International Journal of Neuropsychopharmacology (2000), **3**, 285–296.

ARTICLE

The Schedule for the Assessment of Drug-Induced Movement Disorders (SADIMoD): test–retest reliability and concurrent validity

Anton J. M. Loonen[1], Cees H. Doorschot[2], Dianne A. van Hemert[3], Marc C. J. M. Oostelbos[4], Arnold E. S. Sijben[5] and the MASEAS Team*

[1] *Delta Psychiatric Hospital, PO Box 800, NL-3170 DZ Poortugaal, The Netherlands*
[2] *GGZ 's-Hertogenbosch, PO Box 10.150, NL-5260 GB Vught, The Netherlands*
[3] *Tilburg University, Department of Psychology, PO Box 90.153, NL-5000 LE Tilburg, The Netherlands*
[4] *Two Way bv, Muntelbolwerk 1, NL-5213 SZ 's-Hertogenbosch, The Netherlands*
[5] *ABC/Opes bv, Zutphensestraatweg 17, NL-6881 WN Velp, The Netherlands*

Abstract

The SADIMoD is a newly developed instrument, consisting of a compilation of rating scales, to measure the severity of drug-induced movement disorders: dystonia, dyskinesia, Parkinsonism, akathisia, ataxia, and several types of tremors. The test–retest reliability of this scale and the concurrent validity with the Abnormal Involuntary Movement Scale (AIMS), the Simpson–Angus Scale (SEE) and the Barnes Akathisia Scale (BAS) was assessed in 31 patients [20 male/11 female; 57.1 ± 16.5 yr (mean ± S.D.)] with a variety of movement disorders by six teams of investigators. The teams were trained by means of a standard package of instruction material to such an extent that a single member of the team could represent the entire team. Each patient was rated according to the AIMS, SEE and BAS and recorded on videotape according to the SADIMoD Schedule. These video-recordings were scored twice; first, in consensus by the entire team and secondly [110.3 ± 58.0 d (mean ± S.D.) later] by a single representative of that team. One team underwent a major change between scoring and was excluded from this analysis. Despite these difficult circumstances, these first and second ratings correlated to a highly significant degree with Spearman's correlation coefficients of 0.57 to 0.88 (median 0.69). The homogeneity of the applied scales was good (Cronbach's $\alpha = 0.75$–0.94). Convergent validity was found between the SADIMoD dyskinesia and (to a lesser extent) dystonia scales and the AIMS as well as between the akathisia sub-scales and the BAS, with divergent validity with the other sub-scales. The SEE discriminated less well between the Parkinsonism sub-scale and the other sub-scales.

Received 16 March 2000; Reviewed 17 May 2000; Revised 28 June 2000; Accepted 29 June 2000

Key words: Rating scales, movement disorders – drug induced, basal ganglia diseases, ataxia.

Introduction

Chronic psychiatric patients often suffer from a variety of movement disorders that may be related to their psychiatric disorder, this may be psychotropic drug-induced but may also result from the long-term abuse of alcohol or street drugs (Loonen and Doorschot, 1998). Sometimes drug-induced movement disorders occur acutely and in a dose-dependent fashion, they may also have a tardive character and augment upon a dose reduction. Overviewing the neuro-psychiatric and neuro-psychopharmacological literature, several types of drug-induced movement disorders can be distinguished (acute or tardive occurring): dyskinesia, dystonia, Parkinsonism, akathisia,

Address for correspondence: Dr A. J. M. Loonen, Delta Psychiatric Hospital, PO Box 800, NL-3170 DZ Poortugaal, The Netherlands.
Tel.: +31 10 503 13 13 *Fax*: +31 10 503 13 88 *E-mail*: Anton.Loonen@wxs.nl

* The MASEAS Team consists of the examiners and/or raters and the camera-operators making the video-recordings of the six participating centres. **Delta Psychiatric Hospital, Poortugaal**: B. Resida, rater; M. Kwakkelstein, rater; C. Sterrenburg, rater; L. Timmerman, investigator; Th. van Zonneveld, rater; R. Keulemans, assistant. **GGZ 's-Hertogenbosch OP; 's -Hertogenbosch**: E. Delescen, investigator; J. van der Veer, rater; W. Garenfeld, rater; G. Ackermans, assistant. **GGZ 's- Hertogenbosch VP, Vught**: C. Doorschot, investigator; W. van Rooij, rater; Ch. Maton, rater; H. Kaethoven, rater/assistant. **GGZ Midden-Brabant, Tilburg**: W. Standaert, investigator; M Pepping, assistant. **GGZ Oost-Brabant, Rosmalen**: B. Jogems, investigator; F. Schenkelaars, rater; M. van Gool, rater/assistant; J. Sturkenboom, rater/assistant. **PCA Zuid/Nieuw-West, Amsterdam**: L. Perquin, investigator; R. Halkema, rater; J Wesselius, rater; K. Hanekroot, rater; E. Beyaert, rater; P. Kempe, assistant.

tremor, ataxia, myoclonus, tics, asterixis, and convulsions (American Psychiatric Association, 1994; Ayd, 1961: Lingjærde et al., 1987). The first 6 of these are usually continuously present and can be assessed by examining the patient at regular intervals. The last 4 are rarely drug-induced or occur in a paroxysmal manner (Loonen and Doorschot, unpublished observations). Interviewing the patient, or observation by a family member or nurse, better assesses them. The incidence of the extrapyramidal movement disorders reported in the literature varies enormously, this is probably the result of the application of different diagnostic criteria (Lingjærde et al., 1987; Van Harten, 1998). In the literature, a vast majority of papers deal with extrapyramidal movement disorders, while hardly any describe ataxia as a psychotropic drug-induced disorder. We consider this unjustified as ataxia often occurs in chronic psychiatric patients. From a total of 112 patients studied, 44 suffered from ataxia of at least mild severity (Loonen et al., unpublished observations). Moreover, different types of tremors should be distinguished, i.e. the (Parkinsonian) rest tremor, the postural tremor and the intention tremor. A postural tremor may occur in Parkinson's disease (Koller et al., 1994) and possibly Parkinsonism, as well as a side-effect of, for example, antidepressant drugs or lithium (Lemus and Lieberman, 1992).

