# SARA captures disparate progression and responsiveness in spinocerebellar ataxias

Emilien Petit, Tanja Schmitz-Hübsch, Giulia Coarelli, Heike Jacobi, Anna Heinzmann, Karla Figueroa, Susan Perlman, Christopher Gomez, George Wilmot, Jeremy Schmahmann, et al.

***SARA captures disparate progression and responsiveness in spinocerebellar ataxias***

Emilien Petit[1],MS, Tanja Schmitz-Hübsch[4], MD, Giulia Coarelli[1], MD, PhD, Heike Jacobi[3], MD, Anna Heinzmann[1], MD, Karla P Figueroa[6], MD, Susan L Perlman[7], MD, Christopher M Gomez[8], MD, PhD, George R Wilmot[9], MD, PhD, Jeremy D Schmahmann[10], MD, Sarah H Ying[11], MD, Theresa A Zesiewicz[12], MD, Henry L Paulson[13], MD, PhD, Vikram G Shakkottai[13], MD, PhD, Khalaf O Bushara[14], MD, Sheng-Han Kuo[15], MD, Michael D Geschwind[16], MD, PhD, Guangbin Xia[17], MD, Stefan M Pulst[6], MD, S H Subramony[17], MD, Claire Ewenczyk[1], MD, Alexis Brice[1], MD, PhD, Alexandra Durr[1], MD, PhD, Thomas Klockgether[2], MD, PhD, Tetsuo Ashizawa[5], MD, Sophie Tezenas du Montcel, MD, PhD,[1]

Corresponding author: Emilien Petit, emilien.petit@icm-institute.org

[1] Sorbonne Université, Paris Brain Institute - ICM, Inserm, CNRS, Inria, AP-HP, Paris, France
[2]German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany
[3]Department of Neurology, University Hospital Heidelberg, Heidelberg, Germany
[4]Experimental and Clinical Research Center, a cooperation of Max-Delbrueck Center of Molecular Medicine and Charité – Universitätsmedizin Berlin, Berlin, Germany.
[5] Weill Cornell Medicine at Houston Methodist Hospital, Houston, Texas, USA
[6]Department of Neurology, University of Utah, Salt Lake City, UT 84132, USA.
[7]Department of Neurology, University of California, Los Angeles, Los Angeles, CA, USA.
[8]Department of Neurology, University of Chicago, Chicago, IL, USA.
[9]Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA.
[10]Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA
[11]Department of Neurology, Johns Hopkins University, Baltimore, MD, USA.
[12]Department of Neurology, University of South Florida, Tampa, FL, USA.
[13]Department of Neurology, University of Michigan, Ann Arbor, MI, USA.
[14]Department of Neurology, University of Minnesota, Minneapolis, MN, USA.
[15]Department of Neurology, College of Physicians and Surgeons, Columbia University, New York, NY, USA.
[16]Department of Neurology, University of California, San Francisco, CA, USA.
[17]Department of Neurology and McKnight Brain Institute, University of Florida, Gainesville, FL, USA.

## ABSTRACT

**Background:** The Scale for Assessment and Rating of Ataxia (SARA) is a widely used clinical scale to assess cerebellar ataxia but faces some criticisms about the relevancy of all its items.

**Objectives:** To prepare for future clinical trials, we analyzed the progression of SARA and its items in several polyQ spinocerebellar ataxias (SCA) from various cohorts.

**Methods:** We included data from patients with SCA1, SCA2, SCA3, and SCA6 from four cohorts (EUROSCA, RISCA, CRC-SCA, and SPATAX) for a total of 850 carriers and 3431 observations. Longitudinal progression of the SARA and its items was measured. Cohort, stage and genetic effects were tested. We looked at the respective contribution of each item to the total scale. Sensitivity to change of the scale and the impact of item removal was evaluated by calculating sample sizes needed in various scenarios.

**Results:** Longitudinal progression was significantly different between cohorts in SCA1, SCA2 and SCA3, the EUROSCA cohort having the fastest progression. Advanced stage patient were progressing slower in SCA2 and SCA6. Items were not contributing equally to the full scale through ataxia severity: gait, stance, hand-movement, and heel-shin contributed the most in early stage, and finger-chase, nose-finger, and sitting in later stages. Few items drove the sensitivity to change of SARA, but changes in the scale structure could not improve its sensitivity in all populations.

**Conclusion:** SARA and its items progression pace showed high heterogeneity across cohorts and SCAs. However, no combinations of items improved the responsiveness in all SCAs or populations taken separately.

3794 words

## INTRODUCTION

Autosomal dominant cerebellar ataxias of type 1, 2, 3, and 6, also known as spinocerebellar ataxias (SCA1, 2, 3, and 6), are clinically heterogeneous neurodegenerative diseases. They are caused by (CAG)n repeat expansions in the *ATXN1*, *ATXN2*, *ATXN3*, and *CACNA1A* genes respectively, resulting in the expansion of a polyglutamine tract in the corresponding proteins (Coarelli et al., 2023; Durr, 2010). SCAs are rare diseases, with a global prevalence of 0.0 to 5.6 per 100,000 (Ruano et al., 2014). The number of CAG repeats determines disease carrier status, the threshold between what is considered a pathological or a normal size varies for each SCA, and the longer the repeats in the expansion the earlier the age at onset (Tezenas du Montcel et al., 2014), and the faster the disease in SCA1 and SCA2 (Jacobi et al., 2011).

In the context of readiness for clinical trials in SCAs, there is a need to understand the natural history over the individual course of disease to refine the inclusion criteria for potential treatments. Cerebellar ataxia (including signs and symptoms) is clinically progressive over the disease course, considered to reflect the progression of the underlying pathology of SCA and it can be assessed using the Scale for Assessment and Rating of Ataxia (SARA) (Schmitz-Hubsch et al., 2006). A previous meta-analysis already studied the progression of the SARA for these 4 SCA subtypes in different cohorts to estimate the annual progression of the SARA score (Diallo et al., 2020). In addition, the responsiveness to change of the SARA has already been studied in several ataxias (Traschütz et al., 2023) where some modifications of the scale by removing items seemed to increase its overall responsiveness. However, the course of individual items and their contribution to the overall scale in terms of responsiveness to change has not been investigated in SCA1, 2, 3, and 6.

In this study, we aim to (1) assess the progression of the total score and of each item and their differences between cohorts and SCAs, (2) describe the contribution of each item to the total scale across all SARA ranges, (3) analyze the responsiveness to change of the total scale by sample size estimates, and find potentially better items combinations to create a more responsive scale.

