

## Reproducibility of Nordic Sleep Questionnaire in spinal cord injured

F Biering-Sørensen MD PhD, M Biering-Sørensen MD, J Hilden MD

*Centre for Spinal Cord Injured, Rigshospitalet, The National University Hospital, Copenhagen, and Statistical Research Unit, University of Copenhagen, Denmark*

A recently proposed Nordic Sleep Questionnaire (NSQ) comprises 26 questions concerning qualitative and quantitative aspects of the respondent's sleep habits. Its reproducibility was evaluated in 32 spinal cord injured individuals (SCI), 24 men and eight women (23–72 years), and 79 normal subjects, 23 men and 56 women (19–77 years). They completed the NSQ twice at a median interval of 15 days (range 10–26) and 27 days (range 4–103) respectively. The group of normal subjects were evenly divided into group 26, i.e. those who completed the two NSQs within 26 days, and group 27 with 27 days or more between their replies. Generally, group 27 showed no worse test–retest agreement than group 26. In addition, the respondents' answers, with a few exceptions, were reasonably stable in terms of test–retest agreement or standard deviation. The SCI group exhibited the same level of reproducibility, although they had more 'pathology' to report and thus more scope for contradicting themselves. The questions in the NSQ generally were satisfactorily reproducible. However, answers to the ordered five-point questions about sleepiness in the morning and during the daytime ought to be interpreted with caution. The same may be said about the number of minutes required to fall asleep, and the duration of daytime naps.

*Keywords:* sleep; sleep disorders; sleep questionnaire; reproducibility; reliability; spinal cord injury; normal subjects.

### Introduction

In spinal cord injured (SCI) individuals, obstructive sleep apnoea may contribute to disruption of sleep and be responsible for many of their daytime symptoms.<sup>1</sup> As this hypothesis has not been subjected to an epidemiological investigation, we found it to be of interest to do so. In reviewing assessment methods of sleep, Closs<sup>2</sup> found subjective rating scales inexpensive and simple to use, producing relevant and useful information when assessing sleep satisfaction. Further, self-administered questionnaires were judged to be reliable on test–retest administrations. It was decided to use the newly developed Nordic Sleep Questionnaire (NSQ)<sup>3</sup> for this purpose. As no reproducibility assessment existed for the NSQ in general, and none at all for an SCI population, we carried out the present reproducibility study, the point being that test–retest evaluations in one population cannot readily be used in another, if the frequencies of

sleep disturbances are different: one runs a risk of overestimating the reliability by relying on data based on normal subjects only.

### Questionnaire

The NSQ contains 21 sleep questions (Tables I and II). Most of the questions can be answered by a tick in the appropriate box, but some must be answered by numbers of minutes, hours, years, or by time of day or night. The last question (not analysed here) is an open one where sleep problems can be described in free text. The questionnaire generally asks the subject to describe his or her situation during 'the last 3 months'.

In addition there are some introductory questions on job situation, intake of stimulants (smoking, alcohol, coffee/tea), together with questions on the respondent's

**Table I** Test–retest reproducibility of those items of the Nordic Sleep Questionnaire (NSQ) which are answered on an ordered five-point scale. The normal subjects have been divided into group 27 with  $\geq 27$  days between the two rounds of NSQ completion, and group 26 with  $\leq 26$  days between the two rounds. In this respect group 26 is comparable to the group of spinal cord injured individuals (SCI group)<sup>a</sup>

Questions	Number with two answers	Mean value of answers 1st round/ 2nd round	% with identical answers	% with identical or adjacent answers
Q1. Have you had any difficulties in falling asleep?				
Group 27	40	1.7/1.5	67.5	92.5
Group 26	38	1.2/1.4	71.1	97.4
SCI group	31	2.1/2.0	61.3	100
Q3. How often do you wake up during the night?				
Group 27	40	2.4/2.3	67.5	92.5
Group 26	39	2.7/2.6	82.1	100
SCI group	31	3.3/2.9	71	90.3
Q4. How many times do you usually wake up in one night? <sup>b</sup>				
Group 27	33	1.7/1.6	87.9	97
Group 26	34	1.9/1.8	76.5	100
SCI group	31	2.3/2.2	77.4	96.8
Q5. How often have you awakened very early in the morning without being able to fall back to sleep again?				
Group 27	39	1.4/1.3	84.6	100
Group 26	39	1.4/1.4	76.9	94.9
SCI group	31	2.0/2.0	67.7	96.8
Q6. How well have you been sleeping? <sup>c</sup>				
Group 27	40	2.0/1.9	67.5	100
Group 26	39	1.9/1.8	82.1	97.4
SCI group	32	2.2/2.2	84.4	100
Q7. Have you used sleeping pills (by prescription)?				
Group 27	37	1.1/1.1	100	100
Group 26	38	1.1/1.1	100	100
SCI group	29	2.1/2.0	86.2	86.2
Q8. Do you feel excessively sleepy in the morning after awakening?				
Group 27	38	2.4/2.3	47.4	92.1
Group 26	39	2.4/2.3	56.4	89.7
SCI group	32	2.5/2.5	53.1	96.9
Q9. Do you feel excessively sleepy during the daytime?				
Group 27	38	2.5/2.3	55.3	92.1
Group 26	39	2.6/2.2	53.8	84.6
SCI group	30	3.1/2.6	46.7	86.7
Q10. Have you suffered from an irresistible tendency to fall asleep while <i>at work</i> ?				
Group 27	38	1.4/1.3	76.3	94.7
Group 26	35	1.4/1.3	85.7	97.1
SCI group	16	1.8/1.8	81.3	100

