Measuring Movement Disorders in Antipsychotic Drug Trials

The Need to Define a New Standard

Anton J.M. Loonen, MD, PharmD, PhD and*
Herman M. van Praag, MD, PhD†

The history of modern antipsychotic drugs began in 1952 with the discovery of the specific calming effects of chlorpromazine (RP4560) in agitated manic patients by Delay and Deniker at La Hôpital Sainte-Anne in France. It was almost immediately recognized that these drugs induced a movement disorder with symptoms similar to Parkinson disease. The first formal report of extrapyramidal side effects was published in 1954.[1] The occurrence of dyskinesias associated with long-term use of neuroleptics was reported by Sigwald et al[2] in 1959.

The term atypical refers to antipsychotic drugs that induce fewer extrapyramidal side effects than classic antipsychotic drugs. By 1970, the 3 available atypical antipsychotic drugs were thioridazine, sulpiride, and clozapine, introduced in 1958, 1968, and 1969. However, in 1975, clozapine was withdrawn from the market in most countries because of reports of agranulocytosis in patients taking the drug. Thioridazine and sulpiride were not very potent, and they were associated with other conditions, such as retinitis pigmentosa in the case of thioridazine and hyperprolactinemia in the case of sulpiride. However, to the merit of Kane et al,[3] the unique properties of clozapine in treatment-resistant schizophrenia were rediscovered. The drug was reintroduced and became widely used, despite its limitations. As it became evident that clozapine fulfilled a specific clinical need, the pharmaceutical industry developed a series of other atypical or modern antipsychotic drugs. All of these drugs share the common property that they induce parkinsonism to a lesser degree than classic antipsychotic drugs. Haloperidol in fairly high doses is a prototype of a classic, parkinsonism-inducing drug,[4] although there is debate as to whether this is true for lower doses. It is even less certain that the atypical drugs are less likely to cause extrapyramidal side effects other than parkinsonian rigidity.

Antipsychotic drugs are indicated in patients with schizophrenia or schizoaffective disorders. However, antipsychotics are also used to treat other psychotic conditions and for the general management of severe agitation in psychiatric patients. During the last few years, atypical antipsychotic drugs have assumed a firm position in the treatment of bipolar disorders,[5,6] and this has resulted in 2 new problems with respect to treatment-induced movement disorders. First, there is some evidence that patients with bipolar disorder are particularly vulnerable to tardive movement disorders.[7,8] Second, the other drugs that are used for treating bipolar disorder, including lithium and anticonvulsant and antidepressant drugs, cause movement disorders other than those associated with the use of classic antipsychotic drugs. These movement disorders are usually not evaluated in clinical trials that evaluate the effects of antipsychotic drugs, even when lithium and/or anticonvulsants are used for comparison.

*Delta Chair on Pharmacotherapy in Psychiatric Patients, University of Groningen, Groningen, The Netherlands; and †Universities of Groningen, Utrecht, Maastricht, The Netherlands, and the Albert Einstein College of Medicine, New York, NY (emeritus).
Address correspondence and reprint requests to Anton J.M. Loonen, MD, PharmD, PhD, University of Groningen, Department of Social Pharmacy, Pharmacoepidemiology and Pharmacotherapy, Antonius Deusinglaan 1, 9713 AV, The Netherlands. E-mail: a.j.m.loonen@rug.nl.

ISSN: 0271-0749/07/2705-0423
DOI: 10.1097/jcp.0b013e31814f1105

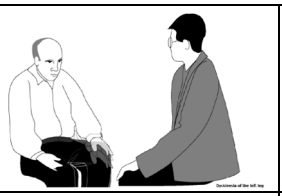
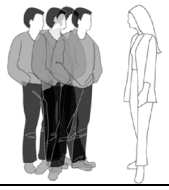
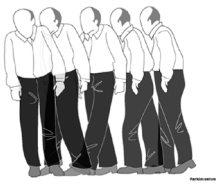
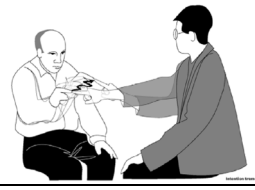
	Dyskinesia	Rapid, irregular, repetitive contractions; motionless intervals between contractions; increases during activity and anxiety.
	Dystonia	Slow, irregular, continuous contractions (spasms); contractions result in slow movements or abnormal postures; contractions continue for more than 2 seconds.
	Akathisia	Increased frequency of regular movements; primarily, but not exclusively affects the legs; an adversity to standing or sitting still.
	Parkinsonism	Reduction and/or slowing of all movement; equable plastic resistance to passive movement; hypersalivation and shiny skin; slow tremors that appear during rest and decrease during activity.
	Tremor	Rhythmical continuous pendular movements of variable amplitude and fixed frequency; no vibration-free intervals.
	Ataxia	Irregular movements while maintaining posture; inappropriate execution of voluntary movements; not a result of muscular weakness.