For the purpose of our study, we pooled data of SCA1, 2, 3, and 6 patients from four cohorts (EUROSCA, CRC-SCA, SPATAX, and RISCA cohorts). The first three are composed of ataxic patients and RISCA only included pre-ataxic participants, allowing us to have patients at all stages of the disease.

## MATERIALS AND METHODS

### *Study population*

Data from the four cohorts of SCA carriers were pooled together. Three cohorts included affected subjects with 2 European cohorts: EUROSCA (Jacobi et al., 2015) and SPATAX (Monin et al., 2015)) and one from the USA: CRC-SCA (Ashizawa et al., 2013). The fourth cohort (RISCA) included European pre-symptomatic expansion carriers (Jacobi et al., 2020). From these cohorts, which included participants with different SCA types, we have selected the subjects with a positive genetic test for *ATXN1*, *ATXN2*, *ATXN3*, or *CANCA1A* gene. Only patients with at least 2 visits with available information on SARA and SARA items were kept for analysis. The characteristics of the four cohorts are given below and in Table 1.

**EUROSCA:** The study was conducted in 17 European centers. Patients were eligible if they had progressive, otherwise unexplained ataxia and a positive molecular genetic test for SCA1, SCA2, SCA3, or SCA6. Patients were consecutively recruited within a predetermined time frame between July 2005 and August 2006. Patients were seen at a baseline visit, followed by annual visits for 3 years. After the initial 3-year observation period,

study participants entered an extension phase in which study assessments were performed in conjunction with routine visits resulting in irregular intervals between visits.

**CRC-SCA:** The study was conducted at 12 U.S. centers. Patients were eligible if they had a positive molecular genetic test for SCA1, SCA2, SCA3, and SCA6 and if they were at least 6 years of age. Subjects with comorbidities affecting SARA and other ataxia measures used in the study were excluded. The study began in April 2010. Clinical assessments were performed at the baseline visit and every six months until two years from the baseline visit or until the end of August 2012. The study continued after August 2012, but we only had data from the first 2 years.

***RISCA***: Between September 13, 2008, and December 1, 2011, offspring or siblings of patients with SCA1, SCA2, SCA3, or SCA6 were enrolled in a prospective, longitudinal observational study at 14 European centers. To be eligible for inclusion in this study, individuals had to have no ataxia and between the ages of 18–50 years if directly related to individuals with SCA1, SCA2, or SCA3, or between the ages of 35–70 years if directly related to individuals with SCA6. Non-converters, i.e. patients without a visit with a SARA > 3 were removed from the database. These patients have irregular visits but some of them are seen by the clinician on an annual basis.

***SPATAX:*** Since 2005 (first visit with SARA assessed), all individuals with spinocerebellar degenerations were referred to the National (French) reference center for neurogenetic diseases and were entered into the SPATAX database (created in 2000, using the REDCap electronic data capture tools for data collection and management), which was hosted and insured by the Paris Brain Institute (Institut du Cerveau, ICM). We included patients with SCA 1, 2, 3, and 6. We removed the overlapping individuals from EUROSCA and RISCA because SPATAX participated in these studies.

***Outcomes***

To assess the cerebellar ataxia, we used the SARA score (Schmitz-Hubsch et al., 2006). In addition, we analyzed the SARA items, grouped as follows: the four axial SARA items (*gait* (0-8), *stance* (0-6), *sitting* (0-4) and *speech disturbance* (0-6)) and the four appendicular SARA items (*finger-chase* (0-4), *nose-finger* (0-4), *fast-alternating hand movements* (0-4), and *heel-shin slide* (0-4)). Appendicular items are calculated as the average of both sides, resulting in half-point values. Participants were stratified into three groups according to their SARA at baseline: Early-stage patients (<10), intermediate-stage stage (10-25), and advanced stage (>25). These thresholds are arbitrary but were chosen to detect potential non-linear effects of the baseline SARA.

Age of onset is defined as the patient-reported time of disease onset. For RISCA participants who converted during the study, the age at onset was not reported and was taken to be the age at the first visit with a SARA greater than 3. The estimated age at onset was calculated using the CAG repeats as proposed for SCA1, 2, and 6 (Tezenas du Montcel et al., 2014) and for SCA3 (Peng et al., 2021). The Time to Onset (TTO) was defined as the time to the estimated age at onset at a given time.

### *Statistical analysis*

Descriptive statistics

Baseline data are described using frequencies and percentages for qualitative data and means and standard deviations (SD) for quantitative variables. To compare the sample selected for analysis (at least two visits) with participants not selected for analysis (no follow-up) and to compare the characteristics of the patients in each cohort we used a chi-squared test for qualitative variables and ANOVA for quantitative variables.

Modeling of the SARA score progression

We modeled the change from baseline in the full SARA score and the eight SARA items. We used linear mixed models estimated by maximum likelihood. We first tested cubic, quadratic, and linear progressions to detect potential floor or ceiling effects. The significance of the cubic and quadratic effects was assessed using the likelihood ratio test. Because we were examining the change from baseline, no fixed intercepts were included in the model, only fixed slopes. The random effect included the individual's intercept and slope.

For full SARA models, the effect of the cohort, stage group (early, intermediate, and advanced stage), and normalized expanded CAG repeat length was assessed by adding interaction with the slope for these covariables. Non-significant covariables with the lowest mean sum of squares were iteratively removed from the model. CAG repeat length was standardized within all SCAs and the effect of CAG on the slope was expressed as the annual point of SARA per standard deviation of CAG repeat length. For item-level models, we adjusted all models for cohort, stage group, and CAG repeat length.

Transversal Item contribution to the full-scale

Each SARA item was divided by its maximum score, normalizing them between 0 and 1 to make them comparable. For this analysis, the SARA score was divided into increments of 5 between 0 and 40 and, within each division, the mean and the 95% CI interval for each item were calculated, allowing us to compare the relative contribution of each item to the full scale. The mean value was also compared to a theoretical x=y straight line representing what an equal contribution of the items to the scale would look like.

Sample size calculation

To estimate the sensitivity of SARA and items combination, we simulated data at 1 year after inclusion and calculated the mean and SD of the variation. Data were simulated after determining whether the progression included quadratic and cubic effects. We then calculated the sample size needed to detect a 50% reduction in progression with 80% power. In each subset by SCA or by cohort, we iteratively added items that most increased (or less decreased) the sample size to build a *stepwise-optimized SARA* (SO-SARA), inspired by (Traschütz et al., 2023).

Statistical tests were performed at the conventional two-tailed type I error of 0.05. Data were analyzed using R version 4.2.0 (R Core Team, 2018).

