Developing a pictorial Epworth Sleepiness Scale

Ramesh Ghiassi, Kevin Murphy, Andrew R Cummin, Martyn R Partridge

ABSTRACT

Objective The Epworth Sleepiness Scale (ESS) was designed to be self-completed by the patient. However, it may not be understood by all, and unrecognised problems with literacy can impair the process. The ESS has been translated into a pictorial version for use in those with normal or diminished literacy skills.

Methods An evaluation of the patients’ ability to self-complete the ESS was undertaken in sleep and non-sleep respiratory clinics. Errors or problems encountered were recorded on a standard questionnaire. With the aid of a medical artist, pictorial representations of the eight ESS questions were developed and the new pictorial ESS was offered to patients alongside the traditional ESS. The two scales were compared for agreement with a kappa statistic, and patients were asked to record a preference for either the written or the pictorial scale.

Results Evaluation of the traditional ESS showed that 33.8% (27/80) of ESS-naive patients made errors and 22.5% (18/80) needed help completing the questionnaire. The translated pictorial ESS showed good agreement with the traditional ESS on most questions; median kappa score 0.63, IOR 0.04. Fifty-five per cent reported a preference for the pictorial scale compared with the standard written ESS. Despite the fact that errors were frequently made on the traditional ESS, 96.8% of participants in the second study reported both scales to be easy to complete. More patients (75.6%) reported the pictorial ESS to be very easy, in comparison with 64.6% who found the traditional written ESS easy to complete. Pictures with words have been shown to enhance the understanding and translation of medical information, and a pictorial translation of the ESS produces scores comparable with the traditional ESS and may be a suitable alternative for those with normal or diminished literacy.

INTRODUCTION

Obstructive sleep apnoea syndrome (OSAS) is estimated to affect 5% of the adult population. A key component of the disorder is excessive sleepiness in the daytime (EDS). In 1991, the Epworth Sleepiness Scale (ESS) has become one of the most commonly used questionnaires to measure EDS. Designed to be self-administered, it asks the user to rate their level of sleepiness in eight different situations. Despite translation into several languages (including Chinese, German, Greek and Spanish), understanding of the scale is dependent on being able to read, write and interpret medical information. Our anecdotal experience suggests that patients have difficulty understanding and accurately self-completing the ESS.

Translating the ESS into a pictorial ESS (pESS)

Working with a medical artist, a team of health professionals with expertise in sleep medicine produced pictures representing the eight domains on the traditional ESS. A series of silhouette images depicting the sleepy person with an altered body position, with increased likelihood of sleepiness (ie, slumping forward in a chair) as in figure 1 were developed. Question 8 (Q8) required a decision as to whether the driver or the passenger was the sleepy character and two versions of the pESS representing this difference were produced and tested (studies 1 and 2). The sleep patients were then asked in study 3 to report a preference for the pESS or the traditional ESS.

Study 1

Patients attending the sleep centre or the lung function laboratory were asked to complete...
a traditional ESS and a pESS, with the order of completion reversed on alternate days. Those attending the lung function laboratory were having routine lung function tests prior to operations, or for diagnosis or follow-up of a lung condition, and were not thought to have OSAS. Total ESS and pESS scores for each patient were compared for variability using the Wilcoxon signed rank test and a Cohen $\kappa$ statistic to measure agreement between responses to the eight different pictorial and traditional ESS questions. Total sleepiness scores and responses to individual questions on both the traditional ESS and pESS are summarised in figure 2. The scores on the pESS (Q1–9) were calculated using (1) $\sum$ pESS Qs 1–8 (sleepy passenger) and (2) $\sum$ pESS Qs 1–7 + Q 9 (sleepy driver) and compared with the respondent’s traditional ESS scores. Driving status (driver= respondent holding a provisional, full UK or international driving licence) was recorded and differences in response to the pESS and traditional ESS among those holding a driving licence and those who reported they could not drive was evaluated using a Wilcoxon signed rank test.

Study 3
To evaluate patient acceptability of the pESS and traditional ESS, 82 additional patients attending the sleep centre (who had not taken part in earlier studies) were issued the two scales (described above). Patients were asked to choose one of four statements reflecting ease of completion for each scale: very easy, fairly easy, fairly difficult or very difficult, and to express a preference for the traditional ESS or pESS.

RESULTS
A total of 564 patients have contributed to the evaluation of the ESS and development of a pESS.

Evaluation of patients self-completing the ESS (n=201)
A total of 201 patients were evaluated completing the ESS (80 naive and 121 serial users of the questionnaire). Results show that 53.8% (27/80) of naive and 15.7% (19/121) of routine users made errors, the most frequent being writing an answer between scores, that is 1–2 or as a fraction 1 ½ (13.8% naive, 5.8% previous users). Of those that made errors, other problems included annotating answers or writing ‘yes’ or ‘no’ (7.5% naive, 10.8% previous users), leaving questions blank (10.0% naive, 3.3% previous users), or putting a tick or a cross against the questions rather than scoring them (3.8% naive, 1.7% previous users). A further group volunteered that they could not read or write (3.8% naive, 0.8% previous users), or stated that they had left their spectacles at home (2.5% naive, 1.7% previous users). In total 28/201 (13.9%) needed additional help from either a family member, friend or health professional to enable them to complete the scale.