FIGURE 1. Examples and characteristics of the most significant drug-induced movement disorders.

In 1989, we used the Abnormal Involuntary Movement Scale (AIMS)[9] to assess the therapeutic potential of the calcium antagonist diltiazem in patients with tardive dyskinesia.[10] We were confronted with limitations of the scale that decrease its ability to measure changes in the severity of dyskinesias (discussed later). Moreover, in many studies of the effects of new antipsychotic drugs, the Simpson-Angus Scale (SEE) is used to measure parkinsonism.[11] Unfortunately, we have to agree with Cunningham Owens[12] who stated that "the Simpson-Angus Scale was a noble pioneer, but now deserves a decent burial. Few clinicians—and no researchers—should shed any tears at its passing." However, that is not what happened, and the AIMS and SEE scales continue to hold an exclusive place in the American Psychiatric Association's Handbook of Psychiatric Measures.[13] Our wish to include this point on the agenda of clinical research design has not been easy to implement. This is not surprising because it is difficult to discard a standard that has been in use for so many years. Moreover,

it is pointless to utter criticisms without offering better alternatives.

In this review, we discuss the rating scales that are used to measure the most outstanding drug-induced movement disorders (Fig. 1). The advantages and disadvantages of the scales are described. We argue that one of the main flaws of the scales is the absence or scarcity of adequately conducted biometrical studies analyzing their strengths and weaknesses. Subsequently, we present an alternative that incorporates the 2 most often used scales, the AIMS and the Barnes Akathisia Rating Scale (BARS).[14] The review is intended to initiate an open discussion on how to best measure side effects in future trials of treatments for psychotic, mood, and anxiety disorders.

INSTRUMENTS TO MEASURE MOVEMENT DISORDERS

Several types of instruments have been evaluated for their suitability to assess drug-induced movement disorders in clinical trials.[15–17] These instruments can be broadly divided into instrumental techniques, frequency counting techniques, and rating scales. Sophisticated instrumental techniques to measure movement disorders have been described in the literature[15,16] and are generally suitable to quantify one or only a few aspects of movement disorders, such as severity of rigidity[18–20] or the frequency and amplitude of a resting hand tremor.[21] Moreover, special equipment, computer programs, and trained personnel are necessary to collect and analyze movement data. Therefore, instrumental techniques are seldom suitable for application in ordinary clinical trials or in daily practice on a routine basis. More appropriate for the latter purpose are the so-called rating scales. A variety of global or multi-item rating scales attempts to measure specific movement disorders by quantifying clinical observations under standardized conditions. Some rating scales are developed to assess only 1 type of movement disorder.[15,16] In clinical trials, combinations of rating scales are used to measure multiple movement disorders. Other rating scales intend to quantify a set of disorders separately and simultaneously.[22] Generally speaking, rating scales have better validity than the aforementioned, more objective, instrumental techniques, but their reliability is less certain.[16] The reliability of rating scales can be improved by providing precise instructions on how to examine the patient, how to score the individual movement disorders, and how to rate the severity of their various components. Moreover, by videotaping the standardized examination, the movement disorder can be scored by several independent evaluators.

CHARACTERISTICS OF THE MOST FREQUENTLY USED SCALES

For research purposes, it is common practice to use a combination of the SEE,[11] the BARS,[14] and the AIMS.[9]

Abnormal Involuntary Movement Scale

The AIMS is a rating scale for dyskinesias constructed by the Psychopharmacology Research Branch of the National Institutes of Mental Health.[16] The AIMS rates 7 dyskinesias (ie, face, lips, jaw, tongue, arms, legs, and trunk) on a 5-point scale. In addition, global severity of abnormal movements are rated by (1) the observer, (2) the patient's reaction to the abnormal movements, and (3) the incapacitation that results from the abnormal movements. Two additional items consider the patient's dental status. The examination protocol is described in detail. A complex differentiation is made between spontaneous and activated movements. The severity scores of dyskinesias that occur during activity are reduced by 1 point to obtain the final score. Thereafter, the higher of 2 ratings is chosen: that of dyskinesias that occur while the patient is active and that of dyskinesias that occur while the patient is at rest. This procedure makes it difficult to score the movements reliably. Moreover, the validity of the dental status items is doubtful, and the usefulness of the global ratings can be disputed. Consequently, most authors consider only the first 7 or 8 items of the scale.[10] The face validity of the AIMS 1 to 7 is good. The first 4 items measure orofacial dyskinesias, and items 5 to 7 measure peripheral dyskinesias. Unfortunately, there are no published studies that examined the construct validity of the AIMS. However, there are at least 8 reports on the cross-sectional interrater reliability of the AIMS.[23] Pearson r and the intraclass correlation coefficient (ICC) were approximately 0.8. It can be concluded from the study by Bergen et al[24] that the interrater reliability of individual AIMS items is poor (0.3–0.6). The interrater reliability of the total 1 to 7 score is far better, especially when the evaluators are intensively trained as a single group.[25] However, the AIMS cannot be considered suitable to assess dyskinesias in long-term trials without videotaping the patients. Tracy et al[26] observed that without periodic joint training, the longitudinal interrater reliability decreased from 0.87 to 0.60 within a few months.