**Table I** (Cont)

Questions	Number with two answers	Mean value of answers 1st round/ 2nd round	% with identical answers	% with identical or adjacent answers
Q11. Have you suffered from an irresistible tendency to fall asleep during <i>free time</i> (leisure time)?				
Group 27	38	2.4/2.3	55.3	92.1
Group 26	36	2.0/1.9	61.1	91.7
SCI group	28	2.5/2.4	60.7	89.3
Q15a. How often do you have a nap during daytime?				
Group 27	40	2.2/1.9	72.5	97.5
Group 26	39	2.3/2.2	74.4	97.4
SCI group	32	3.0/2.9	71.9	93.8
Q16. Do you snore while sleeping (ask other people)?				
Group 27	37	1.8/1.6	70.3	97.3
Group 26	34	2.2/2.0	70.6	97.1
SCI group	30	3.0/2.8	63.3	90.0
Q17. In what way do you snore (ask other people about the quality of your snoring)? <sup>d</sup>				
Group 27	31	1.7/1.6	71	100
Group 26	32	2.0/1.9	71.9	96.9
SCI group	29	2.5/2.5	82.8	93.1
Q18. Have you had breathing pauses (sleep apnea) during sleep (have other people noticed that you have pauses in respiration when you sleep)?				
Group 27	38	1.0/1.0	100	100
Group 26	30	1.1/1.1	100	100
SCI group	25	1.5/1.6	84	88

<sup>a</sup>The basic scale for the answer-alternatives is as follows (for exceptions see below):

1 never or less than once per month; 2 less than once per week; 3 on 1-2 days per week; 4 on 3-5 days per week; 5 daily or almost daily. 'Nights' instead of 'days' if more appropriate. The questions generally ask about the past 3 months.

<sup>b</sup>Answer-alternatives: 1 usually I don't wake up at night; 2 once per night; 3 2 times; 4 3-4 times; 5 at least 5 times per night.

<sup>c</sup>Answer-alternatives: 1 well; 2 rather well; 3 neither well nor badly; 4 rather badly; 5 badly.

<sup>d</sup>Answer-alternatives: 1 I don't snore; 2 my snoring sounds regular and it is of low voice; 3 it sounds regular but rather loud; 4 it sounds regular but it is very loud (other people hear my snoring in the next room); 5 I snore very loudly and intermittently (there are silent breathing pauses when snoring is not heard and at times very loud snorts with gasping).

height and weight. These introductory questions will not be discussed further in this presentation.

In its original form the NSQ is in English, but for the present purpose it was translated into Danish.

### Participants

In all, 111 persons were requested to fill in the NSQ. They were all asked without

consideration being given to possible sleep disturbances. Two groups were studied. One comprised 32 SCI individuals, 24 men and eight women, aged 23-72 years (mean 42, standard deviation (SD) 13 years). They had spinal cord lesions between C4 and L1. The other group comprised 79 non-SCI individuals (normal group), 23 men and 56 women, aged 19-77 years (mean 40, SD 13 years; age was unknown for two). No special selection criteria were enforced,

**Table II** Test-retest reproducibility of quantitative aspects of the Nordic Sleep Questionnaire (NSQ). Group 27, group 26 and SCI group as in Table I