Translation of the traditional ESS into a pESS (n=145)
A total of 145 patients (93 sleep centre and 52 lung function laboratory patients) completed the traditional written ESS and the pESS (figure 1). Resulting ESS scores ranged from 0 to 24, with a group (n=145) mean total ESS score of 8.80 (± 4.72) and a group mean total ESS score of 8.46 (±4.43); full tabulated total pESS and ESS results are available online, in supplementary table 1. A Cohen $\kappa$ statistic showed good agreement (median $\kappa$ score Q1–8 = 0.63, IQR 0.04) between the responses to the individual questions on the traditional ESS and pESS. Kappa scores are summarised in figure 2 and supplementary online figure 3. In figure 2, responses of our patients to each of the individual questions on both the traditional ESS and pESS are displayed. On the same chart are displayed the data from Professor Johns’ original study population in 1992. Wilcoxon signed rank test showed no significant difference between the overall score obtained from the traditional ESS compared with the scores on the pESS (p=0.086). Ranking item scores from low to high (online data supplement table 2 and table 3) show similarities in response across the three groups, with some items (such as Q5) scoring consistently high while others (such as Q8 and Q6) produced consistently low results.

Evaluation of the depiction of driving in the pESS (n=136)
A total of 136 patients (82 sleep centre and 54 lung function laboratory patients) completed the questionnaires. The mean sleepiness scores recorded were: 8.68 (± 5.2) pESS Q1–8 (Q.8 representing the person in the vehicle as the sleepy passenger), 8.54 (± 5.04) pESS Q1–7 + Q.9 (Q.9 representing person in the image as the sleepy driver) and 8.63 (± 5.5) the standard.
ESS; full tabulated results for total pESS and ESS scores are available online, in supplementary table 4. Item scores for the three different versions of Q8 were analysed separately. No differences in item scores across the three versions of Q8 were observed in those who reported they could not drive (n = 49). However, respondents who held a driving licence (n = 87) were significantly more likely to report sleepiness when the pESS depicted a sleepy passenger in the vehicle than when the pESS depicted a sleepy driver in the vehicle (p = 0.003). A similar difference was seen in comparison with Q8 of the traditional ESS (p = 0.02) and this was so whether the respondent came from the sleep or lung function group of patients (table 1).

**Patient preferences for the traditional ESS or pESS (sleep n = 82)**
Eighty-two other patients attending the sleep centre, who had not taken part in the previous studies, completed evaluations of the pESS and traditional ESS questionnaires with the following mean sleepiness scores: pESS 9.24 (± 4.48) and traditional ESS 9.56 (± 4.99); full online results, table 5. A total of 54.9% (45/82) reported a preference for the pESS compared with 45.1% (37/82) for the traditional word-only ESS. The majority of patients, 96.8%, reported both scales to be easy to complete. More people, 75.6%, recorded the pESS as very easy to complete in contrast to the word ESS questionnaire, 64.6%.

**DISCUSSION**
Quantification of sleepiness remains an essential tool in sleep medicine. Objective laboratory measurements such as the multiple sleep latency test are useful but labour intensive and costly, and for this reason subjective sleepiness records are more commonly employed. While the traditional written ESS has received international acceptance, it has limitations. In our practice we observed patient difficulties in completing the form and quantified this finding. We have demonstrated how one-third (33.8%) of new users and one-sixth (15.7%) of routine users of the ESS make errors or need assistance completing of the

---

**Figure 2** Pictorial and traditional Epworth Sleepiness Scale (ESS) responses for each of the eight questions for 145 patients (93 sleep centre and 52 lung function laboratory). Shown in grey are the original ESS scores reported by Johns in 1992 (n = 144, 104 third year medical students and 40 sleepy patients). Values are mean responses with 1 SE bars. Above each SE bar is the corresponding kappa statistic comparing pictorial ESS and ESS item scores.

ESS; full tabulated results for total pESS and ESS scores are available online, in supplementary table 4. Item scores for the three different versions of Q8 were analysed separately. No differences in item scores across the three versions of Q8 were observed in those who reported they could not drive (n = 49).

**Table 1** Comparing item scores for two pictorial derivations of question 8 of the Epworth Sleepiness Scale (ESS); one showing a sleepy passenger (Q8) and the other showing a sleepy driver (Q9) in the vehicle, and question 8 of the traditional ESS, among total group (n = 136), sleep centre (n = 82) and lung function laboratory (n = 54) patients