## RESULTS

### *Data description*

Patients with only 1 visit, who were not included in the further analysis, had higher SARA at baseline and longer disease duration (Supplementary Table 1), suggesting a lower probability of follow-up when at more advanced disease stages. A total of 850 mutation carriers were included in the analysis: 186 SCA1 (107 EUROSCA, 37 in CRC-SCA, 16 in SPATAX, and 26 in RISCA), 248 SCA2 (146 EUROSCA, 51 CRC-SCA, 29 SPATAX, and 22 RISCA), 272 SCA3 (120 EUROSCA, 91 CRC-SCA, 50 SPATAX, and 11 RISCA) and 144 SCA6 (86 EUROSCA, 52 CRC-SCA, 4 SPATAX, and 2 RISCA), for a total of 3431 observations. The median number of visits per patient was 4 (3, 5) and the mean time between two visits was 14.1±8.7 months. These figures were heterogeneous between the cohorts: SPATAX participants had the lowest mean number of visits (2.9±1.1) and EUROSCA ones the highest (4.8±1.7). CRC-SCA had the shortest time between two visits (6.6±1.6 months) and RISCA had the longest (25.8±8.5 months).

The cohorts differed in the following characteristics (Supplementary Table 2, 3, 4, and 5): RISCA subjects were younger than the other three cohorts due to their inclusion criteria (SARA <3 at the baseline visit). In SCA1 and SCA3, there were significant differences in CAG repeat length between the cohorts. In SCA1, CRC-SCA participants had significantly fewer CAG repeats than SPATAX and EUROSCA participants (46.0±3.6 vs 47.7±5.7 and 50.3±6.8, p=0.021 and p=0.028 respectively) and later age of onset. In SCA3, CRC-SCA participants had longer CAG repeat lengths than EUROSCA ones (70.4±3.6 vs 68.8±4.2, p=0.0037), but this was not associated with a difference in ages of onset.

Among the cohorts with ataxic patients at baseline, CRC-SCA participants had lower SARA at baseline than SPATAX participants in SCA2 and SCA3 (p=0.017 and p=0.038 respectively).

***Modeling of the SARA score progression from baseline***

The models of progression from baseline are described in Table 2. All models included quadratic or cubic effects. The cohort interaction with slope was significant for SCA1 (p=0.001), SCA2 (p=0.0007), and SCA3 (p<0.0001). EUROSCA participants always had the higher linear component with 2.19±0.13, 1.69±0.11, and 1.52±0.12 in SCA1, SCA2, and SCA3, respectively. RISCA subjects had the slowest progression in SCA1 (1.32±0.21), SPATAX subjects in SCA2 (0.84±0.23), and CRC subjects in SCA3 (0.68±0.18). The effect of CAG repeat length on the slope was significant for SCA1 (0.36±0.09, p <0.0001) and SCA2 (0.27±0.07, p = 0.0002), after adjustment for the cohort. Patients in the advanced stage had a slower progression in SCA2 and SCA6, with a reduction in the linear progression of -1.05±0.24 and -0.90±0.33, respectively, compared to patients at an intermediate stage.

***Items contribution to the total score***

Throughout the progression of the total SARA, the contribution of the items was variable depending on the disease stage and the SCA (Figure 1). The *sitting* item had a low contribution in the early and intermediate stages in each SCA as its mean normalized value was lower than any other item in each group with SARA < 20. In SCA1 and SCA2, items *gait*, *stance*, *hand-movement,* and *heel-shin* had higher mean normalized scores than *finger-chase* and *nose-finger* for groups with SARA between 15 and 30. In SCA2 specifically, *heel-shin* had a significantly higher value than any other item in the 0-5 SARA visit group. In SCA3, the differences in dynamics between items were larger: *gait* was significantly higher than any other item in visit groups with SARA from 15 to 30.

***Modeling of the item progression from baseline***

The effect of cohort, CAG repeat length, and SARA at baseline on item progressions were very heterogeneous among SCAs (Figure 2). The cohort effect was significant for *gait* in SCA1, SCA2, and SCA3, with EUROSCA having the fastest progression, followed by CRC-SCA. SCA3 had the strongest cohort effect, significant for all items except *sitting* and *speech*, and CRC-SCA patients had the slowest progression, for all appendicular items with a significant improvement for the *nose-finger* item*)* . CAG repeat length had a significant effect only in SCA1 and SCA2, for all items but *speech* and *sitting*. For all SCAs, advanced stage patients had slower progression in *gait* and *stance* compared to early and intermediate stages.

***Stepwise Optimized SARA by SARA at baseline***

For intermediate and early-stage patients, the sample size required to assess a 50% reduction in total SARA progression over one year was 120 and 126, respectively (Figure 3A). The best sensitivity of the SO-SARA was achieved with five items representing a sample size reduction of 4% and 10% in early and intermediate-stage patients, respectively, compared to the full SARA (n=115 and n=113). In the advanced-stage patients, the best sample size was also achieved with five items, but increased significantly with additional items, resulting in a sample size reduction of 47% compared to full SARA. In the next analysis, we decided to remove advanced patients due to their differences in progression compared to early and intermediate.

***Stepwise Optimized-SARA by SCA***

The sample size required to detect a 50% reduction in SARA progression over 1 year in the pooled cohort was the lowest in SCA1 (n=81) due to faster progression (Figure 3B). The total SARA was never the most sensitive scale in any SCA. The best sensitivity was achieved with three, six, four, and two items for SCA1, SCA2, SCA3, and SCA6, respectively. Compared to total SARA, it reduces the sample size by 24 (30%), 3(3%), 39 (26%) and 265 (55%) respectively. *Gait* was the best item overall, ranked 1$^{st}$ in all SCAs except SCA1. *Speech* and *hand-movements* were well ranked, and always improved the sensitivity of the SO-SARA (except *Hand-movements* in SCA1). In the total dataset, i.e. with all SCAs pooled, the best sensitivity was achieved with 4 items (*Gait, stance, Speech and Hand-movements* included), but this reduced the sample size by only 8 (6%) patients compared to the full scale.

***Stepwise Optimized SARA by cohort***

The sample size required to detect a 50% reduction in full SARA progression over 1 year was the highest with the progression estimation from the CRC-SCA cohort (n=273) and the lowest with the progression estimation from the EUROSCA cohort (n=108) (Figure 3C). The most sensitive items were heterogeneous between cohorts, but *gait* was the most sensitive in all cohorts except SPATAX (ranked 3$^{rd}$), where the best item was *speech*.

## DISCUSSION

In this retrospective multi-cohort analysis, we have shown that there are large discrepancies in SARA progression between cohorts and SCAs at both the full scale and the item level. Analysis of sensitivity to change of various combination of items highlighted that no item removal can reliably increase the sensitivity of the scale in every SCAs or population.

