Scale for the assessment and rating of ataxia (SARA)

Endorsed by ERN-RND: 26.07.2018

Introduction to the European Reference Network for Rare Neurological Diseases (ERN-RND):

ERN-RND is a European Reference Network established and approved by the European Union. ERN-RND is a healthcare infrastructure which focuses on rare neurological diseases (RND). The three main pillars of ERN-RND are (i) network of experts and expertise centres, (ii) generation, pooling and dissemination of RND knowledge, and (iii) implementation of e-health to allow the expertise to travel instead of patients and families.

ERN-RND unites 32 of Europe's leading expert centres in 13 Member States and includes highly active patient organizations. Centres are located in Belgium, Bulgaria, Czech Republic, France, Germany, Hungary, Italy, Lithuania, Netherlands, Poland, Slovenia, Spain and the UK.

The following disease groups are covered by ERN-RND:

- Ataxias and Hereditary Spastic Paraplegias
- Atypical Parkinsons' Disease
- Dystonia, Paroxysmal Disorder and Neurodegeneration with Brain Ion Accumulation
- Frontotemporal Dementia
- Huntingtons' Disease and other Choreas
- Leukodystrophies

Specific information about the network, the expert centres and the diseases covered can be found at the networks web site www.ern-rnd.eu.

Recommendation for clinical use:

The European Reference Network for Rare Neurological Diseases strongly recommends the use of the Scale for the Assessment and Rating of Ataxia (SARA) as best clinical practice for the assessment and rating of Ataxia patients.

This document was supported by and done in the framework of the European Reference Network for Rare Neurological Diseases (ERN-RND). ERN-RND is one of the 24 European Reference Networks (ERNs) approved by the ERN Board of Member States. The ERNs are co-funded by the European Commission (ERN-RND: 3HP 767231).

Disclaimer:

Clinical practice guidelines, practice advisories, systematic reviews and other guidance published, endorsed or affirmed by ERN-RND are assessments of current scientific and clinical information provided as an educational service. The information (1) should not be considered inclusive of all proper treatments, methods of care, or as a statement of the standard of care; (2) is not continually updated and may not reflect the most recent evidence (new information may emerge between the time information is developed and when it is published or read); (3) addresses only the question(s) specifically identified; (4) does not mandate any particular course of medical care; and (5) is not intended to substitute for the independent professional judgement of the treating provider, as the information does account for individual variation among patients. In all cases, the selected course of action should be considered by the treating provider in the context of treating the individual patient. Use of the information is voluntary. ERN-RND provided this information on an "as is" basis, and makes no warranty, expressed or implied, regarding the information. ERN-RND specifically disclaims any warranties of merchantability or fitness for a particular use or purpose. ERN-RND assumes no responsibility for any injury or damage to persons or property arising out of or related to any use of this information or for any errors or omissions.

METHODOLOGY

The endorsement process has been performed by the Disease group for Ataxia and Hereditary Spastic Paraplegias of ERN-RND.

Disease group for Ataxia and Hereditary Spastic Paraplegias:

Disease group coordinators:

Caterina Mariotti[16;] Rebecca Schuele-Freyer[14]

Disease group members:

Healthcare professionals:

Segolene Ayme[1]; Enrico Bertini[2]; Kristl Claeys[3]; Maria Teresa Dotti[4]; Alexandra Durr[1]; Antonio Federico[4]; Josep Gámez[5]; Paola Giunti[6]; David Gómez-Andrés[5]; Kinga Hadziev[7]; York Hellenbroich[8]; Jaroslav Jerabek[9]; Mary Kearney[10]; Jiri Klempir[11]; Thomas Klockgether[12]; Thomas Klopstock[13]; Norbert Kovacs[7]; Ingeborg Krägeloh-Mann[14]; Berry Kremer[15];
Alfons Macaya[5]; Bela Melegh[7]; Maria Judit Molnar[8]; Isabella Moroni[16]; Alexander Münchau[8]; Esteban Muñoz[17]; Lorenzo Nanetti[16]; Andrés Nascimento[17]; Mar O'Callaghan[17]; Damjan

This document was supported by and done in the framework of the European Reference Network for Rare Neurological Diseases (ERN-RND). ERN-RND is one of the 24 European Reference Networks (ERNs) approved by the ERN Board of Member States. The ERNs are co-funded by the European Commission (ERN-RND: 3HP 767231).

