



Progression of fragile X-associated tremor/ataxia syndrome revealed by subtype and stage inference

Journal:	<i>Brain Communications</i>
Manuscript ID	Draft
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Morrison, Douglas; University of California Davis School of Medicine, Department of Public Health Sciences Ponzini, Matthew; University of California Davis School of Medicine, Department of Public Health Sciences Santos, Ellery; University of California Davis, MIND Institute Biag, Hazel; University of California Davis, MIND Institute; University of California Davis Medical Center, Department of Pediatrics Espinal, Glenda; University of California Davis, MIND Institute Tassone, Flora; University of California Davis, MIND Institute; University of California Davis School of Medicine, Department of Biochemistry and Molecular Medicine Rivera, Susan; University of Maryland College of Behavioral & Social Sciences, Psychology Hessl, David; University of California Davis, MIND Institute; University of California Davis Medical Center, Department of Psychiatry and Behavioral Sciences Schneider, Andrea; University of California Davis Medical Center, MIND Institute Bourgeois, James; University of California Davis Medical Center, Department of Psychiatry and Behavioral Sciences Hagerman, Randi; University of California Davis, MIND Institute; University of California Davis Medical Center, Department of Pediatrics Kim, Kyoungmi; University of California Davis School of Medicine, Department of Public Health Sciences
Keywords:	fragile X-associated tremor/ataxia syndrome, subtype and stage inference, sequential events, disease development and progression, sex differences, latent subgroups

SCHOLARONE™
Manuscripts

Progression of fragile X-associated tremor/ataxia syndrome revealed by subtype and stage inference

Douglas E. Morrison¹, Matthew Ponzini¹, Ellery R. Santos², Hazel Maridith Barlahan Biag^{2,6}, Glenda Espinal², Flora Tassone^{2,4}, Susan M. Rivera³, David Hessl^{2,5}, Andrea Schneider², James A. Bourgeois⁵, Randi Hagerman^{2,6}, and Kyoungmi Kim¹

Abstract

The fragile X-associated tremor/ataxia syndrome (FXTAS) is a neurodegenerative disorder caused by the premutation (55 to 200 CGG repeats) in the fragile X messenger ribonucleoprotein-1 (*FMR1*) gene. An open question is: In what sequential order do FXTAS symptoms typically appear, and how does that sequence vary among patients and between males and females?

We applied the ordinal-outcomes version of the Subtype and Stage Inference algorithm (“Ordinal SuStaIn”) to identify the sequential events of clinical and brain MRI changes in cross-sectional data collected during baseline visits from a longitudinal cohort of FXTAS patients at stages 0-5. We included 31 neurodegenerative symptoms collected from 253 premutation carriers (101 females and 152 males) and 44 controls (7 females and 37 males), aged 40 - 86 years old at entry, who participated in two longitudinal studies, with entry dates between 2008 - 2023.

We found substantial differences in order of events depending on sex, and possibly in combination of sex and CGG repeats. The main finding is the predominance of the psychiatric co-morbidities that occur early in females (often before the onset of tremor and ataxia) compared to males. Females appear to experience early cognitive problems involving executive dysfunction and memory in comparison to males.

These findings suggest that the sequence of neuropsychiatric symptoms for FXTAS is different in females compared to males, particularly for early symptoms in disease development and progression. This could lead to sex-specific modifications of the FXTAS diagnostic stages.

Author affiliations:

1 Department of Public Health Sciences, School of Medicine, University of California, Davis,
2 Davis CA, United States

3 MIND Institute, University of California, Davis, Sacramento CA, United States

4 Department of Psychology, College of Behavioral and Social Sciences, University of
5 Maryland, College Park, MD, United States

6 Department of Biochemistry and Molecular Medicine, School of Medicine, University of
7 California, Davis, Sacramento CA, United States

8 Department of Psychiatry and Behavioral Sciences, University of California, Davis Medical
9 Center, Sacramento CA, United States

10 Department of Pediatrics, University of California, Davis Medical Center, Sacramento CA,
11 United States

12 Correspondence to:
13 Douglas Ezra Morrison
14 Medical Sciences 1C #113
15 One Shields Ave
16 Davis, CA 95616
17 demorrison@ucdavis.edu

18 **Running title:** Progression of FXTAS symptoms

19 **Keywords:** fragile X-associated tremor/ataxia syndrome; subtype and stage inference;
20 sequential events; disease development and progression; sex differences; latent subgroups

Introduction

The fragile X-associated tremor/ataxia syndrome (FXTAS) is a neurodegenerative disorder caused by the premutation (55 to 200 CGG repeats) in the fragile X messenger ribonucleoprotein-1 (*FMRI*) gene.¹ The premutation is associated with elevation of the *FMRI* mRNA. The RNA toxicity due to the elevated expression levels of the *FMRI* mRNA leads to oxidative stress, mitochondrial dysfunction, calcium dysregulation, and sequestration of proteins important to neuronal function.²⁻¹² Eosinophilic, tau-negative intranuclear inclusions are present in neurons and astrocytes throughout the CNS and in the peripheral nervous system (PNS).^{13,14} FXTAS is a late-onset disorder, usually occurring in the 60s though some patients have an earlier onset, and its signs and symptoms worsen with age. As premutation carriers age, the prevalence of FXTAS increases from approximately 40% in males in their 60s to 75% in their 80s.¹⁵ Females with the premutation are also at risk for FXTAS, but the prevalence is lower; approximately 20% may develop FXTAS, and it is usually less severe than in males.¹⁶ In general for males with FXTAS, the higher the CGG repeat number within the premutation range, the earlier onset of FXTAS, and earlier death.¹³ The prevalence of the premutation is estimated at one in 148 to 200 females and one in 290 to 855 males in the general population.^{17,18}

The symptoms of FXTAS appear and develop over many years in adulthood, usually after age 50, and include tremor, gait ataxia, neuropathy, cognitive decline, and depressive and anxiety disorders.¹ Both the type and severity of FXTAS symptoms vary among patients. Some patients with FXTAS have multiple symptoms that progress rapidly in severity, while others have few symptoms that remain mild over many years. The sequence in which those symptoms typically appear is still understudied, but it could help us to understand the disease process and its variability across patients. Therefore, the present study attempts to address the following questions: 1) In what sequential order do FXTAS symptoms and brain changes typically appear? 2) How does that sequence vary among patients? and 3) Do specific baseline characteristics predict the order of presentation?

Here, we applied the ordinal-outcomes version of the Subtype and Stage Inference algorithm (“Ordinal SuStaIn”)^{19,20} to model sequential orders of clinically manifest FXTAS symptoms in a cohort of FXTAS patients. SuStaIn is a data-driven analytic approach which was originally developed to analyze cross-sectional biomarker data to simultaneously (1) estimate the ordering of onset for biomarkers of disease progression events and (2) cluster study participants

1
2
3 into latent subtypes with different event orderings. The SuStaIn modeling approach has been
4 applied to Alzheimer's disease,^{19–22} amyotrophic lateral sclerosis,²³ Parkinson's disease,^{24,25}
5 primary tauopathies,²⁶ frontotemporal dementia,¹⁹ multiple sclerosis,²⁷ chronic obstructive
6 pulmonary disease (COPD),²⁸ and knee osteoarthritis,²⁹ but not previously for FXTAS. In this
7 study we applied it to a cross-sectional data set of premutation carriers diagnosed with FXTAS
8 at different clinical stages to study the sequential ordering of FXTAS symptomatic events.
9
10
11
12
13
14
15

Materials and methods

Data

Study cohorts

Research participants were recruited between 2008 and 2023 for two longitudinal cohorts: the Genotype-Phenotype Study in Fragile X Families (“GP”; funded by NICHD R01HD036071; PI Hagerman) and the Trajectories and Markers of Neurodegeneration in Fragile X Premutation Carriers study (“TRAX”; funded by NINDS NS110100; PIs Hessl and Rivera) conducted at the University of California, Davis MIND Institute. A written informed consent form was obtained from all participants according to the procedures approved by the University of California, Davis Institutional Review Board.

The GP longitudinal study is an ongoing effort focused on adult male and female carriers of the *FMR1* premutation who have neurological symptoms or have been diagnosed with FXTAS and healthy control individuals without the fragile X premutation, aged 40-85 at entry, with follow-up visits approximately every two years. The TRAX longitudinal study is an ongoing effort studying adult male premutation carriers ranging from 40-82 years and male healthy controls ages 40-75 at baseline visit. The TRAX participants return for follow up visits with varying intervals between visits averaging 2.5 years to assess phenotypic progression over time.

In both studies, each evaluation includes a detailed medical history, neurological examination, neuropsychological testing including the Wechsler Adult Intelligence Scale, Fourth Edition (WAIS IV), Behavior Dyscontrol Scale (BDS-2), Mini-Mental State Exam (MMSE) and Cambridge Automated Neuropsychological Test Battery (CANTAB); Structured Clinical Interview for DSM-IV Disorders (SCID-I/NP), motor testing, and brain MRI.^{30–32} After the

1
2
3 clinical and MRI evaluation, each patient received a stage designation of FXTAS ranging from
4 0 to 6 according to stages identified by Bacalman *et al.* 2006³³ based on tremor and ataxia
5 severity. Stage 1 represents subclinical or uncertain tremor; Stage 2 is mild tremor without
6 significant interference with activities of daily living (ADLs); Stage 3 is significant tremor that
7 interferes with ADLs and significant ataxia; Stage 4 is significant ataxia needing a cane or
8 walker; Stage 5 is requiring a wheelchair; and Stage 6 is bedridden (no participants at Stage 6
9 were included in the study).

10
11
12
13
14
15 The primary sequence analysis of symptomatic events included baseline visit data from 297
16 participants from the GP and TRAX cohorts, consisting of 253 premutation carriers and 44
17 controls. [Table 1](#) provides demographic information about the study participants. Data from
18 unaffected control participants were used to estimate the distribution of the analyzed symptoms
19 among non-FXTAS individuals as reference in SuStIn modeling (more details in section
20 [Ordinal SuStIn model](#)). Then the cases' data were used to estimate FXTAS event sequences,
21 using the estimated control distributions to account for random variation in observed symptoms
22 unrelated to the underlying event sequence.

31 Symptoms of neurodegenerative events

32
33 We analyzed measurements of 31 ordinal symptoms, each with between two and six levels,
34 listed in [Table 2](#). Herein the term "symptoms" refers to a broad range of medical signs or
35 indications of medical state observed from patients. Each "clinically elevated" ordinal level
36 (above the first-listed, reference level) constitutes an outcome event in the disease progression
37 modelling analysis (see [Statistical analysis](#)).

38
39
40 We created a composite variable named "any autoimmune disorder" combining systemic lupus
41 erythematosus, rheumatoid arthritis, multiple sclerosis, anti-nuclear antibody (ANA) positivity,
42 Sjogren's syndrome, Raynaud's syndrome, and pulmonary fibrosis, since these conditions
43 were too rare to analyze separately (details in Supplementary material, [Any autoimmune](#)
44 [disorder](#)). We also created composite variables for four domains of the Structured Clinical
45 Interview for DSM-IV (SCID-IV): mood disorders, substance use disorders, anxiety disorders,
46 and somatoform disorders; there were no participants with psychotic disorders in our data
47 (more details in Supplementary material, [SCID composite variables](#)). We also created
48 composite variables for MRI variables for cerebral and cerebellar abnormalities; we did not
49 combine the variables representing corpus callosum MRI abnormalities, since these variables
50 were created using Likert scales that differed from each other (details in Supplementary
51 material).
52
53
54
55
56
57
58
59
60

material, [MRI variables](#)). For the Cambridge Neuropsychological Test Automated Battery (CANTAB) variables, we used categorization cutoffs taken from Talebi et al 2020.³⁴

Supplementary Tables [1](#) - [11](#) summarize each of the symptoms that we included in the analysis models, stratified by CGG repeat size (CGG <55, CGG 55-99, CGG 100-199) and sex.

Statistical analysis

Ordinal SuStaIn model

We applied a discrete event-based disease progression model³⁵⁻³⁷ to our data using the Ordinal Subtype and Stage Inference (“SuStaIn”) algorithm²⁰ to estimate event orderings and subtypes for FXTAS patients. SuStaIn is a data-driven statistical modeling algorithm that combines event-based disease progression modelling³⁵⁻³⁷ and latent-cluster finite mixture modeling³⁸⁻⁴⁰ to model event sequences using cross-sectional samples of patient and control populations. The algorithm simultaneously clusters individuals into latent subtypes and characterizes the event ordering that best defines each subtype, thus capturing heterogeneity in both disease subtype and disease stage.

Ordinal SuStaIn²⁰ is a version of the SuStaIn modeling algorithm,¹⁹ adapted for analyzing ordinal-valued data. Ordinal SuStaIn uses a “scored events model”, which assumes that for each symptom, there is a discrete set of underlying ordinal severity levels, but the measured versions of the symptoms may contain some amount of random noise. For example, a patient who was really at ataxia severity level 2 may be incorrectly assessed as being at ataxia severity level 1, depending on the patient’s temporary disease status on the day of the exam or inter-rater differences. The first step in applying the Ordinal SuStaIn algorithm is to determine, for each symptom, the probability that an individual is “correctly scored” at their “true underlying level.”²⁰ In this analysis, we assumed that all controls are truly at the reference levels for each symptom, and we estimated the percentage of correctly scored individuals as the percentae of controls who were assessed as being at the reference level ([Supplementary Table 12](#)). It is crucial to allow some possibility of incorrect scoring, so we capped the estimated correct scoring probabilities at 95%.

Ordinal SuStaIn then uses Markov Chain Monte Carlo (MCMC) sampling⁴¹ to estimate the Bayesian posterior probability of each possible event sequence for each subtype given the training dataset, assuming a uniform prior distribution over the set of all possible orderings.