A total of 850 patients with ataxia or close to ataxia onset with pathologic repeats in ATXN1, ATXN2, ATXN3, and CACNA1A were analyzed for ataxia progression over 3431 visits with a mean follow-up of 4.15±2.98 years. At baseline, apart from the obvious differences in the RISCA participants compared to the ataxia cohorts, there were only slight differences in CAG repeat length in the SCA1 and SCA3 cohorts, with CRC-SCA patients having shorter repeats than EUROSCA patients. Ataxia progression from baseline was assessed as the change in SARA score. We found that the progression from baseline had a slight ceiling effect (except for SCA6, where the quadratic effect was positive), but was almost linear in the early stages. This ceiling effect could be explained by the slower progression observed in the advanced stage patients, but also by slower progression in patients with longer follow-up (Jacobi et al., 2015). The early and intermediate stage patients showed no differences in progression, suggesting that SARA increases steadily after it exceeds 3. In ataxic patients and converters, the pace of SARA progression since enrollment was consistent with previous findings, with linear component progression in cohorts ranging from 1.32-2.19, 0.84-1.69, 0.68-1.52, and 0.8 for SCA1, 2, 3, and 6, respectively (Jacobi et al., 2011, p. 6). Large differences in the rate of SARA progression between cohorts were noted for SCA1, SCA2 and SCA3 patients, with EUROSCA participants progressing faster than any other cohort. These differences were also present at the item level and cannot be explained by geographic or ethnic reasons, as SPATAX participants (French cohort) had a progression more similar to CRC-SCA participants (US cohort) than EUROSCA participants (European cohort). These differences cannot be explained by disease stage or genetic factors as they were similar between cohorts. One bias could be that SPATAX contributed to EUROSCA and RISCA, and thus the remaining SCA patients in SPATAX have a different profile than the inclusion criteria of EUROSCA/RISCA. Also, there were no SARA training tool available at the time of the data collection, but EUROSCA raters were the ones involved in development of the scale, resulting in higher experience in the use of the scale compared to other cohorts. In addition, the follow-up time of the SPATAX or CRC-SCA cohort was shorter than that of the EUROSCA cohort, which may lead to a higher misestimating of SARA progression.

In the analysis of the contribution of each item to the SARA scale, the contribution of the items to the total scale was heterogeneous throughout disease progression, with *gait*, *stance*, and *heel-shin* having higher contribution in the early stage and *sitting*, *finger-chase,* and *nose-finger* in the later stage. In addition, the progression of the items was heterogeneous; *gait* and *stance* (respectively ranging from 0-8 and 0-6) items were progressing faster, as would be expected with their wider range. The *speech* item showed moderate progression despite ranging from 0-6. Similar to previous report (Jacobi et al., 2011), the expanded CAG repeat length was associated with faster SARA progression in SCA1 and SCA2, and even at the item level.

In the context of a clinical trial using SARA as the primary endpoint, SCA1 patients would require a smaller sample size due to faster progression, with 81 patients required for a 1-year trial with a 50% efficacy of the intervention tested. The SO-SARA analysis showed that a few items mainly drive the sensitivity to change of SARA, and the addition of less sensitive items leads to a very small decrease (or even increase) of the sample size. However, even though the axial items were often the most sensitive ones, the sensitivity ranking of items was heterogeneous among SCAs and cohorts. This shows that even if total SARA was never the best combination of items in terms of sensitivity, there is no combination of items that would improve the responsiveness in all SCAs or populations taken separately. In addition, the differences between cohorts

highlight that these conclusions are highly dependent on the data used and that the absolute value of sample size estimate must be taken with caution.

The estimated sample sizes remain very large for such a rare disease as our estimations are based on a 1-year trial and with a 50% reduction on the slope, namely a relatively strong treatment effect. Moreover, one limitation in our analysis is that we did not consider any placebo effect, assuming that the placebo arm would progress similarly to the natural history, which is unlikely (Choi et al., 2022; Coarelli et al., 2022).

Although SARA remains a good tool for assessing the ataxia severity, the relatively slow progression of the patients with SCA diseases limits the relevance of using a clinical scale to show moderate treatment effects in short-term clinical trials. This is very clear in SCA6, the slowest progressing SCA, where the best combination of items reduced by 55% the sample size required compared to full SARA. Clinical trials should focus on detecting changes in biomarkers levels, especially in the pre-ataxic stage which has been shown to be a relevant time window for treating patients. Modification of biomarkers modifications (NFLs (Faber et al., 2023; Tezenas Du Montcel et al., 2023), MRI (Chandrasekaran et al., 2022)) already occurs at this stage while the SARA remains at very low levels. The READISCA project (for which longitudinal data will soon be available), which includes both pre-ataxic and ataxic SCA1 and SCA3 participants in a multi-continental study, will help to clarify these early changes.

## DECLARATIONS

### Funding and conflicts of interests

Dr. Tezenas du Montcel receives research support from Biogen. Dr Subramony receives research support from National Ataxia Foundation, Biohaven, NIH, FDA, MDA, Wyck Foundation, FARA, Reata, PTC therapeutics, Retrotope, Avidity Biosciences, Fulcrum therapeutics, Reneo Pharma and AAVANTIBio, and serves on the Scientific Advisory Board for Reata, Avidity and Dyne therapeutics.B.V., serves on the Scientific Advisory Board of BrainSpec Inc. and receives research support from Biogen. Dr Ashizawa received grants from NAF and Biogen and participates in Biohaven clinical trials NCT03952806 and NCT03701399. Dr. Schmitz-Hübsch received unrestricted research grants from Celgene/bms, speakers honoraria from Bayer, Biogen and AbbVie.

This work was supported by the National Institute of Neurological Disorders and Stroke (NINDS) grant U01 NS104326 to T.A., A.D., and T.K. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The CRC-SCA data were obtained using NIH RC1 NS068897 grant to TA.

## References

1. Ashizawa, T., Figueroa, K.P., Perlman, S.L., Gomez, C.M., Wilmot, G.R., Schmahmann, J.D., Ying, S.H., Zesiewicz, T.A., Paulson, H.L., Shakkottai, V.G., Bushara, K.O., Kuo, S.-H., Geschwind, M.D., Xia, G., Mazzoni, P., Krischer, J.P., Cuthbertson, D., Holbert, A., Ferguson, J.H., Pulst, S.M., Subramony, S., 2013. Clinical characteristics of patients with spinocerebellar ataxias 1, 2, 3 and 6 in the US; a prospective observational study. Orphanet J. Rare Dis. 8, 177. https://doi.org/10.1186/1750-1172-8-177
2. Chandrasekaran, J., Petit, E., Park, Y.-W., Tezenas du Montcel, S., Joers, J.M., Povazan, M., Banan, G., Valabregue, R., 2022. Clinically meaningful MR endpoints sensitive to preataxic and early ataxic stages of SCA1 and SCA3.

