Movement Disorders
CLINICAL PRACTICE
REVIEW

Review of Severity Rating Scales for Restless Legs Syndrome: Critique and Recommendations

Arthur S. Walters, MD,[1,] Birgit Frauscher, MD,[2] Richard Allen, PhD,[3] Heike Benes, MD,[4,5] K. Ray Chaudhuri, MD,[6] Diego Garcia-Borreguero, MD,[7] Hochang B. Lee, MD,[8] Daniel L. Picchietti, MD,[9] Claudia Trenkwalder, MD,[10] Pablo Martinez-Martin, MD, PhD,[11] Anette Schrag, MD,[12] Glenn Stebbins, PhD[13]*

Abstract: Over the last decade, research in restless legs syndrome (RLS; also known as Willis-Ekbom disease) has increased dramatically. The International Parkinson and Movement Disorder Society commissioned a task force to formally evaluate the available evidence on severity rating scales in RLS. A literature search retrieved instruments specific to RLS. Each scale was evaluated by three criteria: (1) use in RLS; (2) use by research or clinical groups other than the group that developed the scale; and (3) formal validation and adequate clinimetric properties. Scales were then qualified as "recommended" when all three criteria were met, "suggested" when used for RLS but only one of the other criteria was met, and "listed" when only used in RLS. Details regarding the development, use, and psychometric properties of each instrument and the recommendations of the committee are summarized. The scale of the International Restless Legs Syndrome Study Group for rating the severity of RLS (International Restless Legs Scale or IRLS) and the Augmentation Severity Rating Scale fulfilled criteria for "recommended" instruments to assess severity. Future endeavors should include a validation of the Pediatric RLS Severity Scale, the only available instrument for evaluation of the severity of pediatric RLS, and a validation of a patient version of the IRLS that will not require the interface of a live interviewer.

Restless legs syndrome (RLS; also known as Willis-Ekbom disease [WED]) is a common disorder affecting 7.2% of the population, with at least one third of these severe enough to warrant treatment. The impact on quality of life of these RLS sufferers approaches that observed in type 2 diabetes mellitus, depression, or hypertension with osteoarthritis.[1]

A number of rating scales have also been developed to address the overall severity of RLS. The purpose of the current review is to evaluate the strengths and weaknesses of the current available severity rating scales for RLS, educate the reader as to the circumstances under which each scale is to be used, and make an overall recommendation for use of each of the scales. This critique does not focus on RLS diagnostic instruments or on quality-of-life scales or sleep scales with items assessing RLS. These are reviewed elsewhere (A.S. Walters, B. Frauscher & R. Allen, Submitted-a; A.S. Walters, B. Frauscher & R. Allen, Submitted-b).

Methods

Administrative Organization and Critique Process

This project has been organized under the leadership and guidance of the International Parkinson and Movement Disorder Society (MDS). The senior author of the project (A.S.W.) was

[1]Department of Neurology, Vanderbilt University School of Medicine, Nashville, Tennessee, USA; [2]Department of Neurology, Medical University of Innsbruck, Innsbruck, Austria; [3]Johns Hopkins University, Baltimore, Maryland, USA; [4]Institute for Medical Research and Sleep Medicine, Schwerin, Germany; [5]Department of Neurology, University of Rostock, Rostock, Germany; [6]National Parkinson Foundation Center of Excellence, Kings College, London, United Kingdom; [7]Sleep Research Institute, Madrid, Spain; [8]Department of Psychiatry, Yale University, New Haven, Connecticut, USA; [9]University of Illinois School of Medicine and Carle Foundation Hospital, Urbana, Illinois, USA; [10]Department of Neurology, Georg-August University, Goettingen, Germany; [11]National Center of Epidemiology and CIBERNED, Carlos III Institute of Health, Madrid, Spain; [12]UCL Institute of Neurology, University College, London, United Kingdom; [13]Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois, USA

[*]**Correspondence to:** Dr. Arthur S. Walters, Department of Neurology, Vanderbilt University School of Medicine, MCN A-0118, 1161 21st Avenue South, Nashville, TN 37232, USA; E-mail Arthur.Walters@Vanderbilt.edu
Keywords: restless legs syndrome, Willis-Ekbom disease, evaluative and discriminative instruments, severity rating scales.
Members of the MDS Committee on Rating Scales are listed in the Appendix.
Relevant disclosures and conflicts of interest are listed at the end of this article.
Received 14 July 2014; revised 19 August 2014; accepted 19 August 2014.
Published online 30 September 2014 in Wiley InterScience (www.interscience.wiley.com). DOI:10.1002/mdc3.12088

doi:10.1002/mdc3.12088

asked to head and form the working group on the current submission with the goal of evaluating severity rating scales specific to RLS.