Osredkar[18]; Massimo Pandolfo[19]; Joanna Pera[20]; Borut Peterlin[18];; Maria Salvadó[5]; Ludger Schöls[14]; Deborah Sival[15]; Matthis Synofzik[14]; Franco Taroni[16]; Sinem Tunc[8]; Bart van de Warrenburg[21];; Judith van Gaalen[21]; Martin Vyhnálek[9]; Michèl Willemsen[21]; Ginevra Zanni[2]; Judith Zima[7]; Alena Zumrová[9]

Patient representatives:

Lori Renna Linton[10], Cathalijne van Doorne[10]

[1] Assistance Publique-Hôpitaux de Paris, Hôpital Pitié-Salepétrière, France: Reference Centre for Rare Diseases 'Neurogenetics'; [2] Pediatric hospital Bambino Gesù, Rome, Italy; [3] University Hospitals Leuven, Belgium; [4] AOU Siena, Italy; [5] Hospital Universitari Vall d'Hebron, Spain; [6] University College London Hospitals NHS Foundation Trust, United Kingdom; [7] University of Pécs, Hungary; [8] Semmelweis University, Hungary;[8] Universitätsklinikum Schleswig-Holstein, Germany; [9] Motol University Hospital, Czech Republic; [10] Patient representative; [11] General University Hospital in Prague, Czech Republic; [12] Universitätsklinikum Bonn, Germany; [13] Klinikum der Universität München, Germany; [14] Universitätsklinikum Tübingen, Germany; [15] University Medical Center Groningen, Netherlands; [16]–IRCCS-Foundation Neurological Institute Carlo Besta – Milan, Italy; [17] Hospital Clínic i Provincial de Barcelona y Hospital de Sant Joan de Déu, Spain; [18] University Medical Centre Ljubljana, Slovenia; [19] Université libre de Bruxelles, Belgium; [20] University Hospital in Krakow, Poland; [21] Stichting Katholieke Universiteit, doing business as Radboud University Medical Center Nijmegen, Netherlands;

This document was supported by and done in the framework of the European Reference Network for Rare Neurological Diseases (ERN-RND). ERN-RND is one of the 24 European Reference Networks (ERNs) approved by the ERN Board of Member States. The ERNs are co-funded by the European Commission (ERN-RND: 3HP 767231).

Endorsement process:

- Mapping of used disease scales by disease group – June – December 2017
- Proposal for endorsement of rating scale by ERN-RND disease group coordinators – 15/05/2018
- Discussion in ERN-RND disease group during annual meeting – 08/06/2018
- Consent on endorsement of disease scale during ERN-RND annual meeting 2018 – 08/06/2018
- Consent on endorsement by whole disease group – 13/07/2018

This document was supported by and done in the framework of the European Reference Network for Rare Neurological Diseases (ERN-RND). ERN-RND is one of the 24 European Reference Networks (ERNs) approved by the ERN Board of Member States. The ERNs are co-funded by the European Commission (ERN-RND: 3HP 767231).

SCALE

1) Gait		2) Stance	
Proband is asked (1) to walk at a safe distance parallel to a wall including a half-turn (turn around to face the opposite direction of gait) and (2) to walk in tandem (heels to toes) without support. **0 Normal, no difficulties in walking, turning and walking tandem (up to one misstep allowed)** **1 Slight difficulties, only visible when walking 10 consecutive steps in tandem** **2 Clearly abnormal, tandem walking >10 steps not possible** **3 Considerable staggering, difficulties in half-turn, but without support** **4 Marked staggering, intermittent support of the wall required** **5 Severe staggering, permanent support of one stick or light support by one arm required** **6 Walking > 10 m only with strong support (two special sticks or stroller or accompanying person)** **7 Walking < 10 m only with strong support (two special sticks or stroller or accompanying person)** **8 Unable to walk, even supported**		Proband is asked to stand (1) in natural position, (2) with feet together in parallel (big toes touching each other) and (3) in tandem (both feet on one line, no space between heel and toe). Proband does not wear shoes, eyes are open. For each condition, three trials are allowed. Best trial is rated. **0 Normal, able to stand in tandem for > 10 s** **1 Able to stand with feet together without sway, but not in tandem for > 10s** **2 Able to stand with feet together for > 10 s, but only with sway** **3 Able to stand for > 10 s without support in natural position, but not with feet together** **4 Able to stand for >10 s in natural position only with intermittent support** **5 Able to stand >10 s in natural position only with constant support of one arm** **6 Unable to stand for >10 s even with constant support of one arm**	
Score		**Score**	

This document was supported by and done in the framework of the European Reference Network for Rare Neurological Diseases (ERN-RND). ERN-RND is one of the 24 European Reference Networks (ERNs) approved by the ERN Board of Member States. The ERNs are co-funded by the European Commission (ERN-RND: 3HP 767231).

3) Sitting

Proband is asked to sit on an examination bed without support of feet, eyes open and arms outstretched to the front.

0 **Normal, no difficulties sitting >10 sec**

1 **Slight difficulties, intermittent sway**

2 **Constant sway, but able to sit > 10 s without support**

3 **Able to sit for > 10 s only with intermittent support**

4 **Unable to sit for >10 s without continuous support**

Score	

4) Speech disturbance

Speech is assessed during normal conversation.

