (using the Hosmer and Lemeshow test) proved that the analysis model fitted the data well ($p = 0.455$).

The Epworth scoring was able to explain 7 to 10 per cent of the variance in the probability of occurrence of OSAS (Cox and Snell $R^2 = 0.07$, Nagelkerke $R^2 = 0.10$), with only 69.4 per cent of the cases being correctly predicted by the model. The odds ratio for the Epworth score was 1.118 (95 per cent CI 1.068–1.171). This implied that when the Epworth score increased by one point, the odds of occurrence of OSAS increased nearly 1.12 times.

A receiver operating curve (Figure 5) was used to analyze the predictive potential and to validate the results. Essentially, this is a graph that plots sensitivity on the y axis and the false positive rate (1-specificity) on the x axis, to determine the association between sensitivity and the false positive rate. Any curve above the reference line shown in Figure 5 implies a positive association. Calculating the area under the curve is a good way of determining the predictive potential.[23] An area under the curve of 0.5 to 0.7 implies a marginally useful test, 0.7 to 0.9 a good test and greater than 0.9 an excellent test.[23]

The area under the curve for the Epworth score was 0.672, making it only a marginally useful test. This also implies that Epworth scoring would only be able to correctly predict the occurrence of OSAS in 67.2 per cent of cases. This is close to the estimation made by logistic regression analysis, whereby nearly 69 per cent of cases were correctly classified, thus validating the analysis.

## Discussion

Numerous studies have assessed the association of the Epworth Sleepiness Scale with OSAS, but few have compared the effect sizes of the relative groups. This has led to confusing conclusions, many made indirectly. Our study analyses indicate that, although the mean Epworth score was significantly greater in the OSAS group, this does not translate into a clinically significant cause effect association due to the low effect size of the Epworth score. This is a unique conclusion which contrasts with other authors' interpretations of their findings.

The studies we reviewed used various statistical measures, and a large number of variables were included in their analyses. This distracted attention from the primary question of whether the Epworth Sleepiness Scale is a good predictor of OSAS. Effect size estimation was not strongly emphasised in the studies we reviewed, despite the fact that this is the only measure that will indicate the strength of the relationship between Epworth scoring and OSAS. Moreover, with one exception (Smith *et al.*),[10] our review studies used inclusion criteria for their other variables which were neither sufficiently clear nor consistently validated on statistical grounds. In our study,

| TABLE II | | |
|---|---|---|
| EPWORTH SCORE BY AHI | | |
| AHI | Pts (*n*) | ESS (mean) |
| None (0–5) or mild (5–15) | 238 | 7.73 |
| Moderate (15–30) | 56 | 10.9 |
| Severe (>30) | 49 | 12.4 |

AHI = Apnoea–Hypopnoea Index; Pts = patients; ESS = Epworth Sleepiness Scale score

other variables were not added into the regression model, as this would have affected the intercorrelation between variables and subsequently altered the findings of our analysis.

To our best knowledge, ours is the only study to provide a dual validation of the association of Epworth scoring with OSAS on a one-to-one basis, along with effect size estimation. Our findings add to the growing body of evidence questioning the importance of the Epworth Sleepiness Scale as a tool in the evaluation of sleep apnoea.

A large number of ENT centres depend heavily on Epworth scoring to screen patients with sleep apnoea.[7,24,25] Use of the Epworth Scale as a screening tool has been justified in a recent study that suggested that only patients with an Epworth score of more than 11 should be referred for a sleep study.[25] Although the Epworth Scale is a relatively simple scoring system and can be easily completed in the out-patients department, relying solely on Epworth scores for OSAS screening could exclude a significant proportion of patients with OSAS (Table III).

Since the first introduction of the Epworth Sleepiness Scale by Johns[1] in 1991, there have been numerous studies on its validation. Johns used a rigorous exploratory factor analysis to justify the inclusion of the eight points in the analysis. A recent study by Smith and colleagues,[10] however, refutes Johns' inclusion criteria by means of a confirmatory factor analysis, which provides a more stringent validation of Johns' results. In addition, a six-point scale (instead of an eight-point scale) was found to have more favourable inter-reliability statistics.[10]

Various studies have shown that there is a substantial amount of confusion regarding the validity of the Epworth scale. Some good quality type one studies have relied upon comparison of mean Epworth scores in patients with and without sleep apnoea. This difference has been noted to be significant, and the same result is reproduced in our study. In our patients, the mean Epworth score was 7.73 in the non-OSAS group and 10.73 in the OSAS group. This difference was statistically significant ($p = 0.003$).