Simpson-Angus Scale

The SEE, published by Simpson and Angus in 1970,[11] was actually the second edition of a scale published in 1964.[27] The original scale was expanded to 10 items, including tremor and salivation, rated on a 5-point scale. The evaluation of the trunk muscles on the original scale was omitted from the SEE, as it was too difficult to quantify. An important criticism of the SEE is that in 1 way or another, 6 of the 10 items deal with rigidity in the neck and extremities. Furthermore, the examination instructions are quite brief and sometimes unclear. For example, different examiners were found to execute and interpret the glabella tap in a deviant manner, and the authors' advice was sought regarding examination technique and interpretation of results. Frequently, a patient goes from a rating of 0 to 4 without an intermediate score (G. M. Simpson, written communication). In addition, 2 items on the scale, head dropping and leg pendulousness, assume that an examination table is readily accessible. As this is often not the case, the examiner is obliged to alter the examination procedure. These deviations are expected to decrease the interrater reliability. Although the authors modified the scale to eliminate the problem, the uncorrected SEE is still used.

Besides the original publication of Simpson and Angus,[11] only 1 article addressed the cross-sectional interrater reliability of the SEE. Using data from 10 patients assessed by 4 examiners, Sweet et al[25] calculated an ICC of 0.79 for the total score. For the separate items, the ICC varied from 0.33 to 1.00. Simpson and Angus[11] reported a correlation coefficient of 0.87 for the interrater reliability of 2 physicians' scores of 14 patients. However, their result may be biased because the raters were probably not sufficiently independent. Recently, 2 additional studies of the biometric properties of the original SEE were published.[28,29] Janno et al[28] studied 99 patients with schizophrenia to establish the internal consistency of the SEE and its ability to identify drug-induced parkinsonism. The authors reported an internal consistency of 0.79 (Cronbach α) and concluded that the SEE's case-identifying properties converged with an experienced clinician's diagnosis based on *Diagnostic and Statistical Manual of Mental Disorders-IV* criteria. However, the SEE score did not correlate with lower limb motor activity as measured with another method; all rigidity items could be accounted for by elbow rigidity; and the items "glabella tap" and "salivation" did not discriminate between cases with and without parkinsonism. The latter may be caused by the presence of 20 clozapine users in their patient sample. Clozapine is known to cause salivation independent of parkinsonism. These results suggest that the SEE reliably measures only rigidity. Calvo-Gómez et al[29] studied the cross-sectional interrater reliability of a Spanish translation of the SEE in 86 psychiatric inpatients. They calculated an ICC of 0.81 for 3 raters for a 15-patient subgroup. A factor analysis indicated that 97% of the variability was attributable to a single factor, rigidity, and the internal consistency was too high (Cronbach α, 0.93, is well above the acceptable upper limit). Only the item glabella tap seemed to vary independently of the other scores. These results lead us to the same conclusions as those of Janno et al.[28]

In conclusion, the SEE is a well-documented rating scale that measures rigidity in the neck and extremities. The item glabella tap does not add to its positive characteristics. All items can be replaced with elbow rigidity. The cross-sectional reliability is reasonable for the scale as a whole, but its cross-sectional reliability for separate items is not. The longitudinal reliability of the SEE has not been determined. Furthermore, the examination procedure of the original scale is impractical, and for that reason, it is often modified.