0 **Normal**

1 **Suggestion of speech disturbance**

2 **Impaired speech, but easy to understand**

3 **Occasional words difficult to understand**

4 **Many words difficult to understand**

5 **Only single words understandable**

6 **Speech unintelligible / anarthria**

Score	

5) Finger chase

Rated separately for each side
Proband sits comfortably. If necessary, support of feet and trunk is allowed. Examiner sits in front of proband and performs 5 consecutive sudden and fast pointing movements in unpredictable directions in a frontal plane, at about 50 % of proband´s reach. Movements have an amplitude of 30 cm and a frequency of 1 movement every 2 s. Proband is asked to follow the movements with his index finger, as fast and precisely as possible. Average performance of last 3 movements is rated.

0 **No dysmetria**

1 **Dysmetria, under/ overshooting target <5 cm**

2 **Dysmetria, under/ overshooting target < 15 cm**

3 **Dysmetria, under/ overshooting target > 15 cm**

4 **Unable to perform 5 pointing movements**

Score	**R**ight	**L**eft
mean of both sides (R+L)/2		

6) Nose-finger test

Rated separately for each side
Proband sits comfortably. If necessary, support of feet and trunk is allowed. Proband is asked to point repeatedly with his index finger from his nose to examiner's finger which is in front of the proband at about 90 % of proband's reach. Movements are performed at moderate speed. Average performance of movements is rated according to the amplitude of the kinetic tremor.

0 **No tremor**

1 **Tremor with an amplitude < 2 cm**

2 **Tremor with an amplitude < 5 cm**

3 **Tremor with an amplitude > 5 cm**

4 **Unable to perform 5 pointing movements**

Score	**R**ight	**L**eft
mean of both sides (R+L)/2		

This document was supported by and done in the framework of the European Reference Network for Rare Neurological Diseases (ERN-RND). ERN-RND is one of the 24 European Reference Networks (ERNs) approved by the ERN Board of Member States. The ERNs are co-funded by the European Commission (ERN-RND: 3HP 767231).

<table>
<tr>
<td colspan="3">7) Fast alternating hand movements

Rated separately for each side
Proband sits comfortably. If necessary, support of feet and trunk is allowed. Proband is asked to perform 10 cycles of repetitive alternation of pro- and supinations of the hand on his/her thigh as fast and as precise as possible. Movement is demonstrated by examiner at a speed of approx. 10 cycles within 7 s. Exact times for movement execution have to be taken.

0 Normal, no irregularities (performs <10s)
1 Slightly irregular (performs <10s)
2 Clearly irregular, single movements difficult to distinguish or relevant interruptions, but performs <10s
3 Very irregular, single movements difficult to distinguish or relevant interruptions, performs >10s
4 Unable to complete 10 cycles</td>
<td colspan="3">8) Heel-shin slide

Rated separately for each side
Proband lies on examination bed, without sight of his legs. Proband is asked to lift one leg, point with the heel to the opposite knee, slide down along the shin to the ankle, and lay the leg back on the examination bed. The task is performed 3 times. Slide-down movements should be performed within 1 s. If proband slides down without contact to shin in all three trials, rate 4.

0 Normal
1 Slightly abnormal, contact to shin maintained
2 Clearly abnormal, goes off shin up to 3 times during 3 cycles
3 Severely abnormal, goes off shin 4 or more times during 3 cycles
4 Unable to perform the task</td>
</tr>
<tr>
<td>Score</td>
<td>Right</td>
<td>Left</td>
<td>Score</td>
<td>Right</td>
<td>Left</td>
</tr>
<tr>
<td colspan="2">mean of both sides (R+L)/2</td>
<td></td>
<td colspan="2">mean of both sides (R+L) / 2</td>
<td></td>
</tr>
</table>

Annex - Original publication

Scale for the assessment and rating of ataxia

Development of a new clinical scale

T. Schmitz-Hübsch, MD; S. Tezenas du Montcel, MD, PhD; L. Baliko, MD; J. Berciano, MD; S. Boesch, MD; C. Depondt, MD; P. Giunti, MD; C. Globas, MD; J. Infante, MD; J.-S. Kang, MD; B. Kremer, MD; C. Mariotti, MD; B. Melegh, MD, PhD; M. Pandolfo, MD; M. Rakowicz, MD; P. Ribai, MD; R. Rola, MD; L. Schöls, MD; S. Szymanski, MD; B.P. van de Warrenburg, MD; A. Dürr, MD; and T. Klockgether, MD

