Predicting individual disease progression including parameter uncertainty in rare neurodegenerative diseases: the example of Autosomal-Recessive Spastic Ataxia Charlevoix Saguenay (ARSACS)

Niels Hendrickx[@,1], France Mentré[1], Rebecca Schüle[3,4], Cynthia Gagnon[5], ARCA Study Group, EVIDENCE-RND consortium,

Andreas Traschütz [3,4], Matthis Synofzik[3,4], Emmanuelle Comets[1,2]

[@]niels.hendrickx@inserm.fr, [1]Université Paris Cité, IAME, Inserm, F-75018, Paris, France, [2]Univ Rennes, Inserm, EHESP, Irset - UMR_S 1085, 35000, Rennes, France.,
[3]Division Translational Genomics of Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research (HIH), University of Tübingen, Tübingen, Germany, [4]German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany,
[5]Centre de recherche du CHUS et du Centre de santé et des services sociaux du Saguenay–Lac-St-Jean, Faculté de médecine, Université de Sherbrooke, Québec, Canada

Context

- Genetic cerebellar ataxias are progressive ultra-rare neurodegenerative diseases affecting the cerebellum, causing debilitating impairment of gait, balance, speech and fine motor skill
- Over a hundred ataxia diseases are autosomal-recessive cerebellar ataxias (ARCA), often starting in early childhood or early adulthood
- We use Autosomal Recessive Spastic Ataxia of Charlevoix-Saguenay (ARSACS), one of the more common ataxias, as a case study
- There is an urgent need to develop robust statistical methodology that allows to predict progression trajectories[1]
- We develop a disease progression model to describe the individual evolution of the Scale for the Assessment and Rating of Ataxia (SARA) score versus Time Since Onset of symptoms (TSO) and its uncertainty

Modelling SARA score

Data description

- ARCA Registry[2]: longitudinal study with 173 ARSACS patients (1-7 visits, median 2)
- Core dataset including age, age of onset (AoO) as reported by the patients and SARA score
- Missing AoO imputed to the median AoO in the population (2 years)
- Covariates at inclusion: BMI, INAS score (Inventory of Non Ataxia Signs), genotype of the mutation (missence or loss of function), sex and Age of Onset

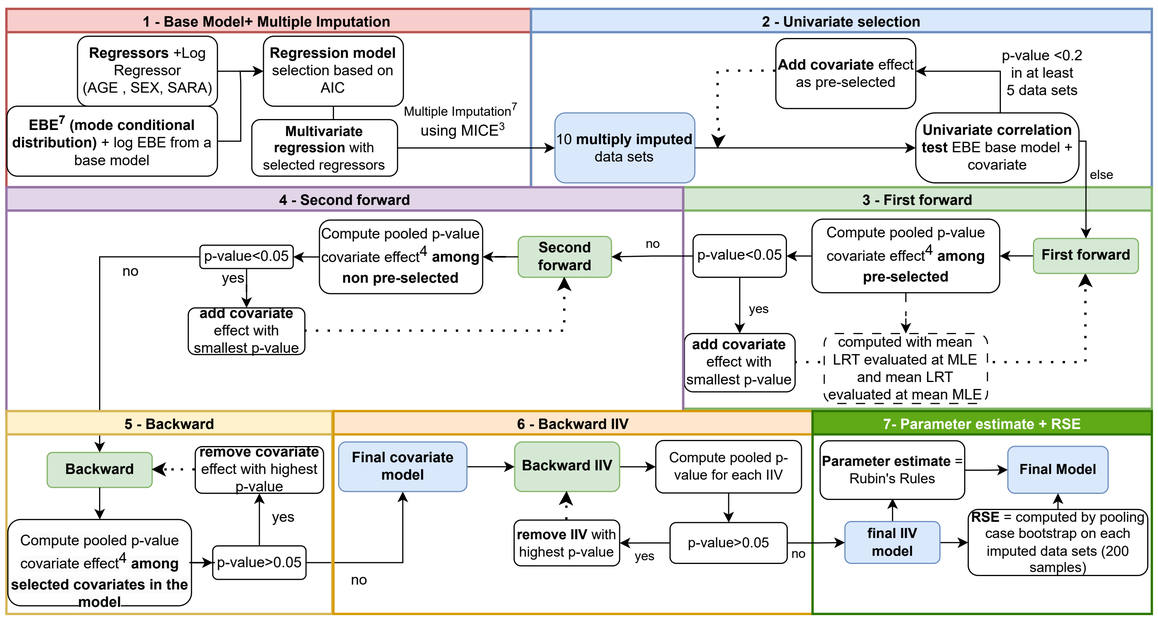
Non linear mixed effect model

- Structural model: 4 parameter logistic function
- $f(TSO) = S_0 + (S_0 - S_m) \times \frac{1}{1+\frac{e^{\alpha TSO}-1}{e^{\alpha T_{50}}-1}}$
- constant error model, lognormal distribution for parameters, diagonal covariance matrix, covariate effect: linear on log parameter, (and log covariate for continuous)