Barnes Akathisia Rating Scale

The BARS was derived from an examination of the signs and symptoms exhibited by 104 consecutively admitted acute psychiatric inpatients who had received antipsychotic drugs, and 89 chronic psychiatric outpatients with long-term antipsychotic use.[14,30,31] The BARS has 3 items, which makes it difficult to establish its internal consistency. Its face validity is, however, very good. The scale rates observable characteristic restless movements and a combination of the patient's awareness of the restlessness and the patient's distress related to the restlessness on a 4-point scale. The BARS also includes a clearly defined 6-point global severity rating scale. The global assessment offers the opportunity to distinguish pseudoakathisia. The BARS provides brief examination instructions. Barnes[14,30] studied the interrater reliability of 2 examiners in a sample of 42 drug-treated, inpatients with schizophrenia. The interrater reliability, expressed as linearly weighted Cohen κ, ranged from 0.74 to 0.95 for the 4 items. Sweet et al[25] estimated an ICC of 0.93 for the total score and ICCs that ranged from 0.83 to 0.94 for the 4 items. Edson et al[23] reported an ICC of 0.73 for the total score from 9 raters who reviewed videotapes of 10 subjects. However, the video material may not be suitable for assessing patients with the BARS. The longitudinal reliability of the BARS has not been studied. Its concurrent validity with other akathisia scales has been examined, but as most of the comparisons were made with variants of the BARS, the results are not very meaningful. Still, it can be concluded that in general, the BARS is very useful for rating akathisia.

Extrapyramidal Symptom Rating Scale

The Extrapyramidal Symptom Rating Scale (ESRS) was developed by Chouinard, a psychiatrist-pharmacologist, and Ross-Chouinard, a neurologist. It was first used in clinical trials in 1976. The validity and reliability of the ESRS were reported to have been studied but were not described. Recently, Chouinard and Margolese[32] detailed some characteristics of the scale. The scale includes a subjective questionnaire for parkinsonian, dystonic, and dyskinetic symptoms. However, precise instructions for interpreting patient responses are missing. A positive point is the inclusion of a standard examination procedure that includes observing the patient executing a set of well-described tasks. The ESRS consists of 7-point multi-item ratings for parkinsonism, acute torsion dystonia, nonacute or chronic torsion dystonia, and dyskinetic movements. A separate subscale for akathisia is lacking, and this movement disorder is scored as a single item on the parkinsonism subscale. The ESRS includes 9-point scales of clinical global impressions of severity for dyskinesia, parkinsonism, and dystonia, as well as a rating of the parkinsonism stage according to Hoehn and Yahr.[33] A weakness of the ESRS is that in the parkinsonism subscale, the rigidity of each limb is counted separately, and tremors in as many as 8 body areas are scored separately. In addition, the distinction that the ESRS makes between acute and nonacute torsion dystonia has doubtful validity. The examiner seldom observes acute torsion dystonia himself/herself because this occurs unexpectedly and not during a planned examination, and the description is quite detailed. The most important methodological objection to the ESRS is the procedures for rating tremors and dyskinetic movements. These are rated on 2 axes or dimensions: the amplitude of the movements and the frequency of their occurrence. It is not clear that this is a valid method for measuring the severity of these movement disorders, and the point has not been addressed in the literature. In addition, the method of calculating factors and total scores is complex. Six factors (ie, hypokinetic parkinsonism, orofacial dyskinesia, trunk/limb dyskinesia, akathisia, tremor, and tardive dystonia) were identified by

assessment of 305 neuroleptic-treated, outpatients with chronic schizophrenia by a single investigator. High concordance between the AIMS and ESRS ratings of dyskinesia was found in 374 patients.[34] However, details of the other subscales are absent, as are data on long-term reliability of the ESRS. In conclusion, the ESRS is too complex to be practical, its validity has been insufficiently studied, and its face validity is low.

Sct. Hans Rating Scale

The Sct. Hans Rating Scale (SHRS) is a multidimensional scale developed in the 1970s. A preliminary version of the hyperkinesia subscale was taken as a comparator by Chien et al[19] in 1977. Gerlach[35] published the final version of the SHRS in 1979. The scale consists of 4 subscales that rate hyperkinesia, parkinsonism, dystonia, and akathisia.[36,37] The hyperkinesia scale scores dyskinesias in 8 body parts. Furthermore, a global score is included. Movement is scored while patients are sitting and relaxed (passive phase) and while active (active phase). The dystonia subscale consists of a single global item. The SHRS includes a standardized examination procedure. Originally, the SHRS included a special section for the detailed analysis of oral dyskinesia, but this part is only useful for specialized study. The interrater reliability of the hyperkinesia and parkinsonism scales was determined by evaluation of 30 psychiatric patients by 7 examiners.[37] A slightly modified AIMS was used for comparison. The interrater and test-retest reliability was generally high for experienced examiners (ICC, 0.82–0.98) but was considerably lower for less experienced examiners. Convergent validity was found for the dyskinesia scales and the AIMS, and divergent validity was found for the other scales. The parkinsonism subscale had high construct validity, but the dyskinesia subscale did not. The latter finding was attributed to the individual distribution of peripheral and head/face hyperkinesias, independent of the severity of the syndrome. Nevertheless, the face validity of this subscale is good. It can be concluded that the characteristics of the SHRS are incompletely studied, but the characteristics that have been studied are trustworthy.