Abstract—*Objective:* To develop a reliable and valid clinical scale measuring the severity of ataxia. *Methods:* The authors devised the Scale for the Assessment and Rating of Ataxia (SARA) and tested it in two trials of 167 and 119 patients with spinocerebellar ataxia. *Results:* The mean time to administer SARA in patients was 14.2 ± 7.5 minutes (range 5 to 40). Interrater reliability was high, with an intraclass coefficient (ICC) of 0.98. Test-retest reliability was high with an ICC of 0.90. Internal consistency was high as indicated by Cronbach's α of 0.94. Factorial analysis revealed that the rating results were determined by a single factor. SARA ratings showed a linear relation to global assessments using a visual analogue scale, suggesting linearity of the scale ($p < 0.0001$, $r^2 = 0.98$). SARA score increased with the disease stage ($p < 0.001$) and was closely correlated with the Barthel Index ($r = -0.80$, $p < 0.001$) and part IV (functional assessment) of the Unified Huntington's Disease Rating Scale (UHDRS-IV) ($r = -0.89$, $p < 0.0001$), whereas it had only a weak correlation with disease duration ($r = 0.34$, $p < 0.0002$) *Conclusions:* The Scale for the Assessment and Rating of Ataxia is a reliable and valid measure of ataxia, making it an appropriate primary outcome measure for clinical trials.

NEUROLOGY 2006;66:1717–1720

With continuous progress in the understanding of the molecular pathogenesis of many ataxia disorders, novel therapies are on the horizon that will need to be evaluated in clinical trials. An essential prerequisite for such trials is the availability of validated neurologic assessment methods to measure the severity of ataxia. Although a number of ataxia rating scales have been proposed,[1-4] there have been only few attempts to validate these scales. A widely used ataxia rating scale is the International Cooperative Ataxia Rating Scale (ICARS). Although interrater reliability of this scale was shown to be high,[5] more detailed evaluations revealed problems concerning practicability and subscale structure that questioned the usefulness of ICARS for future interventional trials.[6] Recently, a new scale, Friedreich's Ataxia Rating Scale, was devised and validated.[4] However, this scale is not readily applicable to cerebellar ataxias because it was designed as a disease-specific scale for the evaluation of Friedreich ataxia and thus also considers other symptoms than ataxia.

We therefore devised a new scale named Scale for the Assessment and Rating of Ataxia (SARA) that is based on a semiquantitative assessment of cerebellar

Additional material related to this article can be found on the *Neurology* Web site. Go to www.neurology.org and scroll down the Table of Contents for the June 13 issue to find the title link for this article.

See also page 1711

From the Department of Neurology (T.S.-H., T.K.), University Hospital of Bonn, Bonn, Germany; Department of Biostatistics and Medical Informatics (S.T.M.), Hôpital de la Pitié-Salpêtrière, Assistance Publique-Hôpitaux, Paris, France; Modeling in Clinical Research (S.T.M.), EA 3974, University Pierre et Marie Curie, Paris, France; Department of Neurology and Stroke (L.B.), County Hospital, Veszprém, Hungary; Department of Neurology (J.B., J.I.), University Hospital Marqués de Valldecilla, Santander, Spain; Department of Neurology (S.B.), University of Innsbruck, Innsbruck, Austria; Department of Neurology (C.D., M.P.), Hôpital Erasme, Université Libre de Bruxelles, Brussels, Belgium; Department of Biochemistry and Genetics (R.F., C.M.), Istituto Nazionale Neurologico C. Besta, Milan, Italy; Department of Molecular Neuroscience (P.G.), Institute of Neurology, Queen Square, London, United Kingdom; Department of Neurology and Hertie-Institute for Clinical Brain Research (C.G., L.S.), University of Tübingen, Tübingen, Germany; Department of Neurology (J.-S.K.), University of Frankfurt, Frankfurt/M, Germany; Department of Neurology (B.K., B.P.W.), Radboud University Nijmegen Medical Center, Nijmegen, Netherlands; Department of Medical Genetics and Child Development (M.B.), University of Pécs, Pécs, Hungary; Institute of Psychiatry and Neurology (M.R., R.R.), Warsaw, Poland; INSERM U679 and Department of Genetics (P.R., A.D.), Cytogenetics and Embryology, Hôpital de la Pitié-Salpêtrière, Paris, France; Department of Neurology (S.S.), St. Josef Hospital, University Hospital of Bochum, Bochum, Germany.

Supported by grant EUROSCA/LSHM-CT-2004-503304 from the European Union, grant GeneMove/01 GM 0503 from the German Ministry of Education and Research, and grant 3 PO5B 019 24 from the Polish Ministry of Scientific Research and Information Technology.

Disclosure: The authors report no conflicts of interest.

Received October 10, 2005. Accepted in final form March 6, 2006.