Literature Search and Selection of Scales

A PubMed literature search was conducted. The literature search included all publications on the appropriate topics to July 2013. When the terms "restless legs syndrome" or "RLS" were crossed with the term "sleep scale," 106 articles were identified, of which two were pertinent. When the terms "restless legs syndrome severity scale" or "RLS severity scale" were employed, 246 articles were identified, of which 14 were judged to be pertinent. When the terms "pediatric restless legs syndrome" or "pediatric RLS" were used, there were 66 articles identified, of which two were pertinent. When the terms "restless legs syndrome" or "RLS" were crossed with the term "augmentation," 123 articles were identified, of which three were thought to be relevant. These overlapping strategies identified five scales pertinent to severity in RLS.

Review Criteria

The following criteria were used to designate a scale as "recommended," "suggested," or "listed" for use in RLS (Table 1). If an RLS severity rating scale has been applied to RLS-related populations, has been employed by investigators other than the group that originally developed it, and has satisfactory clinimetric properties, that instrument is "recommended;" if it has been applied to RLS, but meets only one of the other two requirements, it is "suggested;" and if it has only been applied to RLS, but does not meet either of the other two criteria, it is "listed". It is cautioned that the tested clinimetric attributes of the scales can only be applied in the circumstances and to the populations in which they were validated.

For each scale, the determination as to whether the scale (1) has been used in RLS populations, (2) has been employed by investigators other than those that developed it, and (3) has adequate clinimetric properties is summarized under the "key criteria" section for each scale.

In general, the elements considered for proper evaluation of the scales included sensitivity, specificity, positive predictive value, negative predictive value, inter-rater reliability, Cronbach's alpha, factor analysis, criterion validity, convergent validity, divergent validity, known groups validity, and responsiveness to change. The group did not require a minimum number of these variables. Rather, the group made an overall decision about each validation study based upon an overview of the number of the variables and the quality of the data. Determination as to the strength of the statistical results obtained from the studies, for example, whether correlation coefficients were weak, moderate, or high and whether effect sizes were small, moderate, or large, were done in consultation with one of our coauthors (P.M.M.) and by standard methodology.[2]

TABLE 1 Definition of the categories for evaluation of RLS Instruments

Category	Criteria
Recommended	(1) Instrument has been applied to RLS populations. (2) Other groups beyond the original developing group have published data on the scale and its clinical utility. (3) Psychometric studies are available, which conclude that the scale is valid and reliable.
Suggested	Instrument has been applied to RLS populations, but only one other criterion (2 or 3) from the above recommended category applies.
Listed	Instrument has been applied to RLS populations, but no further criterion is met.

Results

Based on this search strategy, five instruments pertinent to severity in RLS/WED were identified. One scale, the International Restless Legs Scale (IRLS), has been used in almost all treatment trials of RLS. The Johns Hopkins Restless Legs Severity Scale (JHRLSS) and the RLS-6, as with the IRLS, focus on the severity of adult RLS under normal baseline and therapeutic conditions, but other scales focus on different aspects of RLS severity. The Augmentation Severity Rating Scale (ASRS) focuses on paradoxical worsening of RLS symptoms under dopaminergic treatment. The Pediatric RLS Severity Scale (P-RLS-SS) focuses on RLS severity in children where verbal skills may not be adequate to complete the adult severity rating scales. Each of these rating scales is discussed with summary information found in Table 2 (determination of the MDS Task Force criteria for each individual rating scale and the final determination of the level of endorsement) and Table 3 (summaries of the analyses for each individual instrument).

RLS Severity in Adults

The International Restless Legs Syndrome Study Group rating scale for the severity of restless legs syndrome (IRLS)

The IRLS was developed by the International RLS Study Group by soliciting questions from its membership. Several versions were created with multiple inputs from multiple members of the group until a final version was established. It is reflective of clinical practice where patients must have a known diagnosis of RLS before administration of the scale. The IRLS consists of 10 questions rated from 0 to 4. The scale is validated under conditions of a face-to-face interview with the patient where clarifications regarding the questions can be made to the patient. The scale takes approximately 10 minutes to complete. The patients selected for the validation study had a diagnosis of RLS

doi:10.1002/mdc3.12088

TABLE 2 Overall evaluation of different severity rating scales for RLS

Instruments for RLS	Use in RLS	Use by Other Investigators	Adequate Clinimetrics	Appropriate Population	Conclusions
IRLS	Yes	Yes	Yes	All RLS	Recommended Baseline and Rx
JHRLSS	Yes	Yes	Yes	RLS Sx >5 days/week	Recommended Baseline
RLS-6	Yes	Yes	Abstract only	All RLS	Suggested Baseline and Rx
RLS severity in children P-RLS-SS	Yes	No	No	RLS ages 6–17 years	Listed
Augmentation severity ASRS	Yes	Yes	Yes	All RLS	Recommended Baseline and Rx

Baseline, scale may be used at a single point in time; Rx, scale may be used to determine responsiveness to change under treatment conditions.