	1st round Mean/SD/n	2nd round Mean/SD/n	Error SD/r/n
<b>Q2. For how long (how many minutes as an average) do you stay awake in bed before you fall asleep (after lights out)?</b>			
a. working days:			
Group 27	8.9/ 7.6/38	9.3/ 6.7/39	2.66/0.95/38
Group 26	10.0/ 8.3/34	10.7/ 8.5/32	3.46/0.83/32
SCI group	21.9/28.6/19	21.6/26.3/21	18.76/0.42/18
b. during free time			
Group 27	10.6/ 8.6/37	10.6/ 7.3/38	4.83/0.87/37
Group 26	10.4/ 8.3/35	11.8/ 9.2/33	3.57/0.84/33
SCI group	32.4/35.8/30	27.2/26.1/30	17.53/0.72/29
<b>Q12. How many hours do you usually sleep per night?</b>			
Group 27	6.8/ 0.7/40	6.8/ 0.8/39	0.28/1.00/39
Group 26	6.9/ 0.9/39	6.8/ 0.9/39	0.27/0.91/39
SCI group	6.9/ 1.4/32	6.9/ 1.4/32	0.40/0.91/32
<b>Q13. At what time do you usually go to bed (in order to sleep)?</b>			
a. during a working week: <sup>a</sup>			
Group 27	23.2/ 0.8/39	23.1/ 0.8/38	0.29/0.86/38
Group 26	23.5/ 0.9/35	23.4/ 0.9/36	0.29/0.89/35
SCI group	22.9/ 1.0/22	23.0/ 0.8/20	0.31/0.73/19
b. during free days: <sup>a</sup>			
Group 27	0.2/ 0.9/39	23.9/ 0.7/39	0.53/0.56/38
Group 26	0.2/ 1.1/38	24.0/ 1.1/36	0.36/0.88/36
SCI group	23.4/ 1.2/32	23.5/ 1.0/32	0.30/0.94/32
<b>Q14. At what time do you usually wake up?</b>			
a. during working week: <sup>a</sup>			
Group 27	6.3/ 0.7/37	6.3/ 0.6/37	0.18/0.93/36
Group 26	6.3/ 0.7/34	6.3/ 0.6/36	0.20/0.92/34
SCI group	6.7/ 0.6/19	6.8/ 0.9/21	0.15/0.94/19
b. during free days: <sup>a</sup>			
Group 27	8.2/ 1.1/38	8.1/ 1.1/39	0.38/0.89/38
Group 26	8.1/ 1.0/36	8.0/ 1.0/34	0.24/0.94/33
SCI group	7.5/ 1.5/31	7.3/ 1.5/32	0.46/0.90/31
<b>Q15b. If you take a nap, how long does it usually last? (minutes)</b>			
Group 27	52.4/24.1/30	50.5/26.6/29	16.15/0.92/26
Group 26	48.3/29.8/27	55.0/35.5/26	10.25/0.89/24
SCI group	58.8/27.7/25	55.0/23.2/25	15.50/0.66/24
<b>Q19a. If you snore at least 1-2 times per week, how many years have you been snoring (ask other people if you don't know)?</b>			
Group 27	19.4/16.3/7	28.3/15.1/4	5.11/0.98/4
Group 26	11.3/11.4/9	13.0/11.4/10	2.53/0.96/8
SCI group	19.1/13.8/14	17.3/10.2/13	2.40/0.96/11

**Table II** (Cont)

	1st round Mean/SD/n	2nd round Mean/SD/n	Error SD/r/n
Q19b. Age when you started to snore?			
Group 27	33.2/20.3/6	10.0/14.1/2	not calc. 1
Group 26	35.6/16.0/7	31.4/16.3/10	2.74/0.97/6
SCI group	29.6/11.5/14	29.5/11.5/13	1.92/0.98/11
Q20. How many hours of sleep do you need per night (how many hours would you sleep if you had the possibility to sleep as long as you need to)?			
Group 27	7.8/1.0/40	7.9/1.0/39	0.55/1.00/39
Group 26	7.6/1.0/39	7.6/0.9/38	0.42/0.80/38
SCI group	7.7/1.1/28	7.7/1.2/31	0.45/0.85/27

<sup>a</sup>Hours between zero and 24; thus 21.5 = 9:30 pm, 6.3 = 6:18 am, etc.

SD: standard deviation.

n: number of answers or number of observations included in the analysis.

Error SD: SD of the test–retest differences divided by  $\sqrt{2}$ .

r: correlation coefficient.

not calc: not calculable.

either for the SCI or the normal individuals, i.e. no strict representation is claimed. This was because the purpose was to test the reproducibility of the NSQ and not to make an epidemiological investigation at this point.