We conducted subgroup analyses by fitting models stratified by sex ([Sex differences in sequential order](#)), CGG repeats (<100 vs ≥ 100) (Supplementary material, [Analyses stratified by CGG repeats](#)), and combinations of sex and CGG repeats (<100 vs ≥ 100) (Supplementary material, [Comparing sexes stratified by CGG level](#)). In these subgroup analyses, we did not search for latent clusters.

Imputation of missing data

We assumed that missing symptom data were missing at random (MAR).⁴³ As the longitudinal cohorts evolved over time, new instruments were adopted and added; much of the missingness was due to adding additional measures to the study protocols in later time. The MAR assumption thus seems plausible. We substituted missing outcome event data by assigning a probability distribution across the possible values of the missing variable that matched the marginal distribution of observed data among the cases. For example, 33 of the 253 cases (13.0%) had missing values for ataxia severity, and 220 (87.0%) had recorded values, distributed among ataxia severity levels 0-4 (139 (63.2%), 37 (16.8%), 20 (9.1%), 16 (7.3%), and 8 (3.6%), respectively). For the 33 cases with missing values, we assigned probabilities of 63.2%, 16.8%, 9.1%, 7.3%, and 3.6% to ataxia severity levels 0, 1, 2, 3, and 4, respectively, and missing values were imputed under the marginal distribution of the observed data.

Statistical hypothesis tests

To test for statistical significant evidence of differences in event sequences between males and females and between lower (CGG <100) and higher premutation levels (CGG 100-199), we implemented a permutation test⁴⁴ to calculate a p-value that is the probability of observing a difference at least as extreme as the test statistic given that the null hypothesis of no difference is true. We first created 1000 permuted datasets in which we randomly shuffled the variable being tested. We computed the mean log-likelihood of the data for each permuted dataset (averaging across MCMC samples and summing across the strata being compared) and compared the distribution of permuted mean log-likelihoods to the observed log-likelihood calculated from the original (unpermuted) dataset. We computed the empirical p-value by first computing the percentile of the observed mean log-likelihood relative to the empirical distribution of the permuted mean log-likelihoods, subtracting that percentile from 1 if larger than 0.5, and then multiplying by two to calculate a two-sided test statistic. We declared significance if the p-value was less than or equal to 0.05.

Latent subtype clustering

We also fitted the model on the full dataset (not stratified by sex or CGG) for 2 - 8 latent subtypes, each with their own ordering. We determined the optimal number of latent subtypes for this dataset using the Cross-Validation Information Criterion (CVIC) and the Out-Of-Fold Log-Likelihood (OOFL) criterion.¹⁹ both criteria quantify how well models containing a given number of subtypes extend to new data not used in training. We performed 10-fold cross-validation on the unstratified data and calculated the CVIC and OOFL for 1-8 latent subtypes.

Visualizing modeling results

We visualized the results of Ordinal SuStaIn analysis using “positional variance diagrams” (PVDs). PVDs are heatmaps with symptomatic events on the y-axis and sequence positions of events on the x-axis. Each event is estimated to occur at some point in continuing time *relative to* other events in sequence. It should be noted that the exact onset time of each event cannot be determined by the method we used for this analysis. The PVD’s color scale indicates the Bayesian posterior probability that a particular event (y-axis) appears at a particular position along the progression sequence (x-axis). The different colors (red, blue, purple, green, magenta) indicate the ordinal levels of symptom progression. Color intensity presents the likelihood of sequence position. That is, a brighter color indicates a more probable sequence position, and a paler color indicates a less probable position.

Software

The Ordinal Sustain analysis was performed using the Python programming language, version 3.9⁴⁵ with the pySuStaIn package.⁴⁶ Data pre-processing and results post-processing were performed in R version 4.4.2 (2024-10-31)⁴⁷ using the tidyverse packages.⁴⁸ The reticulate package⁴⁹ was used to create an application programming interface between Python and R. The code used to perform this analysis is available at <https://github.com/UCD-IDDRC/fxtas>.

Results

[Table 1](#) describes summary statistics of patient characteristics included in the analysis. Our data included: 253 fragile X premutation carriers and 44 controls; 189 males and 108 females; and 60, 34, 43, 61, 21, and 9 carriers at FXTAS stages 0, 1, 2, 3, 4, and 5, respectively. We found no differences between males and females in age at baseline visit, ethnicity/race, FXTAS

stage, or CGG repeat length. [Table 2](#) lists the symptoms and their ordinal levels and reports percentages of clinically elevated levels at baseline visit, stratified by sex. We found significant sex differences in percentages of clinically elevated (non-reference) symptom levels at baseline visit for “the CANTAB subtest of Spatial Working Memory (SWM) Between errors” (10% of females versus 38.2% of males, p-value = 0.020), “parkinsonian increased tone” (6.5% of females versus 21.2% of males, p-value = 0.034), “parkinsonian features” (14.3% of females versus 32.8% of males, p-value = 0.028), “SCID: anxiety disorders” (87.5% of females versus 61.5% of males, p-value < 0.001), “SCID: mood disorders” (66.7% of females versus 40.6% of males, p-value < 0.001), “SCID: somatoform disorders” (15.5% of females versus 2.1% of males, p-value < 0.001), “Mini-Mental State Examination (MMSE) total score < 21 (mild impairment or worse)” (5.3% of females versus 20% of males, p-value = 0.010), “any autoimmune disorder” (20.7% of females versus 7.7% of males, p-value = 0.004), and “head tremor” (41% of females versus 21.3% of males, p-value = 0.043).

Sex differences in sequential orders

[Figure 1](#) shows the estimated sequential orders of FXTAS symptoms in males and females, and [Figure 2](#) shows differences in those orders. We found statistically significant evidence of a difference in sequences of symptomatic events between males and females ($p = 0.018$). It appears that female premutation carriers developed SCID mood disorder symptoms, both sub-threshold-level and threshold-level, prior to FXTAS Stage 1 diagnosis, whereas male carriers experienced these symptoms much later (often between FXTAS Stages 3 and 4) ([Figure 2](#)). Ataxia occurred prior to Stage 1 for females and after Stage 2 for males. However, ataxia severity level 1 occurred earlier for males (between Stages 2-3) than for females (between Stages 3-4). Advanced degrees of ataxia severity also occurred earlier in the event sequence for males than females. MRI biomarkers of decline and parkinsonian features also appear to begin at earlier FXTAS stages for males than for females.

When we conducted subgroup analyses by CGG repeat size, we found statistically significant evidence of a difference between males and females among those with $\text{CGG} < 100$ ($p = 0.05$) ([Supplementary Fig. 7](#)). Several psychiatric disorders (as assessed by SCID) occurred prior to FXTAS Stage 1 among females with $\text{CGG} < 100$, but at later FXTAS stages for males with $\text{CGG} < 100$. MRI biomarkers appeared to occur at later stages in females compared to males. We did not find statistically significant evidence of a difference between males and females among those with $\text{CGG} \geq 100$ ($p = 0.202$) ([Supplementary Fig. 8](#)).

We found significant evidence of an overall difference between CGG <100 and CGG \geq 100 among males ($p = 0.008$) ([Supplementary Fig. 9](#)). SWM between errors occurred prior to Stage 1 in males with CGG repeats <100, whereas it occurred between stage 3 and 4 in males with CGG repeats \geq 100 ([Supplementary Fig. 9 \(b\)](#)). Postural, resting, and intermittent tremor occurred in earlier FXTAS stages in those with CGG repeats <100 compared to those with CGG repeats \geq 100. Several psychiatric disorders (as measured by SCID) occurred later in the event sequence in participants with CGG repeats <100 compared to those with CGG repeats \geq 100.

However, we did not find a significant difference between CGG <100 and CGG \geq 100 among females ($p = 0.596$) ([Supplementary Fig. 10](#)). It should be noted that this comparison lacked statistical power given the limited sample size of only 25 females with CGG \geq 100.

Subtype clustering

We conducted latent subtype classification analysis to cluster participants that are relatively homogeneous within the same cluster and heterogeneous from other clusters based on similarities and difference in sequential patterns. Based on the CVIC criterion (Supplementary material, [Detecting latent subtypes](#)), 8 subtypes were suggested to be the optimal number for the full, unstratified dataset ([Supplementary Fig. 2 \(a\)](#)). The out-of-fold log-likelihood (“OOFL”) (Supplementary material, [Detecting latent subtypes](#)) showed substantial fold-to-fold variation ([Supplementary Fig. 2 \(b\)](#)). Between four and eight latent subgroups, the distribution of OOFL appears to be approximately unchanging and the CVIC appears approximately flat, and thus for easier clinical interpretation, we chose to classify participants into 4 subtypes. The subtypes are numbered 1-4 in order of estimated relative frequency (that is, in order of how many participants were clustered into each subtype).

[Figure 3](#) shows the estimated sequential event orders for each of the subtype clusters. [Supplementary Fig. 4](#) shows differences in event sequences between subtypes. [Table 3](#) shows the demographics of the patients clustered in each subtype. Thirty-six patients had experienced too few events to be accurately classified into a subtype and were excluded from [Table 3](#). Subtype clustering analysis showed that subtypes were mainly classified by sex and CGG repeat size, supporting the results above. Subtype 1 appeared more biased towards males (65.1%) and lower CGG repeat size (83.2 ± 15.8) compared to the overall study population (64.5% of males; CGG repeat size = 87.3 ± 18.9) and other subtypes; Subtype 2 appeared biased towards females (53.1% of males) and higher CGG repeat size (90.2 ± 22.3); Subtype

3 appeared over-represented by males (82% of males) and high CGG repeat size (92.4 ± 17.4);
and Subtype 4 was biased towards females (60% of males) and lower CGG repeat (82.7 ± 17.3).
Using the FXTAS Stage transitions as reference points, we examined the differences in the
appearing symptoms relative to these FXTAS Stage transitions. Roughly speaking: Subtype 1
appeared to experience early mood disorders (before FXTAS Stage 1), ataxia between Stages
2 and 3, head tremors between Stages 3 and 4, parkinsonian features between Stages 3 and 4,
and white matter disease between Stages 4 and 5. Compared to Subtype 1, Subtype 2 appeared
to experience white matter disease phenotypes earlier (between Stages 2 and 4), memory
impairment (assessed by SWM) earlier (Stage 3-4 versus 4-5), and psychiatric disorders later
(between Stages 3 and 5). Subtype 2 might be described as “memory-first”. Subtype 3 appeared
to experience white matter disease phenotypes early (between Stages 1 and 3), and mood
disorders late (Stages 4-5). Subtype 4 had the fewest participants and was difficult to interpret.
It may represent a mixture of smaller latent subtypes and/or include outlier individuals with
unusual event sequences.

Discussion

This study attempted to arrange the eventual symptoms of FXTAS in their possible appearing sequential orders and determine risk factors that impact on the sequential order of presentation among patients. The main findings of this analysis were substantial differences in sequential order of eventual symptoms between males and females and in combination of sex and CGG repeat size. An important general conclusion of the study is that there is substantial variation across premutation carriers in the sequence of emergence of symptoms associated with FXTAS.

The prevalences of lifetime anxiety disorders, mood disorders, and somatoform disorders in this study cohort are higher than usual population norms for these illnesses as noted in previous studies.⁵⁰⁻⁵³ This study design did not fractionate specific lifetime psychiatric illnesses (e.g., major depressive disorder, panic disorder, somatization disorder), but these *overall* lifetime prevalence figures are notable. Prior studies of mood and anxiety disorders have demonstrated a higher-than-expected rate of these illnesses in premutation carriers.^{54,55} The same increase in lifetime psychiatric illness burden is seen in other neurodegenerative illnesses (e.g., Parkinson’s disease, multiple sclerosis).⁵⁶⁻⁶⁴ As many psychiatric illnesses have a strong

1
2
3 genetic predisposition, the lifetime prevalences of psychiatric illness in this study persuasively
4 suggest that the premutation carrier state *itself* increases the risk of psychiatric illness.
5
6

7 As the SCID assesses *lifetime* illness risk (as well as current illness), it is probable that the
8 index episode of psychiatric illness *antedated* the onset of tremor, ataxia, and major
9 neurocognitive disorder seen in more advanced premutation conditions like FXTAS. As such,
10 the high lifetime burden of psychiatric illness in this population cannot be plausibly attributed
11 solely (perhaps substantially) to an “emotional reaction” to the loss of function associated with
12 these functional impairments, but rather should be understood as at least partially a component
13 of the premutation carrier state *itself*. It is likely that this intrinsic vulnerability to psychiatric
14 illness is a significant component of the carrier state.
15
16

17 Furthermore, as only a small proportion of the patients had MMSE scores in the impaired range,
18 it is also unlikely that the high rates of depressive, anxiety (and other) psychiatric illness could
19 be attributed *solely* to complications of FXTAS dementia, as most patients had normal MMSE
20 scores. For example, depressive disorders are a common psychiatric complication of dementia,
21 but few of the depressive disorders found here could be attributed to dementia, as they
22 antedated any evidence of cognitive disorders. A common sequence of illnesses seen clinically
23 in carriers is depressive and/or anxiety disorders with young to middle age adult onset, then
24 tremor and ataxia in the 40s or 50s, followed by FXTAS dementia in the 60s or later.^{55,65} Seen
25 in this light, and knowing the ultimate CNS vulnerability of the carrier state, retrospectively
26 the initial presentation of depressive and/or anxiety disorders could be regarded as a psychiatric
27 prodrome of later full spectrum illness. This is analogous to the depressive or psychotic illness
28 seen in SLE before rheumatologic findings and the depressive disorder that is commonly seen
29 in the five years before the motor signs leading to a diagnosis of multiple sclerosis.
30
31

32 This data shows more significant motor involvement in males than females, as predicted, but a
33 remarkable finding is the predominance of the psychiatric illnesses that occur early in females
34 and often before the onset of tremor and ataxia compared to the males. These symptoms,
35 particularly mood disorders, can often occur even before the diagnosis of FXTAS in the
36 females, and the stress involved with mood disorders may be a precipitating feature for the
37 onset of FXTAS. In addition, females may experience early cognitive impairment, particularly
38 involving executive dysfunction and memory. From a clinical perspective, we have seen
39 significant psychiatric illness comorbid with some cognitive deficits in memory and EF
40 abilities, but these females may not meet the criteria of FXTAS diagnosis, because of the
41
42

absence of significant tremor and/or ataxia. However, if white matter disease emerges on MRI, particularly in the splenium of the corpus callosum or in the periventricular area, then this demonstrates the onset of the neuropathology of FXTAS, even if there is no tremor or ataxia yet.