TABLE 3 Detailed analysis of different severity rating scales for RLS

Instruments for RLS	SENS, SPEC	IRR by Kappa or ICC or Other, Test-Retest	CA, FA	CRIT, CONV, DIV, KG	Response ES
RLS severity in adults					
IRLS		IRR ICC = 0.97 Test-retest ICC = 0.87	CA = 0.95 2 factors with 64.3% of variance	Crit = 0.73. Div $P < 0.001$. Conv = 0.78. Item total all >0.4	Response Max $r = 0.75$; $P < 0.0001$
JHRLSS		IRR = 0.91 by Spearman's rank and 0.87 by Cramer's V		Conv $R = 0.60$ with SE and $R = 0.45$ with PLMS	
RLS-6			Two factors	Crit $r = 0.69$ max with IRLS. Conv $r = 0.23$ max with PLMS arousal index. Mean inter correlation of items $r = 0.37$	Response 95% CI max ES (−6.01; −1.80)
RLS severity in children					
P-RLS-SS			Strong face and content validity, but scale not yet validated		
Augmentation severity					
ASRS	Sens = 82% Spec = 92%	IRR = 0.94 Test-retest = 0.72	CA = 0.62 1 factor	Crit = 0.72. Div $P < 0.0001$	Measurement of Augmentation before and after levodopa

Sens, sensitivity; Spec, specificity; IRR, inter-rater reliability (as measured by Cohen's kappa or ICC or other); CA, Cronbach's alpha (which is a measure of the internal consistency of the instrument); FA, factor analysis (which is a method applied to inform about the dimensionality of the instrument [construct validity]); Crit, criterion validity (which means comparison of the instrument to a gold standard); Conv, convergent validity (which refers to the correlation of the instrument with another measure of RLS or its comorbidities); Div, divergent or discriminant validity (which refers to the correlation of the instrument with another measure not related to RLS); KG, known groups (discriminative) validity (which refers to the ability of the instrument to detect differences between groups at a point in time); Response, responsiveness (which a measure of change of an instrument over time with or without treatment); ES, effect size; CC, correlation coefficient; CI, confidence interval; SE, sleep efficiency; PLMS, periodic limb movements in sleep.

made in the clinic by an expert on RLS who was well aware of the diagnostic criteria for RLS and its differential diagnosis.

Key Criteria. The IRLS was designed for use in RLS, has been widely used by other investigators, and is the primary instrument used to determine RLS severity in pharmaceutical and nonpharmaceutical studies of therapeutic agents for RLS. In regard to clinimetric properties, an initial factor analysis suggested two factors, but, because several items loaded on both factors, it was decided to use a single-factor solution.[3] However, the scale was constructed with both Severity and Life Impact factors in mind and a second factor analysis showed a more discrete division into these categories, with a total of 64.3% of the variance explained.[4] The scale shows good test-retest properties at 2 weeks (intraclass correlation coefficient [ICC] = 0.87; $P < 0.001$), good inter-rater reliability performed under blinded conditions (ICC = 0.93 for the first administration and 0.97 for the second administration 2 weeks later; $P < 0.001$), good internal consistency with a Cronbach's alpha of 0.93 and 0.95 for the two administrations, satisfactory criterion validity when tested against a Clinician's Global Impression (CGI; $r = 0.74$ on the first administration and 0.73 on the second administration; $P < 0.001$), and good discriminant validity in that the scale showed markedly more-severe IRLS scores, when compared to those from either a normal control group or a sleep-disordered control group ($P < 0.001$ for both).[3] The scale also showed

doi:10.1002/mdc3.12088

high concurrent or convergent validity with a Patient Global Impression (PGI; 0.82 on the first administration and 0.78 on the second administration; $P < 0.001$ for both). Corrected item-total correlations were always higher than 0.40.[3] In a subsequent publication, a detailed item-response analysis showed that each item of the IRLS was correctly scaled to assess the overall range of RLS severity.[5] There was a floor effect that was observed with item 3 (relief by activity), perhaps because this is a diagnostic criterion for RLS. In yet another publication employing two matching placebo-controlled studies of ropinirole for treating RLS where patients were evaluated several times during a 12-week period, there was excellent responsiveness.[6] There was high correlation between the overall change in the IRLS total score and subscale scores with the overall change in CGI (maximum $r = 0.75$). This responsiveness was further documented by changes in total IRLS scores and subscale scores that paralleled changes in the seven levels of CGI, ranging from much worse to very much improved ($P < 0.0001$; effect size range: −0.23 to −3.67).[6] Improvements in the total IRLS score and subscale scores were larger in those patients who were more improved on the CGI.