Schedule for the Assessment of Drug-Induced Movement Disorders

The original Dutch version of the Schedule for the Assessment of Drug-Induced Movement Disorders (SADIMoD) appeared in 1994,[38] and the most recent English version was published in 2000.[39] The SADIMoD can be considered an expansion of the SHRS. It consists of subscales to quantify the severity of dyskinesias (separate passive and active phases), dystonia, parkinsonism, akathisia, 3 types of tremor (postural, rest, and intention), ataxia, and 4 mental symptoms (sedation, psychosis, depression, anxiety). Moreover, each subscale has a total score and a global score, with the latter offering the examiner the opportunity to express his/her personal opinion concerning the nature and severity of the disorder. To complete the score form of the SADIMoD, the patient is videotaped while undergoing a strictly standardized examination. Additional information is acquired verbally. In short-term drug trials, the video recording can be omitted without decreasing the validity and sensitivity of the procedure (AJML, 2003 unpublished data). The face validity of the SADIMoD is good, but the scale is complex. Inexperienced raters are sometimes deterred by the highly detailed definitions of items. The validity and cross-sectional and longitudinal reliability of the SADIMoD have been studied extensively.[22,40] Six teams of investigators assessed the test-retest and interrater reliability of the SADIMoD and the concurrent validity with the AIMS, SEE, and BARS. In another data set, the homogeneity of the SADIMoD dyskinesia and dystonia subscales showed a significant correlation. Analysis of the repeated ratings of clinically stable patients at 2-week intervals revealed that they were stable and homogeneous for each subscale. It was concluded that the SADIMoD is suitable for assessing the long-term course of drug-induced movement disorders.[22,40]

Conclusion

The currently available instruments to measure movement disorders have certain advantages and limitations. The BARS has the most desirable characteristics, and the SEE has the greatest limitations. The AIMS shows inconsistencies that limit its reliability, especially in long-term trials. All of the rating scales are limited in that the symptoms of different disorders may overlap, which decreases their validity when proper precautions are not taken to correctly interpret the ratings. In this respect, a composite scale may be more appropriate for assessing drug-induced motor disorders if specific guidelines for dealing with overlapping symptoms are provided. Sufficient guidelines and definitions are not included in the ESRS and SHRS but are included in the SADIMoD.

LIMITATIONS OF THE CURRENT PRACTICE

The limitations of the instruments used to assess movement disorders would not be of concern if the shortcomings were clinically insignificant. That, however, is not the case. A decreased tendency to induce extrapyramidal side effects is one of the major claims about modern antipsychotic drugs and is used as an argument to support their use in schizophrenia and bipolar disorders. A low tendency to induce movement disorders was concluded from the results of comparisons of atypical antipsychotics with haloperidol, which is generally used in fairly high doses as the gold standard in clinical trials. However, the Clinical Antipsychotic Trials in Intervention Effectiveness investigation[41] showed that the classic antipsychotic perphenazine is only slightly worse with respect to extrapyramidal side effects than its atypical counterparts. During such comparisons, relevant clinical effects can be overlooked when an unsatisfactory measuring instrument is used. This may be especially true when the AIMS is used to measure dyskinesias, as can be illustrated by the Veterans Affairs Cooperative Study No. 394,[26] which showed that the effect of 1600 IU/d vitamin E was comparable to the placebo effect during long-term treatment of tardive dyskinesia.[42] During the trial, the interrater reliability of the AIMS decreased

from 0.87 to 0.60. This interrater problem may have caused the approximately 50% variation coefficient in the AIMS scores. That standard deviation, which was far larger than the one used in the power and sample size calculation, may have resulted in an inadequate powering of the study and an unacceptable type 2 error. What is true for vitamin E (ie, no significant change in severity of tardive dyskinesia was established) may also be true for modern antipsychotic drugs when the AIMS is used to measure such changes. The AIMS is particularly unsuitable when it is used by different examiners in long-term studies when examinations are not videotaped. However, video recording of movement disorders is not a common practice in long-term trials of the safety and tolerability of psychotropic drugs.