This suggests that the diagnostic criteria for FXTAS may need to be somewhat altered in females compared to males, particularly for early cases. Indeed, the diagnostic criteria for FXTAS were determined with the study of males only.⁶⁶ Subsequent criteria for the diagnosis of FXTAS were modified⁶⁷ to include the involvement of the splenium, which is seen in the majority of females with FXTAS.^{53,68} Hall *et al.*⁶⁷ also added the feature of neuropathy, which occurs in the majority of FXTAS patients, but it is a common finding in the elderly with many causes, so it is just a minor criterion for FXTAS. Although the diagnostic criteria have been modified somewhat for the females, there may be subsequent changes in the early stages of FXTAS as further follow-up studies are able to separate the psychiatric illness in females seen earlier here from more significant neuropathology associated with extended features of FXTAS over time.^{69,70} As new treatments for FXTAS develop, it is hoped that the earlier the diagnosis is established, the more likely the treatment will be effective.

When we compared between 55-99 vs 100-200 CGG repeats in males and females combined, no significant difference was found ($p = 0.176$) ([Supplementary Fig. 5](#)). The finding of only minimal differences between 55-99 vs 100-200 CGG repeats is surprising, since several studies have found that the higher the CGG repeats, the earlier the onset and the faster the progression of FXTAS.^{13,16,71} Perhaps the cut-off of 100 is too high for this distinction to be made.

This study has several limitations. In the subtype clustering analysis, it was challenging to concisely summarize the differences between subtypes, especially given the relatively small sample sizes per subtype and the correspondingly large amounts of uncertainty in their event orderings. We had substantial amounts of missing data. As described in the methods ([Imputation of missing data](#)), we used the marginal distribution of the observed data for each symptom among cases to impute the underlying values of these missing data. The missing data could contribute to uncertainty in the results. Our data contained both continuous and ordinal variables. In order to apply the Ordinal SuStaIn algorithm, we categorized some symptoms that were originally measured as continuous values. In doing so, we likely sacrificed some granular information. There are other variations of the SuStaIn algorithm, such as z-score SuStaIn,¹⁹ which are designed to be used with only continuous measurements. Further development is

warranted to combine Ordinal SuStaIn and z-score SuStaIn to fit an event-based model with both continuous and ordinal data.

Our data come from the GP and TRAX studies, which are longitudinal cohort studies with infrequent follow-up visits. These studies do not collect precise information about timing of symptom onsets, so it is not feasible to use the time-to-event data to confirm the results from the Ordinal SuStaIn analysis. In our analysis approach, the event onsets are not modeled as a function of participant age; instead, each event's onset timing is only modeled relative to the other events. Therefore, the sequence differences between the subtypes are all in relative terms: if one event moves earlier in the sequence when comparing one subtype to another, other events are pushed later in the sequence, even if only one event timing changes relative to age. Additional research including longitudinal data collecting precise ages of onset would be needed to verify our findings.

The data used in this analysis came from patients and controls who participated in longitudinal studies at the UC Davis MIND Institute and may not be directly generalizable to other diverse populations. The participants predominantly reported their primary race/ethnicity as "White", and we did not have data from enough patients with other diverse racial/ethnic identities to accurately estimate stratified models by race/ethnicity. A larger study including a more diverse set of participants is warranted to support analyses with increased external validity.

Data availability

The de-identified data used in this analysis may be provided upon request from the principal investigators of the two cohorts.

Funding

This study was supported by the National Institute of Child Health and Human Development (NICHD, HD036071) and UC Davis MIND Institute's Intellectual and Developmental Disabilities Research Center (IDDRC) (P50HD103256) and the National Institute of Neurological Disorders and Stroke (NINDS, NS110100).

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at Brain online.

References

1. Hagerman RJ, Leehey M, Heinrichs W, et al. Intention tremor, parkinsonism, and
2. generalized brain atrophy in male carriers of fragile X. *Neurology*. 2001;57(1):127-130.
3. Napoli E, Ross-Inta C, Wong S, et al. Altered zinc transport disrupts mitochondrial
4. protein processing/import in fragile X-associated tremor/ataxia syndrome. *Human molecular
5. genetics*. 2011;20(15):3079-3092. doi:[10.1093/hmg/ddr211](https://doi.org/10.1093/hmg/ddr211)
6. Song G, Napoli E, Wong S, et al. Altered redox mitochondrial biology in the
7. neurodegenerative disorder fragile X-tremor/ataxia syndrome: Use of antioxidants in precision
8. medicine. *Molecular Medicine*. 2016;22:548-559. doi:[10.2119/molmed.2016.00122](https://doi.org/10.2119/molmed.2016.00122)
9. Napoli E, Song G, Wong S, Hagerman R, Giulivi C. Altered bioenergetics in primary
10. dermal fibroblasts from adult carriers of the FMR1 premutation before the onset of the
11. neurodegenerative disease fragile X-associated tremor/ataxia syndrome. *The Cerebellum*.
12. 2016;15(5):552-564. doi:[10.1007/s12311-016-0779-8](https://doi.org/10.1007/s12311-016-0779-8)
13. Giulivi C, Napoli E, Tassone F, Halmai J, Hagerman R. Plasma metabolic profile
14. delineates roles for neurodegeneration, pro-inflammatory damage and mitochondrial
15. dysfunction in the FMR1 premutation. *Biochemical Journal*. 2016;473(21):3871-3888.
16. doi:[10.1042/BCJ20160585](https://doi.org/10.1042/BCJ20160585)
17. Wang JY, Napoli E, Kim K, McLennan YA, Hagerman RJ, Giulivi C. Brain atrophy
18. and white matter damage linked to peripheral bioenergetic deficits in the neurodegenerative
19. disease FXTAS. *International Journal of Molecular Sciences*. 2021;22(17).
20. doi:[10.3390/ijms22179171](https://doi.org/10.3390/ijms22179171)

- 1
2
3 7. Alvarez-Mora M, Rodriguez-Revenga L, Madrigal I, Guitart-Mampel M, Garrabou G,
4 Mila M. Impaired mitochondrial function and dynamics in the pathogenesis of FXTAS.
5 *Molecular Neurobiology*. 2017;54:6896-6902. doi:[10.1007/s12035-016-0194-7](https://doi.org/10.1007/s12035-016-0194-7)
- 6
7 8. Alvarez-Mora MI, Podlesniy P, Gelpi E, et al. Fragile X-associated tremor/ataxia
8 syndrome: Regional decrease of mitochondrial DNA copy number relates to clinical
9 manifestations. *Genes, Brain and Behavior*. 2019;18(5):e12565. doi:[10.1111/gbb.12565](https://doi.org/10.1111/gbb.12565)
- 10
11 9. Gohel D, Sripada L, Prajapati P, et al. FMRpolyG alters mitochondrial transcripts level
12 and respiratory chain complex assembly in fragile X associated tremor/ataxia syndrome
13 [FXTAS]. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*.
14 2019;1865(6):1379-1388. doi:[10.1016/j.bbadi.2019.02.010](https://doi.org/10.1016/j.bbadi.2019.02.010)
- 15
16 10. Gohel D, Berguerand NC, Tassone F, Singh R. The emerging molecular mechanisms
17 for mitochondrial dysfunctions in FXTAS. *Biochimica et Biophysica Acta (BBA)-Molecular
18 Basis of Disease*. 2020;1866(12):165918.
- 19
20 11. Gohel D, Sripada L, Prajapati P, et al. Expression of expanded FMR1-CGG repeats
21 alters mitochondrial miRNAs and modulates mitochondrial functions and cell death in cellular
22 model of FXTAS. *Free Radical Biology and Medicine*. 2021;165:100-110.
- 23
24 12. Robin G, López JR, Espinal GM, Hulsizer S, Hagerman PJ, Pessah IN. Calcium
25 dysregulation and Cdk5-ATM pathway involved in a mouse model of fragile X-associated
26 tremor/ataxia syndrome. *Human Molecular Genetics*. 2017;26(14):2649-2666.
27 doi:[10.1093/hmg/ddx148](https://doi.org/10.1093/hmg/ddx148)
- 28
29 13. Greco CM, Berman RF, Martin RM, et al. Neuropathology of fragile X-associated
30 tremor/ataxia syndrome (FXTAS). *Brain*. 2006;129(1):243-255. doi:[10.1093/brain/awh683](https://doi.org/10.1093/brain/awh683)
- 31
32 14. Hunter JE, Sherman S, Grigsby J, Kogan C, Cornish K. Capturing the fragile X
33 premutation phenotypes: A collaborative effort across multiple cohorts. *Neuropsychology*.
34 2012;26(2):156.
- 35
36 15. Jacquemont S, Hagerman RJ, Leehey MA, et al. Penetrance of the fragile X-associated
37 tremor/ataxia syndrome in a premutation carrier population. *JAMA*. 2004;291(4):460-469.
38 doi:[10.1001/jama.291.4.460](https://doi.org/10.1001/jama.291.4.460)
- 39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 16. Tassone F, Protic D, Allen EG, et al. Insight and recommendations for fragile X-
4 premutation-associated conditions from the fifth international conference on FMR1
5 premutation. *Cells*. 2023;12(18):2330.
- 6
7 17. Maenner MJ, Baker MW, Broman KW, et al. FMR1 CGG expansions: Prevalence and
8 sex ratios. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*.
9 2013;162(5):466-473.
- 10
11 18. Hunter J, Rivero-Arias O, Angelov A, Kim E, Fotheringham I, Leal J. Epidemiology
12 of fragile X syndrome: A systematic review and meta-analysis. *American journal of medical*
13 *genetics Part A*. 2014;164(7):1648-1658.
- 14
15 19. Young AL, Marinescu RV, Oxtoby NP, et al. Uncovering the heterogeneity and
16 temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. *Nature*
17 *communications*. 2018;9(1):4273. doi:[10.1038/s41467-018-05892-0](https://doi.org/10.1038/s41467-018-05892-0)
- 18
19 20. Young AL, Vogel JW, Aksman LM, et al. Ordinal SuStIn: Subtype and stage inference
21 for clinical scores, visual ratings, and other ordinal data. *Frontiers in artificial intelligence*.
22 2021;4:613261. <https://doi.org/10.3389/frai.2021.613261>
- 23
24 21. Vogel JW, Young AL, Oxtoby NP, et al. Four distinct trajectories of tau deposition
25 identified in Alzheimer's disease. *Nature Medicine*. 2021;27:871-881. doi:[10.1038/s41591-021-01309-6](https://doi.org/10.1038/s41591-021-01309-6)
- 26
27 22. Collij LE, Salvadó G, Wottschel V, et al. Spatial-temporal patterns of β -amyloid
28 accumulation: A subtype and stage inference model analysis. *Neurology*. 2022;98(17):e1692-
29 e1703. doi:[10.1212/WNL.0000000000200148](https://doi.org/10.1212/WNL.0000000000200148)
- 30
31 23. Young AL, Vogel JW, Robinson JL, et al. Data-driven neuropathological staging and
32 subtyping of TDP-43 proteinopathies. *Brain*. 2023;146(7):2975-2988.
- 33
34 24. Zhou C, Wang L, Cheng W, et al. Two distinct trajectories of clinical and
35 neurodegeneration events in Parkinson's disease. *npj Parkinson's Disease*. 2023;9(1):111.
36 doi:[10.1038/s41531-023-00556-3](https://doi.org/10.1038/s41531-023-00556-3)
- 37
38 25. Sakato Y, Shima A, Terada Y, et al. Delineating three distinct spatiotemporal patterns
39 of brain atrophy in Parkinson's disease. *Brain*. 2024;147(11):3702-3713.
40 doi:[10.1093/brain/awae303](https://doi.org/10.1093/brain/awae303)
- 41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 26. Scotton WJ. *Disease Progression and Genetic Risk Factors in the Primary*
4 *Tauopathies*. PhD thesis. University of London, University College London (United
5 Kingdom); 2023.
- 6
7 27. Eshaghi A, Young AL, Wijeratne PA, et al. Identifying multiple sclerosis subtypes
8 using unsupervised machine learning and MRI data. *Nature Communications*. 2021;12:2078.
9 doi:[10.1038/s41467-021-22265-2](https://doi.org/10.1038/s41467-021-22265-2)
- 10
11 28. Young AL, Bragman FJ, Rangelov B, et al. Disease progression modeling in chronic
12 obstructive pulmonary disease. *American journal of respiratory and critical care medicine*.
13 2020;201(3):294-302.
- 14
15 29. Li M, Lan L, Luo J, Peng L, Li X, Zhou X. Identifying the phenotypic and temporal
16 heterogeneity of knee osteoarthritis: Data from the osteoarthritis initiative. *Frontiers in Public*
17 *Health*. 2021;9:726140.
- 18
19 30. Johnson D, Santos E, Kim K, et al. Increased pain symptomatology among females vs.
20 Males with fragile X-associated tremor/ataxia syndrome. *Frontiers in Psychiatry*. 2022;12.
21 doi:[10.3389/fpsyg.2021.762915](https://doi.org/10.3389/fpsyg.2021.762915)
- 22
23 31. Winarni TI, Hwang YH, Rivera SM, et al. Apolipoproteine and KLOTHO gene variants
24 do not affect the penetrance of fragile X-associated tremor/ataxia syndrome. *International*
25 *Journal of Molecular Sciences*. 2024;25(15):8103.
- 26
27 32. Elias-Mas A, Wang JY, Rodríguez-Revenga L, et al. Enlarged perivascular spaces and
28 their association with motor, cognition, MRI markers and cerebrovascular risk factors in male
29 fragile X premutation carriers. *Journal of the Neurological Sciences*. 2024;461:123056.
- 30
31 33. Bacalman S, Farzin F, Bourgeois JA, et al. Psychiatric phenotype of the fragile X-
32 associated tremor/ataxia syndrome (FXTAS) in males: Newly described fronto-subcortical
33 dementia. *Journal of Clinical Psychiatry*. 2006;67(1):87-94.
- 34
35 34. Talebi M, Majdi A, Kamari F, Sadigh-Eteghad S. The cambridge neuropsychological
36 test automated battery (CANTAB) versus the minimal assessment of cognitive function in
37 multiple sclerosis (MACFIMS) for the assessment of cognitive function in patients with
38 multiple sclerosis. *Multiple Sclerosis and Related Disorders*. 2020;43:102172.
39 doi:<https://doi.org/10.1016/j.msard.2020.102172>
- 40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 35. Fonteijn HM, Clarkson MJ, Modat M, et al. An event-based disease progression model
4 and its application to familial Alzheimer's disease. In: *Information Processing in Medical*
5 *Imaging: 22nd International Conference, IPMI 2011, Kloster Irsee, Germany, July 3-8, 2011.*
6 *Proceedings* 22. Springer; 2011:748-759.
7
8 36. Fonteijn HM, Modat M, Clarkson MJ, et al. An event-based model for disease
9 progression and its application in familial Alzheimer's disease and Huntington's disease.
10 *NeuroImage*. 2012;60(3):1880-1889.
11
12 37. Oxtoby NP. Data-driven disease progression modeling. In: Colliot O, ed. *Machine*
13 *Learning for Brain Disorders*. Springer; 2023:511-532.
14
15 38. Pearson K. Contributions to the mathematical theory of evolution. *Philosophical*
16 *Transactions of the Royal Society of London A*. 1894;185:71-110.
17
18 39. Lazarsfeld PF. The logical and mathematical foundation of latent structure analysis.
19 *Studies in social psychology in world war II Vol IV: Measurement and prediction*. Published
20 online 1950:362-412.
21
22 40. McLachlan GJ, Lee SX, Rathnayake SI. Finite mixture models. *Annual review of*
23 *statistics and its application*. 2019;6:355-378. doi:[10.1146/annurev-statistics-031017-100325](https://doi.org/10.1146/annurev-statistics-031017-100325)
24
25 41. Gelfand AE, Smith AF. Sampling-based approaches to calculating marginal densities.
26 *Journal of the American statistical association*. 1990;85(410):398-409.
27
28 42. Rubin DB. Inference and missing data. *Biometrika*. 1976;63(3):581-592.
29 doi:[10.2307/2335739](https://doi.org/10.2307/2335739)
30
31 43. Little RJ, Rubin DB. *Statistical Analysis with Missing Data*. Vol 793. John Wiley &
32 Sons; 2019.
33
34 44. Welch WJ. Construction of permutation tests. *Journal of the American Statistical*
35 *Association*. 1990;85(411):693-698.
36
37 45. Van Rossum G, Drake FL. *Python 3 Reference Manual*. CreateSpace; 2009.
38
39 46. Aksman LM, Wijeratne PA, Oxtoby NP, et al. pySuStaIn: A python implementation of
40 the subtype and stage inference algorithm. *SoftwareX*. 2021;16:100811. <https://github.com/ucl-pond/pySuStaIn>