Strengths and Weaknesses. The IRLS has excellent clinimetric properties and has been validated in cross-sectional studies as well as pre-/post-treatment studies investigating responsiveness to change. It is the most extensively used of the RLS severity scales in research studies of all types. One criticism of the development of the IRLS is that it did not use patient input or focus groups as part of the development process. In addition, it does not capture all the possible severity measurements of RLS, such as time of day of usual onset of symptoms, latency of time to symptom onset when sitting, or a determination of what body parts other than arms or legs might be involved in RLS. Other severity scales, such as the JHRLSS or the ASRS, have been more successful in capturing these aspects.[7,8] It is to be cautioned that the IRLS can only be reliably predicted to give excellent results in a clinical setting where the interviewer is present to clarify any misunderstandings the patients may have regarding the 10 items.[3] Based upon observed clinical experience, however, the IRLS is often given to the patient as a self-rating scale. Whether the results can be reliably used in this way has yet to be proven. Plans for the validation of the IRLS as a self-rating scale, however, are now under way. Other studies have indicated that the IRLS has a large placebo effect.[9] However, the placebo effect may, in addition, be a property intrinsic to RLS itself rather than simply a scale property. For example, the spontaneous fluctuations in symptoms commonly observed in RLS may imply regression to the mean. In addition, RLS symptoms are possibly more suggestible than symptoms from other neurological conditions.

Conclusions. The IRLS was specifically developed for RLS, it is the most extensively used RLS instrument, serving as the primary endpoint for virtually all therapeutic studies of RLS, and it shows excellent clinimetric properties in a stable clinic population. It also shows excellent response to change in a clinical trial setting. It has now been used as the gold-standard outcome measure for treatment trials in RLS. The IRLS has been translated into multiple languages and has been used in multiple settings by multiple groups. The majority of the translations have been performed by MAPI Research Trust, which holds the patent to the scale for the International RLS Study Group. Back translations were performed in each case to assure accuracy of the information transmitted. There have been independent published validations of the scale in both Japanese and Portuguese.[10,11] Therefore, it is "recommended" for use in office setting or in research (e.g., clinical trials) for the evaluation of severity and as a measure of therapeutic effect (Table 2).

The Johns Hopkins Restless Legs Severity Scale (JHRLSS)

This scale was the first validated severity scale for RLS and is comprised of a single question that probes the usual time of onset of RLS symptoms for at least 50% of days. Patients are asked, "At what time of day do these feelings usually start?" Symptoms are then rated as follows: 0 = no symptoms; 1 = bedtime symptoms after or within an hour of going to bed; 2 = evening and bedtime symptoms starting at or after 6:00 PM; and 3 = day and night symptoms starting before 6:00 PM.[7] The rationale for using 6:00 PM as an exact time of reference is not given in the publication, but is compatible with the well-established circadian rhythmicity of the disorder codified in the obligatory diagnostic criteria for RLS, which includes a requirement that the symptoms be worse later in the day or at night. In later versions, an expanded scale was sometimes used where onset before noon was given a score of 4 and less than almost daily RLS was assigned a score of 0.5.[12] Modified versions of the questions from the scale have also been incorporated as part of a larger questionnaire whose purpose is to diagnose RLS.[12] The scale was validated by chart reviews of patients who had been diagnosed in the clinic by RLS experts fully aware of the diagnostic criteria and differential diagnosis of RLS. The scale takes approximately 3 minutes to complete.

Key Criteria. The JHRLSS was developed for use in RLS and has been used by groups other than the one that developed it. In regard to clinimetric properties, inter-rater reliability of the JHRLSS scores was 0.91 ($P < 0.05$; Spearman's rank coefficient). Cramer's V for inter-rater agreement was 0.87 ($P < 0.05$). Each patient had an all-night polysomnogram, and the JHRLSS shows high correlation with sleep efficiency ($R = 0.60$; $P < 0.01$) and moderate correlation with Periodic Limb Movements of Sleep (PLMS) per hour of sleep ($R = 0.45$; $P < 0.01$).[7]

Strengths and Weaknesses. The clinimetric properties of this scale are good and it has the advantage of being easy to administer given that it consists of only one question. The validation was done on patients who had symptoms at least 5 days a week

doi:10.1002/mdc3.12088

(R.A., personal communication), and it is not at all clear that this scale is valid for patients with less-frequent symptoms. It has not yet been validated for treatment response.

