


Development of SARAhome, a New Video-Based Tool for the Assessment of Ataxia at Home

Marcus Grobe-Einsler, MD,1,2* Arian Taheri Amin,1,2 Jennifer Faber, MD,1,2 Tamara Schaprian, PhD,1 Heike Jacobi, MD,1,3 Tanja Schmitz-Hübsch, MD,4 Alhassane Diallo, PhD,5,6 Sophie Tezenas du Montcel, PhD,7 and Thomas Klockgether, MD,1,2

1 German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany 2 Department of Neurology, University Hospital Bonn, Bonn, Germany 3 Department of Neurology, University Hospital of Heidelberg, Heidelberg, Germany 4 Experimental and Clinical Research Center, A Cooperation of Charité—Universitätsmedizin Berlin and Max-Debreuck Center for Molecular Medicine, Berlin, Germany 5 INSERM U1137-IAME, Department of Biostatistical Modeling, Clinical Investigation, Pharmacometrics in Infectious Diseases, University Paris Diderot, Paris, France 6 Department of Epidemiology, Biostatistics and Clinical Research, APHP Bichat-Claude-Bernard Hospital, Paris, France 7 Sorbonne Université, Institut, Pierre Louis d’Épidémiologie et de Santé Publique, Assistance Publique-Hôpitaux de Paris, Institut National de la Santé et de la Recherche Médicale, University Hospital Pitié-Salpêtrière, Paris, France

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

*Correspondence to: Marcus Grobe-Einsler, German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany, Venusberg Campus 1, 53127 Bonn, Germany; E-mail: marcus.grobe-einsler@dzne.de

Relevant conflicts of interest/financial disclosures: There are no financial conflicts of interest to disclose.

Received: 9 October 2020; Revised: 20 November 2020; Accepted: 3 December 2020

Published online 12 January 2021 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.28478

ABSTRACT: Background: Clinical scales such as the Scale for the Assessment and Rating of Ataxia (SARA) cannot be used to study ataxia at home or to assess daily fluctuations. The objective of the current study was to develop a video-based instrument, SARAhome, for measuring ataxia severity easily and independently at home.

Methods: Based on feasibility of self-application, we selected 5 SARA items (gait, stance, speech, nose-finger test, fast alternating hand movements) for SARAhome (range, 0–28). We compared SARAhome items with total SARA scores in 526 patients with spinocerebellar ataxia types 1, 2, 3, and 6 from the EUROSCA natural history study. To prospectively validate the SARAhome, we directly compared the self-applied SARAhome and the conventional SARA in 50 ataxia patients. To demonstrate feasibility of independent home recordings in a pilot study, 12 ataxia patients were instructed to obtain a video each morning and evening over a period of 14 days. All videos were rated offline by a trained rater.

Results: SARAhome extracted from the EUROSCA baseline data was highly correlated with conventional SARA (r = 0.9854, P < 0.0001). In the prospective validation study, the SARAhome was highly correlated with the conventional SARA (r = 0.9254, P < 0.0001). Five of 12 participants of the pilot study obtained a complete set of 28 evaluable videos. Seven participants obtained 13–27 videos. The intraindividual differences between the lowest and highest SARAhome scores ranged from 1 to 5.5.

Conclusion: The SARAhome and the conventional SARA are highly correlated. Application at home is feasible. There was a considerable degree of intra-individual variability of the SARAhome scores. © 2021 The Authors. Movement Disorders published by Wiley Periodicals LLC on behalf of International Parkinson and Movement Disorder Society.

Key Words: ataxia; digital assessment; SARA; home assessment

Ataxias refer to a group of degenerative diseases of the nervous system in which progressive ataxia is the core feature of clinical manifestation. They comprise diseases of both genetic and nongenetic origin. In most of them, there is prominent damage of the cerebellar cortex with consecutive cerebellar atrophy, but other parts of the nervous system, such as the spinal cord, basal ganglia, brain stem, and peripheral nerves, can be involved.

Proper assessment of the severity of ataxia is essential for clinical care, but even more for observational studies and interventional trials. This is done using clinical
scales, such as the Scale for the Assessment and Rating of Ataxia (SARA), an extensively validated scale that is applied by trained examiners in a clinical or research environment.\(^1\) Such scales cannot be used to study the severity of ataxia at home or to assess day-to-day and within-day fluctuations, which are reported by almost all ataxia patients. It is therefore important to develop new instruments based on digital technology that allow measuring severity of ataxia at home.\(^2,3\) Digital technologies are providing new opportunities for this purpose. The urgent need for instruments that allow remote assessment of ataxia is underlined by the current COVID-19 pandemic that impeded patient care and led to interruption of observational and interventional trials in ataxia.

We here present a video-based assessment, SARA\(^{\text{home}}\), that can be easily done independently of the presence of an examiner applied by ataxia patients themselves at home.