- 1
2
3 47. R Core Team. *R: A Language and Environment for Statistical Computing*. R
4 Foundation for Statistical Computing; 2024. <https://www.R-project.org/>
- 5
6 48. Wickham H, Averick M, Bryan J, et al. Welcome to the tidyverse. *Journal of Open*
7 *Source Software*. 2019;4(43):1686. doi:[10.21105/joss.01686](https://doi.org/10.21105/joss.01686)
- 8
9 49. Ushey K, Allaire J, Tang Y. *Reticulate: Interface to 'Python'*; 2024. <https://CRAN.R-project.org/package=reticulate>
- 10
11 50. Bourgeois JA, Seritan AL, Casillas EM, et al. Lifetime prevalence of mood and anxiety
12 disorders in fragile X premutation carriers. *The Journal of Clinical Psychiatry*.
13 2011;71(2):175-182. doi:[10.4088/JCP.09m05407blu](https://doi.org/10.4088/JCP.09m05407blu)
- 14
15 51. Seritan AL, Ortigas M, Seritan S, A Bourgeois J, J Hagerman R. Psychiatric disorders
16 associated with FXTAS. *Current psychiatry reviews*. 2013;9(1):59-64.
17 doi:[10.2174/157340013805289699](https://doi.org/10.2174/157340013805289699)
- 18
19 52. Seritan A, A Bourgeois J, Schneider A, Mu Y, J Hagerman R, V Nguyen D. Ages of
20 onset of mood and anxiety disorders in fragile X premutation carriers. *Current psychiatry*
21 *reviews*. 2013;9(1):65-71. doi:[10.2174/157340013805289662](https://doi.org/10.2174/157340013805289662)
- 22
23 53. Schneider A, Summers S, Tassone F, et al. Women with fragile X-associated
24 tremor/ataxia syndrome. *Movement Disorders Clinical Practice*. 2020;7(8):910-919.
- 25
26 54. Hagerman RJ, Protic D, Rajaratnam A, Salcedo-Arellano MJ, Aydin EY, Schneider A.
27 Fragile X-associated neuropsychiatric disorders (FXAND). *Frontiers in Psychiatry*. 2018;9.
28 doi:[10.3389/fpsyg.2018.00564](https://doi.org/10.3389/fpsyg.2018.00564)
- 29
30 55. Roberts JE, Bailey Jr. DB, Mankowski J, et al. Mood and anxiety disorders in females
31 with the FMR1 premutation. *American Journal of Medical Genetics Part B: Neuropsychiatric*
32 *Genetics*. 2009;150B(1):130-139. doi:<https://doi.org/10.1002/ajmg.b.30786>
- 33
34 56. Rickards H. Depression in neurological disorders: Parkinson's disease, multiple
35 sclerosis, and stroke. *Journal of Neurology, Neurosurgery & Psychiatry*. 2005;76(suppl 1):i48-
36 i52. doi:[10.1136/jnnp.2004.060426](https://doi.org/10.1136/jnnp.2004.060426)
- 37
38 57. Fusar-Poli P, Borgwardt S, Bechdolf A, et al. The psychosis high-risk state: A
39 comprehensive state-of-the-art review. *JAMA Psychiatry*. 2013;70(1):107-120.
40 doi:[10.1001/jamapsychiatry.2013.269](https://doi.org/10.1001/jamapsychiatry.2013.269)
- 41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 58. Benussi A, Alberici A, Samra K, et al. Conceptual framework for the definition of
4 preclinical and prodromal frontotemporal dementia. *Alzheimer's & Dementia*.
5 2022;18(7):1408-1423. doi:<https://doi.org/10.1002/alz.12485>
- 6
7 59. Noyce AJ, Lees AJ, Schrag AE. The prediagnostic phase of Parkinson's disease.
8 *Journal of Neurology, Neurosurgery & Psychiatry*. 2016;87(8):871-878. doi:[10.1136/jnnp-2015-311890](https://doi.org/10.1136/jnnp-2015-311890)
- 9
10 60. Wijnands JM, Zhu F, Kingwell E, et al. Five years before multiple sclerosis onset:
11 Phenotyping the prodrome. *Multiple Sclerosis Journal*. 2019;25(8):1092-1101.
12 doi:[10.1177/1352458518783662](https://doi.org/10.1177/1352458518783662)
- 13
14 61. Govoni M, Bortoluzzi A, Padovan M, et al. The diagnosis and clinical management of
15 the neuropsychiatric manifestations of lupus. *Journal of Autoimmunity*. 2016;74:41-72.
16 doi:<https://doi.org/10.1016/j.jaut.2016.06.013>
- 17
18 62. Silvagni E, Inglese F, Bortoluzzi A, et al. Longitudinal changes in cerebral white matter
19 microstructure in newly diagnosed systemic lupus erythematosus patients. *Rheumatology*.
20 2020;60(6):2678-2687. doi:[10.1093/rheumatology/keaa677](https://doi.org/10.1093/rheumatology/keaa677)
- 21
22 63. Marrie RA, Maxwell CJ, Rotstein DL, Tsai C-C, Tremlett H. Prodromes in
23 demyelinating disorders, amyotrophic lateral sclerosis, Parkinson disease, and Alzheimer's
24 dementia. *Revue Neurologique*. 2024;180(3):125-140.
25 doi:<https://doi.org/10.1016/j.neurol.2023.07.002>
- 26
27 64. Sloan M, Bourgeois JA, Leschziner G, et al. Neuropsychiatric prodromes and symptom
28 timings in relation to disease onset and/or flares in SLE: Results from the mixed methods
29 international INSPIRE study. *eClinicalMedicine*. Published online 2024.
30 doi:[10.1016/j.eclim.2024.102634](https://doi.org/10.1016/j.eclim.2024.102634)
- 31
32 65. Cordeiro L, Abucayan F, Hagerman R, Tassone F, Hessl D. Anxiety disorders in fragile
33 X premutation carriers: Preliminary characterization of probands and non-probands.
34 *Intractable & Rare Diseases Research*. 2015;4(3):123-130. doi:[10.5582/irdr.2015.01029](https://doi.org/10.5582/irdr.2015.01029)
- 35
36 66. Jacquemont S, Hagerman RJ, Leehey M, et al. Fragile X premutation tremor/ataxia
37 syndrome: Molecular, clinical, and neuroimaging correlates. *The American Journal of Human
38 Genetics*. 2003;72(4):869-878. doi:<https://doi.org/10.1086/374321>
- 39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 67. Hall DA, Birch RC, Anheim M, et al. Emerging topics in FXTAS. *Journal of*
4 *Neurodevelopmental Disorders*. 2014;6:1-10. doi:[10.1186/1866-1955-6-31](https://doi.org/10.1186/1866-1955-6-31)
- 5
6 68. Apartis E, Blancher A, Meissner WG, et al. FXTAS: New insights and the need for
7 revised diagnostic criteria. *Neurology*. 2012;79(18):1898-1907.
8 doi:[10.1212/WNL.0b013e318271f7ff](https://doi.org/10.1212/WNL.0b013e318271f7ff)
- 9
10 69. Johnson D, Santos E, Kim K, et al. Increased pain symptomatology among females vs.
11 Males with fragile X-associated tremor/ataxia syndrome. *Frontiers in Psychiatry*.
12 2022;12:762915.
- 13
14 70. Loesch DZ, Tassone F, Atkinson A, et al. Differential progression of motor dysfunction
15 between male and female fragile X premutation carriers reveals novel aspects of sex-specific
16 neural involvement. *Frontiers in Molecular Biosciences*. 2021;7:Article 577246.
17 doi:[10.3389/fmoleb.2020.577246](https://doi.org/10.3389/fmoleb.2020.577246)
- 18
19 71. Leehey MA, Berry-Kravis E, Goetz CG, et al. FMR1 CGG repeat length predicts motor
20 dysfunction in premutation carriers. *Neurology*. 2008;70(16_part_2):1397-1402.
21 doi:[10.1212/01.wnl.0000281692.98200.f5](https://doi.org/10.1212/01.wnl.0000281692.98200.f5)
- 22
23 72. Cleveland WS, Grosse E, Shyu WM. Local regression models. In: *Statistical Models in*
24 *s*. Routledge; 1992:309-376.
- 25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1: Descriptive statistics of patient characteristics by Sex and CGG repeat level.

Characteristic	Male		Female		M vs. F (all CGG combined)
	CGG < 55 N = 37 ^a	CGG ≥ 55 N = 152 ^a	CGG < 55 N = 7 ^a	CGG ≥ 55 N = 101 ^a	p-value ^b
Age at visit					
Mean (SD)	55.4 (9.87)	63.5 (9.49)	55.1 (11.49)	60.7 (11.45)	0.210 ^c
Median [Min, Max]	54 [40, 75]	64 [40, 86]	54 [44, 78]	61 [40, 85]	
Primary Race/Ethnicity					
White	29 (81%)	122 (90%)	3 (60%)	76 (96%)	
Hispanic	4 (11%)	7 (5.1%)	0 (0%)	1 (1.3%)	
Black	2 (5.6%)	0 (0%)	1 (20%)	0 (0%)	
Other	1 (2.8%)	7 (5.1%)	1 (20%)	2 (2.5%)	
Missing	1	16	2	22	
FXTAS Stage					
0	36 (100%)	30 (20%)	3 (100%)	30 (37%)	
1	0 (0%)	28 (19%)	0 (0%)	6 (7.4%)	
2	0 (0%)	28 (19%)	0 (0%)	15 (19%)	
3	0 (0%)	39 (27%)	0 (0%)	22 (27%)	
4	0 (0%)	16 (11%)	0 (0%)	5 (6.2%)	
5	0 (0%)	6 (4.1%)	0 (0%)	3 (3.7%)	
Missing	1	5	4	20	
CGG					
Mean (SD)	29.3 (6.07)	86.9 (19.80)	33.6 (8.18)	87.0 (19.60)	
Median [Min, Max]	30 [20, 47]	85 [55, 185]	31 [29, 52]	86 [55, 167]	

^an (%)^bp-values represent tests for sex differences in distributions of characteristics, all CGG repeat levels.^cp-value for significance of sex difference by Wilcoxon rank sum test^dp-value for significance of sex difference by Fisher's exact test

Table 2: **Symptoms included in analysis.** The first level listed in each row of column “Defined Ordered Levels” is the “reference level” for the corresponding symptom, and subsequent levels are considered “clinically elevated levels”. Columns “Female” and “Male” list percentages of clinically elevated levels at baseline visit, stratified by sex.