Conclusions. The JHRLSS was developed for RLS, has been used by groups other than the group that designed it, and it shows good inter-rater reliability and correlation with other known measures of RLS severity. It is therefore "recommended" for use for baseline evaluation of RLS in a clinical setting for patients with symptoms 5 days a week or more (Table 2). However, the recommendation comes with limitations because it has not been formally tested for responsiveness to change under treatment conditions. Because it has not been validated for responsiveness to treatment, the overall recommendation for the JHRLSS is "suggested".

The RLS-6

The RLS-6 has six items rated from 0 to 10 covering the previous week that probe satisfaction with sleep, severity of RLS (at falling asleep, during the night, during the day when sitting or lying down, and during the day when active), and a final item probing daytime sleepiness.[13] For the validation study, 259 patients with idiopathic RLS (68% female and mean age 57 years of age) were evaluated at baseline in a cabergoline study. The cohort employed for the validation was comprised of RLS subjects who had undergone an office-based clinical diagnosis by a clinical expert in RLS with full knowledge of the diagnostic criteria for RLS and its differential diagnosis. The scale takes approximately 5 to 10 minutes to complete.

Key Criteria. The RLS-6 was developed for use in RLS and has been employed by groups other than the one that validated it. In regard to clinimetric properties, factor analysis yielded two factors with two items of daytime severity in one factor and the remaining items in the other factor. The mean inter-item correlation was $r = 0.37$. Correlation between the IRLS total score and the RLS-6 items was highest for severity at night ($r = 0.69$). The correlation of the two daytime severity items was $r = 0.45$ at rest and $r = 0.41$ when active. In a second cabergoline trial[13] using polysomnography, the correlation between the PLMS arousal index and the RLS items varied between $r = -0.11$ (daytime sleepiness) and $r = +0.23$ (severity at bedtime). The scale showed good responsiveness to treatment after 5 days with either 2 mg of cabergoline or placebo for four of the six questions, with the exception of severity at bedtime and tiredness during the day. The maximum effect size (95% confidence interval [CI]) was observed for "severity during the night" (−6.01; −1.80), which are considered large effect sizes.

Strengths and Weaknesses. Although the RLS-6 has been used widely in other research publications, the results of the validation study cannot be fully evaluated because they have been published only in abstract form.

Conclusions. The RLS-6 was developed for RLS and has been used extensively in large- and small-scale pharmaceutical trials, but the validation of the scale, to our knowledge, is only published in abstract form. This scale is therefore "suggested" (Table 2).

RLS severity in Children

The Pediatric RLS Severity Scale (P-RLS-SS)

Childhood RLS has a significant impact on functioning in up to 0.5% of children and 1% of adolescents.[14] For this reason, it was important to develop the P-RLS-SS, which was carefully constructed by a multistep iterative process between children with RLS, their parents, and clinicians.[15] The scale was developed by scale experts and clinicians with expertise in the diagnosis of pediatric RLS and its differential diagnosis in children. This scale has strong face and content validity, but is not further considered because it has not been validated. It is therefore "listed".

Augmentation Severity

The Augmentation Severity Rating Scale (ASRS)

Augmentation remains a significant problem arising from dopaminergic treatment in RLS.[16] Briefly put, augmentation is a paradoxical worsening of RLS symptoms from dopaminergic therapy. The significance of this problem and the importance of measuring it in therapeutic trials led to the development of the ASRS.[8] The ASRS can only be employed if it is administered before starting dopaminergic therapy as well as after because baseline data are needed in order to demonstrate worsening of symptoms with dopaminergic treatment. It was developed by an interactive process involving professional and patient input. The final scale has three items and probes symptoms over the past week. The first item asks the time at which symptoms began in the last week. The second item probes how quickly symptoms begin when sitting at various times of the day during the past week, and the third item probes which body parts were involved in the past week. Each item has a total of 8 points for a 24-point total scale. The full validation was done on a cohort of patients diagnosed by RLS experts with full knowledge of the diagnostic criteria for RLS and its differential diagnosis and with full knowledge of the criteria for augmentation. The scale was validated in face-to-face interviews with the patients. The scale takes around 10 minutes to complete with an additional 5 minutes to calculate a score based upon the scoring algorithm.

