

Developing a pictorial Epworth Sleepiness Scale

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

► Additional figures and tables are published online only. To view these files please visit the journal online (<http://thorax.bmj.com>).

Imperial College London, NHLI division at Charing Cross Hospital, London, UK

Correspondence to

Ramesh Ghiassi, Imperial College London, 5 North Charing Cross Hospital, Hammersmith, London W68RF, UK; r.ghiassi@imperial.ac.uk

Received 17 March 2010
Accepted 29 July 2010
Published Online First
26 October 2010

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-naïve 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, IQR 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 people (75.6%) reported the pictorial ESS to be very easy, in comparison with (64.6%) the worded ESS questionnaire.

Conclusion Errors are common when patients self-complete the traditional written ESS. 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.¹ A key component of the disorder is excessive sleepiness in the daytime (EDS). Developed in 1991, the Epworth Sleepiness Scale (ESS) has become one of the most commonly used questionnaires to measure EDS.³ 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.^{4–7} Our anecdotal experience suggests that patients have difficulty understanding and accurately self-completing the ESS.

Difficulty with completing forms can occur for many reasons. The written word and numerals are not universally understood; problems with func-

tional literacy in medicine are both common and underestimated^{8–11}. Pictures or pictorial aids are a useful adjunct to medical information and aid the transfer and comprehension of written and spoken information. Recall of medication and understanding of instructions have been shown to be improved when pictures are used in conjunction with, or instead of, the written word, even in those with normal literacy skills.^{12–14} Such interventions are important as they can improve both understanding and compliance with medical treatments.^{12 15 16}

This report describes an evaluation of patients' ability to self-complete the ESS and a study of the translation of the standard ESS into a pictorial scale to produce a questionnaire that is suitable for all, irrespective of literacy.

METHODS

An evaluation of patients' ability to self-complete the ESS

For each of the eight questions in the ESS, the scorer (the patient) is required to rate their likelihood of sleepiness from 0 to 3: [0=no, 1=slight, 2=moderate and 3=high chance of dozing].³ The total score is marked out of a maximum of 24, and a score ≥ 11 is considered abnormal.

We evaluated the ease of completing the ESS in Epworth-naïve patients (new patients attending a sleep clinic, n=80) and serial users of the scale (patients on treatment for OSAS attending a follow-up clinic, n=121). Consecutive patients were given the ESS at the start of their clinic attendance and observed completing the form. Staff used a checklist (online figure 1) to record whether the ESS was correctly completed, whether help was provided and what errors, if any, were made.

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

Pictorial Epworth Sleepiness Scale				
Name: _____ Date: ____/____/____ Hospital No: _____ Date of Birth: ____/____/____				
In contrast to just feeling tired, how likely are you to doze off or fall asleep in the following situations? Even if you have not done some of these things recently, try to work out how they would affect you. Use the following scale to choose the most appropriate number for each situation.				
Situation <input type="checkbox"/> Please tick box	0 No chance of dozing	1 Slight chance	2 Moderate chance	3 Definitely would doze
Sitting and reading				
Watching TV				
Sitting inactive in a public place (e.g. Theatre or a meeting)				
As a passenger in a car for an hour without a break				
Lying down to rest in the afternoon when circumstances permit				
Sitting and talking to someone				
Sitting quietly after lunch without alcohol				
In a car, while stopped for a few minutes in traffic				

Figure 1 The pictorial Epworth Sleepiness Scale. Each scenario has a series of pictograms depicting increased likelihood of sleepiness; participants are asked to tick the image that best represents their level of daytime sleepiness. (During the trial a ninth question portraying a sleepy driver was also included.)

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 κ statistic to measure agreement between responses to the eight different pictorial and traditional ESS questions. Total sleepiness scores and responses to individual questions were compared for interest with those of the subjects reported by Murray Johns in one of his original studies.¹⁷

Study 2

To test whether depiction of the sleepy person in Q8 as a passenger or driver altered the pESS score, a further set of patients (sleep n=82 and lung function n=54) repeated the methods above but their pESS carried a ninth question: Q8 showing a sleepy passenger in the vehicle and Q9 a sleepy driver in the vehicle (online supplement figure 2). Total pESS scores 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 33.8% (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 1/2 (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 pESS score of 8.80 (\pm 4.72) and a group mean total ESS score of 8.46 (\pm 4.43); full tabulated total pESS and ESS results are available online, in supplementary table 1. A Cohen κ statistic showed good agreement (median κ 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 (\pm 5.2) pESS \sum Q1–8 (Q.8 representing the person in the vehicle as the sleepy passenger), 8.54 (\pm 5.04) pESS \sum Q1–7 + Q.9 (Q.9 representing person in the image as the sleepy driver) and 8.63 (\pm 5.3) the standard

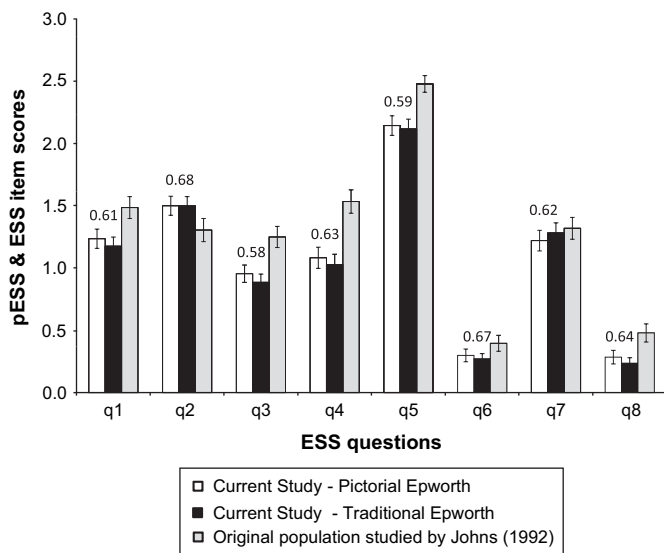


Figure 2 Pictorial and traditional Epworth Sleepiness Scale (ESS) responses for each of the eight questions for 145 patients (93 sleep centre and 52 lung function laboratory). Shown in grey are the original ESS scores reported by Johns in 1992 (n=144, 104 third year medical students and 40 sleepy patients). 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).

