ELSEVIER

Sleep Medicine 4 (2003) 137–141

SLEEP MEDICINE

www.elsevier.com/locate/sleep

Original article

# The Johns Hopkins telephone diagnostic interview for the restless legs syndrome: preliminary investigation for validation in a multi-center patient and control population

Wayne A. Hening[a,*], Richard P. Allen[b], Stacey Thanner[b], Tinna Washburn[b], Debbie Heckler[b], Arthur S. Walters[a,c], Christopher J. Earley[b]

[a]*Department of Neurology, University of Medicine and Dentistry of New Jersey/RW Johnson Medical School, New Brunswick, NJ, USA*
[b]*Department of Neurology, Johns Hopkins School of Medicine, Baltimore, MD, USA*
[c]*NJ Neuroscience Institute at JFK Medical Center, Edison, NJ, USA*

Received 15 August 2002; received in revised form 6 January 2003; accepted 8 January 2003

**Abstract**

**Study objectives**: To develop and validate a telephone diagnostic interview (the Johns Hopkins telephone diagnostic interview for restless legs, abbreviated TDI) for diagnosis of the restless legs syndrome (RLS). **Design and methods**: Using the International RLS Study Group diagnostic criteria, specific questions were developed reflecting the diagnostic features of RLS. Seventy-five subjects (37 previously diagnosed RLS patients and 38 controls self-reported to be free of RLS) were interviewed by three expert interviewers blinded to each others' interviews and the patient's clinical status. The interviewers diagnosed each subject based on responses to the TDI. **Results**: The interviewers overall correctly diagnosed 72 of 75 individuals. Minimum interviewer sensitivity and specificity were 97 and 92%, respectively. The intraclass correlation coefficient (ICC) was used to quantify inter-rater reliability for the three interviewers. The ICC for diagnosis was 0.95. The ICC calculated on other key items in the interview exceeded 0.80 in all cases. The patients were predominantly older individuals with long-standing RLS; 19 of them scored at the highest level of severity on the Johns Hopkins Restless Legs Severity Scale (JHRLSS). The interviewers had more difficulty with assessing the controls accurately, some of whom were probably incorrectly self-categorized as not having RLS. **Conclusions**: The TDI is a sensitive, specific, and reliable instrument for diagnosing RLS by experienced interviewers in a brief, anonymous telephone encounter.
© 2003 Elsevier Science B.V. All rights reserved.

*Keywords:* Restless legs; Diagnosis; Validation; Questionnaire; Reliability; Sensitivity; Specificity

## 1. Introduction

The 'gold standard' for diagnosing the restless legs syndrome (RLS) remains the expert clinical interview. In order to conduct reliable and valid epidemiological studies, investigators need instruments, which are capable of diagnosing RLS efficiently and accurately without sustained clinical interaction. A telephone diagnostic interview (TDI) developed at the Johns Hopkins Center for RLS is one such instrument. It was based on extensive international clinical experience, which is reflected in the published consensus of the International RLS Study Group (IRLSSG) [1]. The four specific diagnostic features are all based on the patient's history:

1. the need or urge to move usually based on uncomfortable sensations, primarily if not exclusively in the legs;
2. motor restlessness (actual movement that successfully relieves symptoms);
3. provocation of the first two symptoms by rest (sitting, lying, inactivity) with relief by activity; and
4. circadian variation of symptoms with symptoms usually worst in the evening and at night in the hours before and during the normal time in bed.

The primary purpose of this instrument then is to facilitate diagnosis without a face-to-face interview. The current study is an initial validation of the instrument performed on a clinical population of RLS patients and non-patient controls.

* Corresponding author. 321 Avenue C (1-F), New York City, NY 10009-1643. Tel.: +1-212-477-7527; fax: +1-410-550-8078.
*E-mail address:* waheningmd@aol.com (W.A. Hening).