<table>
<thead>
<tr>
<th>Drivers and non-drivers</th>
<th>Q8 (In a car, while stopped for a few minutes in traffic)</th>
<th>Q8 (In a car, while stopped for a few minutes in traffic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total group Sleep Lung Function</td>
<td>Total group Sleep Lung Function</td>
<td>Total group Sleep Lung Function</td>
</tr>
<tr>
<td>Mean score</td>
<td>0.31</td>
<td>0.40</td>
</tr>
<tr>
<td>SD</td>
<td>0.19</td>
<td>0.64</td>
</tr>
<tr>
<td>SE</td>
<td>0.00</td>
<td>0.09</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Drivers (n = 87)</th>
<th>Q8 (In a car, while stopped for a few minutes in traffic)</th>
<th>Q8 (In a car, while stopped for a few minutes in traffic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total group Sleep Lung Function</td>
<td>Total group Sleep Lung Function</td>
<td>Total group Sleep Lung Function</td>
</tr>
<tr>
<td>Mean score</td>
<td>0.40</td>
<td>0.44</td>
</tr>
<tr>
<td>SD</td>
<td>0.83</td>
<td>0.87</td>
</tr>
<tr>
<td>SE</td>
<td>0.09</td>
<td>0.10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-drivers (n = 49)</th>
<th>Q8 (In a car, while stopped for a few minutes in traffic)</th>
<th>Q8 (In a car, while stopped for a few minutes in traffic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total group Sleep Lung Function</td>
<td>Total group Sleep Lung Function</td>
<td>Total group Sleep Lung Function</td>
</tr>
<tr>
<td>Mean score</td>
<td>0.14</td>
<td>0.18</td>
</tr>
<tr>
<td>SD</td>
<td>0.50</td>
<td>0.60</td>
</tr>
<tr>
<td>SE</td>
<td>0.07</td>
<td>0.12</td>
</tr>
</tbody>
</table>

*p < 0.003 between drivers’ responses: pictorial ESS Q8 vs pictorial ESS Q9.
*p < 0.02 pictorial ESS Q8 vs ESS Q8.
scale. For international use it requires translation and it is likely to be less useful in those with reduced literacy skills\textsuperscript{4–7} Our study suggests that this is a highly prevalent problem. People have problems filling in forms for many reasons. The addition of pictures to words has been shown to be useful and is one method of broadening the accessibility of medical questionnaires or information. A South African group have previously developed a cartoon faces sleepiness scale for children, which measures ‘point’ and ‘here and now’ sleepiness (as opposed to propensity to sleepiness associated with behaviours over time).\textsuperscript{21} We worked with a medical artist to produce pictorial translations of the ESS questions and coupled images with words to produce a pESS. When tested with the traditional ESS the pictorial scale achieved an acceptable level of equivalence and comparable total ESS scores. Two of the item scores (Q5 and Q9) were in the moderate $k$ range; however, overall median $k$ results showed good agreement between the scales with an IQR of 0.04. John’s original study obviously used a different population from ours, but the distribution of scores between questions shows a similar pattern to that in our study, with patients being more likely to report excessive daytime sleepiness, for example, when watching television or when resting in the afternoon, than when sitting and talking to someone or when in a car stopped for a few moments in traffic.

Depicting the character of the sleepy person as the driver or the passenger in the vehicle resulted in different responses according to whether the respondent held a driving licence or not. Subjects with a driving licence were significantly less likely to record feeling sleepy on the traditional ESS Q8 and the pESS Q9 when the image depicted a sleepy driver in a vehicle, and reported a higher likelihood of sleepiness when the pESS showed a sleepy passenger in a vehicle. This difference in response was not observed among patients who were non-drivers. Measuring driver sleepiness is a sensitive and complicated issue, and the interpretation of this finding is difficult. There may be reluctance to admit to sleepiness while behind the wheel for legal reasons or for fear of losing one’s driving licence, particularly among people who rely on driving for their occupation. Alternatively, this difference may be the result of driving producing a sufficient stimulus to reduce sleepiness. For the purpose of eliciting a valid response to sleepiness in a vehicle, depicting the passenger as the sleepy character may be the preferred option and has been chosen for the final pESS.

Among patients who rated the two scales, both were reported to be easy to complete. However, more people rated the pictorial scale as very easy, and a preference for the pESS was found in 55% of users.

In summary, the addition of pictures to the ESS to produce a translated pESS did not detract from the original intent of the scale and we have shown that it is possible to devise a pESS which has good agreement with the traditional written ESS. Using pictures in medicine should not be to the exclusion of words. To elicit the best possible understanding, combining both pictures and words to maximise intended meaning is important and necessary\textsuperscript{12,15,16} The authors make the pESS shown in this paper freely available and it may be downloaded at tinyurl.com/pictorialESS. Comments, feedback and discussion on the pESS are welcomed at this site.

**Acknowledgements** The authors are grateful to Anne Wadmore, medical artist, for her valuable help in the development of the pictorial ESS.

**Ethics approval** This study was conducted with the approval of the Riverside Ethics Committee.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**REFERENCES**

Developing a pictorial Epworth Sleepiness Scale

Ramesh Ghiassi, Kevin Murphy, Andrew R Cummin and Martyn R Partridge

Thorax 2011 66: 97-100 originally published online October 26, 2010
doi: 10.1136/thx.2010.136879

Updated information and services can be found at:
http://thorax.bmj.com/content/66/2/97

These include:

Supplementary Material
Supplementary material can be found at:
http://thorax.bmj.com/content/suppl/2010/08/24/thx.2010.136879.DC1

References
This article cites 21 articles, 2 of which you can access for free at:
http://thorax.bmj.com/content/66/2/97#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/