Category	Biomarker	Defined Ordered Levels	Female^a	Male^a	p-value^b
Ataxia	Ataxia	No, Yes	52.7%	43.2%	0.203
Ataxia	Ataxia: severity	0, 1, 2, 3, 4	32.6%	32.2%	>0.999
CANTAB	PAL Total errors (adjusted)	≤ 13, > 13	71.4%	65.9%	0.803
CANTAB	RTI 5-choice movement time	≤ 368.57, > 368.57	27.3%	32.3%	0.805
CANTAB	SWM Between errors	≤ 26, > 26	10%	38.2%	0.020
MRI	Corpus Callosum-Thickness	Normal, Thin	23.1%	45.3%	0.059
MRI	Genu WM Hyperintensity	No, Yes	41.7%	50%	0.629
MRI	MRI: Cerebellar	None, Mild, Moderate, Severe	48.1%	64.8%	0.177
MRI	MRI: Cerebral	None, Mild, Moderate, Severe	92.6%	81.6%	0.234
MRI	Splenium WM Hyperintensity	None, Mild, Moderate, Severe	66.7%	66.7%	>0.999
Parkinsonian	Increased tone	No, Yes	6.5%	21.2%	0.034
Parkinsonian	Masked faces	No, Yes	5.7%	9.8%	0.428
Parkinsonian	Pill-rolling tremor	No, Yes	4.8%	5.5%	>0.999
Parkinsonian	Stiff gait	No, Yes	6.4%	9.3%	0.755
Parkinsonian	Parkinsonian features	No, Yes	14.3%	32.8%	0.028
Parkinson's	Parkinson's disease	No, Yes	21.1%	16.7%	0.720
SCID	Anxiety disorders	Absent, Sub-Threshold, Threshold	87.5%	61.5%	<0.001
SCID	Mood disorders	Absent, Sub-Threshold, Threshold	66.7%	40.6%	<0.001
SCID	Somatoform disorders	Absent, Sub-Threshold, Threshold	15.5%	2.1%	<0.001
SCID	Substance use disorders	Absent, Sub-Threshold, Threshold	16.7%	28.2%	0.066
Scores	BDS-2 Total Score	≥ 20, < 20	22.1%	31.5%	0.169
Scores	MMSE total score	Normal (26-30), Mildly impaired (20-25), Moderately impaired (10-19)	5.3%	20%	0.010
Thyroid	Hyperthyroid	No, Yes	2.9%	1.7%	0.632
Thyroid	Hypothyroid	No, Yes	20.8%	11.7%	0.104
Thyroid	Any autoimmune disorder	No, Yes	20.7%	7.7%	0.004
Tremors	Head tremor	No, Yes	41%	21.3%	0.043
Tremors	Intention tremor	No, Yes	51.4%	51.7%	>0.999
Tremors	Intermittent tremor	No, Yes	28.6%	26.4%	0.741
Tremors	Postural tremor	No, Yes	24.3%	25.4%	>0.999
Tremors	Resting tremor	No, Yes	15.7%	13.8%	0.685

^a% of participants with clinically elevated levels

^bp-value for significance of sex difference by Fisher's exact test

1
2
3
4 **Table 3: Demographics of 4 latent subtype clusters identified by Ordinal SuStaIn.** 36
5 patients had experienced too few events to be accurately classified into a subtype and were
6 excluded from this table.
7
8

Characteristic	Overall N = 217^a	Type 1 N = 63^a	Type 2 N = 64^a	Type 3 N = 50^a	Type 4 N = 40^a	p-value
CGG, Mean (SD)	87.3 (18.9)	83.2 (15.8)	90.2 (22.3)	92.4 (17.4)	82.7 (17.3)	0.010 ^b
CGG 55-99, n (%)	162 (75%)	49 (78%)	44 (69%)	36 (72%)	33 (83%)	0.393 ^c
Male, n (%)	140 (65%)	41 (65%)	34 (53%)	41 (82%)	24 (60%)	0.014 ^c
Primary Race/Ethnicity, n (%)						0.920 ^d
White	177 (92.7%)	51 (91.1%)	53 (91.4%)	41 (93.2%)	32 (97.0%)	
Hispanic	7 (3.7%)	3 (5.4%)	3 (5.2%)	1 (2.3%)	0 (0.0%)	
Other	7 (3.7%)	2 (3.6%)	2 (3.4%)	2 (4.5%)	1 (3.0%)	
Missing	26	7	6	6	7	

19 ^an (column %)
20 ^bGroup comparison was done by One-way analysis of means (not assuming equal variances)
21 ^cGroup comparison was done by Pearson's Chi-squared test
22 ^dGroup comparison was done by Fisher's exact test

or Review Only

Figure 1: Event sequences of FXTAS symptoms stratified by sex. The different colors (red, blue, purple, green, magenta) indicate the ordinal levels of symptom progression. Color gradient intensity represents the likelihood of sequence position. The brighter the color, the more likely that the corresponding symptom event occurs in that position in the sequence.

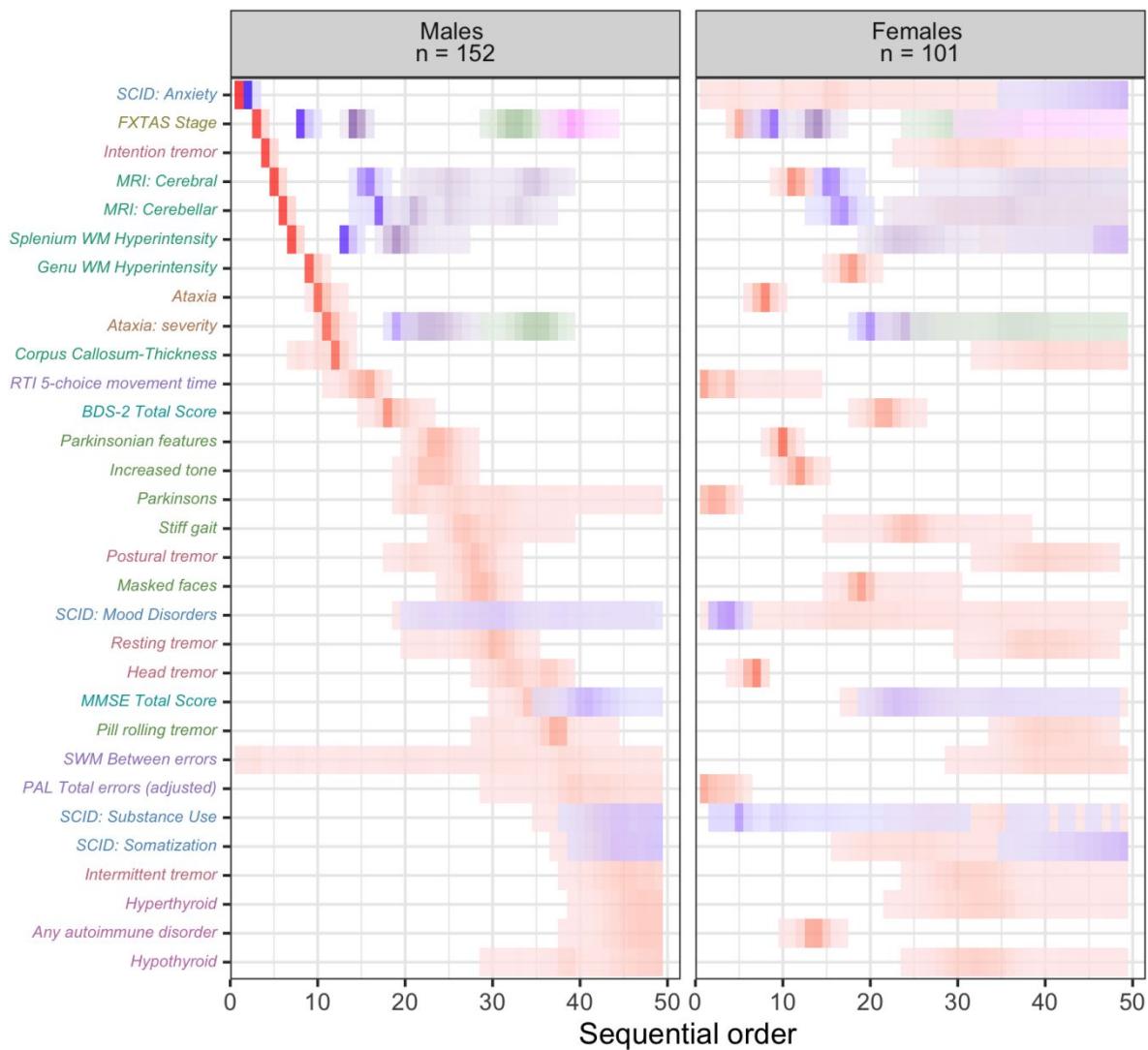


Figure 2: Positional differences in estimated event sequence between males and females. Red lines indicate symptoms that moved to later positions between the left-hand subgroup and the right-hand subgroup. Blue lines indicate symptoms that moved to earlier positions. Line opacity levels indicate the number of positions changed (higher opacity represents more positions changed).

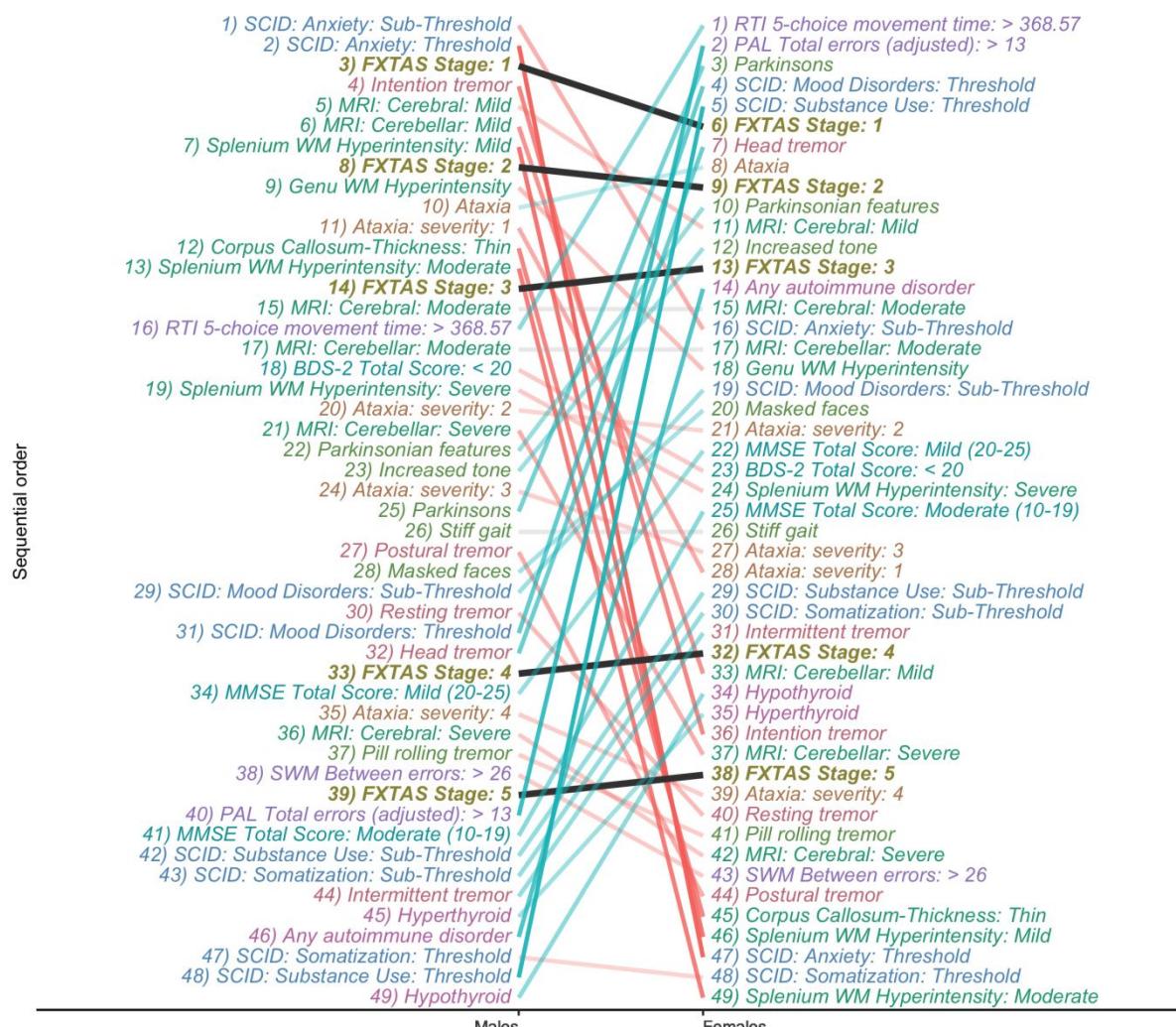
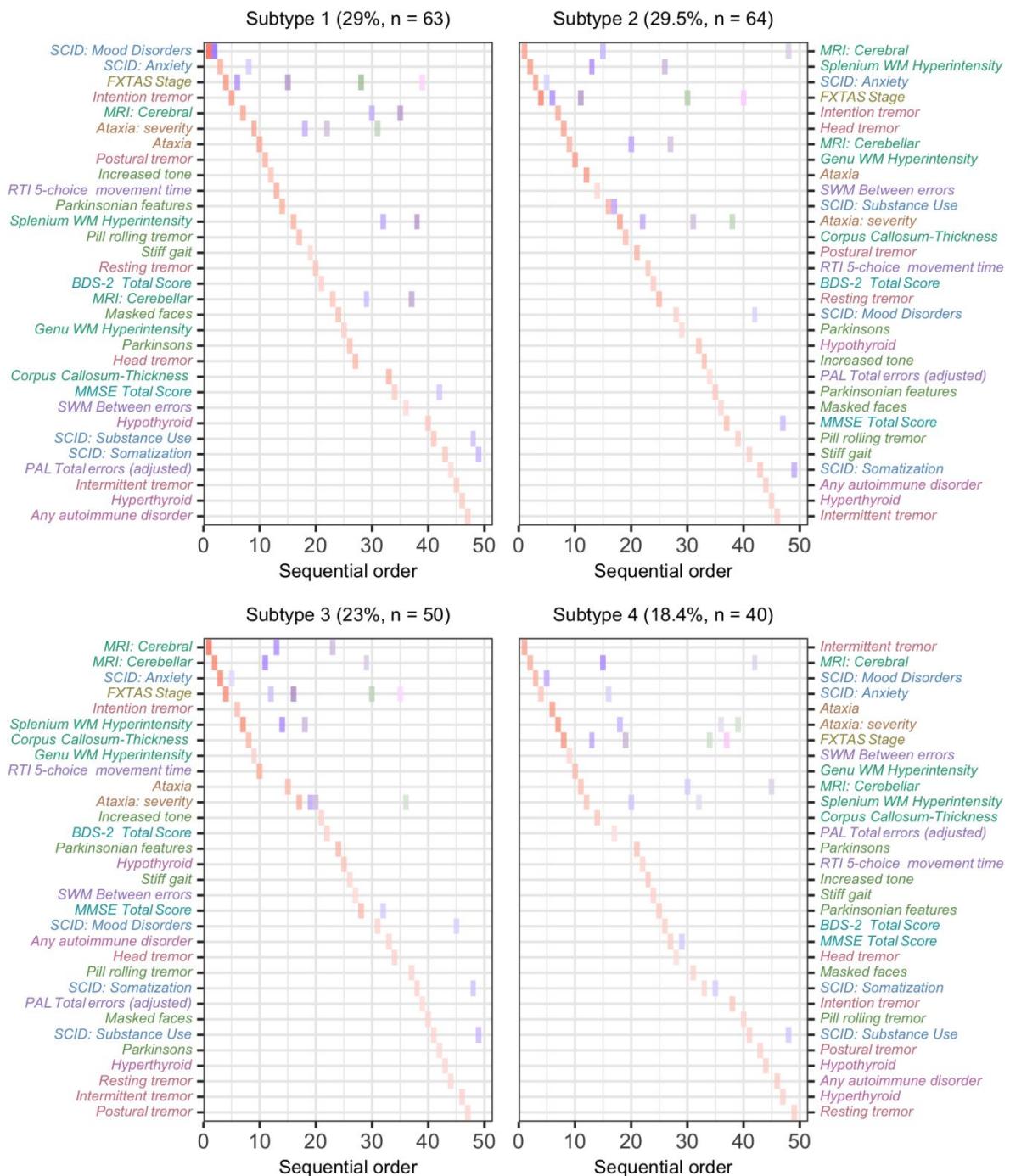