Several types of instruments have been examined for their suitability to assess drug-induced movement disorders in clinical trials (Gardos et al., 1977; Loonen et al., 1997). These instruments can be broadly divided into instrumentational techniques, frequency counting techniques and rating scales. Rating scales can be considered as the most practical and most clinically relevant instruments applied in clinical trials. Presently, many authors use a combination of the Simpson and Angus (1970) Rating Scale for Extrapyramidal Side Effects (SEE), the Barnes (1989) Rating Scale for Drug-Induced Akathisia (BAS) and the Abnormal Involuntary Movement Scale (AIMS) of the NIMH (Guy, 1976). Others apply composite scales such as the Extrapyramidal Symptom Rating Scale (ESRS) of Chouinard et al. (1984), or the Sct. Hans Rating Scale for Extrapyramidal Syndromes (SHRS) of Gerlach (1979, 1983). Two of the authors (A.J.M.L., C.H.D.) have searched the literature in order to identify rating scales for the assessment of drug-induced movement disorders. Moreover, detection of reports dealing with the nomological characteristics of the most commonly applied scales was attempted. Twenty-five scales were found. Three of them dealt with akathisia, 6 with dyskinesia, 1 with dystonia and 7 with Parkinsonism. Eight scales measured multiple disorders. Scales intended to measure the severity of genuine Parkinson's disease were not extensively covered by this search. Eight studies were found dealing with the inter-rater reliability of the AIMS (Edson et al., 1997), 1 with that of the Simpson–Angus SEE (Sweet et al., 1993), 3 dealing with that of the Barnes Rating Scale (Barnes, 1989, 1992; Edson et al., 1997; Sweet et al., 1993) and 1 with that of the Sct. Hans Rating Scale (Gerlach et al., 1993). Not a single report described the characteristics of the ESRS. It may be concluded that none of these scales alone is suitable to for use in long-term trials on the course of drug-induced movement disorders. They should be combined with videotape recordings of the movement disorders. However, even under those conditions the Simpson–Angus Rating Scale and the AIMS perform poorly. The utilization of the AIMS may, for example, have resulted in the negative results of a large-scale, 9-site long-term trial of up to 2 yr treatment with d-vitamin E in 150 subjects with tardive dyskinesia (Adler et al., 1999). It was observed that the inter-rater reliability of the AIMS was rapidly and dramatically decreasing during the course of this study (Tracy et al., 1997). They tried to solve this problem by organizing additional training sessions, but this may not have been succesful.

The SADIMoD was developed as an instrument to collect data on the presence of adverse clinical symptoms emerging during the usage of psychotropic drugs. It belongs to and is a part of the Multi-Axial Side Effect Assessment System (MASEAS) (Loonen and Doorschot, 1994; Loonen et al., 1997). The SADIMoD is a compilation of rating scales. It consists of a standardized examination schedule, a questionnaire to assess subjective complaints, a writing test (Haase, 1977), a rating form and a glossary. This glossary contains instructions for administering the SADIMoD, the forms and the criteria for the classification of different movement disorders and the definitions of severity scores (Doorschot et al., 1993; Loonen et al., 1993). In creating the SADIMoD, the Sct. Hans Rating Scale for Extrapyramidal Syndromes (SHRS) served as a model to which several well-known involuntary movement assessment scales were added:

- Abnormal Involuntary Movement Scale (AIMS) (Kief, 1980);
- Fahn–Marsden Dystonia Movement Scale (Burke et al., 1985);
- Barnes Akathisia Scale (BAS);
- Webster's Parkinson's Disease Rating Scale for Symptoms and Signs (Marsden and Schachter, 1981; Webster, 1968);
- Dillen–Roach Scale of Rigidity (Van Dillen and Roach, 1988).

Although it was intended to maintain the structure and item definitions of these scales, some of the scales incorporated had to be slightly adapted. To this set 3

Table 1. Major advantages and disadvantages of various rating scales

	Advantages	Disadvantages
Simpson–Angus Rating Scale (SEE)	Old and well-known scale	Characteristics insufficiently studied Instructions and definitions unclear Examination table needed Rigidity is to heavily counted
Barnes Akathisia Rating Scale (BAS)	Inter-rater reliability adequately studied Clear item and severity definitions Brief instructions for examination present	Only measures akathisia
Abnormal Involuntary Movement Scale (AIMS)	Old and well-known scale Inter-rater reliability adequately studied Clear examination schedule	Only measures dyskinesias Inter-rater reliability is rather poor Effect of training rapidly disappears Not suitable for long-term trials
Extrapyramidal Symptom Rating Scale (ESRS)	Measures different movement disorders Clear examination schedule	Characteristics insufficiently known Akathisia not separately measured Structure is of disputable calibre Function of questionnaire is unclear Invalid method of quantifying tremors and dyskinesias
Sct. Hans Rating Scale (SHRS)	Measures different movement disorders Inter-rater reliability adequately studied Clear examination schedule Suitable for assessing long-term changes	Measures dystonia only globally No clear definitions are given Inexperienced raters perform poorly No specific training material available
SADIMoD	Measures different movement disorders separately and independently Includes modified BAS and AIMS Well-standardized procedures lead to concordant scorings Clear item and severity definitions Suitable for retrospective, independent scoring Suitable for assessing long-term changes	Phenomenologically movement disorders may overlap Does not assess daily life situation Has a quite complex structure Needs the availability of trained personnel

newly developed sub-scales were added. One scale was intended to assess various tremors, classified according to Hallett (1991); 1 assessed ataxia and 1 assessed relevant psychiatric syndromes (psychosis, depression, anxiety and drowsiness), derived from the Brief Psychiatric Rating Scale (Overall and Gorman, 1962) or the UKU Side Effect Rating Scale (Lingjærde et al., 1987). When the rating scales used did not include strict definitions for severity scores, these were adopted from the UKU Side Effect Rating Scale. In order to be able to complete the score form of the SADIMoD, the patient is videotaped while submitting to a strictly standardized examination schedule (Doorschot et al., 1994; Loonen et al., 1994). Not only the sequence in which the movements are to be carried out, but also the examination itself is standardized. After the conclusion of this examination schedule, further information is elicited verbally. The whole procedure takes approx. 25 min, whereas approx. 14 min are taped.