3. Choi, J.-H., Shin, C., Kim, H.-J., Jeon, B., 2022. Placebo response in degenerative cerebellar ataxias: a descriptive review of randomized, placebo-controlled trials. J. Neurol. 269, 62–71. https://doi.org/10.1007/s00415-020-10306-7
4. Coarelli, G., Coutelier, M., Durr, A., 2023. Autosomal dominant cerebellar ataxias: new genes and progress towards treatments. Lancet Neurol. 22, 735–749. https://doi.org/10.1016/S1474-4422(23)00068-6
5. Coarelli, G., Heinzmann, A., Ewenczyk, C., Fischer, C., Chupin, M., Monin, M.-L., Hurmic, H., Calvas, F., Calvas, P., Goizet, C., Thobois, S., Anheim, M., Nguyen, K., Devos, D., Verny, C., Ricigliano, V.A.G., Mangin, J.-F., Brice, A., Tezenas du Montcel, S., Durr, A., 2022. Safety and efficacy of riluzole in spinocerebellar ataxia type 2 in France (ATRIL): a multicentre, randomised, double-blind, placebo-controlled trial. Lancet Neurol. 21, 225–233. https://doi.org/10.1016/S1474-4422(21)00457-9
6. Diallo, A., Jacobi, H., Tezenas du Montcel, S., Klockgether, T., 2020. Natural history of most common spinocerebellar ataxia: a systematic review and meta-analysis. J. Neurol. https://doi.org/10.1007/s00415-020-09815-2
7. Durr, A., 2010. Autosomal dominant cerebellar ataxias: polyglutamine expansions and beyond. Lancet Neurol 9, 885–94. https://doi.org/10.1016/S1474-4422(10)70183-6
8. Faber, J., Berger, M., Carlo, W., Hübener-Schmid, J., Schaprian, T., Santana, M.M., Grobe-Einsler, M., Onder, D., Koyak, B., Giunti, P., Garcia-Moreno, H., Gonzalez-Robles, C., Lima, M., Raposo, M., Vieira Melo, A.R., De Almeida, L.P., Silva, P., Pinto, M.M., Van De Warrenburg, B.P., Van Gaalen, J., Jeroen De Vries, J., Oz, G., Joers, J.M., Synofzik, M., Schöls, L., Riess, O., Infante, J., Manrique, L., Timmann, D., Thieme, A., Jacobi, H., Reetz, K., Dogan, I., Onyike, C., Povazan, M., Schmahmann, J., Ratai, E.-M., Schmid, M., Klockgether, T., 2023. Stage-dependent biomarker changes in spinocerebellar ataxia type 3 (preprint). Neurology. https://doi.org/10.1101/2023.04.21.23287817
9. Jacobi, H., Bauer, P., Giunti, P., Labrum, R., Sweeney, M.G., Charles, P., Dürr, A., Marelli, C., Globas, C., Linnemann, C., 2011. The natural history of spinocerebellar ataxia type 1, 2, 3, and 6 A 2-year follow-up study. Neurology 77, 1035–1041.
10. Jacobi, H., du Montcel, S.T., Bauer, P., Giunti, P., Cook, A., Labrum, R., Parkinson, M.H., Durr, A., Brice, A., Charles, P., Marelli, C., Mariotti, C., Nanetti, L., Panzeri, M., Rakowicz, M., Sulek, A., Sobanska, A., Schmitz-Hübsch, T., Schöls, L., Hengel, H., Baliko, L., Melegh, B., Filla, A., Antenora, A., Infante, J., Berciano, J., van de Warrenburg, B.P., Timmann, D., Szymanski, S., Boesch, S., Kang, J.-S., Pandolfo, M., Schulz, J.B., Molho, S., Diallo, A., Klockgether, T., 2015. Long-term disease progression in spinocerebellar ataxia types 1, 2, 3, and 6: a longitudinal cohort study. Lancet Neurol. 14, 1101–1108. https://doi.org/10.1016/S1474-4422(15)00202-1
11. Jacobi, H., du Montcel, S.T., Romanzetti, S., Harmuth, F., Mariotti, C., Nanetti, L., Rakowicz, M., Makowicz, G., Durr, A., Monin, M.-L., Filla, A., Roca, A., Schöls, L., Hengel, H., Infante, J., Kang, J.-S., Timmann, D., Casali, C., Masciullo, M., Baliko, L., Melegh, B., Nachbauer, W., Bürk-Gergs, K., Schulz, J.B., Riess, O., Reetz, K., Klockgether, T., 2020. Conversion of individuals at risk for spinocerebellar ataxia types 1, 2, 3, and 6 to manifest ataxia (RISCA): a longitudinal cohort study. Lancet Neurol. 19, 738–747. https://doi.org/10.1016/S1474-4422(20)30235-0
12. Monin, M., Tezenas Du Montcel, S., Marelli, C., Cazeneuve, C., Charles, P., Tallaksen, C., Forlani, S., Stevanin, G., Brice, A., Durr, A., 2015. Survival and severity in dominant cerebellar ataxias. Ann. Clin. Transl. Neurol. 2, 202–207. https://doi.org/10.1002/acn3.156
13. Peng, L., Chen, Z., Long, Z., Liu, M., Lei, L., Wang, C., Peng, H., Shi, Y., Peng, Y., Deng, Q., Wang, S., Zou, G., Wan, L., Yuan, H., He, L., Xie, Y., Tang, Z., Wan, N., Gong, Y., Hou, X., Shen, L., Xia, K., Li, J., Chen, C., Qiu, R., Klockgether, T., Tang, B., Jiang, H., 2021. New Model for Estimation of the Age at Onset in Spinocerebellar Ataxia Type 3. Neurology 96, e2885–e2895. https://doi.org/10.1212/WNL.0000000000012068
14. R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
15. Ruano, L., Melo, C., Silva, M.C., Coutinho, P., 2014. The Global Epidemiology of Hereditary Ataxia and Spastic Paraplegia: A Systematic Review of Prevalence Studies. Neuroepidemiology 42, 174–183. https://doi.org/10.1159/000358801
16. Schmitz-Hubsch, T., du Montcel, S.T., Baliko, L., Berciano, J., Boesch, S., Depondt, C., Giunti, P., Globas, C., Infante, J., Kang, J.-S., Kremer, B., Mariotti, C., Melegh, B., Pandolfo, M., Rakowicz, M., Ribai, P., Rola, R., Schols, L., Szymanski, S., van de Warrenburg, B.P., Durr, A., Klockgether, T., 2006. Scale for the assessment and rating of ataxia: Development of a new clinical scale. Neurology 66, 1717–1720. https://doi.org/10.1212/01.wnl.0000219042.60538.92
17. Tezenas du Montcel, S., Durr, A., Bauer, P., Figueroa, K.P., Ichikawa, Y., Brussino, A., Forlani, S., Rakowicz, M., Schöls, L., Mariotti, C., van de Warrenburg, B.P.C., Orsi, L., Giunti, P., Filla, A.,