What is true for measuring dyskinesias is also true for parkinsonism. That is, it is impossible to obtain reliable results when unreliable measuring instruments are used. There are 2 forms of parkinsonism: a bradykinetic form and a hypertonic-hyperkinetic form. The first form is common, but is inconclusively evaluated by most rating scales. For example, the SEE limits itself to an obvious side effect of haloperidol: rigidity. The same is true of the ESRS, which places too much weight on tremor. This means that a major clinically invalidating aspect of drug-induced parkinsonism is not sufficiently measured. Furthermore, investigators puzzle over how to measure the effects of antipsychotic drugs on cognitive dysfunction in schizophrenia, not realizing that most drug-induced effects on executive and other cognitive functions are reflected by slowness of thought (ie, bradyphrenia) that is another aspect of bradykinesia.

A final argument against the current evaluating techniques is the bias that may result from combining separate scales. In this case, the way in which the patient is examined is insufficiently standardized. In addition, the ratings are rarely adapted to each other. For example, the presence of dyskinesias in the lower extremities may result in a false positive score on the objective akathisia scale (pseudoakathisia). In this case, the latter score should be rounded to zero.

It can be concluded that the validity of the scales used to measure parkinsonism (ie, SEE, ESRS) is insufficient to adequately measure this drug-induced extrapyramidal syndrome. The reliability of the AIMS is insufficient to measure hyperkinesias, especially in long-term trials, and the validity of the relevant subscale of the ESRS is questionable. Apart from the ESRS, none of the scales address dystonia, and the validity of the ESRS is doubtful. All scales are unable to measure the typical motor effects of lithium and anticonvulsants. Ataxia provides an example. Ataxia is easily overlooked when it is not specifically probed. This may have been the case in several trials on the treatment of mania. Ataxia may result in falls and other accidents and therefore is a relevant motor side effect. Finally, the combination of a set of scales to measure different disorders decreases the validity to a low level. These limitations decrease the meaningfulness of the results of trials that evaluate the efficacy and tolerability of daily antipsychotic drug use. There is much room for improvement in the current practice of measuring movement disorders in clinical trials.

QUALIFICATIONS FOR A NEW STANDARD OF ASSESSMENT: THE SADIMOD

Because of the limitations mentioned above and others, there is an urgent need to adapt a standard to correct the inaccurate conclusions that may result from the existing approach. The work of Kay et al[43,44] on the assessment of schizophrenia symptoms may serve as an example. When the Positive and Negative Syndrome Scale was constructed, a similar problem existed. In the case of Kay et al,[43,44] it was the Brief Psychiatric Rating Scale that was invariably used in trials on antipsychotic drugs for treating schizophrenia, but that scale inadequately addressed several aspects of the disorder. Kay et al[43,44] solved the problem by incorporating the Brief Psychiatric Rating Scale into the Positive and Negative Syndrome Scale. This is what we attempted to accomplish with the SADIMoD. Specifically, we attempted to incorporate improved versions of some accepted rating scales to measure akathisia, parkinsonism, dyskinesias, and dystonias. In addition, the SADIMoD specifically addresses some movement disorders that are not evaluated by the older instruments.

The BARS required little adaptation. For the sake of uniformity, the specific anchor points for scoring the global subscale of the BARS were omitted in the SADIMoD. Furthermore, a modified version of the AIMS was incorporated as a separate subscale. The inadequacies of the AIMS were deleted inter alia by distinguishing 2 separate subscales to rate dyskinesias (ie, dyskinesias that occur during activity and those that occur during rest). However, the entire AIMS can be reconstructed from individual SADIMoD items.

Because the SEE puts too much weight on rigidity and can be replaced by a single item (ie, rigidity),[28] it was not incorporated into the SADIMoD subscale to measure parkinsonism. Instead, the SADIMoD uses the parkinsonism subscale of the SHRS and includes 4 items concerning bradykinesia, 2 concerning posture and gait, 1 concerning rigidity, 1 concerning resting tremor, and 1 concerning autonomic signs. Thus, it evaluates the separate aspects of parkinsonism more completely than the SEE.

Adaptation of rating scales is not, in itself, a sufficient improvement. Therefore, the examination procedure for the SADIMoD was standardized. Specific guidelines are to be followed when examining patients and recording observations. Furthermore, in long-term trials evaluating the effects of drugs on dyskinesias and dystonias, at least 2 video recordings should be made: one before beginning treatment and one several months after the initiation of treatment. Only then can subtle changes in the severity of the disorder be objectified in a reproducible way.

Finally, we developed a set of guidelines for adapting the different subscales to each other. The earlier example of the rating of pseudoakathisia may serve as an example. The same applies to scoring dyskinesias and dystonias when these movement disorders occur together.