## Methods

All the participants in the study were asked, in a first round, to fill in the self-administered NSQ, and to give the authors some indications about the use of the questionnaire; this was done without the knowledge that they would be asked to fill it in again on a later occasion. The second time they were given the questionnaire, they were specifically asked to fill it in as well as possible without trying to recall what they wrote the first time. The intended interval between the two rounds was 2–3 weeks, but this rule proved impossible to enforce; see below.

## Statistical evaluations

Variables on ordered five-point scales (Table I) were handled by reporting the percentage identical and percentage identical-or-adjacent answers. Mean scores for the first and second questionnaire rounds

are given in order to pinpoint those items where there is little scope for disagreement in the first place because most participants tick the normal alternative (mean score close to '1').

The quantitative variables (Table II) are evaluated in terms of the error SD. It is calculated as the SD of the subjects' test–retest differences, divided by  $\sqrt{2}$ . Means and SDs for the first and second questionnaire rounds, and test–retest correlation coefficients are also reported.

## Results

### Populations

The age distribution in the SCI and the normal groups are comparable, whereas the gender distributions are different, with more women (71%) than men in the normal population and the opposite among the SCI individuals (only 25% females).

### Interval between questionnaires

Within the SCI group the two questionnaires were answered at an interval of 10–26 days (median 15 days).

The individuals in the normal group answered the two questionnaires at an interval of 4–103 days (median 27 days).

This group was divided into two according to the number of days between the questionnaire answers: one within which the number of days was 26 or less (group 26), the other with 27 or more days between the answers of the two NSQs (group 27). Group 26 with 39 individuals was comparable to the SCI group, with the median number of days between the two NSQs being 16 days, whereas it was 41 days among the 40 participants in group 27.

#### *Test-retest findings*

Generally, group 27 showed no worse test-retest agreement than group 26, despite the longer time interval between the NSQ administrations. The SCI group exhibited the same level of reproducibility, although these responders had more 'pathology' to report and thus more scope to contradict themselves. In a few instances the test-retest data are insufficient for evaluation of the item (Q19a and b in particular (Table II)). In addition there are questions related to working days (Q10 (Table I), Q2a, Q13a, Q14a (Table II)), but these have often been omitted by the SCI participants, because fewer in this group have gainful employment.

Table I evaluates those 14 aspects of sleep disturbance for which the NSQ provides a five-point scale. The reproducibility is satisfactory. As may have been expected, the highest percentage identical answers (100%) was obtained for questions with mean values of answers near to one (Q7 and Q18 in group 26 and group 27). On the other hand, the questions with the lowest average percentage identical answers were Q8 and Q9.

In addition to illustrating the moderate sleep problems of SCI individuals, all the data on the quantitative aspects of the subjects' sleep habits (Table II) show that people are able to answer these questions reproducibly in terms of small test-retest differences and high correlations ( $\geq 0.80$ , except for five instances out of 32 tests). The few exceptions include Q2a and b, and Q15b. For the answers to Q2a and b it is probably because the distributions are very right-skewed, especially in the SCI group

making the correlation coefficient unreliable. Two of the remaining three not so high correlations (0.56–0.73) were also found in the SCI group, but as the error SDs are not here higher than in the other groups, this must be a chance feature of the data.

#### **Discussion**

The test-retest procedure is probably the best way to assess the reliability of a questionnaire.<sup>4</sup> In the present study the test-retest analyses have been guided by four considerations. First, the characterisation of the reproducibility of any given item must reflect the structure of its reply alternatives (its scale), in our case ordered or quantitative. Second, the purpose of the reproducibility study is descriptive. There is no need, for instance, to perform any formal tests—because there are no fixed standards of performance. Third, if an item shows what everybody would regard as good reproducibility, there is no need for further calculations.

Fourth, when a participant filling in the questionnaire genuinely sees himself as midway between feeling sleepy during the day 'on 1–2 days per week' and 'on 3–5 days per week', he must make an arbitrary choice, and one cannot then hope for 100% test-retest agreement. One may, however, hope for test and retest answers no more than one level apart. Thus, in the case of ordered five-level scales one would regard identical or adjacent alternatives as acceptable.

Generally, the authors favour direct characterisation of the 'error of measurement' inherent in each item because this is what determines how well the item will discern genuine differences in sleep patterns in future studies where situations, such as before and after some intervention, or patient groups of research interest are being compared. It is the error SDs that go into calculations of the necessary sample sizes. Measurement error is directly reflected in the error SD in the case of a quantitative variable and, somewhat less directly, in disagreement rates in the non-quantitative case.