Address correspondence and reprint requests to Dr. Thomas Klockgether, Department of Neurology, University Hospital of Bonn, Sigmund-Freud-Str. 25, D-53105 Bonn, Germany; e-mail: klockgether@uni-bonn.de

Table 1 *Clinical characteristics of the study population*

	Patients in trial 1	Patients in trial 2	Controls in trial 2
No.	167	119	110
F/M	72/89	61/58	67/43
Age, y	49.2 ± 14.2 (13–80)	50.3 ± 13.2 (22–81)	47.1 ± 15.0 (19–78)
Disease duration, y	10.8 ± 4.4 (0–46)	13.3 ± 8.3 (1–43)	
Barthel Index	88.9 ± 20.0 (5–100)	82.6 ± 25.8 (0–100)	100.0 ± 0.0 (100–100)
UHDRS-IV	n.d.	16.7 ± 8.0 (0–25)	25.0 ± 0.0 (25–25)

Age, disease duration, Barthel Index, and UHDRS-IV are given as mean ± SD. The range is given in parentheses.

UHDRS-IV = Unified Huntington's Disease Rating Scale, Part IV (functional assessment); n.d. = not determined.

ataxia on an impairment level. We report here the results of a large two-step validation trial performed in patients with spinocerebellar ataxia (SCA).

Methods. *Description of the scale.* SARA was devised during a meeting held in February 2004 in Bonn, Germany, by a group of European neurologists. Test items were selected according to specificity for ataxia and the possibility to standardize testing and rating procedures. Moreover, assessment had to be close to standard neurologic examination and not require technical equipment. In its initial form, SARA was composed of nine items including one item related to oculomotor function. Based on the results of Trial 1, SARA was modified and the oculomotor item was omitted. In its final form, SARA has eight items that yield a total score of 0 (no ataxia) to 40 (most severe ataxia); 1: gait (score 0 to 8), 2: stance (score 0 to 6), 3: sitting (score 0 to 4), 4: speech disturbance (score 0 to 6), 5: finger chase (score 0 to 4), 6: nose-finger test (score 0 to 4), 7: fast alternating hand movements (score 0 to 4), 8: heel-shin slide (score 0 to 4). Limb kinetic functions (items 5 to 8) are rated independently for both sides, and the arithmetic mean of both sides is included in the SARA total score (see appendix E-1 on the *Neurology* Web site at www.neurology.org).

Development of the scale. The study population consisted of 167 patients with SCA (SCA1: 23 patients, SCA2: 56 patients, SCA3: 23 patients, SCA6: 18 patients, SCA7: seven patients, SCA14: two patients, SCA17: one patient, unknown mutation: 29 patients) and eight controls in Trial 1, and 119 patients with SCA (SCA1: 15 patients, SCA2: 28 patients, SCA3: 26 patients, SCA6: 19 patients, SCA7: eight patients, SCA14: one patient, SCA17: three patients, SCA23: one patient, unknown mutation: 18 patients) and 110 controls in Trial 2. Clinical characteristics are given in table 1. Inclusion criteria for the patients with SCA were 1) progressive ataxia and 2) positive genetic testing for any SCA mutation or linkage to any SCA locus or clinical evidence of autosomal dominant inheritance. The majority of patients (Trial 1: n = 97, Trial 2: n = 95) were rated independently by two investigators. In Trial 2, a number of patients were rated twice with an interval of 1 to 34 days by as many as three investigators, resulting in a set of 15 intrarater assessments. All investigators were neurologists or senior residents in neurology training. There was a maximum of three investigators at each center. Investigators did not receive formal training, but were asked to closely follow the instructions that are part of the SARA scale (see appendix E-1).

In Trial 1, SARA was compared with ataxia disease stages, ICARS, and the Barthel Index and in Trial 2 with ataxia disease stages, Barthel Index, and the part IV (functional assessment) of the Unified Huntington's Disease Rating Scale (UHDRS-IV).[3,7-9] Disease stages were defined as follows: Stage 0: no gait difficulties; Stage 1: disease onset, as defined by onset of gait difficulties; Stage 2: loss of independent gait, as defined by permanent use of a walking aid or reliance on a supporting arm; Stage 3: confinement to wheelchair, as defined by permanent use of a wheelchair.[7] The ICARS results have been reported elsewhere.[6]

Thirty assessments of Trial 2 were video recorded. Eleven videos that were representative of the entire range of clinical severity were selected and randomly ordered on a video disk. From these videos, SARA rating was done by two examiners who were not involved in the clinical assessments. In addition, the severity of ataxia of the 11 cases presented on video was globally assessed by 14 examiners on a visual analogue scale (VAS) ranging from 0 (no ataxia) to 100 (most severe ataxia).

The study was approved by the ethics committees of the participating centers. Informed and written consent was obtained from all study participants.

Statistical analyses. Data are given as mean ± SD and range. Interrater and test-retest reliability was expressed using intraclass coefficients (ICCs), internal consistency using Cronbach's α coefficient. Calculation of ICCs was based on a model with only random effects. Coefficients exceeding 0.80 are considered acceptable for scales used to make group comparisons. In order to analyze internal validity, a principal component analysis was performed using varimax rotation.