1389-9457/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S1389-9457(03)00006-6

A secondary purpose of the instrument was to characterize the interviewee's RLS, if the diagnosis were positive. The interview queried the age of onset and the time of day when symptoms typically began. The latter is equivalent to one measure of severity, the Johns Hopkins Restless Legs Severity Scale (JHRLSS). The JHRLSS has been validated against objective measures of RLS severity such as sleep efficiency and periodic limb movements and therefore is a useful, one-question assessment of severity in RLS [2]. The age of onset is an important aspect of clinical history because the etiology of RLS may be related to age of onset, with younger RLS subjects more likely to show positive familial clustering, whereas older patients may be more likely to have a sporadic presentation often with an underlying provocative condition [3,4]. Recent investigations have also shown that patients with different ages of onset may differ in their phenotype or some of their biologic deviations, consistent with a differentiation of RLS patients into two subgroups primarily separated by age of onset [5]. We assessed in this study both what information was obtained from these queries and the agreement among interviewers on the information collected.

## 2. Methods

### 2.1. Development of the instrument

By using the IRLSSG diagnostic criteria and the authors' clinical experience, specific questions were developed that allowed determination of the four key diagnostic features of RLS. The instrument was designed to be delivered by a trained interviewer via telephone. To facilitate comprehension, its vocabulary and sentence structure were designed so that they did not exceed an eigth grade reading level. The instrument was revised over a 2-year period on the basis of feedback from preliminary administrations to both patients and controls. There were seven diagnostic questions embedded in a mix of other questions. The diagnostic questions probed for the following information:

Did the subject have uncomfortable or unpleasant leg sensations when sitting or lying down?
Did the subject experience a need or urge to move when sitting or lying down?

These two questions are aimed at the first diagnostic criteria of urge or need to move that occurs due to sensations in the legs and also the third criterion that the symptoms occur at rest.

Were symptoms worse when the interviewee was lying down than when moving around?
Were symptoms relieved by moving around or walking?

These questions were aimed at further fulfilling the third diagnostic criterion, requiring provocation by rest and relief with activity.

Were symptoms worse at night?

This question was aimed at the fourth diagnostic criterion, the circadian accentuation during the night. It was expected that RLS patients would respond positively to the presence of all the features expressed in these questions.

Besides these questions aimed at ascertaining the diagnosis of RLS, the telephone interviews included seven other questions about the clinical status of those diagnosed with RLS. These questions asked whether sensations were painful, age at onset, and at what time of day they began. If patients were on medication, they were asked to indicate the time of day that symptoms would have begun if they were not taking medication. The latter question permits scoring the JHRLSS [2] in which severity of RLS is measured by progressively early times at which symptoms begin (0, no symptom; 1, symptoms start at bedtime or during night; 2, symptoms begin after 6 p.m., but before bedtime; 3, symptoms begin before 6 p.m.).

### 2.2. Subjects

After approval of the protocol by the human studies committees of both Johns Hopkins Bayview Medical Center (JHBMC) and the New Jersey Neuroscience Institute (NJNI) at JFK Medical center, a total of 75 patients and controls participated in the validation. The patients were being followed at the two institutions and had been diagnosed with RLS by board-certified sleep professionals who were expert in diagnosing RLS (R.P.A., C.J.E., A.S.W.). The controls were individuals familiar with the manifestations of RLS, who reported that they were free of RLS. However, they were not screened to be certain they did not have RLS.

### 2.3. Telephone interview

Interviews took place in four sessions with 15–20 subjects interviewed per session. Each subject was interviewed in sequence by each of three clinicians identified above (R.P.A., C.J.E., W.A.H.). The order in which the clinicians interviewed the subjects was constant within a given session, but varied across the four sessions. To ensure anonymity, contact was initiated by an independent staff member at JHBMC, who then transferred the subject to the first interviewing clinician. Each subject was introduced as Mr or Mrs Jones and was instructed to respond to the interviewers with that pseudonym. The subjects were not recognized, even by clinicians who were familiar with them. At the end of each interview, the subject was passed on to the next clinician until the three telephone interviews were completed.