From a conceptual point of view, this new construct is important since it allows a unified, rather in-depth evaluation of movement dysfunctions that may not necessarily stem from the same pathophysiological substrates. From a practical point of view, the SADIMoD offers researchers the possibility to use one instrument instead of a compilation of various scales. The item definitions of the separate sub-scales were slightly adapted in such a manner that the total scale formed a coherent whole. Moreover, the assessments were based on observations during a single, distinct and strictly standardized examination of the patient, which was recorded on videotape. The instructions of other scales are far less clear and sometimes conflict with one another. The examiner is therefore obliged to adapt the examination schedule, this may be another source of variation. When the videorecordings of different examinations covering a large time interval are rated in one session, the long-term course of the movement disorders being investigated can be adequately assessed. The most important advantages and limitations of various rating instruments are summarized in Table 1.

Before using a new instrument in clinical research, the instrumental nomological characteristics should be examined. These characteristics comprise intra- and inter-rater reliability, construct validity (homogeneity), concurrent validity, and sometimes predictive validity (Loonen and Zwanikken, 1987). As previously indicated, this rule has not always been followed with rating scales measuring extrapyramidal side-effects. Moreover, it is necessary to check whether or not the changes made to the original rating instruments, allowing them to be included in the new scale, negatively influence their properties. In order to demonstrate the usefulness of this schedule, the authors undertook a multi-centre study to assess the validity and reliability of the SADIMoD. In this paper the test–retest reliability is described as well as the concurrent validity with AIMS, SEE, and BAS.

Method

Six investigators (5 psychiatrists and 1 specialized clinical researcher) from different centres were each asked to form an investigating team. The team consisted of at least 1 experienced staff member (usually a psychiatrist) to examine the patients, a research nurse to operate the video-camera, and raters to score the patients during the video sessions. Two teams were considered to be experienced in doing SADIMoD examinations, 2 teams were considered less experienced but had worked with the SADIMoD before and 2 teams were completely inexperienced in this respect. Four of the 6 investigators had participated in several clinical trials in which the severity of extrapyramidal side-effects was assessed.

In each centre a study initiation visit was organized during which all members of the team were present. During this visit the study protocol, centre log file and patient record form was explained and discussed, the examiner received the SADIMoD manual, the prescribed examination materials and an instruction video. This instruction video contains some background information on the SADIMoD, an explanation of its structure, some examples of typical movement disorders, and instructions to solve potential problems. In addition, video-recordings of SADIMoD examinations of 3 patients are provided with the corresponding scores in the manual. These patients suffer from a variety of moderately severe movement disorders and the recordings are intended to train the participating raters. The investigator was instructed to use this material to train his team to an acceptable standard and to achieve a consensus, thereby allowing him to consider all individual team members capable of representing the team.

Each investigator selected up to 6 male or female, psychiatric in- or out-patients, that were over 18 years of age, and suffered from at least one mildly severe, relatively stable, movement disorder that was possibly, probably or certainly psychotropic drug-induced. Patients were required to be able to complete the study procedures and give their informed consent to participate in the study, and also to agree to have a video-recording made. Excluded were patients, who, as judged by the investigator, suffered from a movement disorder that was unusual for this patient population or that was probably or certainly not related to the usage of psychotropic drugs. No other selection criteria were applied, while it was the intention that a variety of movement disorders of various degrees of severity were present in the sample.

Each patient was evaluated by the same examiner in 1 session by means of 4 different assessment instruments, namely: SADIMoD, Simpson–Angus Rating Scale for Extrapyramidal Side Effects (SEE) (Simpson and Angus, 1970), Barnes Rating Scale for Drug-Induced Akathisia (BAS) (Barnes, 1989), and the Abnormal Involuntary Movement Scale (AIMS) (Guy, 1976). The examiners were not officially trained in applying the SEE, BAS and AIMS, but the investigator and/or the examiner himself had usually participated in drug trials in which these instruments were used. When it was not possible for practical reasons to complete all assessments in 1 session, the examiner had to ensure that the clinical condition was unchanged during the different examinations.

After having examined all patients, the investigator organized 1–3 assessment sessions during which the complete team scored the video-recordings of his patients. It was established that each case was discussed under referral to the glossary and that consensus was reached concerning every rating. This resulted in the first ratings of the patients of that centre.

After all 6 participating centres had completed the rating of their patients; a further 6 sessions were organized to reassess the videotapes. During these sessions a single representative from each team met at one of the centres and scored the patients of *that* centre. Therefore, 1 member of that centre's team scored every patient of that centre for the second time. This was considered to be the retest score of that patient. The representatives were instructed not to watch the video-recordings between these two sessions and not to look up previous scores. The raters could ask to see fragments of the videotape, but no case discussions were allowed between the representatives of the different teams. Three video-recordings were scored before and three after a 45 min break during which a light meal was served.

Following the last session a meeting of all participating investigators was held during which they described protocol deviations, commented on the protocol and

material supplied, discussed inconsistencies and shortcomings, and gave suggestions for improving the SADIMoD. These comments were used to finalize the manual of the SADIMoD.

This study was undertaken in accordance with the Helsinki Declaration and Dutch legislation concerning the performing of medical research and, in particular, videotaping psychiatric patients. The protocol was submitted and approved by the Toetsingscommissie Zuid-Nederland.

Rating scales

The SADIMoD consists of 8 sub-scales (Figure 1): a 9-item scale for dystonia, an 8-item scale for Parkinsonism, two 7-item scales for dyskinesias during activity and in rest, a 5-item scale for ataxia, a 2-item scale for akathisia, a 3-item scale for tremors and a 4-item scale for sedation, depression, psychosis and anxiety. In the sub-scale for dystonia the severity score (0–4) is multiplied with a provoking factor (0–4) considering the conditions of occurrence and a weight factor considering the body part that is affected. In all other sub-scales a scoring code of 0–4 is applied. With score 1 the rater can express uncertainty. Of the first 7 sub-scales a total score is calculated and for the first 6 sub-scales a global score is added. With this global score the examiner can also express his opinion concerning the true character of the observed movement disorder, e.g. in the case of pseudo-akathisia a global akathisia score of 0 should be given. Details of the construction of this scale are given elsewhere (Loonen and Doorschot, unpublished observations).