In order to test the linearity of SARA, a regression analysis of the video ratings with the VAS assessments as a reference was performed. To make SARA and the VAS values directly comparable, SARA values were multiplied with a factor of 2.5 to yield a range from 0 to 100. As SARA will be used to measure differences between two assessments, i.e., before and after treatment, we also verified the linearity of differences between two measures. To this end, all differences between patient ratings of a given examiner were computed. For both analyses, a linear model was tested. For the SARA scores, a model with no intercept was used, and for the differences, a model with an intercept was used.

To test for correlation between SARA scores and Barthel Index, UHDRS-IV, and disease duration, Pearson's correlation test was used. The relation between SARA scores and disease stages was analyzed with analysis of variance. All tests were two sided; $p < 0.05$ was considered significant. Statistical analyses were performed using the SAS 8.2 statistical package (SAS Institute, Cary, NC).

Results. Analysis of Trial 1 showed high interrater reliability of SARA total score (ICC = 0.97) and of the items related to gait, stance, sitting, fast alternating hand movements, and heel-shin slide. In contrast, the items related to finger chase, nose-finger test, speech, and oculomotor function were less successful, with ICC <0.80 (range 0.70 to 0.79). Internal consistency was high, with Cronbach's α of 0.90; it increased when the oculomotor item was omitted. A factorial analysis showed that the results of all items, with the exception of the oculomotor item, were determined by a single factor. Based on these results, instructions for the items with interrater reliability <0.80 were modified, and the oculomotor item was omitted.

Analysis of Trial 2 determined the metric properties of SARA in its final form (see appendix E-1). In the patient group, the mean SARA score was 15.9 ± 8.5 (range 1.5 to 40). Ceiling effects were negligible, with one patient (0.84%) receiving the maximal possible score. Skewness of the distribution was 0.74. The mean SARA score of the control group was 0.4 ± 1.1 (range 0 to 7.5), and 79% of

Table 2 *Reliability and internal consistency of SARA*

Item	Interrater reliability (ICC)	Test-retest reliability (ICC)	Cronbach's α
SARA 1	0.96644	n.d.	0.930805
SARA2	0.95106	n.d.	0.929620
SARA3	0.92363	n.d.	0.933156
SARA4	0.85085	n.d.	0.934190
SARA5 right	0.85446	n.d.	0.936336
SARA5 left	0.85367	n.d.	0.935705
SARA6 right	0.80507	n.d.	0.932309
SARA6 left	0.75753	n.d.	0.933658
SARA7 right	0.87335	n.d.	0.931609
SARA7 left	0.81716	n.d.	0.933208
SARA8 right	0.80890	n.d.	0.933362
SARA8 left	0.74417	n.d.	0.933398
SARA total	0.97776	0.90032	0.938361

ICC = intraclass coefficient; SARA = Scale for the Assessment and Rating of Ataxia; n.d. = not determined.

controls had a score of 0. Positive ratings in controls were mainly obtained in items 5 and 8 left. Mean time to administer SARA was 14.2 ± 7.5 minutes (range 5 to 40) in patients and 7.2 ± 2.6 minutes (range 3 to 13) in controls.

Based on the results obtained in 95 patients assessed by two examiners, interrater reliability was very high, with an ICC of 0.98. All single items had good interrater reliability, with ICC >0.80 with exception of item 6 left (ICC = 0.76) and item 8 left (ICC = 0.74). Based on a limited set of 15 intrarater assessments, test-retest reliability was high, with an ICC of 0.90. Internal consistency was high, as indicated by Cronbach's α of 0.94; it did not increase when any item was deleted (table 2).

Factorial analysis revealed that the rating results in all items were determined by a single factor with an eigenvalue of 6.4, which explained 80% of the variance. All other factors had eigenvalues far less than 1.

Linearity of SARA was tested by a regression analysis of video ratings of 11 patients with global assessments as a reference. The SARA ratings and the differences between measures fitted a linear model (SARA: $p < 0.0001$, $r^2 = 0.98$; differences: $p < 0.0001$, $r^2 = 0.72$) (figure 1).

The SARA score increased with the disease stage ($p < 0.001$) (figure 2A). The SARA score was closely correlated with the Barthel Index ($r = -0.80$, $p < 0.0001$) and UHDRS-IV ($r = -0.89$, $p < 0.0001$), while it had only a weak correlation with disease duration ($r = 0.34$, $p < 0.0002$) (figure 2, B through D). Further analysis showed that the correlation between SARA score and disease duration was better, when only SCA-1 ($r = 0.74$, $p < 0.003$) or SCA3 patients ($r = 0.59$, $p < 0.002$) were considered. This was not found for SCA2 ($r = 0.08$, $p > 0.68$) and SCA6 ($r = 0.26$, $p > 0.28$).