During the telephone interview, each interviewer read the questions from a work sheet. If the subject was not clear about the question, the question was repeated with an emphasis on the unclear aspect. Efforts were made not to go beyond the question, but the interviewer could go beyond the script in order to clarify a question if the subject had difficulty understanding it as read. The interviewer then marked the subject's answers to each question. At the end of each telephone interview, the interviewer marked a diagnosis of RLS or not-RLS on the interview form. Each interview took between 2 and 10 min to complete. Times between successive subjects were kept to a minimum of 10 min.

### 2.4. Analysis

Sensitivity and specificity of the diagnoses were calculated by determining the match between the individual interviewers' diagnoses and the subjects' diagnoses. The sensitivity and specificity of the raters as a group was determined by comparing the consensus diagnosis to the subjects' diagnoses. The consensus diagnosis was either the unanimous diagnosis of the interviewers or the diagnosis of two out of three interviewers.

Inter-rater agreement among the three interviewers was calculated as an intraclass correlation coefficient (ICC) [6], which is a measure of how much of the total variance of measures can be attributed to differences between the subjects. Thus, the higher the ICC, the more the difference between ratings is due to different subjects and not different raters. Calculations of the ICC were done using SPSS.

## 3. Results

Thirty-seven subjects with RLS (19 women, 18 men) were interviewed, 28 from JHBMC and nine from NJNI. Average age (± standard deviation) was 68.1(± 12.0), range from 32 to 94 years. Thirty-eight control subjects (25 women, 13 men) were interviewed, 31 from JHBMC and seven from NJNI. Average age (± standard deviation) was 65.8 (± 13.1), range from 36 to 88 years.

The individual interviewers had sensitivities of 97, 100, and 100%. The sensitivity of the consensus diagnosis was 100%. The individual interviewers had specificities of 95, 92, and 92%. The sensitivity of the consensus diagnosis was 92%. Each interviewer correctly diagnosed 72 of 75 subjects, as did the consensus diagnosis.

The ICC among interviewers for diagnosis was 0.95. For the questions on the presence of leg sensations and urge to move, the ICC were 0.84 and 0.82, respectively. For the JHRLSS, the ICC among the interviewers was 0.92 for the patient group. For age of onset of RLS for the RLS subjects the ICC among the interviewers was 0.87.

### 3.1. Characteristics of the patient group

All RLS patients were rated as having leg discomfort and a need or urge to move by at least two interviewers (Table 1). Thirty-six of 37 were rated as having leg discomfort and 35 of 37 as having an urge to move by all three interviewers. One of the patients rated as not having an urge to move was diagnosed as not having RLS by one of the raters. All RLS patients were rated by the interviewers as having relief with walking.

Thirty-four RLS subjects were rated by at least two interviewers as having symptoms currently worse in the evening or night (characteristics of RLS patients are summarized in Table 2). Eight RLS subjects were rated as having painful leg discomfort by all three interviewers and six more by a pair of interviewers. Thus, by consensus 14 of 37 RLS subjects (38%) had some degree of pain as a part of their RLS leg discomfort.

Mean age of onset for RLS symptoms was 43.5 years of age. Excellent agreement was found for onset age (within two years for all interviewers) in 33 RLS subjects – for two of the four other RLS subjects, it appears that age of onset was confounded with duration of illness by an interviewer. Approximately half (19/37) of the RLS subjects had early onset (< age 45) [5], but age at onset varied over the entire lifespan (from 5 to 80 years). Mean duration of illness of the RLS subjects was 24.6 years (range 1–59 years).

Thirty-five of the RLS subjects (94%) were rated by at least two interviewers as having daily symptoms and all still had their symptoms in the period of the interview. Nineteen RLS subjects were rated as having symptoms that began before 6 p.m. by at least two interviewers (i.e. the highest level of severity on the JHRLSS). Twelve had symptoms beginning between 6 p.m. and bedtime (JHRLSS = 2) and six had symptoms that began at bedtime or during the night (JHRLSS = 1).