The AIMS is a rating scale devised for the assessment of dyskinesias. This scale consists of separate ratings on a 5-point scale (0–4) of dyskinesias of the face, lips, jaw, tongue, arms, legs and trunk. In addition, 3 global severity ratings of abnormal movements are added: those seen by the observer, the patient's reaction to them, and the incapacitation that results from them. Two additional items (0–1) deal with the dental status. A differentiation is made between spontaneous and activated movements. The rating of dyskinesias that occur during activity should be reduced with 1 point in order to obtain the final score. A total of the first 7 items and a grand total of all items are calculated.

In the SEE, which was published by Simpson and Angus in 1970, 10 items are rated on a 5-point scale (0–4). Six or 7 (depending on the question whether or not the glabella tap reflects something else) of the 10 items, deal with some form of rigidity. In addition, gait, tremors and salivation are assessed. A total of all 10 items is calculated.

The BAS comprises ratings on a 4-point scale of the observable characteristic restless movements, the patient's awareness of this restlessness, and the patient's distress related to the restlessness (0–3). A total of these scores is calculated. In addition, a global severity rating on a 6-point scale is present, with clear definitions of each scale point. This global assessment offers the opportunity to distinguish pseudo-akathisia.

Statistical analysis

All data were screened for irregularities before analysis. Missing data that could be inferred were added. Ten missing values of the 310 possible item scores in the SEE were replaced by the mean value for that individual. The remaining missing data were included as such in subsequent analyses. All analyses were performed with SPSS for Windows 7.5.

Concurrent validity

Spearman correlation coefficients were calculated for all scales to express the degree of convergent (or divergent) validity. These analyses were performed on the first SADIMoD ratings only. A good correlation between AIMS scores and SADIMoD dyskinesia and dystonia sub-scales scores was expected. Moreover, a good correlation was predicted between SEE scores and SADIMoD Parkinsonism scores. Finally, BAS scores were expected to correlate with SADIMoD akathisia scores. In addition, it was expected that the divergent validity with the other sub-scales could be negatively influenced by the mutual correlation of the severity of different, dose-dependent movement disorders, i.e. Parkinsonism, tremors, akathisia and ataxia.

Test–retest reliability

The intra-rater reliability was estimated by calculating Spearman coefficients for all SADIMoD sub-scales. These were calculated for separate raters as well as for all raters together. Also the intra-rater reliability was calculated with and without the least reliable rater. All calculations were performed on first and second ratings for each centre.

Construct validity

The internal consistency of the current data set was measured by the Cronbach α-coefficient (Cronbach, 1951). This coefficient tests the sufficiency with which one item

S A D I M o D

Schedule for the assessment of drug-induced movement disorders

Date
Time
Initials
Patientnr.

DYSTONIA

	Provoking Factor	Severity Factor	Weight Factor	Product
Eyes			0,5	
Mouth			0,5	
Speech/Swallowing			1,0	
Neck			0,5	
Torso			1,0	
R Arm			1,0	
L Arm			1,0	
R Leg			1,0	
L Leg			1,0	
Total				
Dystonia Global				

PARKINSONISM

Facial Expression	
Bradykinesia	
Tremor	
Posture	
Arm Sway	
Gait	
Rigidity	
Salivation	
Total	
Parkinsonism Global	

DYSKINESIA

	Passive	Active
Jaw		
Tongue		
Lips		
Face		
Torso		
Upper Extr.		
Lower Extr.		
Total		
Dyskinesia Global		

ATAXIA

Dysarthria	
Upper Extr.	
Lower Extr.	
Posture	
Gait	
Total	
Ataxia Global	

TREMOR

Tremor at rest	
Postural Tremor	
O Fast:	
O Moderate:	
O Slow:	
Intention Tremor	
Total	

AKATHISIA

Motor (objective)	
Psychic (subjective)	
Total	
Akathisia Global	

PSYCHIC SYMPTOMS

Sedation	
Depression	
Psychosis	
Anxiety	

SCORING CODE

0	Absent
1	Dubious
2	Mild
3	Moderate
4	Severe

Figure 1. Example of the scoring form used (© A. J. M. Loonen and C. H. Doorschot).

can substitute for the other. A Cronbach coefficient of 0.70 or higher is usually considered acceptable (Nunnally, 1970).

Results

Protocol deviations

The most important protocol deviations concerned the SEE examinations. Many examiners did not have an examination table at their disposal and therefore felt unable to adequately assess the items 'leg pendulousness' and 'head dropping'. Moreover, many examiners felt the instructions on doing the glabella tap unclear. Therefore 2, 5, and 2 missing scores, respectively, had to be replaced by an average score for that individual. The deficient SADIMoD data (out of a possible 217 scores) included: 7 for the global Parkinsonism scale, 9 scores for the global dystonia scale, 4 for the global ataxia scale, 2 for both the global active phase and the global passive phase dyskinesia scale, and 1 for the global akathisia scale. These data were considered 'missing' in all further analyses. Teams D and E were changed between the first and second SADIMoD assessment of patients as one of the raters had left the hospital and had been replaced by a less trained rater. In the case of team D an experienced psychiatrist left.

Patients

Thirty-one patients (20 male, 11 female) were examined and videotaped. Three centres made recordings of 6 patients, two of 5 patients and one of only 3 individuals. The patients were 57.1 ± 16.5 yr (mean $\pm$ S.D.; range: 29–83 yr). Four patients suffered from dystonia, 8 from active and 8 from passive phase dyskinesia, 19 from Parkinsonism, 3 from akathisia, and 10 from ataxia of at least mild severity (global score $\geqslant 2$). Seven patients suffered of Parkinsonism combined with either active phase, passive phase or both types of dyskinesias. Active and passive phase dyskinesias occurred together in 7 patients. Three patients showed dystonia as well as ataxia. One patient scored more than 2 on all global scales. Thirteen patients suffered from more than one movement disorder simultaneously. The average time between the first and second assessments measured 110.3 ± 58.0 d (mean $\pm$ S.D.; range 14–231 d).