Szymanski, S., Klockgether, T., Berciano, J., Pandolfo, M., Boesch, S., Melegh, B., Timmann, D., Mandich, P., Camuzat, A., Clinical Research Consortium for Spinocerebellar Ataxia (CRC-SCA), EUROSCA network, Goto, J., Ashizawa, T., Cazeneuve, C., Tsuji, S., Pulst, S.-M., Brusco, A., Riess, O., Brice, A., Stevanin, G., 2014. Modulation of the age at onset in spinocerebellar ataxia by CAG tracts in various genes. Brain J. Neurol. 137, 2444–2455. https://doi.org/10.1093/brain/awu174

18. Tezenas Du Montcel, S., Petit, E., Olubajo, T., Faber, J., Lallemant-Dudek, P., Bushara, K., Perlman, S., Subramony, S.H., Morgan, D., Jackman, B., Paulson, H.L., Öz, G., Klockgether, T., Durr, A., Ashizawa, T., for the READISCA Consortium Collaborators, 2023. Baseline Clinical and Blood Biomarkers in Patients With Preataxic and Early-Stage Disease Spinocerebellar Ataxia 1 and 3. Neurology 100, e1836–e1848. https://doi.org/10.1212/WNL.0000000000207088
19. Traschütz, A., Adarmes- Gómez, A.D., Anheim, M., Baets, J., Brais, B., Gagnon, C., Gburek- Augustat, J., Doss, S., Hanağası, H.A., Kamm, C., Klivenyi, P., Klockgether, T., Klopstock, T., Minnerop, M., Münchau, A., Renaud, M., Santorelli, F.M., Schöls, L., Thieme, A., Vielhaber, S., Van De Warrenburg, B.P., Zanni, G., Hilgers, R., PREPARE Consortium, Synofzik, M., 2023. Responsiveness of the Scale for the Assessment and Rating of Ataxia and Natural History in 884 Recessive and Early Onset Ataxia Patients. Ann. Neurol. ana.26712. https://doi.org/10.1002/ana.26712

**Figure 1: Contribution to the SARA scale of each individual item of the scale**

Each point represents the average (with 95% CI) of a normalized item score for a group of visits in increments of 5 SARA points. For each class, the point is displayed in the middle of the class: for instance, the point referring to the 0-5 SARA group is displayed at SARA = 2.5. Non-overlapping CI are considered significant differences. The black line represents how the curve would be if all items were contributing equally throughout the disease, namely an x=y curve.