Figure 3: Event sequences for 4 latent subtypes. The different colors (red, blue, purple, green, magenta) indicate the ordinal levels of symptom progression. Color gradient intensity represents the likelihood of sequence position. The brighter the color, the more likely that the corresponding symptom event occurs in that position in the sequence.



Supplementary material

Summary data tables

The following supplementary tables 1 - 11 summarize the distributions of the symptoms that we analyzed in this paper, stratified by sex and CGG level.

Supplementary Table 1: Tremors by CGG level and sex

	Male			Female		
	CGG <55 (N=37)	CGG 55-99 (N=116)	CGG 100-199 (N=36)	CGG <55 (N=7)	CGG 55-99 (N=76)	CGG 100-199 (N=25)
Intention tremor						
No	31 (93.9%)	34 (39.5%)	7 (23.3%)	3 (100%)	28 (51.9%)	4 (26.7%)
Yes	2 (6.1%)	52 (60.5%)	23 (76.7%)	0 (0%)	26 (48.1%)	11 (73.3%)
Missing	4 (10.8%)	30 (25.9%)	6 (16.7%)	4 (57.1%)	22 (28.9%)	10 (40.0%)
Resting tremor						
No	32 (94.1%)	71 (86.6%)	22 (75.9%)	3 (100%)	47 (85.5%)	9 (75.0%)
Yes	2 (5.9%)	11 (13.4%)	7 (24.1%)	0 (0%)	8 (14.5%)	3 (25.0%)
Missing	3 (8.1%)	34 (29.3%)	7 (19.4%)	4 (57.1%)	21 (27.6%)	13 (52.0%)
Postural tremor						
No	32 (97.0%)	53 (65.4%)	21 (75.0%)	3 (100%)	42 (79.2%)	8 (57.1%)
Yes	1 (3.0%)	28 (34.6%)	7 (25.0%)	0 (0%)	11 (20.8%)	6 (42.9%)
Missing	4 (10.8%)	35 (30.2%)	8 (22.2%)	4 (57.1%)	23 (30.3%)	11 (44.0%)
Intermittent tremor						
No	27 (93.1%)	45 (61.6%)	23 (85.2%)	3 (100%)	37 (69.8%)	10 (71.4%)
Yes	2 (6.9%)	28 (38.4%)	4 (14.8%)	0 (0%)	16 (30.2%)	4 (28.6%)
Missing	8 (21.6%)	43 (37.1%)	9 (25.0%)	4 (57.1%)	23 (30.3%)	11 (44.0%)

Supplementary Table 2: Head tremors

	Male			Female		
	CGG <55 (N=37)	CGG 55-99 (N=116)	CGG 100-199 (N=36)	CGG <55 (N=7)	CGG 55-99 (N=76)	CGG 100-199 (N=25)
Head tremor						
No	5 (100%)	34 (81.0%)	9 (64.3%)	0 (NA%)	15 (57.7%)	8 (61.5%)
Yes	0 (0%)	8 (19.0%)	5 (35.7%)	0 (NA%)	11 (42.3%)	5 (38.5%)
Missing	32 (86.5%)	74 (63.8%)	22 (61.1%)	7 (100%)	50 (65.8%)	12 (48.0%)

Supplementary Table 3: Ataxia by CGG status

	Male			Female		
	CGG <55 (N=37)	CGG 55-99 (N=116)	CGG 100-199 (N=36)	CGG <55 (N=7)	CGG 55-99 (N=76)	CGG 100-199 (N=25)
Ataxia						
No	29 (90.6%)	50 (56.2%)	9 (26.5%)	3 (100%)	26 (49.1%)	6 (33.3%)
Yes	3 (9.4%)	39 (43.8%)	25 (73.5%)	0 (0%)	27 (50.9%)	12 (66.7%)
Missing	5 (13.5%)	27 (23.3%)	2 (5.6%)	4 (57.1%)	23 (30.3%)	7 (28.0%)
Ataxia severity						
0	32 (94.1%)	66 (64.7%)	18 (51.4%)	3 (100%)	43 (69.4%)	12 (57.1%)
1	2 (5.9%)	21 (20.6%)	3 (8.6%)	0 (0%)	10 (16.1%)	3 (14.3%)
2	0 (0%)	9 (8.8%)	2 (5.7%)	0 (0%)	6 (9.7%)	3 (14.3%)
3	0 (0%)	5 (4.9%)	7 (20.0%)	0 (0%)	1 (1.6%)	3 (14.3%)
4	0 (0%)	1 (1.0%)	5 (14.3%)	0 (0%)	2 (3.2%)	0 (0%)
Missing	3 (8.1%)	14 (12.1%)	1 (2.8%)	4 (57.1%)	14 (18.4%)	4 (16.0%)

Supplementary Table 4: Parkinson's disease by CGG status

	Male			Female		
	CGG <55 (N=37)	CGG 55-99 (N=116)	CGG 100-199 (N=36)	CGG <55 (N=7)	CGG 55-99 (N=76)	CGG 100-199 (N=25)
Parkinson's disease						
No	5 (100%)	13 (81.3%)	7 (77.8%)	0 (NA%)	10 (76.9%)	5 (83.3%)
Yes	0 (0%)	3 (18.8%)	2 (22.2%)	0 (NA%)	3 (23.1%)	1 (16.7%)
Missing	32 (86.5%)	100 (86.2%)	27 (75.0%)	7 (100%)	63 (82.9%)	19 (76.0%)

Supplementary Table 5: FXTAS Stage by CGG level

	Male			Female		
	CGG <55 (N=37)	CGG 55-99 (N=116)	CGG 100-199 (N=36)	CGG <55 (N=7)	CGG 55-99 (N=76)	CGG 100-199 (N=25)
FXTAS Stage						
0	36 (100%)	25 (22.3%)	5 (14.3%)	3 (100%)	24 (40.0%)	6 (28.6%)
1	0 (0%)	23 (20.5%)	5 (14.3%)	0 (0%)	5 (8.3%)	1 (4.8%)
2	0 (0%)	26 (23.2%)	2 (5.7%)	0 (0%)	10 (16.7%)	5 (23.8%)
3	0 (0%)	28 (25.0%)	11 (31.4%)	0 (0%)	13 (21.7%)	9 (42.9%)
4	0 (0%)	8 (7.1%)	8 (22.9%)	0 (0%)	5 (8.3%)	0 (0%)
5	0 (0%)	2 (1.8%)	4 (11.4%)	0 (0%)	3 (5.0%)	0 (0%)
Missing	1 (2.7%)	4 (3.4%)	1 (2.8%)	4 (57.1%)	16 (21.1%)	4 (16.0%)

Supplementary Table 6: Behavior Dyscontrol Scale - Second Edition (BDS-2)

	Male			Female		
	CGG <55 (N=37)	CGG 55-99 (N=116)	CGG 100-199 (N=36)	CGG <55 (N=7)	CGG 55-99 (N=76)	CGG 100-199 (N=25)
BDS-2 total score						
Mean (SD)	22.3 (2.94)	20.7 (4.76)	17.5 (5.95)	23.3 (2.87)	21.3 (4.24)	21.4 (2.66)
Median [Min, Max]	23.0 [15.0, 27.0]	22.0 [2.00, 27.0]	19.0 [1.00, 26.0]	24.5 [19.0, 25.0]	22.0 [4.00, 26.0]	22.0 [17.0, 26.0]
Missing	1 (2.7%)	17 (14.7%)	6 (16.7%)	3 (42.9%)	24 (31.6%)	4 (16.0%)
BDS-2 total score						
≥ 20	30 (83.3%)	69 (69.7%)	14 (46.7%)	3 (75.0%)	42 (80.8%)	15 (71.4%)
< 20	6 (16.7%)	30 (30.3%)	16 (53.3%)	1 (25.0%)	10 (19.2%)	6 (28.6%)
Missing	1 (2.7%)	17 (14.7%)	6 (16.7%)	3 (42.9%)	24 (31.6%)	4 (16.0%)

Supplementary Table 7: Mini-Mental State Examination (MMSE)

	Male			Female		
	CGG <55 (N=37)	CGG 55-99 (N=116)	CGG 100-199 (N=36)	CGG <55 (N=7)	CGG 55-99 (N=76)	CGG 100-199 (N=25)
MMSE total score						
Mean (SD)	29.7 (0.577)	27.4 (4.28)	26.5 (3.41)	29.8 (0.500)	28.6 (2.44)	28.8 (1.51)
Median [Min, Max]	30.0 [29.0, 30.0]	29.0 [13.0, 30.0]	27.0 [16.0, 30.0]	30.0 [29.0, 30.0]	29.0 [16.0, 30.0]	29.0 [25.0, 30.0]
Missing	34 (91.9%)	73 (62.9%)	17 (47.2%)	3 (42.9%)	27 (35.5%)	3 (12.0%)
MMSE total score						
Normal (26-30)	3 (100%)	36 (83.7%)	13 (68.4%)	4 (100%)	46 (93.9%)	21 (95.5%)
Mild impairment (20-25)	0 (0%)	3 (7.0%)	5 (26.3%)	0 (0%)	2 (4.1%)	1 (4.5%)
Moderate impairment (10-19)	0 (0%)	4 (9.3%)	1 (5.3%)	0 (0%)	1 (2.0%)	0 (0%)
Missing	34 (91.9%)	73 (62.9%)	17 (47.2%)	3 (42.9%)	27 (35.5%)	3 (12.0%)

Supplementary Table 8: Structured Clinical Interview for DSM Disorders (SCID)

	Male			Female		
	CGG <55 (N=37)	CGG 55-99 (N=116)	CGG 100-199 (N=36)	CGG <55 (N=7)	CGG 55-99 (N=76)	CGG 100-199 (N=25)
SCID: mood disorders						
Absent	19 (59.4%)	52 (61.2%)	14 (53.8%)	1 (33.3%)	17 (33.3%)	6 (33.3%)
Sub-Threshold	2 (6.3%)	7 (8.2%)	3 (11.5%)	0 (0%)	2 (3.9%)	0 (0%)
Threshold	11 (34.4%)	26 (30.6%)	9 (34.6%)	2 (66.7%)	32 (62.7%)	12 (66.7%)
Missing	5 (13.5%)	31 (26.7%)	10 (27.8%)	4 (57.1%)	25 (32.9%)	7 (28.0%)
SCID: substance use disorders						
Absent	26 (83.9%)	59 (69.4%)	17 (65.4%)	3 (100%)	43 (84.3%)	14 (77.8%)
Sub-Threshold	2 (6.5%)	3 (3.5%)	0 (0%)	0 (0%)	3 (5.9%)	0 (0%)
Threshold	3 (9.7%)	23 (27.1%)	9 (34.6%)	0 (0%)	5 (9.8%)	4 (22.2%)
Missing	6 (16.2%)	31 (26.7%)	10 (27.8%)	4 (57.1%)	25 (32.9%)	7 (28.0%)
SCID: anxiety disorders						
Absent	12 (37.5%)	31 (36.5%)	12 (46.2%)	0 (0%)	7 (13.7%)	2 (11.1%)
Sub-Threshold	8 (25.0%)	16 (18.8%)	1 (3.8%)	0 (0%)	7 (13.7%)	1 (5.6%)
Threshold	12 (37.5%)	38 (44.7%)	13 (50.0%)	3 (100%)	37 (72.5%)	15 (83.3%)
Missing	5 (13.5%)	31 (26.7%)	10 (27.8%)	4 (57.1%)	25 (32.9%)	7 (28.0%)
SCID: somatoform disorders						
Absent	30 (96.8%)	84 (98.8%)	25 (96.2%)	2 (66.7%)	43 (86.0%)	15 (83.3%)
Sub-Threshold	1 (3.2%)	0 (0%)	0 (0%)	0 (0%)	2 (4.0%)	1 (5.6%)
Threshold	0 (0%)	1 (1.2%)	1 (3.8%)	1 (33.3%)	5 (10.0%)	2 (11.1%)
Missing	6 (16.2%)	31 (26.7%)	10 (27.8%)	4 (57.1%)	26 (34.2%)	7 (28.0%)
SCID: psychotic symptoms						
Absent	32 (100%)	85 (100%)	26 (100%)	3 (100%)	51 (100%)	17 (94.4%)
Sub-Threshold	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (5.6%)
Missing	5 (13.5%)	31 (26.7%)	10 (27.8%)	4 (57.1%)	25 (32.9%)	7 (28.0%)