CONCLUSIONS

Drug-induced movement disorders limit the use of psychotropic drugs. This is especially true for new indications, such as the use of antipsychotics for the long-term treatment

of mood and anxiety disorders. Unfortunately, the quality of the current measuring instruments for these side effects in long-term clinical trials of drug efficacy and safety is limited: (1) the SEE only measures rigidity; the AIMS shows inconsistencies that limit its reliability, especially in long-term trials; the validity of various subscales of the ESRS is questionable; and the SHRS lacks sufficiently extensive item definitions; (2) none of these scales are able to measure the typical motor effects of lithium and anticonvulsants; (3) these scales are not adapted to each other when used in combination. Therefore, it is important to adopt a new standard using a compilation of traditional rating scales to assess all relevant movement disorders in a precise and complementary fashion. This should be accompanied by clear instructions on how to examine patients, how to distinguish the different movement disorders when they occur together, and on how to score the disorders in a clear and reproducible manner. We propose the use of the SADIMoD for this purpose.

ACKNOWLEDGMENTS

The authors thank Dr Paul Kretchmer (kretchmer@sfedit.net) at San Francisco Edit for his assistance in editing this manuscript.

REFERENCES

1. Steck H. Le syndrôme extra pyramidal et diencephalique au cours des traitements au Largactil et au Serpasil. *Annales Medico-Psychologiques.* 1954;112:737–744.
2. Sigwald J, Bouttier D, Raymond C, et al. Quatre cas de dyskinesie facio-bucco-linguo-masticatice a evolution prolongée secondaire a un traitement par les neuroleptiques. *Revue Neurologique (Paris).* 1959; 100:751–755.
3. Kane JM, Honigfeld G, Singer J, et al. Clozapine and the treatment-resistant schizophrenic: a double-blind comparison with chlorpromazine. *Arch Gen Psychiatry.* 1988;45:789–796.
4. Geddes J, Freemantle N, Harrison P, et al. Atypical antipsychotics in the treatment of schizophrenia: systematic overview and meta-regression analysis. *BMJ.* 2000;321:1371–1376.
5. Malhi GS, Berk M, Bourin M, et al. Atypical mood stabilizers: a 'typical' role for atypical antipsychotics. *Acta Psychiatr Scand.* 2005;111(suppl 426):29–38.
6. Yatham LN. Atypical antipsychotics for bipolar disorder. *Psychiatr Clin North Am.* 2005;28:325–347.
7. Kane JM. Tardive dyskinesia in affective disorders. *J Clin Psychiatry.* 1999;60(suppl 5):43–47.
8. Keck PE Jr, McElroy SL, Strakowski SM, et al. Antipsychotics in the treatment of mood disorders and risk of tardive dyskinesia. *J Clin Psychiatry.* 2000;61(suppl 4):33–38.
9. Guy W. *ECDEU Assessment Manual for Psychopharmacology*. Washington, DC: US Department of Health, Education, and Welfare; 1976.
10. Loonen AJM, Verwey HA, Roels PR, et al. Is diltiazem effective in treating the symptoms of (tardive) dyskinesia in chronic psychiatric inpatients? A negative, double-blind, placebo-controlled trial. *J Clin Psychopharmacol.* 1992;12:39–42.
11. Simpson GM, Angus JWS. A rating scale for extra pyramidal side effects. *Acta Psychiatr Scand.* 1970;45(suppl 212):11–19.
12. Cunningham Owens DG. *A Guide to the Extra Pyramidal Side-Effects of Antipsychotic Drugs*. Cambridge: Cambridge University Press; 1999.
13. American Psychiatric Association. *Handbook of Psychiatric Measures*. Arlington: American Psychiatric Publishing; 2000.
14. Barnes TRE. A rating scale for drug-induced akathisia. *Br J Psychiatry.* 1989;154:672–676.
15. de Leon J, Simpson GM. Assessment of neuroleptic-induced extra pyramidal symptoms. In: Kane JM, Lieberman JA, eds. *Adverse Effects of Psychotropic Drugs*. New York: The Guildford Press; 1992:218–234.
16. Gardos G, Cole JO, La Brie R. The assessment of tardive dyskinesia. *Arch Gen Psychiatry.* 1977;34:1206–1212.
17. Loonen AJM, Doorschot CH, Jogems-Kosterman BJM, et al. Nieuw instrumentarium voor het meten van bijwerkingen. *COBO Bull.* 1997; 30:31–40.