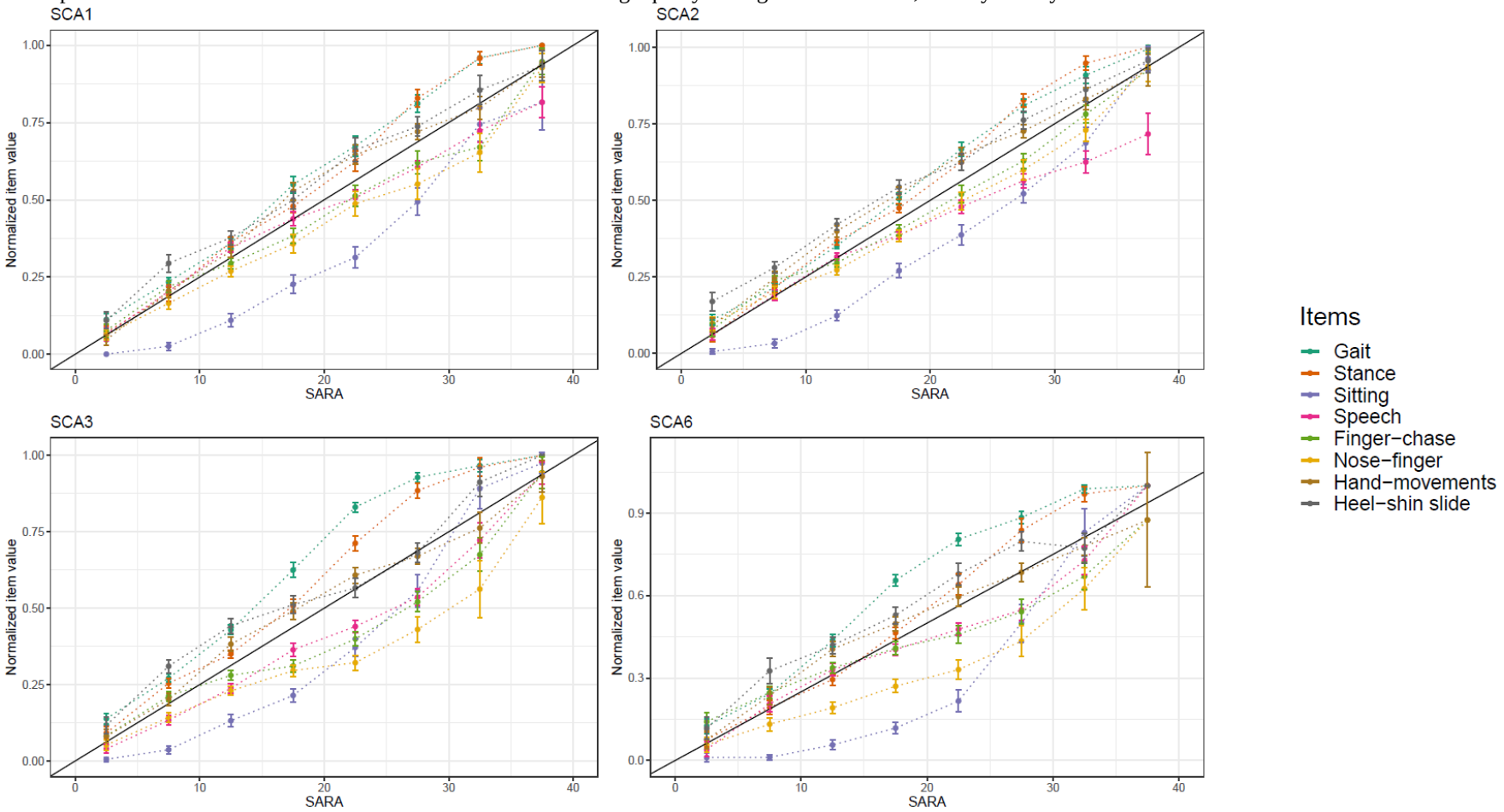

**Figure 2: Cohort, stage and CAG effect on item progression**

The estimations (with 95% CI) of the interaction between the covariables and time in the linear mixed effect model are displayed. Quadratic and cubic effects were tested for all items, but not shown in this graph if significant as their effects are negligible in early years . Significance of effects are represented as *** if p-value < 0.001, ** if p-value < 0.01 and * if p-value < 0.05. For cohort effects, the p-value indicates that it exists a significant difference of the linear progression between cohorts. For CAG, the p-value indicates that the CAG effect is significantly different from 0. For stage groups, the p-value indicates that at least one group is progressing faster or slower than the reference group (Intermediate group).

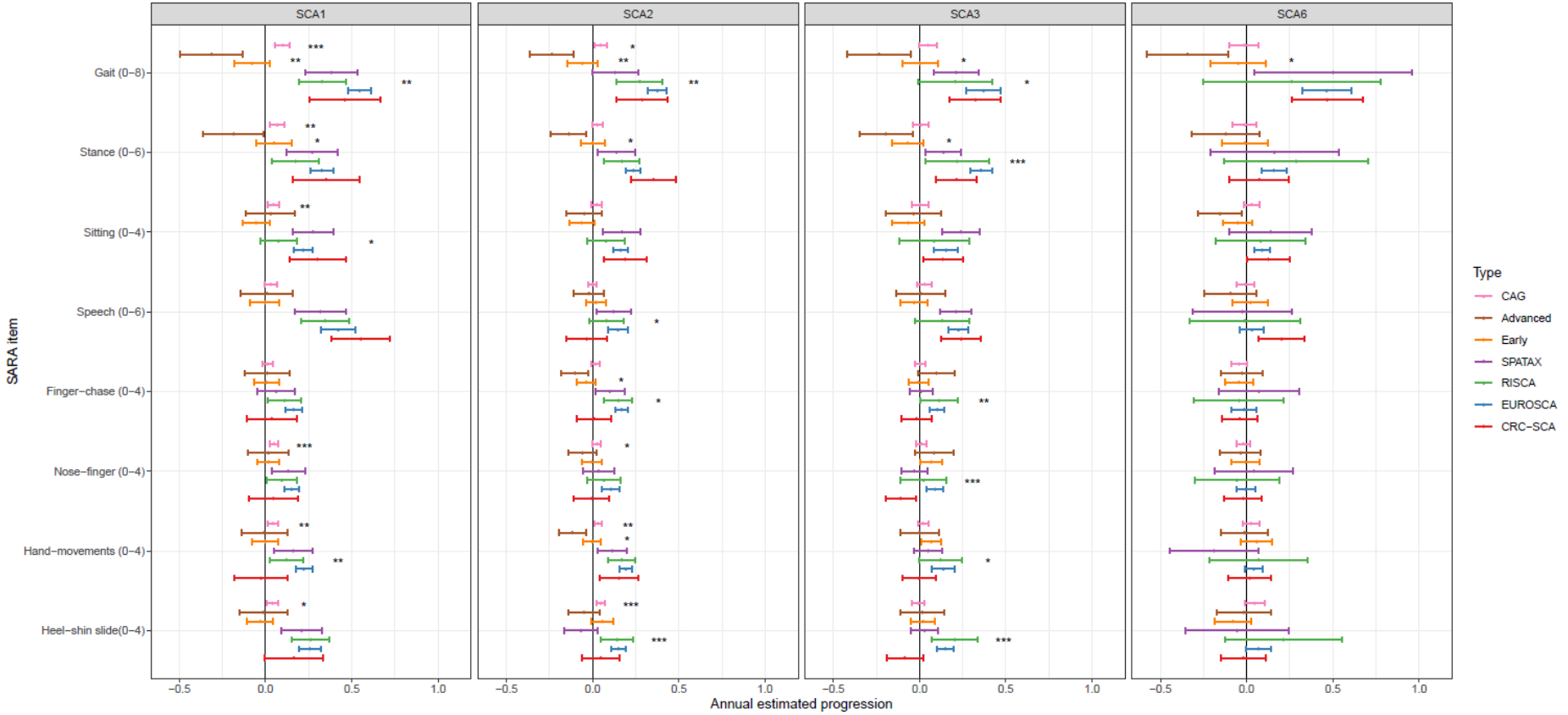

**Figure 3: Stepwise-optimized SARA by stage, SCA and cohort**

The left panels show the sample size estimation for each SO-SARA created. The right panels show in which order the items were added to the SO-SARAS to ensure maximum increase (or minimum decrease) of sensitivity to change.