Key Criteria. The ASRS has been developed for the evaluation of augmentation in RLS and has been employed by groups other than the one that validated it. In terms of clinimetric properties, the validation process involved 63 mostly untreated RLS patients from six centers across Europe. Translation plus back-translation of the scale was done in each language to assure

doi:10.1002/mdc3.12088

accuracy of the translation. The scale was administered before and after treatment with levodopa up to 500 mg/day as clinically needed. Thirty-six (60%) of the patients experienced augmentation. As originally conceptualized, the scale had an additional fourth item ("What was the usual severity of your RLS symptoms?") that was excluded after item analysis because it did not correlate well with the other items. The final three-item ASRS had a high correlation between test and retest of 0.72 off medication, as determined by a repeat assessment after 2 weeks. The ASRS total score inter-rater reliability analysis showed a very high correlation coefficient for two independent raters of 0.94. Construct validity was determined by a factor analysis with sufficient loadings on a single factor. Internal consistency of the scale, as measured by Cronbach's alpha, was 0.62, which is considered slightly weak. Criterion validity or external validation of the scale was demonstrated by a high Spearman's correlation of 0.72 between the worst augmentation score under treatment and the independent rating of an expert as well as by moderate correlations between Augmentation Severity Scores and the CGI for Augmentation. The scale also showed excellent discriminant validity given that the ASRS scores were significantly different between subjects with (mean score: 7.4) and without augmentation (mean score: 2.0; $P < 0.0001$ for total score). A receiver operating characteristic plot for different total scores of the ASRS suggests that a score of 5 optimizes the balance between an acceptable sensitivity (82%) and a satisfactory specificity (92%), and this is the cutoff recommended for augmentation.[8]

Strengths and Weaknesses. The scale has good clinimetric properties and is useful in a research clinical trial setting to prove or disprove the presence of augmentation during the administration of dopaminergic or nondopaminergic RLS medications. Given that, to date, augmentation has been shown to primarily occur with dopaminergic therapy, theoretically, the scale would be useful in a clinical office setting when dopaminergic medications are administered. However, a baseline ASRS must be administered before dopaminergic therapy is even started. Clinicians in a busy practice are not likely to administer this to their patients before therapy for the purposes of evaluating a side effect that may never happen. In addition, it must be pointed out that the ASRS only measures the severity of augmentation and cannot be substituted for a severity measure of RLS, such as the IRLS. It is recommended that the IRLS or other measure of primary RLS severity be administered along with the ASRS.

Conclusions. The scale was developed for RLS, it has been used by groups other than the one that originally designed it, and it has good clinimetric properties. It is therefore "recommended" for the determination of the severity of augmentation in RLS patients who have been treated with dopaminergic agents (Table 2). It is to be cautioned that the scale can only be used if a baseline value is also obtained before the institution of dopaminergic therapy. For logistical reasons therefore, it may more useful as a research tool than in clinical practice.

Discussion

The MDS commissioned a task force to formally evaluate the quality of RLS severity scales (summarized in Table 2). We would caution that the reliability and validity of the instruments cannot be guaranteed for use in circumstances and populations other than those in which they were validated. The severity rating scales that have been validated under baseline and responsiveness to change conditions as well as employed widely by other groups are the IRLS and the ASRS. A baseline value for the ASRS must be obtained before treatment is started.

The IRLS and the ASRS are useful clinically and experimentally and have adequate clinimetric properties for continued use. The group is therefore not recommending the development of new scales at this time. Future endeavors should include a validation of the P-RLS-SS, which is the only available tool for the diagnosis of pediatric RLS. The validation of the IRLS as a self-administered patient scale without the need for the intervention of a clinician is also desirable. The minimal requirement for use in clinical evaluation of RLS patients is a knowledge of the four basic diagnostic criteria (urge to move the legs, worse at rest, relief by activity, and worse later in the day or at night) plus the new fifth criterion (exclusion of the conditions that may mimic RLS).[17] Although some of the validations presented in these studies were done before formalization of the need to do a differential diagnosis, all of the scales evaluated here did employ RLS experts to make the diagnosis of RLS in the cohorts studied with full knowledge of the diagnosis of RLS and its differential diagnosis, as has always been true. A formal evaluation of the severity of RLS as reviewed in this article can be a useful clinical and research accompaniment.

Finally, it should be stated that the use of any of these scales in populations with RLS secondary to peripheral neuropathy, radiculopathy, renal failure, pregnancy, and so on, is limited, although there is every reason to believe that they would perform similarly in these populations. Further studies, of course, would be needed to determine this.

Author Roles

(1) Research Project: A. Conception, B. Organization, C. Execution; (2) Statistical Analysis: A. Design, B. Execution, C. Review and Critique; (3) Manuscript: A. Writing of the First Draft, B. Review and Critique.