However, such a comparison is not in itself a reliable method of establishing a cause–effect relationship, as it suggests that every patient with OSAS is likely to have a high Epworth score, but this does not mean that any patient with a high Epworth score is likely to have OSAS. There are many other reasons why a patient may have daytime somnolence and fatigue. Conversely, a low Epworth score does not exclude sleep apnoea.

Thus, in addition to a simple comparison of means, we must determine the variance in the probability of the occurrence of OSAS that is accounted for by the Epworth score. This was 7 to 10 per cent in our study, suggesting that close to 90 per cent of the variance was due to other factors. The confounding effect of these covariates has rarely been taken into account in other studies.

Effect size is important in showing the strength of a relationship. In our study, this was represented by the odds ratio. The odds of having OSAS went up 1.118 times if the Epworth Sleepiness Scale score went up by one unit. Although this is significant (95 per cent CI 1.068–1.171), the effect size is very low as an odds ratio of one would be consistent with no effect at all.

Figure 1 shows the results of our partial meta-analysis involving studies using mean value comparisons. The total effect size was 0.77, suggesting that one could expect a difference of approximately 80 per cent of a standard deviation between the mean Epworth scores of OSAS and non-OSAS patients. This effect size may be statistically significant, but may be inappropriate for justifying Epworth scoring as a screening test, since the large number of false negatives generated may be considered unacceptable (Table III).

The large effect size was also to some extent reflected in the significant difference in the Epworth scores of OSAS and non-OSAS patients in our study. However, by simply comparing the mean values, the effect of other variables is not included. Since the various studies used different cut-off points to define OSAS, this itself has the potential to introduce an inclusion bias towards larger effect sizes, as some studies included healthy volunteers whereas others only included patients with OSAS. Therefore, one can expect the actual effect size to be smaller, as seen in Figure 2, which shows the correlation coefficients of various studies.

Two studies that used a comparison of mean Epworth scores (Kumar *et al.*[4] and Chung *et al.*[15]), reported relatively large effect sizes. However, in the former the error margin was large (wide confidence intervals), and in the latter a selection bias cannot be ruled out as the 'control' group were people who worked as associates or colleagues in the same institution as the principal investigator. If these studies were excluded from our analysis, the effect size would reduce further.

An interesting observation is the low effect size in terms of the correlation coefficient reported in a few

studies that also reported a standardised mean difference showing a moderate effect. The possible explanation could be the presence of other compounding variables that would lead to a low effect size while measuring correlation. Figure 2 shows the universally low effect sizes obtained for correlations between Epworth scoring and AHI or Respiratory Disturbance Index parameters.

The literature abounds with descriptions of confounding factors presumed to affect the Epworth Sleepiness Scale. Body mass index (BMI) has been shown to be a predictive factor for hypertension in an Asian population.[26] This identifies BMI as one of the many Epworth Scale confounding factors. Another study[27] showed that predictive factors such as age and subjective sleep history, among others, accounted for nearly 12.8 per cent of the observed variance, and reported that almost 33 per cent of subjects had discrepant Epworth scores. Again, this shows the importance of confounding factors.

Some studies used a simple comparison of means to justify the predictive ability of the Epworth Sleepiness Scale.[4,28] However, this does not imply a cause–effect relationship. A few studies[7,13,14,16] used correlation coefficients to study Epworth Scale associations; however, this method does not establish a certain cause–effect relationship either.

Interestingly, only two studies[3,5] used regression analysis, both of which found the regression coefficient and/or the correlation coefficient to be significant but small. A standardised regression coefficient generated by a regression analysis using many predictor variables along with the Epworth score is a measure of a semi-partial correlation between the Epworth score and a measure of the dependent variable (OSAS).

There is no debate that the Epworth Sleepiness Scale is a good, simple, easy screening method for excessive daytime sleepiness, and is also useful in the prediction of certain parameters associated with daytime sleepiness and snoring. Excessive daytime sleepiness irrespective of OSAS is in itself associated with significant morbidity.[29]

There is an abundance of support for the validity of the Epworth Scale in current literature, including studies using a non-English version of the Scale.[30] The correlation between the Epworth Scale and the Mean Sleep Latency Test has been found to be not significant.[30]

A few of our type two studies did not assess the cause–effect relationship. Fong *et al.*[11] used analysis of variance to compare the mean Epworth scores of groups of patients graded according to OSAS severity. They found no significant difference between Epworth scores, and concluded that the Mean Sleep Latency Test varied significantly between these groups.