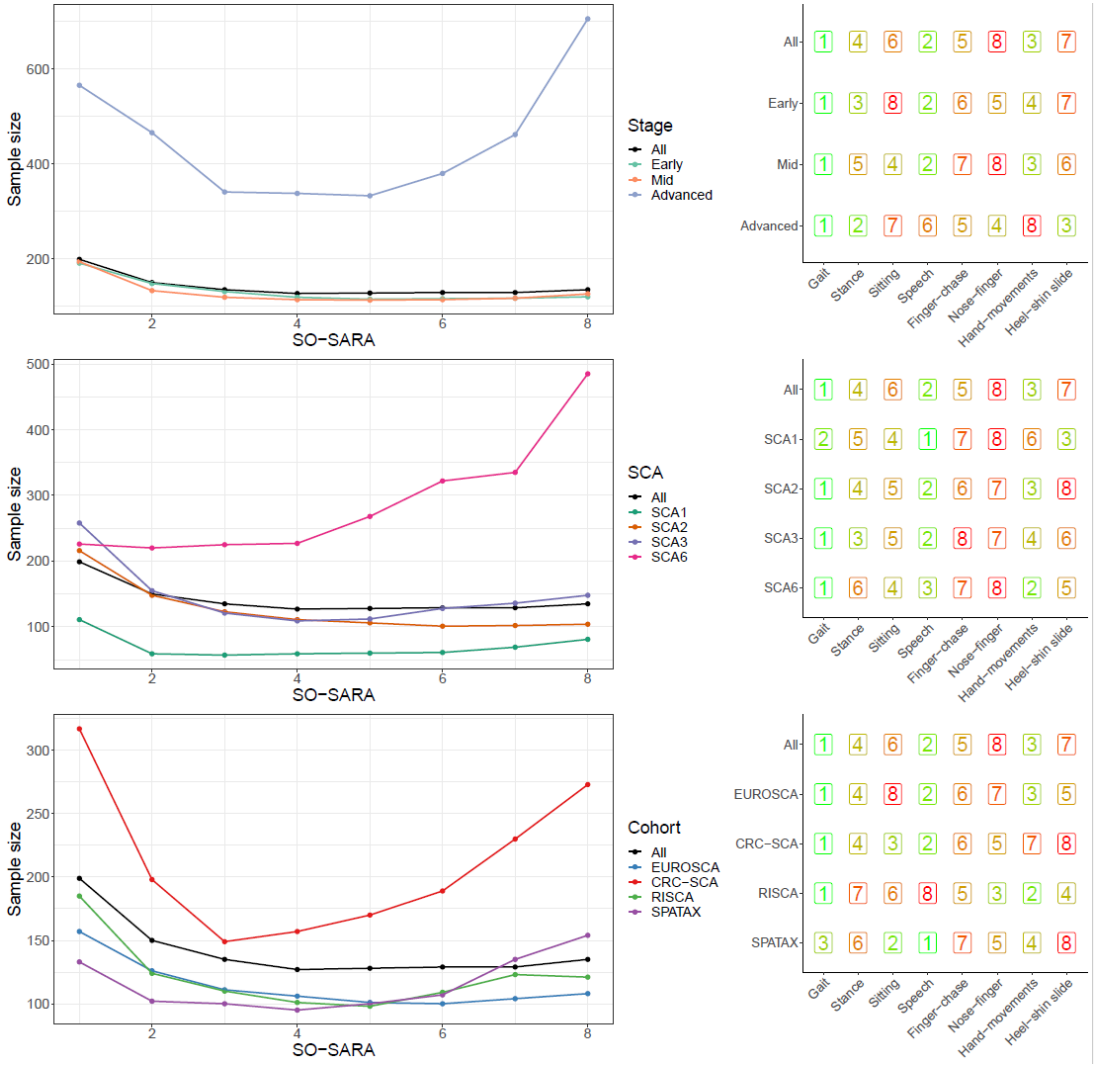

**Table 1: Cohort's Inclusion criteria and follow-up characteristics**

| | **CRC-SCA** | **EUROSCA** | **RISCA** | **SPATAX** |
|---|---|---|---|---|
| Study Population | Patients with SCA 1, 2, 3 and 6 of all races/ethnicities and both genders | Patients with SCA 1, 2, 3 and 6 | Unaffected adults' individuals that descend from SCA1, SCA2, SCA3 and SCA6 patients | Patients with SCA 1, 2, 3 and 6 |
| Inclusion Criteria | | | | |
| Ataxia | Presence of symptomatic ataxic disease | Progressive, otherwise unexplained ataxia | Absence of ataxia (SARA < 3) | |
| Molecular diagnosis | Pathogenic expansions in ATXN1, 2, 3 and CACNA1A either in the subject or another affected family member | Pathogenic expansions in ATXN1, 2, 3 and CACNA1A | Offspring and sibs of an individual with a pathogenic expansion in ATXN1, 2, 3 and CACNA1A | Pathogenic expansions in ATXN1, 2, 3 and CACNA1A |
| Age | 6 years and above. | - | - 18 - 50 years for descendants of SCA1, SCA2 or SC3 patients<br>- 35 - 70 years for descendants of SCA6 patients | >18 years |

| Consent | Willingness to participate in the study and ability to give informed consent | Written informed consent by the patient or his legal agent | Written, informed consent | Written, informed consent |
|---|---|---|---|---|
| Number of visits | 3.0±1.2 | 4.8±1.7 | 4.0±1.0 | 2.9±1.1 |
| Time between 2 visits (months) | 6.6± 1.6 | 14.3 ± 6.8 | 25.9 ± 8.5 | 19.5 ± 15.4 |
| Follow-up time (years) | 1.1±0.5 | 4.5±2.4 | 6.5±2.0 | 3.1±2.0 |

**Table 2: Models of SARA progression since inclusion**

| | SCA1 | | | SCA2 | | | SCA3 | | | SCA6 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Predictors | Estimates | SE | **p-value** | Estimates | SE | **p-value** | Estimates | SE | p-value | Estimates | SE | p-value |
| Cohort effect | | | **0.001** | | | **0.0007** | | | **<0.0001** | | | |
| [CRC] * TIME | 1.75 | 0.33 | **<0.0001** | 1.08 | 0.22 | **<0.0001** | 0.68 | 0.18 | **0.0001** | | | |
| [EUROSCA]* TIME | 2.19 | 0.13 | **<0.0001** | 1.69 | 0.11 | **<0.0001** | 1.52 | 0.12 | **0.0001** | | | |
| [RISCA] * TIME | 1.32 | 0.21 | **<0.0001** | 1.31 | 0.27 | **<0.0001** | 1.11 | 0.30 | **0.0002** | | | |
| [SPATAX] * TIME | 1.60 | 0.30 | **<0.0001** | 0.84 | 0.23 | **0.0004** | 0.81 | 0.18 | **<0.0001** | | | |
| TIME | | | | | | | | | | 0.80 | 0.14 | **<0.0001** |
| TIME^2 | -0.026 | 0.016 | 0.10 | -0.037 | 0.007 | **<0.0001** | 0.028 | 0.030 | 0.35 | 0.039 | 0.014 | **0.0056** |
| TIME^3 | | | | | | | -0.004 | 0.002 | **0.03** | | | |
| Stage group effect | | | | | | **<0.0001** | | | | | | **0.028** |
| [EARLY] * TIME | | | | -0.09 | 0.17 | 0.60 | | | | -0.19 | 0.22 | 0.39 |
| [ADVANCED] * TIME | | | | -1.05 | 0.24 | **<0.0001** | | | | -0.90 | 0.33 | **0.008** |
| [CAG] * TIME | 0.36 | 0.09 | **<0.0001** | 0.27 | 0.07 | **0.0002** | | | | | | |

Models were selected by first determining the presence of a quadratic or cubic effect. Covariate effects were then tested, and nonsignificant covariates with the smallest sum of squares were iteratively removed. For the stage group effect, intermediate participants were used as the reference. *: If there is a significant cohort effect, the linear term is reported for each of the respective cohorts ([cohort]*TIME).