Concurrent validity

The concurrent validity was expressed as Spearman's correlation coefficients for the sub-scales of SADIMoD, AIMS, Simpson–Angus SEE, and BAS, respectively (Table 2). As was expected the AIMS scores showed the highest correlation with the dyskinesia sub-scale scores of the SADIMoD. A lower, but yet significant, correlation existed with the dystonia ratings of the SADIMoD. As expected, the correlation with the other sub-scales was generally low. The BAS correlated highly significantly with the akathisia sub-scales, although the correlation coefficients are somewhat lower than those in the case of AIMS scores. Moreover, BAS discriminated less well between akathisia and the other sub-scales of the SADIMoD. As shown in Table 3, the lower correlation coefficients are probably mainly due to the ratings on the items awareness and distress of the BAS in comparison to the item motor (objective) of the SADIMoD. The correlation between the other items of these scales is much higher. In addition, the correlation between the subjective BAS scores (i.e. awareness and distress) and the global akathisia SADIMoD scores is apparent. A rather low, but significant, correlation existed between total BAS score and total dystonia score on the SADIMoD (Table 2). The total SEE score correlated highly significantly with total Parkinsonism score, and significantly with the global Parkinsonism and ataxia sub-scale scores. Apart from the ataxia sub-scale, the correlation with the scores of the other sub-scales was low.

Test–retest reliability

The test–retest variability was examined by calculating the Spearman correlation coefficients for the first and second SADIMoD ratings for each centre/rater (Table 4). In order to compare the performance of individual centre/raters, the first scoring of that centre was also compared with the averaged scoring of that patient by all 6 raters during the second session (data not shown). It appeared that centre D, in particular, performed rather poorly. None of the SADIMoD scores during the first session from this centre correlated significantly with those of the second session, nor with the averaged second ratings. Nevertheless, either with or without including the ratings of this centre, the first and second ratings of all patients correlated to a highly significant extent. The only exceptions were the postural and intention tremor scores, which had rather low correlation coefficients. However, the total tremor scorings correlated very well.

Internal consistency

The internal consistency of the SADIMoD sub-scales and the other rating scales for the 31 assessed patients was expressed with Cronbach's α-coefficients. These coefficients amounted to 0.75–0.94 (median 0.83) for the

Table 2. Concurrent validity expressed as Spearman's correlation coefficient, for subscales of the Schedule for the Assessment of Drug-Induced Movement Disorders (SADIMoD), the Abnormal Involuntary Movement Scale (AIMS), the Simpson–Angus Rating Scale for Extrapyramidal Side Effects (SEE), and the Barnes Rating Scale for Drug-Induced Akathisia (BAS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
(1) Dystonia (total)	1																			
(2) Dystonia (global)	0.96**	1																		
(3) Parkinsonism	0.26	0.23	1																	
(4) Parkinsonism (global)	0.37*	0.32	0.90**	1																
(5) Dyskinesia (passive)	0.47**	0.38*	0.04	0.03	1															
(6) Dyskinesia (passive, global)	0.44*	0.36	0.06	0.06	0.97**	1														
(7) Dyskinesia (active)	0.57**	0.53**	0.11	0.08	0.91**	0.88**	1													
(8) Dyskinesia (active, global)	0.56**	0.55**	0.08	0.08	0.86**	0.90**	0.94*	1												
(9) Ataxia	0.46**	0.51**	0.29	0.36*	0.26	0.19	0.15	0.09	1											
(10) Ataxia (global)	0.34	0.36	0.38*	0.39*	0.10	0.08	0.00	−0.04	0.89**	1										
(11) Tremor at rest	−0.08	−0.22	0.25	0.17	0.10	0.01	0.10	−0.02	−0.03	−0.09	1									
(12) Postural tremor	0.36*	0.32	0.40*	0.31	0.16	0.09	0.25	0.22	0.18	0.14	0.46**	1								
(13) Intention tremor	0.21	0.36	−0.07	−0.17	0.15	0.12	0.17	0.18	0.18	0.10	0.11	0.41*	1							
(14) Akathisia	0.44*	0.39*	0.16	0.21	0.24	0.18	0.18	0.16	0.49**	0.28	0.04	−0.04	−0.08	1						
(15) Akathisia (global)	0.36*	0.33	0.05	0.10	0.19	0.13	0.19	0.17	0.37*	0.15	0.12	0.00	−0.13	0.93**	1					
(16) AIMS (total)	**0.43***	**0.33**	0.07	0.09	**0.70****	**0.72****	**0.77****	**0.76****	0.07	0.03	0.09	0.27	−0.03	0.00	0.03	1				
(17) AIMS (4 perioral items)	**0.52****	**0.53****	−0.03	−0.01	**0.75****	**0.72****	**0.84****	**0.83****	0.13	0.04	0.07	0.31	0.23	0.01	0.02	0.79**	1			
(18) AIMS (7 items)	**0.56****	**0.49****	0.07	0.10	**0.74****	**0.72****	**0.84****	**0.82****	0.14	0.07	0.12	0.31	0.01	0.08	0.12	0.95**	0.89**	1		
(19) BAS (total)	0.37*	0.24	0.19	0.25	0.26	0.24	0.17	0.14	0.35	0.18	0.11	0.07	−0.24	**0.66****	**0.60****	0.20	−0.09	0.16	1	
(20) SEE (total)	0.16	0.20	**0.48****	**0.42***	−0.04	0.06	0.01	0.10	0.27	0.45*	−0.16	0.15	0.19	0.14	0.08	0.12	−0.01	0.05	0.02	1

SADIMoD ratings of first session only; $^{*}p < 0.05$, $^{**}p < 0.01$.
Figures in **bold** indicate most relevant coefficients.