### 3.2. Characteristics of the control subjects

Ten control subjects (26%) were rated as having leg discomfort by at least two raters (seven unanimous) (Table 1). Six control subjects (16%) were rated as having a need or urge to move by at least two raters (four unanimous). In contrast, 23 of the controls (60%) were unanimously rated

Table 1
Ratings of patients and controls: diagnosis and sensory symptoms

| | Patients | | Controls | |
|---|---|---|---|---|
| Rating | Yes | No | Yes | No |
| Diagnosis of RLS | 37 (100%) | 0 (0%) | 3 (8%) | 35(92%) |
| Leg sensations | 37 (100%) | 0 (0%) | 10 (26%) | 28(74%) |
| Urge to move | 37 (100%) | 0 (0%) | 6 (16%) | 32(84%) |

The ratings are the interviewers' rating of the RLS diagnosis and the two prinicipal sensory symptoms as present or absent; the consensus rating is used: either a unanimous rating or that of two out of three raters.

Table 2
Characteristics of rls patient group

| | |
|---|---|
| Average age at onset (SD) | 43.5 (20.9) |
| Symptoms worse in evening or night | 34 Patients (92%) |
| Painful leg discomfort | 14 Patients (38%) |
| Daily symptoms | 35 Patients (95%) |
| Symptoms begin before 6 p.m. | 19 Patients (51%) |

Average age of onset based on average of interviewers; presence of four bottom features based on consensus – unanimous or agreement of two of three interviewers; in Johns Hopkins RLS Severity scale time of day indicates severity – the most severe patients experience symptoms before 6 p.m.

as having no leg discomfort and 26 of the controls (68%) were unanimously rated as having no urge to move.

All three of the control subjects with a consensus diagnosis of RLS were found unanimously to have leg discomfort and two of three were found by consensus to have an urge to move.

## 4. Discussion

### 4.1. Accuracy

The sensitivity of the TDI proved to be excellent at 97%. The specificity was somewhat lower at 92%. The lower specificity, however, may have been due in part to mischaracterization of the control group, which was not rigorously screened before interview to exclude the possibility of RLS. Indeed, with further contact with one volunteer, it became clear that this individual did have RLS and this control subject had been correctly identified by the interviewers as having RLS. Therefore, the specificity we measured was probably an underestimate of the true specificity achievable with this instrument.

Inter-rater reliability, as measured by ICC, was acceptable for diagnosis and also for the responses to the two diagnostic questions answered by all subjects (presence of discomfort and urge to move) as well as key aspects of RLS manifestation, the timing of symptom onset and the age of symptom onset. ICC for each of these exceeded 0.8, a level considered more than acceptable for a rating instrument [7]. Because the interviewers were sleep specialists who were quite familiar with RLS diagnosis, however, we cannot exclude the possibility that less well trained and expert interviewers might not perform to the same standard. This interview is designed for use by interviewers knowledgeable about RLS and not by untrained interviewers. Additional sensitive and specific instruments that can be administered by those without a knowledge of RLS or can be completed by a member of the lay public await future development.

The high ICC for determination of both the JHRLSS (0.92) and the age of onset (0.87) indicates that key elements of the clinical picture of RLS subjects useful for therapeutic trials or epidemiological study can be obtained reliably with the TDI.