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) a preference 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.¹⁸⁻²⁰ 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

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

	Q8			Q9			Q8		
							'In a car, while stopped for a few minutes in traffic'		
	Total group	Sleep	Lung Function	Total group	Sleep	Lung Function	Total group	Sleep	Lung Function
Drivers and non-drivers									
Mean score	0.31	0.40	0.17	0.17	0.24	0.06	0.21	0.27	0.11
SD	0.74	0.84	0.50	0.54	0.64	0.30	0.59	0.69	0.37
SE	0.06	0.09	0.07	0.05	0.07	0.04	0.05	0.08	0.05
Drivers (n=87)									
Mean score	0.40	0.44	0.25	0.21*	0.25	0.00	0.24†	0.28	0.06
SD	0.83	0.87	0.58	0.59	0.65	0.00	0.66	0.72	0.25
SE	0.09	0.10	0.14	0.06	0.08	0.00	0.07	0.09	0.06
Non-drivers (n=49)									
Mean score	0.14	0.18	0.13	0.10	0.18	0.08	0.14	0.18	0.13
SD	0.50	0.60	0.47	0.42	0.60	0.36	0.41	0.40	0.41
SE	0.07	0.18	0.08	0.06	0.18	0.06	0.06	0.12	0.07

*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.^{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’ or ‘here and now’ sleepiness (as opposed to propensity to sleepiness associated with behaviours over time).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 (Q3 and Q5) were in the moderate κ range; however, overall median κ results showed good agreement between the scales with an IQR of 0.04. Johns’ 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.^{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

1. Young T, Peppard PE, Gottlieb DJ. Epidemiology of obstructive sleep apnea: a population health perspective. *Am J Respir Crit Care Med* 2002;**165**: 1217–39.
2. Young T, Evans L, Finn L, et al. Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. *Sleep* 1997;**20**: 705–6.
3. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991;**14**:540–5.
4. Bloch KE, Schoch OD, Zhang JN, et al. German version of the Epworth Sleepiness Scale. *Respiration* 1999;**66**:440–7.
5. Chen NH, Johns MW, Li HY, et al. Validation of a Chinese version of the Epworth sleepiness scale. *Qual Life Res* 2002;**11**:817–21.
6. Chiner E, Arriero JM, Signes-Costa J, et al. [Validation of the Spanish version of the Epworth Sleepiness Scale in patients with a sleep apnea syndrome] In Spanish. *Arch Bronconeumol* 1999;**35**:422–7.
7. Tsara V, Serasli E, Amfilochiou A, et al. Greek version of the Epworth Sleepiness Scale. *Sleep Breath* 2004;**8**:91–5.
8. Baker DW, Gazmararian JA, Sudano J, et al. The association between age and health literacy among elderly persons. *J Gerontol B Psychol Sci Soc Sci* 2000;**55**: S368–74.
9. Taylor J, Dawson S, Sridhar M, et al. Functional illiteracy amongst those with chronic obstructive pulmonary disease (COPD). *Eur Respir J* 2005;**26**:575.
10. Williams MV, Davis T, Parker RM, et al. The role of health literacy in patient–physician communication. *Fam Med* 2002;**34**:383–9.
11. Williams MV, Parker RM, Baker DW, et al. Inadequate functional health literacy among patients at two public hospitals. *JAMA* 1995;**274**:1677–82.
12. Houts PS, Bachrach R, Witmer JT, et al. Using pictographs to enhance recall of spoken medical instructions. *Patient Educ Couns* 1998;**35**:83–8.
13. Houts P, Witmer J, Egeth H, et al. Using pictographs to enhance recall of spoken medical instructions II. *Patient Educ Couns* 2001;**43**:231–42.
14. Houts P, Doak C, Loscalzo M. The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. *Patient Educ Couns* 2006;**61**:173–90.
15. Katz MG, Kripalani S, Weiss BD. Use of pictorial aids in medication instructions: a review of the literature. *Am J Health Syst Pharm* 2006;**63**:2391–7.
16. Delp C, Jones J. Communicating information to patients: the use of cartoon illustrations to improve comprehension of instructions. *Acad Emerg Med* 1996;**3**:264–70.
17. Johns MW. Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep* 1992;**15**:376–81.
18. Bliwise DL. Is the measurement of sleepiness the Holy Grail of sleep medicine? *Am J Respir Crit Care Med* 2001;**163**:1517–19.
19. Sangal RB. When is sleepiness a disease? How do we measure it? *Sleep Med* 2006;**7**:310–11.
20. Shen J, Barbera J, Shapiro CM. Distinguishing sleepiness and fatigue: focus on definition and measurement. *Sleep Med Rev* 2006;**10**:63–76.
21. Maldonado CC, Bentley AJ, Mitchell D. A pictorial sleepiness scale based on cartoon faces. *Sleep* 2004;**27**:541–8.