Our correlation coefficients were found to

be as high as those in three reliability studies of the sleep section of the Nottingham Health Profile, i.e. 0.80–0.94 (Hunt *et al* 1981<sup>5</sup> (original English version), Wiklund *et al* 1988<sup>6</sup> (Swedish version), Thorsen *et al* 1993<sup>7</sup> (Danish version)). For the Stanford Sleepiness Scale, a preliminary reliability analysis yielded a correlation of 0.88.<sup>8</sup> Ellis *et al*<sup>9</sup> found for the St Mary's Hospital Sleep Questionnaire a test–retest correlation using Kendall's  $\tau$  varying from 0.70 to 0.96. The latter questionnaire was concerned with the previous night's sleep, and the two identical questionnaires were answered within 4 hours on the same day. Jenkins *et al*<sup>10</sup> performed in cardiac surgery patients at an interval of 6 months a test–retest evaluation of a three-item sleep scale and found a significant correlation of 0.59 ( $n = 426$ ).

It should be noted that with questionnaires of this kind good reproducibility may be spurious in that a respondent may give consistent but wrong answers. For instance, he may have taught himself to believe that he sleeps 7 hours every night when in fact he sleeps only  $6\frac{1}{2}$  hours on the average (Q12). Further, recalling when one lost consciousness may be too much to ask of a human

being (Q2a and b). These indications are supported by the study of Carskadon *et al*<sup>11</sup> as they found that, in comparison with laboratory data, most subjects consistently underestimated the amount of time slept and overestimated the amount of time it took them to get to sleep. This was confirmed in a later study;<sup>12</sup> in addition poor sleepers were found to be significantly more inaccurate in their estimations.

It can be concluded that the questions in the NSQ in general were satisfactorily reproducible, in the normal as well as in the SCI population. However, answers to the ordered five-point questions about sleepiness in the morning (Q8), and during the daytime (Q9) ought to be interpreted with caution. Likewise the quantitative data regarding time required to fall asleep (Q2a and b), and the duration of daytime naps (Q15b) should be used with care.

### Acknowledgements

Support has been received from the Danish Medical Research Council, and the Dr Eilif Trier-Hansen and Ane Trier-Hansen Foundation.

### References

- 1 Bonekat HW, Andersen G, Squires J (1990) Obstructive disordered breathing during sleep in patients with spinal cord injury. *Paraplegia* **28**: 392–398.
- 2 Closs SJ (1988) Assessment of sleep in hospital patients: a review of methods. *J Adv Nursing* **13**: 501–510.
- 3 Gislason T, Hansen P, Hetta J *et al* (1988) *Basic Nordic Sleep Questionnaire*. Scandinavian Society for Sleep Research.
- 4 McDowell I, Newell C (1987) *Measuring Health: A Guide to Rating Scales and Questionnaires*. Oxford University Press.
- 5 Hunt SM, McKenna SP, Williams J (1981) Reliability of a population survey tool for measuring perceived health problems: a study of patients with osteoarthritis. *J Epidemiol Community Health* **35**: 297–300.
- 6 Wiklund I, Romanus B, Hunt SM (1988) Self-assessed disability in patients with arthrosis of the hip joint. Reliability of the Swedish version of the Nottingham Health Profile. *Int Disabil Studies* **10**: 159–163.
- 7 Thorsen H, McKenna SP, Gottschalck L (1993) The Danish version of the Nottingham Health Profile: Its adaptation and reliability. *Scand. J Primary Health Care* **11**: 124–129.
- 8 Hoddes E, Dement W, Zarcone V (1972) The development and use of the Stanford Sleepiness Scale (SSS). *Psychophysiology* **9**: 150 (abstract).
- 9 Ellis BW, Johns MW, Lancaster R *et al* (1981) The St. Mary's Hospital Sleep Questionnaire: A study of reliability. *Sleep* **4**: 93–97.
- 10 Jenkins CD, Stanton B-A, Niemcryk SJ, Rose RM (1988) A scale for the estimation of sleep problems in clinical research. *J Clin Epidemiol* **41**: 313–321.
- 11 Carskadon MA, Dement WC, Mitler MM *et al* (1976) Self-reports versus sleep laboratory findings in 122 drug-free subjects with complaints of chronic insomnia. *Am J Psychiatry* **133**: 1382–1388.
- 12 Adam K, Tomeny M, Oswald I (1986) Physiological and psychological differences between good and poor sleepers. *J Psychiatr Res* **20**: 301–316.