The importance of clinical assessment alone (including Epworth scoring) in detecting non-apnoeic snorers has been emphasised in the literature.[31] The sensitivity in this particular study was reported as more than 90 per cent. This implies that a group of simple snorers with high Epworth scores, and therefore excess daytime sleepiness, could be detected by clinical assessment alone, irrespective of their Epworth scoring. Apart from the fact that this requires very rigorous clinical history-taking, the subjective element of such an exercise cannot be ignored. This is adequately justified by Table III, which shows the number of apnoeic patients missed by Epworth Scale assessment. It has been suggested that the Epworth Sleepiness Scale is highly variable when administered sequentially to a population, suggesting that its reproducibility in the clinical setting is not optimum.[22]

Osman and colleagues[14] have suggested that the Epworth Sleepiness Scale is not a good predictor of OSAS, on the basis of a poor correlation observed between the Epworth score and the AHI. They believed this to be due to the fact that excessive daytime sleepiness can be present in simple snorers (caused by unknown mechanisms); this has also been mentioned by Gottlieb and colleagues.[2]

- **The Epworth Sleepiness Scale is only marginally useful in predicting the occurrence of obstructive sleep apnoea syndrome (OSAS)**
- **Epworth scoring should not be used alone to screen patients for OSAS**
- **Patients with loud, constant snoring and a history suggestive of OSAS (or no reliable history) should undergo sleep assessment regardless of their Epworth score**

Overall, the evidence[32] suggests that any patient with a history of possible OSAS should have some form of sleep assessment regardless of their Epworth scoring. The importance of dividing referrals into those with possible sleep apnoea and those with simple snoring has been described.[32] In the UK, this will be particularly important with the introduction of the forthcoming 18-week referral-to-treatment period, so that patients can obtain a cross-specialty referral for a different condition from that of the initial referral if necessary. An exception will be some specialist ENT centres that have the expertise to perform sleep assessments in the same unit.

The importance of pulse oximetry, and its predictive potential, are well documented.[19] This measure is a valid predictor of sleep apnoea.[19] Not all patients in our study underwent full five-channel nocturnal polysomnography, due to the need to optimise resource allocation. Thus, there was the potential for false negative results. However, this was minimised by having a liberal threshold for polysomnography referral. Evidence in the current literature justifies the instigation of therapeutic measures for OSAS simply on the basis of pulse oximetry, suggesting that the possibility

of false negatives during such a process may be statistically insignificant.[19]

## References

1 Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep* 1991;**14**:540–5

2 Gottlieb DJ, Yao Q, Redline S, Ali T, Mahowald MW. Does snoring predict sleepiness independently of apnea and hypopnea frequency? *Am J Respir Crit Care Med* 2000;**162**:1512–17

3 Herzog M, Kühnel T, Bremert T, Herzog B, Hosemann W, Kaftan H. The upper airway in sleep-disordered breathing: a clinical prediction model. *Laryngoscope* 2009;**119**:765–73

4 Kumar P, Bhatia VG, Tripathi M, Srivastava M, Jain AK. Obstructive sleep apnoea: a case-control study. *Neurol India* 2003;**51**:497–9

5 Montoya FS, Bedialauneta JR, Larracochechea UA, Ibarguen AM, Del Rey AS, Fernandez JMS. The predictive value of clinical and epidemiological parameters in the identification of patients with obstructive sleep apnoea (OSA): a clinical prediction algorithm in the evaluation of OSA. *Eur Arch Otorhinolaryngol* 2007;**264**:637–43

6 Ibrahim AS, Almohammed AA, Allangawi MH, A Sattar HA, Mobayed HS, Pannerselvam B *et al*. Predictors of obstructive sleep apnoea in snorers. *Ann Saudi Med* 2007;**27**:421–6

7 Olson L, Cole M, Ambrogetti A. Correlations among Epworth Sleepiness Scale scores, multiple sleep latency tests and psychological symptoms. *J Sleep Res* 1998;**7**:248–53

8 Chervin RD, Aldrich MS. The Epworth Sleepiness Scale may not reflect objective measures of sleepiness or sleep apnoea. *Neurology* 1999;**51**:125–31

9 Rosenthal LD, Dolan DC. The Epworth Sleepiness Scale in identification of obstructive sleep apnoea. *J Nerv Ment Dis* 2008;**196**:429–31

10 Smith SS, Oei TP, Douglas JA, Brown I, Jorgensen G, Andrews J. Confirmatory factor analysis of the Epworth Sleepiness Scale (ESS) in patients with obstructive sleep apnoea. *Sleep Med* 2008;**9**:739–44

11 Fong SYY, Ho CKW, Wing YK. Comparing MSLT and ESS in the measurement of excessive daytime sleepiness in obstructive sleep apnoea syndrome. *J Psychosom Res* 2005;**58**:55–60