Continuous Covariate	median	25-75%	nb missing
BMI ($kg.m^{-2}$)	24.4	21.4-29	36
INAS	7	5-8	67
AoO (yr)	2	1-6	22
Categorical covariate	**Amount**	**percentage**	**nb missing**
genotype of mutation	115 loss of function	82%	32
AoO 0-7	131	87%	-
AoO 8-14	7	4.5%	-
AoO 15-40	13	8.5%	-
Sex	92 females	52%	0

Building covariate model via multiple imputation

- Many covariates have missing values = a specific approach is used[4]



Implementation

- Parameter estimation performed with the SAEM algorithm (saemix[5] in R 4.2.0[6]).

Results

Model	Without Covariates		With covariates	
Parameter	Value	RSE(%)	Value	RSE(%)
S_0	6.9	11	8.1	9
β_{male,S_0}	-	-	-0.28	50
S_m	35.8	2	34.3	2
β_{male,S_m}	-	-	0.08	38
α (yr^{-1})	0.11	13	0.12	12
T_{50} (yr)	36.8	3	38.1	3
$\beta_{AOO>15,T_{50}}$	-	-	-0.43	20
ω_{S_0}	0.27	37	0.19	42
ω_α	0.59	19	0.58	22
$\omega_{T_{50}}$	0.21	11	0.16	11
σ	1.92	6	1.86	6

- RSE computed with case bootstrap (200 samples)
- 8 covariates selected as pre-selected, 3 final covariates selected during first forward (none have missing values)

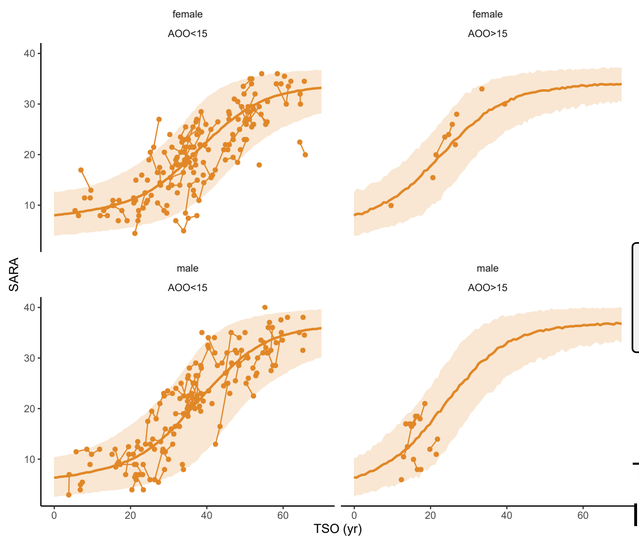
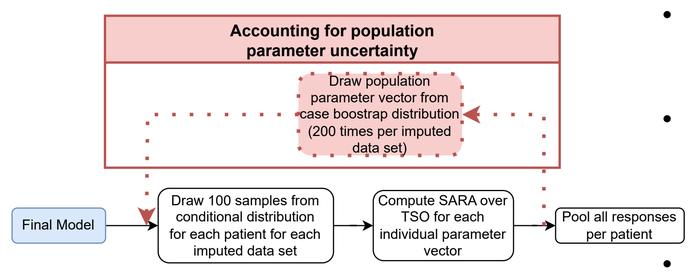


Figure 1: Spaghetti plot with simulations, median + 90% Simulation Interval, stratified by covariates

Individual predictions

Method



- Select all patients with last SARA< 20 (70)
- Compute predicted SARA and width 90% Prediction Interval (PI) at 5 years after last visit
- Compare values with and without population parameter uncertainty