Discussion. SARA, the new ataxia scale reported here, includes eight items reflecting neurologic manifestations of cerebellar ataxia. SARA only rates ataxia-related symptoms and does not consider nonataxia symptoms that often occur in patients with

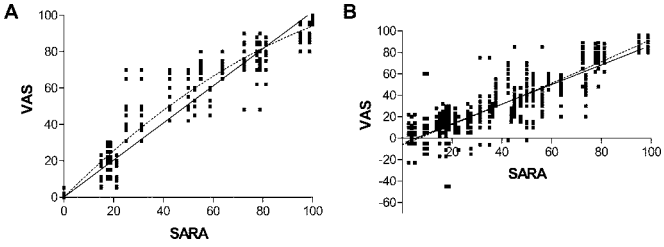

Figure 1. Regression analysis of Scale for the Assessment and Rating of Ataxia (SARA) video ratings of 11 representative patients with global assessments as a reference. SARA rating was done by two examiners who were not involved in the clinical assessments. SARA values were multiplied with a factor of 2.5 to yield a range from 0 to 100. Severity of ataxia was globally assessed by 14 examiners on a visual analogue scale (VAS) ranging from 0 (no ataxia) to 100 (most severe ataxia). (A) SARA ratings ($p < 0.0001$, $r^2 = 0.98$). (B) Differences between SARA ratings of a given examiner ($p < 0.0001$, $r^2 = 0.72$).

SCA. Therefore, it is possible that disease severity in certain diseases with extracerebellar features might not be faithfully reflected in the SARA score. However, consideration of nonataxia symptoms would have precluded the development of a scale that meets basic biometric requirements. To account for nonataxia symptoms, we have devised an Inventory of Non-Ataxia Symptoms (INAS) that provides structured and semiquantitative information on nonataxia symptoms. The items selected for SARA follow the steps of a standard neurologic examination. Oculomotor functions, which were part of the initial SARA version, are not considered in the final SARA version because the first validation trial indicated two underlying constructs, one globally reflecting cerebellar ataxia and the second reflecting solely oculomotor dysfunction.

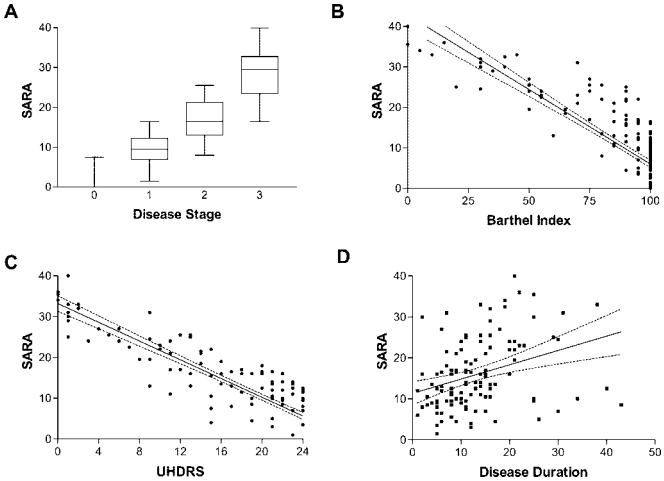

Figure 2. Validation of Scale for the Assessment and Rating of Ataxia (SARA) with external measures. (A) SARA scores at different ataxia disease stages. Data are given as box plots. (B) Correlation of the SARA score with Barthel Index ($r = -0.80$, $p < 0.0001$). (C) Correlation with UHDRS-IV ($r = -0.89$, $p < 0.0001$). (D) Correlation with disease duration ($r = 0.34$, $p < 0.0002$).

SARA underwent a rigorous validation procedure involving two clinical trials in large groups of patients with SCA and controls. The number of patients with SCA included in these trials is far greater than the number included in any previous clinical study of patients with SCA. Clinical presentation of most patients with SCA is characterized by prominent cerebellar ataxia. They therefore represent an appropriate study population for the validation of an ataxia scale. To demonstrate that SARA is also appropriate for other groups of ataxia patients, we are currently conducting further clinical studies.

There was a clear separation of the SARA ratings in patients and controls. Nevertheless, about 20% of controls had positive ratings in at least one item. The items producing the majority of positive ratings were related to kinetic functions of the nondominant side. These observations show that SARA is not an ideal clinical instrument to detect disease onset. Elimination of these items would increase the discriminatory power of SARA but reduce its sensitivity.