### 4.2. Characteristics of the patient and control groups

Although they have a wide range of ages, the RLS subjects, relatively well matched to the controls, were on average in the later middle aged to elderly group. This is typical of patients who present for clinical treatment of RLS. Their disease duration, on average 25 years, suggests that for most of these RLS subjects, the problem has been long-standing, which has been noted previously as typical of RLS patients [8]. In this sense, they may differ as a group from those individuals in the general population who might be diagnosed with RLS. They, like their selected controls, may have been more knowledgeable about RLS and able to answer questions more accurately and decisively. The positive responses of control subjects to questions about leg symptoms also indicates that control subjects can present with a variety of leg complaints that are not caused by RLS. Some of these positive answers likely derive from conditions such as cramps [9], hypotensive akathisia [10], positional discomfort, peripheral neuropathy, or local pathology in the legs (in the skin, muscle, vasculature, bone, joint, etc.) that can be confused with RLS. Cramps, in particular, can occur in an individual at night while at rest, provoke an urge to move, and can be relieved by movement or walking. These symptoms that mimic those in RLS present the greatest challenge for RLS diagnosis, but the TDI seems to have accurately discriminated these individuals from RLS patients. We are currently examining whether adding explicit questions to the TDI to exclude such conditions will improve diagnostic accuracy in a general population.

### 4.3. Future developments

It became clear in the course of this study that at least one presumed control subjects did have RLS, which was actually identified by the TDI. In future studies on RLS comparing patients and controls, it will be important to ensure that control subjects are carefully screened to exclude those with unknown or undisclosed RLS. It was also clear that the subjects for this study, especially the patients but also the controls, were familiar with RLS and thus could readily answer the specific questions. This may be less true in populations where those with RLS are unaware of their diagnosis and most, if not all, are unfamiliar with the syndrome. Determination of the performance of the TDI when applied to such a population needs to be assessed. The authors are now undertaking such an assessment in a less knowledgeable sample.

### 4.4. Addendum

Since these studies were conducted, a workshop was held at the National Institutes of Health (NIH) (May 1–3, 2002) under the auspices of the National Institute of Aging, the Restless Legs Syndrome Foundation, and other NIH institutes in order to revise the diagnostic criteria for

RLS (publication of formal report elsewhere in this issue, pages 11–30). Several of the authors of the current paper attended this meeting and contributed to the revisions (R.P.A., C.J.E., W.A.H., A.S.W.). In brief, the relevant decisions of the workshop were a revision of the 1995 diagnostic criteria [1] to achieve greater clarity. The major changes are deletion of a requirement of motor restlessness, because this criterion had caused substantial misunderstanding, and the division of the relief with movement and provocation by rest into two separate criteria. These changes should not impact on the validity of the TDI, since it does not include a specific question on motor restlessness and also addresses the issues of both relief with movement and provocation by rest.

## References

[1] The International Restless Legs Syndrome Study Group (Arthur S., Walters M.D. – Group Organizer and Correspondent), Towards a better definition of the restless legs syndrome. Mov Disord 1995;10: 634–42.

[2] Allen RP, Earley CJ. Validation of the Johns Hopkins restless legs severity scale. Sleep Med 2001;2(3):239–42.

[3] Ondo W, Jankovic J. Restless legs syndrome: clinicoetiologic correlates. Neurology 1996;47(6):1435–41.

[4] Winkelmann J, Wetter TC, Collado-Seidel V, et al. Clinical characteristics and frequency of the hereditary restless legs syndrome in a population of 300 patients. Sleep 2000;23(5):597–602.

[5] Allen RP, Earley CJ. Defining the phenotype of the restless legs syndrome (RLS) using age-of-symptom-onset. Sleep Med 2000;1(1): 11–19.

[6] Shrout PE, Fleiss JL. Intraclass correlations: use in assessing rater reliability. Pschol Bull 1979;86:420–8.

[7] Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use, 2nd ed. New York, NY: Oxford University Press, 1995.

[8] Walters AS, Hickey K, Maltzman J, et al. A questionnaire study of 138 patients with restless legs syndrome: the 'night-walkers' survey. Neurology 1996;46(1):92–5.

[9] Abdulla AJ, Jones PW, Pearce VR. Leg cramps in the elderly: prevalence, drug and disease associations. Int J Clin Pract 1999;53(7): 494–6.

[10] Cheshire WP. Hypotensive akathisia: autonomic failure associated with leg fidgeting while sitting. Neurology 2000;55(12):1923–6.