Results

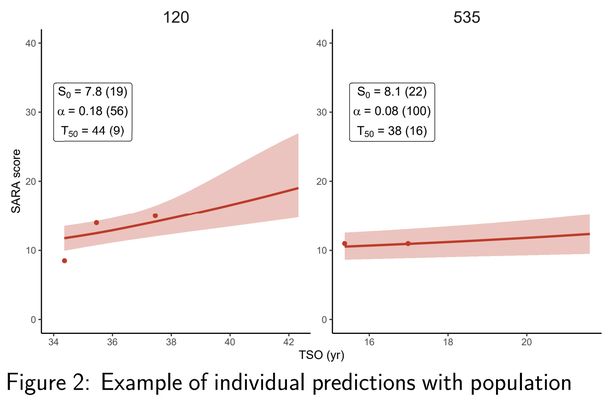


Figure 2: Example of individual predictions with population parameter uncertainty (median + 90% PI) and individual parameter estimates (estimated as the median + RSE (%) of conditional distributions)

- Parameter standard deviation depends on the number of observation (ex: patient 120) but the size of the PI is more influenced by the fact that TSO is close to the inflexion point ($T_{50} = 44$)
- Median (90% PI) increase in SARA score at 5 years of +2.9 points (0.33-5.1)

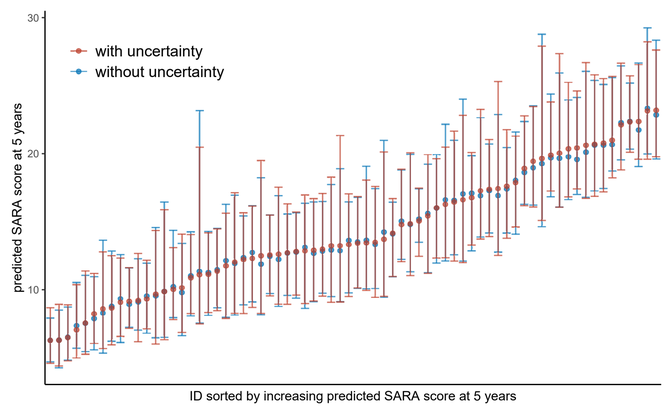


Figure 3: Distribution of the predicted SARA score and its 90% PI at 5 years, with and without population parameter uncertainty

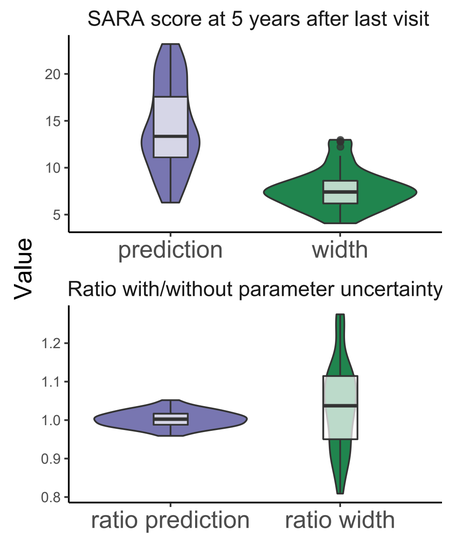


Figure 4: Violin plot of the predicted SARA score and its width at 5 years, and its ratio with/without uncertainty

- Median (90% PI) SARA score/width at 5 years of 13/7.4 (7.3-22)/(4.7-12)
- Population parameter uncertainty: little impact, likely due to the small RSE in the model

Conclusion

- Disease progression in ARSACS could be described using a logistic growth function. The score at onset of disease (S_0)/maximum SARA score (S_m) was estimated at 8.1/34.3 for females, 6.1/37.2 for males. Model predicts faster progression for late onset patients
- We used a method to select a covariate model with Multiple Imputation, accounting for missing values in covariates. In the final model, none of the selected covariates had missing values
- Parameter uncertainty did not change the predicted individual SARA score and width of PI
- Individual predictions while accounting for parameter uncertainty will be implemented in smaller populations in the registry (and most likely with higher RSE due to the small amount of patients)

[1] Synofzik et al. *J Nucleic Acid Ther* 2022; [2] Traschütz et al. *Front Neurol* 2021; [3] van Buuren et al. *J Stat Soft* 2011; [4] Meng et al. *Biometrika* 1992; [5] Comets et al. *J Stat Soft* 2017; [6] R core team et al. *R Foundation for Statistical Computing* 2022. [7] Johansson et al. *AAPS J* 2013.

Acknowledgements

This work was supported by the European Joint Programme on Rare Diseases (EJPRD) WP20 Innovation Statistics project "EVIDENCE-RND") and the EJPRD PROSPAX consortium (DFG, German Research Foundation, No 441409627), both under the EJP RD Grant Agreement (n°825575). It was further supported, in part, by the Clinician Scientist program "PRECISE.net" funded by the Else Kröner-Fresenius-Stiftung.