A.S.W.: 1A, 1B, 1C, 2C, 3A, 3B
B.F.: 1A, 1B, 1C, 2C, 3B
R.A.: 1B, 2C, 3B
H.B.: 1B, 2C, 3B
K.R.C.: 1B, 2C, 3B
D.G.-B.: 1B, 2C, 3B
H.B.L.: 1B, 2C, 3B
D.L.P.: 1B, 2C, 3B
C.T.: 1B, 2C, 3B
P.M.-M.: 1A, 1B, 1C, 2C, 3B
A.S.: 1A, 1B, 1C, 2C, 3B
G.S.: 1A, 1B, 1C, 2C, 3B

doi:10.1002/mdc3.12088

Acknowledgments

The authors thank Patience Bridges, CAP, for her help in the preparation of this manuscript. This article is dedicated to the memory of our dear and recently departed colleague, Dr. Ralf Kohnen, who was the developer and validation leader for many of the scales and evaluative instruments we employ in RLS/WED research today, including the widely used RLS-6 and RLS Quality of Life scales. For the RLS/WED community, he was the principal person to whom we all turned for advice and help in the development and validation of these instruments. He will be much missed.

Disclosures

Funding Sources and Conflicts of Interest: The authors report no source of funding and no conflicts of interest.

Financial Disclosures for previous 12 months: Arthur S. Walters served as a consultant for and has received research funding from UCB Pharma and Mundi Pharma. He has also received funding from the USA National Institutes of Health.He has also served as an expert witness in a single legal case involving Restless Legs Syndrome. Birgit Frauscher MD received research funding from the Austrian Science Fund, a Schroedinger Fellowship Abroad and the National Bank of Austria Anniversary Fund. Richard Allen, PhD served as a consultant for Luitpold Pharmaceuticals, Pfizer, Pharmacosmos, Impax Pharmaceuticals and UCB Pharma. He also received research support from Glaxo Smith Kline, Pharmacosmos and the USA National Institutes of Health. Heike Benes, M.D. has received payment for lectures and has served on the medical advisory boards of UCB Pharma and MundiPharma. K Ray Chaudhuri has done consulting for UCB Pharma and Abbott, honoraria fromBoehringer-Ingelheim, Britannia, Abbott, Cephalon, and UCB, and research grants from Parkinson's UK, PDNMG, and the UK Department of Health. Diego Garcia-Borreguero, M.D. has received honoraria for advisory boards or lectures from Xenoport, UCB Pharma, Pfizer, Impax Pharmaceuticals and MSD. Hochang Benjamin Lee, M.D. has no financial disclosures to report. Daniel Picchietti, M.D. has served on the advisory board of the Willis-Ekbom Disease Foundation and has received royalities for updating a book chapter. Claudia Trenkwalder, M.D. has received honoraria from Boehringer-Ingelheim, UCB Pharma, Glaxo Smith Kline, MundiPharma, Desitin and Novartis. Glenn T. Stebbins, PhD has received consulting and has Advisory Board Membership with honoraria from Acadia Pharmaceuticals, Adamas Pharmaceuticals, Inc, Ceregene, Inc, CHDI Management, Inc. Ingenix Pharmaceutical Services (i3 Research), and Neurocrine Biosciences, Inc. He has research grants from the National Institutes of Health, the Michael J. Fox Foundation for Parkinson's Research and the Dystonia coalition. He has received honoraria from the Movement Disorder Society, the American Academy of Neurology, and the Michael J. Fox foundation for Parkinson's Research. Anette Schrag, PhD has done consulting for Novartis Pharmaceuticals, Merck Pharmaceuticals, Boehringer-Ingelheim Pharmaceuticals. She has served on the Medical Advisory Board of and has received honoraria from Boehringer-Ingelheim Pharmaceuticals. She has received grant funding from AMGEN Pharmaceuticals, Acadia Pharmaceuticals, the EU Commission, Parkison's UK, and GE Healthcare. She has received royalties from Oxford University Press for publications on Rating Scales in Parkinson Disease. Pablo Martinez-Martin, MD, PhD has received honoraria from speaking engagements at scientific meetings of Italfarmaco and TEVA, and from serving in a scientific advisory board of AbbVie. He has grants from the Carlos III Institute of Health (FIS), IMSERSO, the Michael J. Fox Foundation and the Reina Sofia Foundation.