Table 3. Concurrent validity expressed as Spearman's correlation coefficient, for the akathisia sub-scale of the Schedule for the Assessment of Drug-Induced Movement Disorders (SADIMoD), and the sub-scales of the Barnes Rating Scale for Drug-Induced Akathisia (BAS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Awareness (BAS)	1							
(2) Distress (BAS)	0.71**	1						
(3) Global (BAS)	0.71**	0.68**	1					
(4) Objective (BAS)	0.46**	0.50**	0.76**	1				
(5) BAS (total)	0.87**	0.85**	0.92**	0.72**	1			
(6) Motor (SADIMoD)	0.20	0.24	0.51**	0.60**	0.43*	1		
(7) Psychic (SADIMoD)	0.68**	0.61**	0.70**	0.61**	0.73**	0.42*	1	
(8) Akathisia global (SADIMoD)	0.43*	0.41*	0.65**	0.65**	0.60**	0.83**	0.75**	1

SADIMoD ratings of first session only; * $p < 0.05$, ** $p < 0.01$.

Table 4. Intra-rater reliability expressed as Spearman's correlation coefficient, for SADIMoD sub-scales per centre and in total

SADIMoD scales	A ($n = 6$)	B ($n = 3$)	C ($n = 6$)	D ($n = 5$)	E ($n = 5$)	F ($n = 6$)	Total ($n = 31$)	Total without D ($n = 26$)
Dystonia	0.22	1.00**	0.62	0.00	0.50	0.42	0.52**	0.57**
Dystonia (global)	0.69	1.00**	–[a]	–[b]	–[a]	0.63	0.61**	0.66**
Parkinsonism	0.70	0.87	0.74	0.30	0.70	0.59	0.54**	0.69**
Parkinsonism (global)	0.82*	1.00**	1.00**	0.25	0.65	1.00**	0.69**	0.88**
Dyskinesia (passive)	0.87*	0.50	0.66	0.10	0.26	0.76	0.72**	0.78**
Dyskinesia (passive, global)	0.87*	0.50	0.72	0.14	−0.89*	0.94**	0.60**	0.66**
Dyskinesia (active)	0.91*	1.00**	0.41	0.20	0.46	0.91*	0.67**	0.74**
Dyskinesia (active, global)	0.91*	0.87	0.49	0.14	0.06	0.95**	0.63**	0.69**
Ataxia	0.27	1.00**	0.89*	0.80	0.29	0.90*	0.77**	0.81**
Ataxia (global)	0.66	0.50	0.78	0.30	0.00	0.87*	0.66**	0.73**
Rest tremor	0.71	–[a]	0.32	0.83	0.49	0.48	0.41*	0.52**
Postural tremor	0.31	−0.50	–[a]	0.06	0.73	0.42	0.33	0.38
Intention Tremor	0.98**	–[a]	−0.18	0.35	0.89*	0.78	0.46*	0.49*
Tremor total	0.92**	−0.95	−0.24	0.54	0.79	0.58	0.54**	0.59**
Akathisia	0.24	1.00**	0.03	0.75	0.98**	0.65	0.62**	0.63**
Akathisia (global)	0.00	1.00**	0.26	0.75	0.95*	0.78	0.68**	0.68**

Coefficients are based on first and second ratings per centre.
n, Number of patients considered.
[a] Restriction of range in one of the ratings.
[b] Three out of five ratings missing.
* $p < 0.05$; ** $p < 0.01$.

SADIMoD sub-scales (excluding and including global scores). They measured 0.82, 0.88 and 0.80 for AIMS, BAS and SEE total scores, respectively.

Discussion

Background and rationale

This paper describes the construction and initial validation of a new rating instrument for various movement disorders. More than a decade ago, two of the authors (A.J.M.L., C.H.D.) decided to develop this instrument, as they were dissatisfied with the performance of existing scales. They felt, that these instruments could easily miss subtle, albeit important, shifts from one movement disorder to another. Therefore, they developed a new scale that measures a complete set of movement disorders that may be induced by psychotropic drugs and that, according to their clinical experience, frequently occurred in (chronic) psychiatric patients. Advantages and

limitations of existing scales and the SADIMoD were discussed.

Methods and results

In this study, we followed a 'most pure scenario' for the assessment of the intra-rater variability and the concurrent validity; this means that we attempted to obtain an impression of the magnitude of the relevant parameters that would apply under rather poor conditions. In each centre a team of raters was formed, who were trained together. Each member of this team could represent the principal investigator during the second rating sessions. Moreover, two teams (centres D and E) were changed between times as an important participant left the hospital. The time that had elapsed between the first and second sessions was rather long, with an average of approx. 3.5 months. In spite of this, the test–retest reliability is good. Excluding the tremor scores and the ratings of centre D, the Spearman correlation coefficients vary from 0.57 to 0.88 (median 0.69). In their study on the intra-rater reliability of Sct. Hans Rating Scale for Extrapyramidal Syndromes, Gerlach et al. (1993) reached correlation coefficients for their hyperkinesia sub-scales of individual raters of 0.65–0.97 (median 0.84). For their Parkinsonism sub-scales these coefficients measured 0.68–0.98 (median 0.90), and for their akathisia sub-scales 0.45–0.94 (median 0.73). However, they had video-recordings of their patients assessed by the same raters with a time interval of only 2 wk. Moreover, the SADIMoD is not intended to be used in such a manner as investigated by us. Under normal circumstances patients will be examined according to the SADIMoD schedule with specified time intervals and the video-recordings made in these circumstances will be rated in one session by the same rater(s). Under these conditions, the intra-rater reliability will probably be very high. However, the intra-rater variability cannot be measured under such conditions, because the ratings cannot be considered to be independent. In fact, this independence is also disputable in the study of Gerlach et al. (1993). The somewhat lower test–retest reliability of the tremor scores cannot be entirely explained. According to the instructions in the glossary the tremor score should only be based on the severity of tremors occurring in the upper extremity. Some raters did not notice this sufficiently. Moreover, difficulties in adequately assessing the rate of the postural tremor may play a role. According to SADIMoD the rater should decide whether the postural tremor is fast (8 Hz), moderate (6 Hz) or slow (3 Hz). However, most raters commented that it is almost impossible to discriminate between these three types of postural tremor. This instruction may have made adequate estimation of the postural tremor score so difficult that it has also negatively affected the reliability of the other tremor scores. Therefore, this specification will be dropped in the final SADIMoD. The team at centre D did not perform well because it underwent an important change between the two sessions. We decided not to exclude this team from the other analyses, as this would flatter the results. However, despite their low correlation coefficients, the intra-rater reliability still remains good when the scorings of this centre are included.