### Patients and Methods

#### Description of SARA\(^{\text{home}}\)

Based on a detailed analysis of the feasibility of self-application at home, we selected 5 of the 8 SARA items (gait, stance, speech, nose-finger test, fast alternating hand movements) for SARA\(^{\text{home}}\). The SARA\(^{\text{home}}\) sum score ranges from 0 to 28, with 0 indicating absence of ataxia and 28 being the most severe degree of ataxia. To facilitate implementation, we made small changes to the instructions including reduction of the walking distance to 5 m (1 way), performance of fast-alternating hand movements and nose-finger test on a chair allowing support of the feet for every patient and replacement of the investigator’s finger that serves as the target in the nose-finger test by a tape-mark on the wall (Fig. 1). After instruction by an experienced rater, patients performed SARA\(^{\text{home}}\) without further assistance in front of a tablet camera. Severely affected patients were allowed support by a second person during gait and stance, in analogy to conventional SARA. If an item could not be performed, patients were asked to verbally share that information on video to be used for rating. SARA\(^{\text{home}}\) was performed in the hospital during the validation study and at home during the pilot study. The tablets were supplied by the study-site for the duration of the study and returned after the last recording for download and rating of videos.

#### Study Participants

To compare the results of the 5 SARA items selected for SARA\(^{\text{home}}\) with those of the complete SARA, we retrospectively analyzed the baseline data of 526 patients with spinocerebellar ataxia types 1, 2, 3, and 6 from the EUROSCA natural history study.\(^4,5\) To calculate progression and performance over time, we analyzed an additional 415 SARA scores from follow-up after 1 year, 416 after 2 years, and 336 after 3 years. To prospectively validate SARA\(^{\text{home}}\), we directly compared the results of SARA\(^{\text{home}}\) with those of conventional complete SARA in 50 ataxia patients. To demonstrate feasibility of independent home recordings, we performed a pilot study in 12 patients. Inclusion criteria in both studies were clinical diagnosis of cerebellar ataxia and capability to comply with study protocol. Study participants of the prospective validation study and the pilot study were consecutively recruited from ongoing ataxia studies of the German Center for Neurodegenerative Diseases (DZNE), 4 patients from the prospective validation also participated in the pilot study. Patient characteristics are given in Table 1. All studies were approved by the responsible ethics committee. Written informed consent was obtained from all study participants at enrollment.

#### Prospective Validation Study

The prospective validation study was performed in the DZNE clinical trial unit in Bonn, Germany. The SARA score was obtained from all study participants by a trained and experienced examiner. Subsequently, study participants received oral instructions on how to perform SARA\(^{\text{home}}\). Subjects then performed SARA\(^{\text{home}}\) independently without any further assistance using a tablet device. Only subjects at risk of falling received support, if necessary. Videos were rated by a trained and experienced rater (T.K.) who had no knowledge of the conventional SARA score.

#### Pilot Study

To demonstrate feasibility under real-life conditions at home and collect first data about daily fluctuation and training effects, 12 ataxia patients performed SARA\(^{\text{home}}\) twice daily (morning and evening) for a period of 14 days in a pilot study. Initial instructions were given by a trained investigator either during a study visit (n = 7) or via a video call (n = 5). Therefore, baseline SARA was not always available in the pilot study. In addition, subjects received detailed instructions for each item via a printout or digitally presented on the tablet device supplied by the DZNE that was also used to record SARA\(^{\text{home}}\) performance on video (Fig. S1). All videos of the pilot study were rated by an experienced rater (M.G.E.).

#### Statistical Analysis

A linear model was used to compare SARA and SARA\(^{\text{home}}\) items in the retrospective and prospective analysis, and Pearson’s correlation was calculated. To see if the differences in the scores of single items were similar
across the spectrum of severity, we performed a Bland–Altman analysis of the respective 5 items of SARA with SARA\textsubscript{home} scores from the prospective validation study. To calculate progression and performance over time between SARA and SARA\textsubscript{home} (retrospective data), a bivariate linear mixed model was used. In the pilot study, a 2-way repeated-measures analysis of variance was performed to evaluate the effect of different times (morning/evening) and day of assessment on SARA\textsubscript{home} score. To determine how many recording days are required to obtain a representative estimate of the entire 14-day period, we calculated mean values and confidence intervals for cumulative days. All analyses were done using GraphPad Prism version 7.02 and R versions 3.4.3 and 3.5.1.

**Results**

The SARA\textsubscript{home} score extracted from the EUROSCA baseline data was highly correlated with the SARA score ($r = 0.985$, $P < 0.0001$; Fig. S2a). Progression of SARA\textsubscript{home} and SARA were also highly correlated ($r = 0.935$, $P < 0.0001$).