Appendix

MDS Committee on Rating Scales Development Members

Pablo Martinez-Martin, Chair: pmartinez@isciii.es; Glenn Stebbins, Co-Chair: Glenn_Stebbins@rush.edu; Esther Cubo: esthercubo@gmail.com; Christopher G. Goetz: Christopher_Goetz@rush.edu; Deborah Hall: Deborah_A_Hall@Rush.Edu; Sheng Luo; Laura Marsh: lmarsh2@comcast.net; Bart Post; Cristina Sampaio: Cristina.Sampaio@chdifoundation.org; Anette Schrag, Past-Chair: a.schrag@ucl.ac.uk; Daniel Weintraub: Daniel.Weintraub@uphs.upenn.edu.

References

1. Allen RP, Walters AS, Montplaisir J, Hening W, Myers A, Bell TJ, Ferini-Strambi L. Restless legs syndrome prevalence and impact: REST general population study. *Arch Intern Med* 2005;165:1286–1292.
2. Martinez-Martin P, Forjaz MJ. How to evaluate validation data. In: Sampaio C, Goetz CG, Schrag A, eds. *Rating Scales in Parkinson's Disease*. Oxford: Oxford University Press; 2012:16–41.
3. Walters AS; Group Organizer and Correspondent: The International Restless Legs Syndrome Study Group. Validation of the International Restless Legs Syndrome Study Group rating scale for restless legs syndrome. *Sleep Med* 2003;4:121–132.
4. Allen RP, Kushida CA, Atkinson MJ; RLS QOL Consortium. Factor analysis of the International Restless Legs Syndrome Study Group's scale for restless legs severity. *Sleep Med* 2003;4:133–135.
5. Wunderlich GR, Evans KR, Sills T, et al. The International Restless Legs Syndrome Study Group. An item response analysis of the international restless legs syndrome study group rating scale for restless legs syndrome. *Sleep Med* 2005; 6: 131–139.
6. Abetz L, Arbuckle R, Allen RP, et al. The reliability, validity and responsiveness of the international restless legs syndrome study group rating scale in a clinical trial setting. *Sleep Med* 2006;7:340–349.
7. Allen RP, Earley CJ. Validation of the Johns Hopkins restless legs severity scale. *Sleep Med* 2001;2:239–242.
8. Garcia-Borreguero D, Kohnen R, Hogl B, et al. Validation of the Augmentation Severity Rating Scale (ASRS): a muticentric, prospective study with levodopa on Restless Legs Syndrome. *Sleep Med* 2007;8:455–463.
9. Fulda S, Wetter TC. Where dopamine meets opioids: a meta-analysis of the placebo effect in restless legs syndrome treatment studies. *Brain* 2008;131:902–917.
10. Inoue Y, Oka Y, Kagimara T, Kuroda K, Hirata K. Reliability, validity and responsiveness of the Japanese version of the International Restless Legs Syndrome Study Group rating scale for restless legs syndrome in a clinical trial setting. *Psychiatr Clin Neurosci* 2013;67:412–419.
11. Masuko AH, Carvalho LB, Machado MA, Morais JF, Prado LB, Prado GF. Translation and validation into the Brazilian Portuguese of the rest-

doi:10.1002/mdc3.12088

less legs syndrome rating scale of the International Restless legs Syndrome Study Group. *Arq Neuropsiquiatr* 2008;66:832–836.

12. Hening WA, Allen RP, Washburn M, Lesage S, Earley CJ. Validation of the Hopkins telephone diagnosic interview for Restless Legs Syndrome. *Sleep Med* 2008;9:283–289.
13. Kohnen R, Oertel WH, Stiasny-Kolster K, Benes H, Trenkwalder C. Severity rating of Restless Legs Syndrome: Validation of the RLS-6 scales. *Sleep* 2004;27:A304.
14. Picchietti D, Allen RP, Walters AS, Davidson JE, Myers A, Ferini-Strambi L. Restless legs syndrome: prevalence and impact in children and adolescents – The Peds REST study. *Pediatrics* 2007;120:253–266.
15. Arbuckle R, Abetz L, Durmer JS, et al. Development of the Pediatric Restless Legs Syndrome Severity Scale (P-RLS-SS): a patient-reported outcome measure of pediatric RLS symptoms and impact. *Sleep Med* 2010;11:897–906.
16. Garcia-borreguero D, Allen RP, Kohnen R, et al. On behalf of the International Restless Legs Syndrome Study Group. Diagnostic standards for dopaminergic augmentation of Restless Legs Syndrome: report from a World Association of Sleep Medicine-International Restless Legs Syndrome Study Group consensus conference at the Max Planck Institute. *Sleep Med* 2007;8:520–530.
17. IRLSSG. Consensus Diagnostic Criteria for RLS/WED. 2012. Available at: http://irlssg.org/diagnostic-criteria/ Accessed (June 2014).

doi:10.1002/mdc3.12088