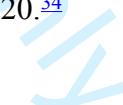
Supplementary Table 9: MRI variables by Sex and CGG level

	Male			Female		
	CGG <55 (N=37)	CGG 55-99 (N=116)	CGG 100- 199 (N=36)	CGG <55 (N=7)	CGG 55-99 (N=76)	CGG 100- 199 (N=25)
MRI: Cerebellar						
None	0 (NA%)	30 (42.3%)	2 (10.0%)	1 (100%)	10 (58.8%)	3 (33.3%)
Mild	0 (NA%)	23 (32.4%)	7 (35.0%)	0 (0%)	6 (35.3%)	6 (66.7%)
Moderate	0 (NA%)	12 (16.9%)	7 (35.0%)	0 (0%)	1 (5.9%)	0 (0%)
Severe	0 (NA%)	6 (8.5%)	4 (20.0%)	0 (0%)	0 (0%)	0 (0%)
Missing	37 (100%)	45 (38.8%)	16 (44.4%)	6 (85.7%)	59 (77.6%)	16 (64.0%)
MRI: Cerebral						
None	0 (NA%)	15 (21.7%)	1 (5.6%)	1 (100%)	0 (0%)	1 (11.1%)
Mild	0 (NA%)	35 (50.7%)	6 (33.3%)	0 (0%)	12 (70.6%)	5 (55.6%)
Moderate	0 (NA%)	15 (21.7%)	7 (38.9%)	0 (0%)	5 (29.4%)	3 (33.3%)
Severe	0 (NA%)	4 (5.8%)	4 (22.2%)	0 (0%)	0 (0%)	0 (0%)
Missing	37 (100%)	47 (40.5%)	18 (50.0%)	6 (85.7%)	59 (77.6%)	16 (64.0%)
Splenium (CC)-WM Hyperintensity						
None	0 (NA%)	21 (39.6%)	1 (7.7%)	1 (100%)	7 (41.2%)	1 (11.1%)
Mild	0 (NA%)	18 (34.0%)	1 (7.7%)	0 (0%)	6 (35.3%)	5 (55.6%)
Moderate	0 (NA%)	8 (15.1%)	6 (46.2%)	0 (0%)	3 (17.6%)	3 (33.3%)
Severe	0 (NA%)	6 (11.3%)	5 (38.5%)	0 (0%)	1 (5.9%)	0 (0%)
Missing	37 (100%)	63 (54.3%)	23 (63.9%)	6 (85.7%)	59 (77.6%)	16 (64.0%)
Genu (CC)-WM Hyperintensity						
No	0 (NA%)	28 (58.3%)	2 (16.7%)	1 (100%)	9 (60.0%)	4 (50.0%)
Yes	0 (NA%)	20 (41.7%)	10 (83.3%)	0 (0%)	6 (40.0%)	4 (50.0%)
Missing	37 (100%)	68 (58.6%)	24 (66.7%)	6 (85.7%)	61 (80.3%)	17 (68.0%)
Corpus Callosum-Thickness						
Normal	0 (NA%)	34 (65.4%)	1 (8.3%)	1 (100%)	13 (81.3%)	6 (66.7%)
Thin	0 (NA%)	18 (34.6%)	11 (91.7%)	0 (0%)	3 (18.8%)	3 (33.3%)
Missing	37 (100%)	64 (55.2%)	24 (66.7%)	6 (85.7%)	60 (78.9%)	16 (64.0%)

1
2
3
4 **Supplementary Table 10: Cambridge Neuropsychological Test Automated Battery**
5 **(CANTAB)**
6
7

	Male			Female		
	CGG <55 (N=37)	CGG 55-99 (N=116)	CGG 100- 199 (N=36)	CGG <55 (N=7)	CGG 55-99 (N=76)	CGG 100- 199 (N=25)
SWM Between errors						
Mean (SD)	31.3 (18.9)	27.5 (21.0)	32.8 (25.2)	12.5 (17.7)	15.8 (7.71)	17.3 (8.38)
Median [Min, Max]	26.0 [4.00, 74.0]	22.0 [0, 97.0]	25.0 [0, 81.0]	12.5 [0, 25.0]	16.5 [0, 27.0]	19.0 [5.00, 27.0]
Missing	14 (37.8%)	37 (31.9%)	15 (41.7%)	5 (71.4%)	64 (84.2%)	19 (76.0%)
SWM Between errors						
≤ 26	12 (52.2%)	51 (64.6%)	13 (61.9%)	2 (100%)	11 (91.7%)	5 (83.3%)
> 26	11 (47.8%)	28 (35.4%)	8 (38.1%)	0 (0%)	1 (8.3%)	1 (16.7%)
Missing	14 (37.8%)	37 (31.9%)	15 (41.7%)	5 (71.4%)	64 (84.2%)	19 (76.0%)
PAL Total errors (adjusted)						
Mean (SD)	22.5 (17.5)	28.1 (22.0)	21.4 (14.0)	22.5 (27.6)	29.0 (21.2)	27.9 (19.4)
Median [Min, Max]	18.0 [2.00, 76.0]	23.0 [2.00, 100]	18.0 [2.00, 48.0]	22.5 [3.00, 42.0]	22.5 [7.00, 63.0]	15.0 [6.00, 54.0]
Missing	14 (37.8%)	38 (32.8%)	14 (38.9%)	5 (71.4%)	64 (84.2%)	18 (72.0%)
PAL total errors						
≤ 13	8 (34.8%)	26 (33.3%)	8 (36.4%)	1 (50.0%)	4 (33.3%)	1 (14.3%)
> 13	15 (65.2%)	52 (66.7%)	14 (63.6%)	1 (50.0%)	8 (66.7%)	6 (85.7%)
Missing	14 (37.8%)	38 (32.8%)	14 (38.9%)	5 (71.4%)	64 (84.2%)	18 (72.0%)
RTI Five-choice movement time						
Mean (SD)	241 (49.1)	344 (119)	417 (175)	309 (7.78)	358 (60.4)	286 (39.3)
Median [Min, Max]	235 [161, 344]	319 [139, 703]	401 [189, 950]	309 [303, 314]	369 [241, 459]	290 [235, 336]
Missing	14 (37.8%)	37 (31.9%)	14 (38.9%)	5 (71.4%)	63 (82.9%)	18 (72.0%)
RTI Five-choice movement time						
≤ 368.57	23 (100%)	51 (64.6%)	10 (45.5%)	2 (100%)	7 (53.8%)	7 (100%)
> 368.57	0 (0%)	28 (35.4%)	12 (54.5%)	0 (0%)	6 (46.2%)	0 (0%)
Missing	14 (37.8%)	37 (31.9%)	14 (38.9%)	5 (71.4%)	63 (82.9%)	18 (72.0%)

39 We used categorization cutoffs taken from Talebi et al 2020.³⁴
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Supplementary Table 11: Thyroid and autoimmune diseases

	Male			Female		
	CGG <55 (N=37)	CGG 55-99 (N=116)	CGG 100-199 (N=36)	CGG <55 (N=7)	CGG 55-99 (N=76)	CGG 100-199 (N=25)
Hypothyroid						
No	24 (100%)	58 (86.6%)	24 (82.8%)	2 (66.7%)	44 (77.2%)	15 (88.2%)
Yes	0 (0%)	9 (13.4%)	5 (17.2%)	1 (33.3%)	13 (22.8%)	2 (11.8%)
Missing	13 (35.1%)	49 (42.2%)	7 (19.4%)	4 (57.1%)	19 (25.0%)	8 (32.0%)
Hyperthyroid						
No	23 (95.8%)	65 (100%)	26 (96.3%)	3 (100%)	49 (96.1%)	16 (100%)
Yes	1 (4.2%)	0 (0%)	1 (3.7%)	0 (0%)	2 (3.9%)	0 (0%)
Missing	13 (35.1%)	51 (44.0%)	9 (25.0%)	4 (57.1%)	25 (32.9%)	9 (36.0%)
any autoimmune disorder						
No	33 (97.1%)	89 (89.0%)	34 (97.1%)	3 (100%)	49 (77.8%)	17 (81.0%)
Yes	1 (2.9%)	11 (11.0%)	1 (2.9%)	0 (0%)	14 (22.2%)	4 (19.0%)
Missing	3 (8.1%)	16 (13.8%)	1 (2.8%)	4 (57.1%)	13 (17.1%)	4 (16.0%)

Estimates of event scoring accuracy

[Supplementary Table 12](#) shows the estimated percentage of correctly scored individuals for each symptom²⁰ estimated using the controls data.

Supplementary Table 12: **Percentages of controls at baseline levels, and corresponding estimates of probability of correct scoring.** All controls are assumed to be truly at the reference levels for each symptom, and the percentage of correctly scored individuals is thus estimated as the percentage of controls who were assessed as being at the reference level.

Biomarker	# controls with data	# at baseline	% at baseline	Est. Pr(correct)
Head tremor	5	5	100%	95%
Intention tremor	36	34	94.4%	94.4%
Resting tremor	37	35	94.6%	94.6%
Postural tremor	36	35	97.2%	95%
Intermittent tremor	32	30	93.8%	93.8%
Ataxia	35	32	91.4%	91.4%
Ataxia severity	37	35	94.6%	94.6%
FXTAS Stage	39	39	100%	95%
parkinsonian features	6	6	100%	95%
Masked faces	32	32	100%	95%
Increased tone	30	30	100%	95%
Pill-rolling tremor	32	32	100%	95%
Stiff gait	30	30	100%	95%
Parkinson's disease	5	5	100%	95%
MRI: Cerebellar	1	1	100%	95%
MRI: Cerebral	1	1	100%	95%
Splenium (CC)-WM	1	1	100%	95%
Hyperintensity	1	1	100%	95%
Genu (CC)-WM Hyperintensity	1	1	100%	95%
Corpus Callosum-Thickness	1	1	100%	95%
MMSE total score	7	7	100%	95%
BDS-2 total score	40	33	82.5%	82.5%
SCID: mood disorders	35	20	57.1%	57.1%
SCID: substance use disorders	34	29	85.3%	85.3%
SCID: anxiety disorders	35	12	34.3%	34.3%
SCID: somatoform disorders	34	32	94.1%	94.1%
SWM Between errors	25	14	56%	56%
PAL total errors	25	9	36%	36%
RTI Five-choice movement time	25	25	100%	95%
Hypothyroid	27	26	96.3%	95%
Hyperthyroid	27	26	96.3%	95%
any autoimmune disorder	37	36	97.3%	95%

Composite variables

Any autoimmune disorder

Due to the rare nature of individual autoimmune diseases, a composite variable ("any autoimmune disorder") was created for analysis: "Yes" if a patient had any of lupus, rheumatoid arthritis, multiple sclerosis, ANA (anti-nuclear antibody) positive, Sjogren's Syndrome, Raynaud's Syndrome, or pulmonary fibrosis; otherwise, "No" for no autoimmune diseases recorded.

MRI variables

The MRI variables are combined into composite variables taking the most severe score within a brain region.

- “MRI Cerebellar” was composed as the most severe score of Cerebellar atrophy, Cerebellar white matter (WM) hyperintensity, and Middle cerebellar peduncle (MCP) WM hyperintensity.
- “MRI Cerebral” was composed as the most severe score of Cerebral atrophy, Cerebral WM hyperintensity, Pons WM hyperintensity, Sub-insular WM hyperintensity, and Periventricular WM hyperintensity.
- Splenium WM hyperintensity, Genu WM hyperintensity, and Corpus Callosum thickness remained as separate fields, because they had been coded using incompatible Likert scales.

SCID composite variables

Similarly, we combined individual SCID disorders into composite variables, taking the highest level (“Absent”, “Sub-Threshold”, or “Threshold”) among the constituent individual disorders.