The examiners and raters were trained in applying the SADIMoD by means of standardized training material. No specific training was given by any of the authors. Two teams had never worked with the SADIMoD and in only 1 team could each member be considered experienced. The raters were not specifically trained in applying AIMS, BAS and SEE. However, the SADIMoD training may also offer some clues in how to use these rating scales. When considering our results, it becomes apparent that some items of the SEE are unclear. Many examiners left items open, because they did not know how to examine the patient or how to score them. This is probably also true when the SEE is used in clinical trials and should be considered during pre-trial training sessions of the participating investigators when this scale is intended to be used. After concluding this study, it was learned that the original Simpson–Angus scale has been modified in the 1980s to avoid the need for an examination table. The item 'leg pendulousness' has been omitted and 'head dropping' has been changed to 'head rotation'. Surprisingly, despite the fact that these authors never published any material on this modified scale, and its characteristics are unknown, it has been widely used in the USA, including the NIMH Supported Treatment Strategies in Schizophrenia (Lerer B, personal communication: May 2000). As the authors were not aware of the existence of the modified scale, and as the published version (Simpson and Angus, 1970) is, to the best of their knowledge, the one used in clinical trials in Europe, this version was applied in the present study. The lack of training and the uncertainty of the examiners in scoring specific items may have resulted in an increase of the inter-rater variability in the present study. This probably explains why the convergent and divergent validity with the sub-scales is lower than expected. Moreover, different forms of rigidity are heavily weighted in the SEE and represent only one item in the Parkinsonism sub-scale of the SADIMoD. Still, the SEE correlates best with the Parkinsonism sub-scales of the SADIMoD. In addition, there exists a significant correlation with the global ataxia sub-scale. This may indicate that the ataxia in this patient group is related to rigidity and is, therefore, of extra-pyramidal nature. However, as the total ataxia score of the SADIMoD correlates to a far lesser extent with the SEE

score, this remains to be elucidated in a future study. In the present study we found a significant correlation between the different AIMS scores and the dystonia and dyskinesia sub-scale scores of the SADIMoD. This finding was expected, as the AIMS cannot distinguish between dyskinesias and dystonia, which are rapid, respectively slow, irregular involuntary movements. As a matter of fact, some dyskinesias (e.g. frequent eye blinking) are indistinguishable from mild dystonia (blepharospasm). However, the correlation between the AIMS scores and the dyskinesia sub-scales scores is easily the best. There existed a possibility, that the total AIMS scores or the 7-item AIMS scores were correlating with the motor akathisia sub-scale score, as the restless movements of the legs can sometimes be interpreted as dyskinesias as well as (pseudo)akathisia. However, this appeared not to be the case. A weak, although significant, correlation was observed between the total BAS score and the total dystonia sub-scale rating (Table 2). Moreover, the motor akathisia sub-scale scores showed the worst correlation with the four examined BAS scores and the other akathisia sub-scale scores (Table 3). In addition, it was found that the global akathisia SADIMoD sub-scale performed less well than the global BAS scoring. In the definition of this BAS item it is specified that pseudo-akathisia should be excluded from the rating. Our findings suggest, that this is an important specification and we have therefore decided to adapt the SADIMoD in this respect.

The primary subject of this study was to document the test–retest reliability and concurrent validity of the SADIMoD. The results also illustrate that the alterations made to the included scales at least do not have a negative impact on the reliability of the final result. The Cronbach α-coefficients of the applied (sub-)scales were calculated in order to verify that their internal consistency in these 31 patients was sufficient to allow conclusions. However, the construct validity of the SADIMoD will be studied with the assessments of a far larger patient population (Loonen et al., unpublished observations).

Conclusion

It can be concluded from the present study, that even under very difficult circumstances, the intra-rater reliability of the SADIMoD is satisfying. Nonetheless, the results also offer clues how to improve the schedule, e.g. by changing the specification of the frequency of postural tremors and by specifying that pseudo-akathisa should be rated zero in the global akathisia scoring. The results indicate that the SADIMoD can also be used with live evaluation of a patient without making video-recordings. However, it is strongly recommended to organize training sessions at regular intervals of 2–3 months in order to maintain the intra-rater reliabilty. In this study, such training sessions were not prescribed and organizing them may well have increased the test–retest reliability. The instruction material, that we have developed, makes the arrangement of such training sessions easy to accomplish.

Acknowledgements

This project was funded by an unrestricted personal grant from Pfizer bv, Capelle aan den IJssel.

References

Adler LA, Rotrosen J, Edson R, Lavori P, Lohr J, Hitzemann R, Raisch D, Caligiuri M, Tracy K, Veterans Affairs Cooperative Study no. 394 Study Group (1999). Vitamin E treatment for tardive dyskinesia. *Archives General Psychiatry 56*, 836–841.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th edn). Washington, DC: American Psychiatric Association.

Ayd Jr. JF (1961). A survey of drug-induced extrapyramidal reactions. *Journal American Medical Association 75*, 1054–1060.

Barnes TRE (1989). A rating scale for drug-induced akathisia. *British Journal of Psychiatry 154*, 672–676.

Barnes TRE (1992). Neuromuscular effects of neuroleptics: akathisia. In: Kane JM, Lieberman JA (Eds.), *Adverse Effects of Psychotropic Drugs* (pp. 201–217). New York: Guilford Press.

Burke RE, Fahn S, Marsden CD, Bressman SB, Moskowitz C, Friedman J (1985). Validity and reliability of a rating scale for the primary torsion dystonias. *Neurology 35*, 73–77.