SARA was easy to administer and required less than 15 minutes per patient. This time is considerably shorter than the time required to complete ICARS.[6] SARA satisfied a number of accepted criteria of reliability. Interrater reliability was high, both for the total score and for most of the items. At the single-item level, SARA performed better than ICARS.[6] Although based on a small number of observations, our data suggest acceptable test-retest reliability of the total SARA score. However, meaningful analysis of the test-retest reliability of each item was not possible. Because the calculation of ICCs was based on a model with only random effects, the results are independent of the investigators and can be generalized to investigators with the same level of training as the investigators of this study. The internal consistency of SARA was high, as indicated by Cronbach's α of 0.94.

Most importantly, SARA results were determined by a single factor, which indicates that SARA measures a common underlying construct, cerebellar ataxia. In this respect, SARA favorably contrasts with ICARS, the results of which were shown to be influenced by four different factors that did not coincide with the ICARS subscales.[6]

Linearity of a scale and particularly of the differences between ratings is an important prerequisite for the use of a scale in clinical trials that include patients with a wide spectrum of clinical severity. To demonstrate linearity, we performed a regression analysis of the SARA ratings of two examiners with global assessments of 14 examiners as a reference. Our analysis showed that SARA ratings and the differences between ratings fitted a linear model. A limitation of this analysis comes from the fact that these ratings were done on videos.

To further demonstrate the validity of SARA, we used external measures. SARA scores increased with ataxia disease stages, as defined in a previous study.[7] In addition, SARA scores decreased with the Barthel Index, which measures activities of daily living and is routinely used to evaluate stroke patients.[8] SARA also decreased with UHDRS-IV, which provides a functional assessment of patients with Huntington disease (HD).[9] Because patients with ataxia and patients with HD face qualitatively similar functional limitations in daily life, it appeared appropriate to use this scale as an external validation criterion. These observations are consistent with the hypothesis that SARA is a valid method to measure the severity of ataxia. Based on the clinical knowledge that SCAs are progressive, but that the progression rate is variable depending on the genotype, repeat length, and unknown genetic and nongenetic factors,[7,10] only a moderate correlation between the SARA score and disease duration was expected. This was indeed found in the present study. Separate analysis of the four most common genotypes showed that the correlation between SARA score and disease duration was better, when only patients with SCA1 or SCA3 were considered. This was not found for SCA2 and SCA6. However, the small number of patients in each subgroup did not allow us to draw any final conclusions.

Acknowledgment

The authors thank the probands for participating in this study. The work of M.E.E.C. Glazowski (Warsaw) in videotaping of assessments is gratefully acknowledged. They also thank all other clinical investigators who contributed to data acquisition: S. Bonato (Brussels); M. Abele, M. Minnerop, B. Viebahn, and I. Vogt (Bonn); S. Romano (Milan); W.F. Abdo and H.J. Schelhaas (Nijmegen); C. Charles (Paris); and J. Zaremba and E. Zdzienicka (Warsaw).

References

1. Fillyaw MJ, Badger GJ, Bradley WG, et al. Quantitative measures of neurological function in chronic neuromuscular diseases and ataxia. J Neurol Sci 1989;92:17–36.
2. Klockgether T, Schroth G, Diener HC, Dichgans J. Idiopathic cerebellar ataxia of late onset: natural history and MRI morphology. J Neurol Neurosurg Psychiatry 1990;53:297–305.
3. Trouillas P, Takayanagi T, Hallett M, et al. International Cooperative Ataxia Rating Scale for pharmacological assessment of the cerebellar syndrome. J Neurol Sci 1997;145:205–211.
4. Subramony SH, May W, Lynch D, et al. Measuring Friedreich ataxia: interrater reliability of a neurologic rating scale. Neurology 2005;64:1261–1262.
5. Storey E, Tuck K, Hester R, Hughes A, Churchyard A. Inter-rater reliability of the International Cooperative Ataxia Rating Scale (ICARS). Mov Disord 2004;19:190–192.
6. Schmitz-Hübsch T, Tezenas du Montcel S, Baliko L, et al. Reliability and validity of the International Cooperative Ataxia Rating Scale (ICARS): A study in 156 spinocerebellar ataxia (SCA) patients. Mov Disord 2006;21:699–704.
7. Klockgether T, Lüdtke R, Kramer B, et al. The natural history of degenerative ataxia: a retrospective study in 466 patients. Brain 1998;121:589–600.
8. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. Md State Med J 1965;14:61–65.
9. Huntington Study Group. Unified Huntington's Disease Rating Scale: reliability and consistency. Mov Disord 1996;11:136–142.
10. Schöls L, Krüger R, Amoiridis G, Przuntek H, Epplen JT, Riess O. Spinocerebellar ataxia type 6: genotype and phenotype in German kindreds. J Neurol Neurosurg Psychiatry 1998;64:67–73.

Scale for the assessment and rating of ataxia: Development of a new clinical scale

T. Schmitz-Hübsch, S. Tezenas du Montcel, L. Baliko, et al.

Neurology 2006;66;1717-1720
DOI 10.1212/01.wnl.0000219042.60538.92