12 Kingshott R, Douglas N, Dreary I. Mokken scaling of Epworth sleepiness scale items in patients with sleep apnoea hypopnoea syndrome. *J Sleep Res* 1998;**7**:293–4

13 Kingshott RN, Sime PJ, Engleman HM, Douglas NJ. Self assessment of daytime sleepiness, patient vs partner. *Thorax* 1995;**50**:994–5

14 Osman EZ, Osborne J, Hill PD, Lee BW. The Epworth Sleepiness Scale: can it be used for sleep apnoea screening among snorers? *Clin Otolaryngol* 1999;**24**:239–41

15 Chung KF. Use of the Epworth Sleepiness Scale in Chinese patients with obstructive sleep apnoea and normal hospital employees. *J Psychosom Res* 2000;**49**:267–372

16 Thong JF, Pang KP. Clinical parameters in obstructive sleep apnoea: are there any correlations? *J Otolaryngol Head Neck Surg* 2008;**37**:894–900

17 Sauter C, Asenbaum S, Popovic R, Bauer H, Lamm C, Klosch G *et al*. Excessive daytime sleepiness in patients suffering from different levels of obstructive sleep apnoea syndrome. *J Sleep Res* 2000;**9**:293–301

18 Cohen J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Hillsdale, New Jersey: Lawrence Erlbaum, 1988

19 Netzer N, Eliasson AH, Netzer C, Kristo DA. Overnight pulse oximetry for sleep-disordered breathing in adults. *Chest* 2001; **120**:625–33

20 Woodhead CJ, Davies JE, Allen MB. Obstructive sleep apnoea in adults presenting with snoring. *Clin Otolaryngol* 1991;**16**: 401–5

21 Lacasse Y, Godbout C, Sériés F. Independent validation of the Sleep Apnoea Quality of Life Index. *Thorax* 2002;**57**:483–8

22 Nguyen AT, Baltzan MA, Small D, Wolkove N, Guillon S, Palayew M. Clinical reproducibility of the Epworth Sleepiness Scale. *J Clin Sleep Med* 2006;**2**:170–4

23 McNeil BJ, Hanley JA. Statistical approaches to the analysis of receiver operating characteristic (ROC) curves. *Med Decis Making* 1984;**4**:137–50

24 Mediano O, Barcelo A, de la Pena M, Gozal D, Agusti A, Barbe F. Daytime sleepiness and polysomnographic variables in sleep apnoea patients. *Eur Respir J* 2007;**30**:110–13

25 Mackay TW. Sleep apnoea diagnosis and treatment. *J R Coll Physicians Edinb* 2008;**38**:32–4

26 Leng PH, Mosharraf-Hossain AK, Chan YH, Tan WC. The clinical predictors of hypertension and sleepiness in an Asian population with sleep-disordered breathing. *Ann Acad Med Singapore* 2006;**35**:6–10

27 Banks S, Barnes M, Tarquinio N, Pierce R, Lack J, Leon C *et al*. Factors associated with maintenance of wakefulness test mean sleep latency in patients with mild to moderate obstructive sleep apnoea and normal subjects. *J Sleep Res* 2004;**13**:71–8

28 Resta O, Carratu P, Carpagnano GE, Maniscalco M, Di Gioia G, Lacedonia D *et al*. Influence of subclinical hypothyroidism and T4 treatment on the prevalence and severity of obstructive sleep apnoea syndrome (OSAS). *J Endocrinol Invest* 2005;**28**:893–8

29 Kessler R, Rodenstein DO. Daytime somnolence. Basic concepts, assessment tools and clinical applications. *Monaldi Arch Chest Dis* 2001;**56**:400–12

30 Beiske KK, Kjelsberg FN, Ruud EA, Stavem K. Reliability and validity of a Norwegian version of the Epworth Sleepiness Scale. *Sleep Breath* 2009;**13**:65–72

31 Lim PV, Curry AR. The role of history, Epworth Sleepiness Scale score and body mass index in identifying non apnoeic snorers. *Clin Otolaryngol* 2000;**25**:244–8

32 West B, Bennett JA, Deegan PC, Merry P, Watson L, Jones NS *et al*. Reducing waiting times for sleep apnoea hypopnoea syndrome and snoring using a questionnaire and home oximetry: results of a second audit cycle. *J Laryngol Otol* 2001;**115**:645–7

Address for correspondence:
Mr A Sil,
Department of ENT,
Monklands Hospital,
Airdrie, Scotland ML6 0JS, UK

Fax: +44 (0) 1236 760015
E-mail: silstat@gmail.com

Mr A Sil takes responsibility for the integrity of the content of the paper
Competing interests: None declared