Chouinard G, Ross-Chouinard A, Gauther S, Annable L, Mercier P (1984). An extrapyramidal rating scale for idiopathic and neuroleptic induced parkinsonism and dyskinesia. *14th CINP Congress*, Florence, Italy.

Cronbach LG (1951). Coefficient alpha and the internal structure of tests. *Psychometrika 16*, 197–334.

Doorschot CH, Doornbos M, Engberink MHA, Loonen AJM (1994). Een nieuw meetinstrument voor bewegingsstoornissen bij psychiatrische patiënten. *Spring Meeting of the Dutch Society of Psychiatry*, Lunteren.

Doorschot CH, Engberink M, Loonen AJM (1993). The stability of drug-induced movement disorders. *9th World Congress of Psychiatry*, Rio de Janeiro.

Edson R, Lavori Ph, Tracy K, Adler LA, Rotrosen J, Veterans Affairs Cooperative Study no. 394 Study Group (1997). Interrater reliability issues in multicenter trials. Part II: Statistical procedures used in department of veterans affairs cooperative study no. 394. *Psychopharmacology Bulletin 33*, 59–67.

Gardos G, Cole JO, La Brie R (1977). The assessment of tardive dyskinesia. *Archives of General Psychiatry 34*, 1206–1212.

Gerlach J (1979). Tardive dyskinesia. *Danish Medical Bulletin 26*, 209–245.

Gerlach J, Korsgaard S (1983). Classification of abnormal involuntary movements in psychiatric patients. *Neuropsychiatrica Clinica 2*, 201–208.

Gerlach J, Korsgaard S, Clemmesen P, Lund Lauersen A-M, Magelund G, Noring U, Povlsen UJ, Bech P, Casey DE (1993). The St. Hans Rating Scale for extrapyramidal syndromes: reliability and validity. *Acta Psychiatrica Scandinavica 87*, 244–252.

Guy W (1976). *ECDEU Assessment Manual for Psychopharmacology*. Washington, DC: US Department of Health, Education, and Welfare.

Haase H-J (1977). *Therapie mit Psychopharmaka und anderen seelisches Befinden beeinflussenden Medikamenten*. Stuttgart: FK Schattauer Verlag.

Hallett M (1991). Classification and treatment of tremor. *Journal of the American Medical Association 266*, 1115–1117.

Kief H (1980). Tardieve dyskinesie: als bijwerking hoofdzaak wordt. *COBO Bulletin 13* (2), 47–51.

Koller WC, Busenbark K, Miner K (1994). The relationship of essential tremor to other movement disorders: report on 678 patients. Essential Tremor Study Group. *Annals of Neurology 35*, 717–723.

Lemus CZ, Lieberman JA (1992). Neuromuscular effects of antidepressants and lithium. In: Kane JM, Lieberman JA (Eds.), *Adverse Effects of Psychotropic Drugs* (pp. 165–174). New York: Guilford Press.

Lingjærde O, Ahlfors UG, Bech P, Dencker SJ, Elgen K (1987). The UKU side effect rating scale. A new comprehensive rating scale for psychotropic drugs and a cross-sectional study of side effects in neuroleptic-treated patients. *Acta Psychiatrica Scandinavica 76* (Suppl. 334), 1–100.

Loonen AJM, Doorschot CH (1994). Klinisch veiligheidsonderzoek van geneesmiddelen: structuur van een documentatiesysteem. *Ziekenhuisfarmacie 10*, 41–44.

Loonen AJM, Doorschot CH (1998). Motorische bijwerkingen bij antipsychoticumgebruik. *Tijdschrift van de Vereniging van Artsen werkzaam in de Zwakzinnigenzorg 16*, 15–20.

Loonen AJM, Doorschot CH, Jogems-Kosterman BJM, Van Hoof JJM, Hulstijn W, Zitman FG (1997). Nieuw instrumentarium voor het meten van bijwerkingen. *COBO Bulletin 30*, 31–40.

Loonen AJM, Engberink MHA, Doornbos M, Doorschot CH (1993). Scale for the Assessment of drug-induced movement disorders. *Meeting of the Dutch Society of Clinical Pharmacology and Biopharmacy*, Nijmegen.

Loonen AJM, Engberink MHA, Doornbos M, Doorschot CH (1994). Scale for the assessment of drug-induced movement disorders (SADIMoD) – optimalization of the video-registration procedure. *XIXth CINP Congress*. Washington, DC, USA.

Loonen AJM, Zwanikken GJ (1987). *Farmaca, Psychofarmaca, Klinisch Onderzoek*. 's-Hertogenbosch: Wulco.

Marsden CD, Schachter M (1981). Assessment of extra-pyramidal disorders. *British Journal Clinical Pharmacology 11*, 129–151.

Nunnally Jr. JC (1970). *Introduction to Psychological Measurement*. New York: McGraw-Hill.

Overall JE, Gorman DR (1962). The brief psychiatric rating scale. *Psychological Reports 10*, 799–812.

Simpson GM, Angus JWS (1970). A rating scale for extrapyramidal side effects. *Acta Psychiatrica Scandinavica 45* (Suppl. 212), 11–19.

Sweet RA, DeSensi EG, Zubenko GS (1993). Reliability and applicability of movement disorder rating scales in the elderly. *Journal Neuropsychiatry Clinical Neurosciences 5*, 56–60.

Tracy K, Adler LA, Rotrosen J, Edson R, Lavori P and the Veterans Affairs Cooperative Study no. 394 Study Group (1997). Interrater reliability issues in multicenter trials. Part 1: Theoretical concepts and operational procedures used in department of veterans affairs cooperative study no. 394. *Psychopharmacology Bulletin 33*, 53–57.

Van Dillen LR, Roach KE (1988). Interrater reliability of a clinical scale of rigidity. *Physical Therapy 68*, 1679–1681.

Van Harten PN (1998). Movement disorders associated with neuroleptics. The Curaçao extrapyramidal syndromes study. University of Utrecht: Utrecht (dissertation).

Webster DD (1968). Critical analysis of the disability in Parkinson's disease. *Modern Treatment 5*, 257–282.