**FIG. 1.** SARA\textsubscript{home} setup. (A) Gait is performed on a 5-m walkway in front of the tablet or smartphone placed on a tripod. Marks on the floor indicate position of the tripod (0 m), 2- and 5-m distances. Patient walks back and forth, as indicated by the dashed line (1). Subsequently, the patient walks from the 5-m mark toward the camera in tandem (2). (B) Stance is performed on the 2-m mark at a 45° angle to the camera. If practicable, 3-foot positions are taken (1-3). (C) Fast alternating hand movements and speech are performed sitting on a chair at the 2-m position in front of the camera. (D) Nose-finger test is performed sitting on a chair at the 2-m position that is rotated by 90° so that the patient faces the wall. Pointing movements are performed toward a mark attached to the wall at eye height.
In the prospective validation study, no falls or injuries occurred during testing. The SARAhome score derived from video rating was highly correlated with the conventional complete SARA score \( (r = 0.9254, P < 0.0001; \text{Fig. S2b}) \). Item-specific analysis showed correlation coefficients ranging from 0.7594 for fast alternating hand movements to 0.9433 with \( P < 0.0001 \) for gait (Table S1). Bias of Bland–Altman analysis was 0.76, and >95% of all data points lay within the 95% limits of agreement, suggesting that SARA and SARAhome are similar across the entire spectrum of severity (Fig. S3).

Five of 12 participants obtained a complete set of 28 recordings, and 7 participants obtained at least 1 recording per day. A comprehensive list of available recordings is given in Table S2. In total, 7 of 293 recordings were rated as missing because of incorrect performance of single items (stance, 2; alternating hand movements, 5). There were no falls or injuries during the home recordings. The intraindividual differences between the lowest and highest SARAhome score ranged from 1 to 5.5 (mean ± SD, 3.3 ± 1.3; see Fig. S4). The time of assessment (morning vs evening) had no statistically significant effect on SARAhome score \( (F_{1,4} = 1.435, P = 0.297) \), and there was no significant training effect during the 14 days of assessments \( (F_{13,52} = 1.234, P = 0.238) \). Interaction between these terms was also not statistically significant \( (F_{13,52} = 0.881, P = 0.577; \text{Table S3}) \). The width of confidence intervals of SARAhome scores dropped from 4.77 for 1-day to 1.25 for 14-day recordings. For all recording periods longer than 3 days, confidence intervals were less than twice the 14-day value (Fig. S5).

**Discussion**

We have developed and validated a video-based assessment of ataxia, SARAhome, that is derived from the original SARA and can be self-applied by patients at home. A retrospective analysis of data from the EUROSCA natural history study and a prospective validation study performed in 50 ataxia patients showed high correlations between SARAhome and conventional SARA and similarity of both scales across the spectrum of severity. We further demonstrated the feasibility of home recordings in a pilot study in 12 ataxia patients.

There have been several attempts to develop digital ataxia assessment tools, which are applicable to home testing.\(^6\)\(^-\)\(^10\) Some of these approaches focused on single SARA items,\(^6\)\(^-\)\(^8\) whereas others introduced items that are not included in SARA.\(^8\)\(^-\)\(^10\) One example is a tool that uses commercially available devices from video games (Microsoft Kinect, Leap Motion Controller) together with a computer screen to record performance of pediatric ataxia patients in 6 of 8 SARA items.\(^10\) Other instruments relied on accelerometers.\(^6\)\(^,\)\(^8\) Real data from home recordings with any of these tools have not been published.

SARAhome has the principal advantage that it is completely video-based without the need for specific hardware or an examiner. SARAhome can be performed with any tablet or smartphone that possesses a camera. In collaboration with a commercial partner, we have developed an app that patients can download on their own devices. The app provides detailed instructions for users and enables recording as well as secure transfer of SARAhome video recordings for off-line rating. As SARAhome is directly derived from SARA, and SARAhome and SARA scores are highly correlated, SARAhome may partly substitute for a conventional SARA assessment in hospitals or research institutions, for example, between scheduled study visits or in situations, such as the COVID-19 pandemic, in which hospital visits are not possible. Currently, SARAhome is still dependent on off-line rating by a trained rater, which eliminates intrarater variability. Intrarater reliability has not been explored in this study, but we expect similar results as those for SARA. In the future, subjective rating may be replaced by automatic video tracking.

Our pilot study in a small group of ataxia patients demonstrated that SARAhome can be safely and reliably performed by patients independently at home.
However, a possible limitation is the applicability in very severely affected patients and patients with major cognitive impairment who may need supervision. For the first time, this study provided data on fluctuation of ataxia severity. Fluctuations of the SARAhome score of at least 1 point were observed in all patients. However, we detected neither systematic differences of ataxia severity between morning and evening nor a training effect. To fully determine the causes of fluctuations, larger trials are required. Based on the analysis of confidence intervals for cumulative days, we suggest that a recording period of 4 days is representative for the entire 14-day period and provides a more meaningful measure of ataxia severity than a single conventional SARA assessment in the hospital.

Acknowledgments: This study was funded by the I2A innovation fund of the DZNE. T.K. is member of the European Reference Network for Rare Neurological Diseases (ERN-RD, project number 739510). We thank the members of the EUROSCA study group for obtaining the data of the EUROSCA study.

References

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site.