18. Caligiuri MP, Bracha HS, Lohr JB. Asymmetry of neuroleptic-induced rigidity: development of quantitative methods and clinical correlates. *Psychiatry Res.* 1989;30:275–284.
19. Chien C-P, Jung K, Ross-Townsend A, et al. The measurement of persistent dyskinesia by piezoelectric recording and clinical rating scales. *Psychopharmacol Bull.* 1977;13:34–36.
20. Teräväinen H, Tsui JKC, Mak E, et al. Optimal indices for testing parkinsonian rigidity. *Can J Neurol Sci.* 1989;16:180–183.
21. Sinnaeve AF. Het meten van tremor. *Techniek Gezondheidszorg.* 1989;5:10–12.
22. Loonen AJM, Doorschot CH, Van Hemert DA, et al. The schedule for the assessment of drug-induced movement disorders (SADIMoD): interrater reliability and construct validity. *Int J Neuropsychopharmacol.* 2001;4:347–360.
23. Edson R, Lavori Ph, Tracy K, et al. Interrater reliability issues in multicenter trials, part II: statistical procedures used in Department of Veterans Affairs Cooperative Study #394. *Psychopharmacol Bull.* 1997; 33:59–67.
24. Bergen JA, Carter NB, Craig J, et al. AIMS ratings—repeatability. *Br J Psychiatry.* 1988;152:670–673.
25. Sweet RA, DeSensi EG, Zubenko GS. Reliability and applicability of movement disorder rating scales in the elderly. *J Neuropsychiatry Clin Neurosci*. 1993;5:56–60.
26. Tracy K, Adler LA, Rotrosen J, et al. Interrater reliability issues in multicenter trials, part 1: theoretical concepts and operational procedures used in department of veterans affairs cooperative study #394. *Psychopharmacol Bull.* 1997;33:53–57.
27. Simpson GM, Amuso D, Blair JH, et al. Aspects of phenothiazine-produced extra pyramidal symptoms. *Arch Gen Psychiatry.* 1964;10: 199–208.
28. Janno S, Holi MM, Tuisku K, et al. Validity of Simpson-Angus Scale (SAS) in a naturalistic schizophrenia population. *BMC Neurology [electronic source].* 2005;5:5.
29. Calvo-Gómez JM, Sánchez-Pedraza R, Jaramillo-González LE, et al. Validación de una Escala para Evaluación de Síntomas Colaterales Extrapiramidales de Simpson-Angus. *Rev Salud pública.* 2006;8:74–87.
30. Barnes TRE. Neuromuscular effects of neuroleptics: akathisia. In: Kane JM, Lieberman JA, eds. *Adverse Effects of Psychotropic Drugs*. New York: The Guildford Press; 1992:201–217.
31. Barnes TR. The Barnes akathisia rating scale—revisited. *J Psychopharmacol.* 2003;17:365–370.
32. Chouinard G, Margolese HC. Manual for the extra pyramidal symptom rating scale (ESRS). *Schizophrenia Res.* 2005;76:247–265.
33. Hoehn MM, Yahr MD. Parkinsonism: onset, progression, and mortality. *Neurology.* 1967;17:427–442.
34. Gharabawi GM, Bossie CA, Lasser RA, et al. Abnormal involuntary movement scale and extra pyramidal symptom rating scale (ESRS): cross-scale comparison in assessing tardive dyskinesia. *Schizophrenia Res*. 2005;77:119–128.
35. Gerlach J. Tardive dyskinesia. *Dan Med Bull.* 1979;26:209–245.
36. Gerlach J, Korsgaard S. Classification of abnormal involuntary movements in psychiatric patients. *Neuropsychiatr Clin.* 1983;2:201–208.
37. Gerlach J, Korsgaard S, Clemmesen P, et al. The St. Hans Rating Scale for extra pyramidal syndromes: reliability and validity. *Acta Psychiat Scand.* 1993;87:244–252.
38. Loonen AJM, Engberink MHA, Doornbos M, et al. Scale for the assessment of drug-induced movement disorders (SADIMoD)—optimalization of the video-registration procedure. Paper presented at: XIXth CINP Congress; Washington, DC, July 1, 1994.
39. Available at: http://WWW.SADIMoD.NL. Accessed January 1, 2007.
40. Loonen AJM, Doorschot CH, Van Hemert DA, et al. The schedule for the assessment of drug-induced movement disorders (SADIMoD): test-retest reliability and concurrent validity. *Int J Neuropsychopharmacol.* 2000;3:285–296.

41. Lieberman JA, Stroup TS, McEvoy JP, et al. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med.* 2005;353:1209–1223.
42. Adler LA, Rotrosen J, Edson R, et al. Vitamin E treatment for tardive dyskinesia. *Arch Gen Psychiatry.* 1999;56:836–841.
43. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bull.* 1987;13:261–276.
44. Kay SR, Opler LA, Lindenmayer JP. The positive and negative syndrome scale (PANSS): rationale and standardisation. *Br J Psychiatry*. 1989;155(suppl 7):59–65.