- “SCID: mood disorders” combines: Bipolar I Disorder (MD01), Lifetime, Bipolar II Disorder (MD02), Lifetime, Other Bipolar Disorder (MD03), Lifetime, Major Depressive Disorder (MD04), Lifetime, Dysthymic Disorder (MD05), Lifetime, Depressive Disorder NOS (MD06), Lifetime, Mood Disorder Due to GMC (MD07), Lifetime, Substance-Induced Mood Dis. (MD08), Lifetime.
- “SCID: substance use disorders” combines Alcohol (SUD17), Lifetime, Sedative-Hypnotic-Anxiolytic (SUD18), Lifetime, Cannabis (SUD19), Lifetime, Stimulants (SUD20), Lifetime, Opioid (SUD21), Lifetime, Cocaine (SUD22), Lifetime, Hallucinogenics/ PCP (SUD23), Lifetime, Poly Drug (SUD24), Lifetime, Other (SUD25), Lifetime.
- “SCID: anxiety disorders” combines Panic Disorder (ANX26), Lifetime, Agoraphobia without Panic (ANX27), Lifetime, Social Phobia (ANX28), Lifetime, Specific Phobia (ANX29), Lifetime, Obsessive Compulsive (ANX30), Lifetime, Posttraumatic Stress (ANX31), Lifetime, Generalized Anxiety (ANX32), Current Only, Anxiety Due to GMC (ANX33), Lifetime, Substance-Induced Anxiety (ANX34), Lifetime, Anxiety

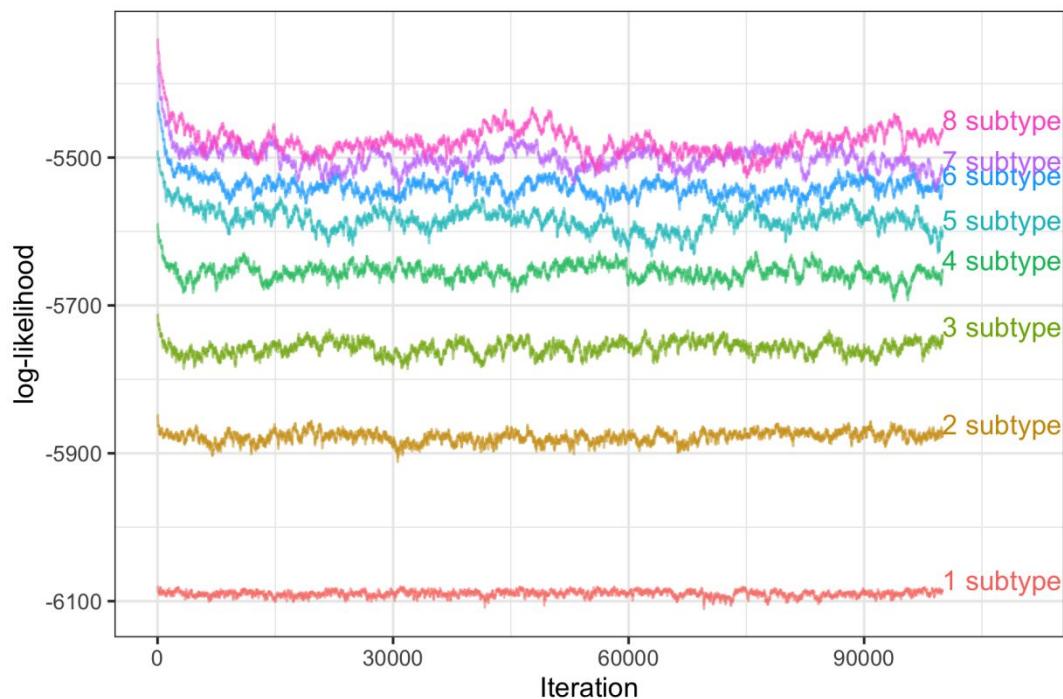
1
2
3 Disorder NOS (ANX35), Lifetime. Note that in the SCID-I/NP for DSM-IV, for
4 generalized anxiety, only the current, not lifetime prevalence, is included.
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19

- “SCID: somatoform disorders” combines Somatization Disorder (SOM36), Pain Disorder (SOM37), Undifferentiated Somatoform (SOM38), Body Dysmorphic (SOM40), Hypochondriasis (SOM39).
- “SCID: psychotic symptoms” consists of only Primary Psychotic Symptoms (PS01), Lifetime.

Detecting latent subtypes

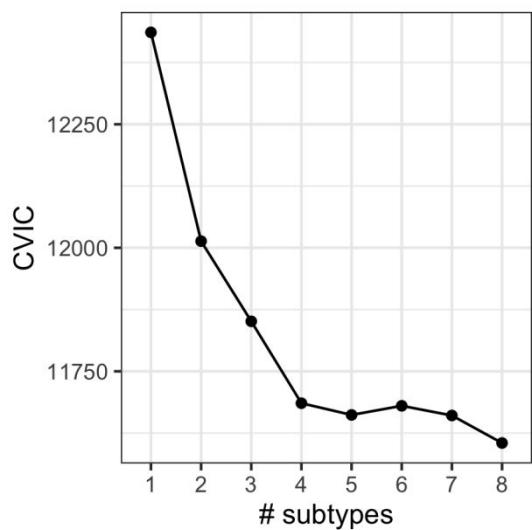
In order to use the Ordinal SuStaIn modeling algorithm, we must specify how many latent subtypes to include in the model. There are several metrics for determining the optimal number of subtypes for a given data set, including the likelihood of the data for the fitted model while varying the number of subtypes used to fit the model and the Cross-Validation Information Criterion (CVIC) described in Young *et al.*¹⁹ We also evaluated the consistency of our cross-validation procedure by looking at the distribution of out-of-fold log-likelihood (“OOFL”) across cross-validation folds (Young *et al.*).¹⁹ [Supplementary Fig. 1](#) shows the distribution of log-likelihoods from the MCMC samples for the full dataset (not stratified by sex or CGG repeats). Adding up to 6 clusters substantially improves the log-likelihood. [Supplementary Fig. 2 \(b\)](#) shows the distribution of the OOFLL statistic as a function of number of latent subgroups. [Supplementary Fig. 3](#) shows estimated disease progression stage (that is, number of events experienced) by age at visit and estimated latent subtype, with a Locally Estimated Scatterplot Smoothing (LOESS) non-parametric regression curve superimposed.⁷² This graph is a model diagnostic; if the model fits the data well, then older patients should on average be classified as being farther along in their disease progression, resulting in an upwards trend. Types 1, 3 and 4 appear to show such an upwards trend. Type 2 has a less clear trend, but does not show substantial evidence of lack of fit. In all four subtypes, patients whose visit occurred at a later age tend to be at a later estimated stage of disease progression. The upwards trend of progression stage with age indicates that the estimated models are plausible.

Supplementary Fig. 1: **log-likelihoods of MCMC samples, by number of subtypes**

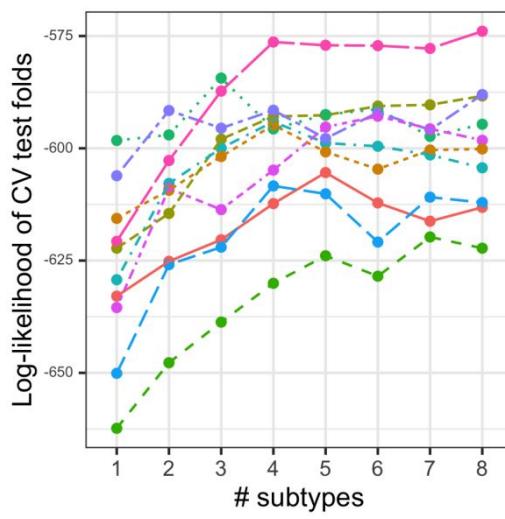


Supplementary Fig. 2: **Selection criteria for number of latent subtypes** (a) Cross-validation information criterion (b) Test set log-likelihood across folds

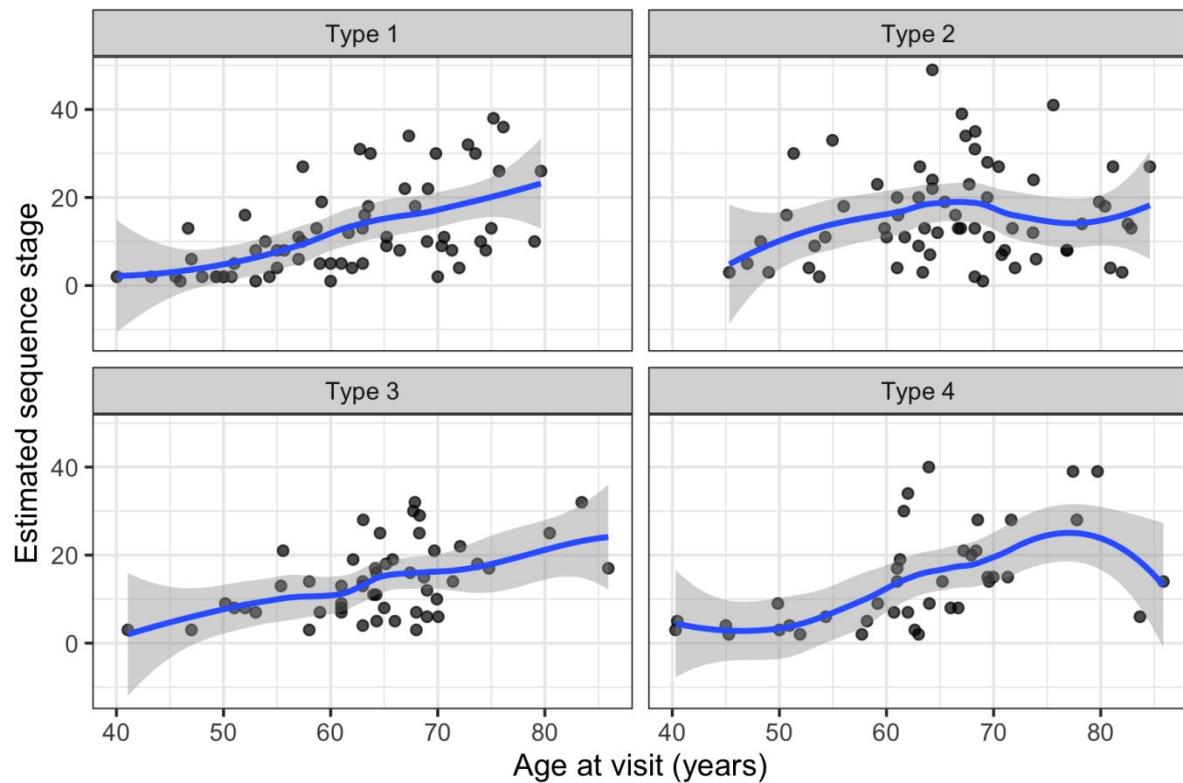
(a)



(b)



Supplementary Fig. 3: **Estimated progression stage by age and latent subtype.** The blue line is a Locally Estimated Scatterplot Smoothing (LOESS) non-parametric regression curve.

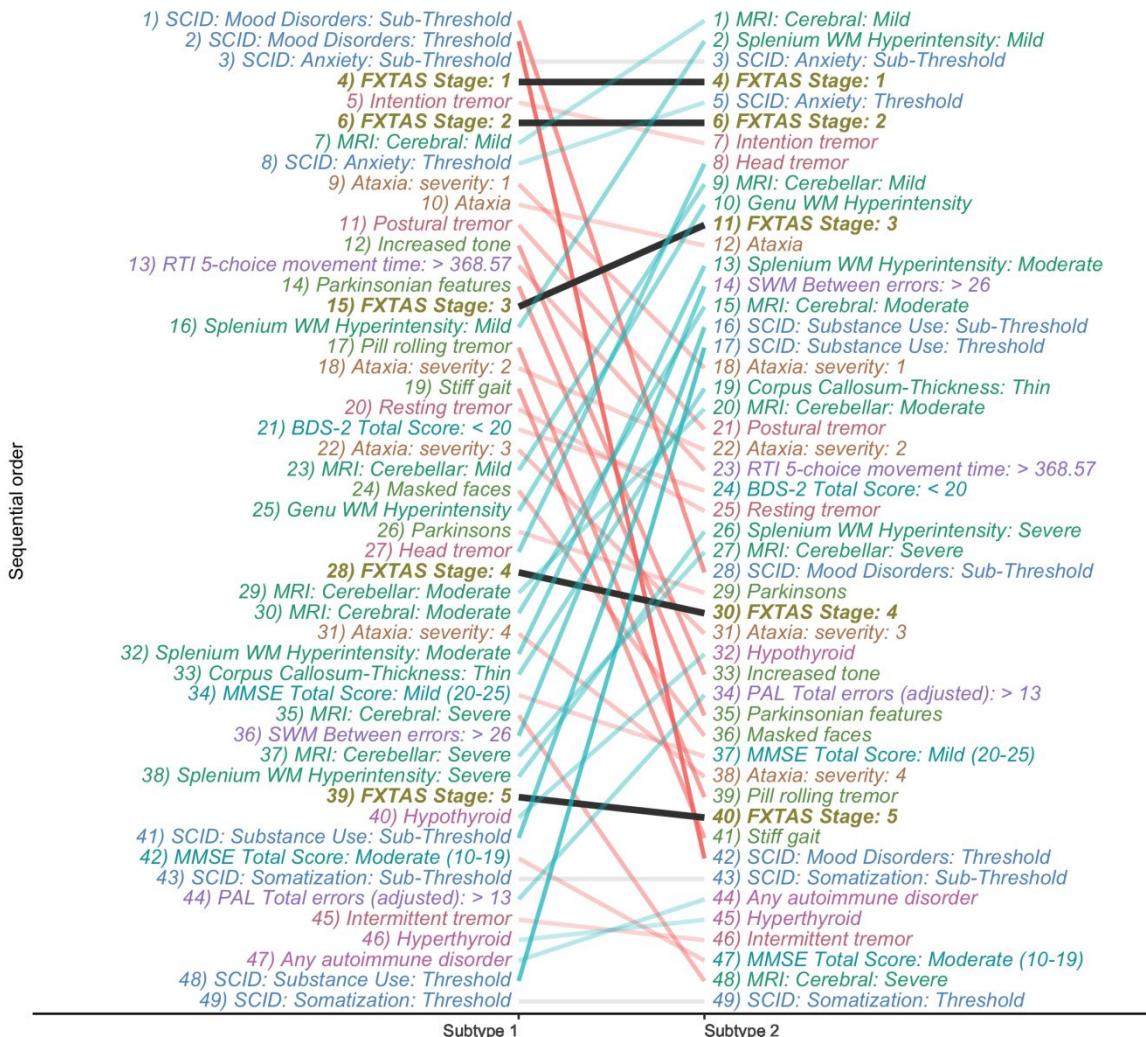


Only

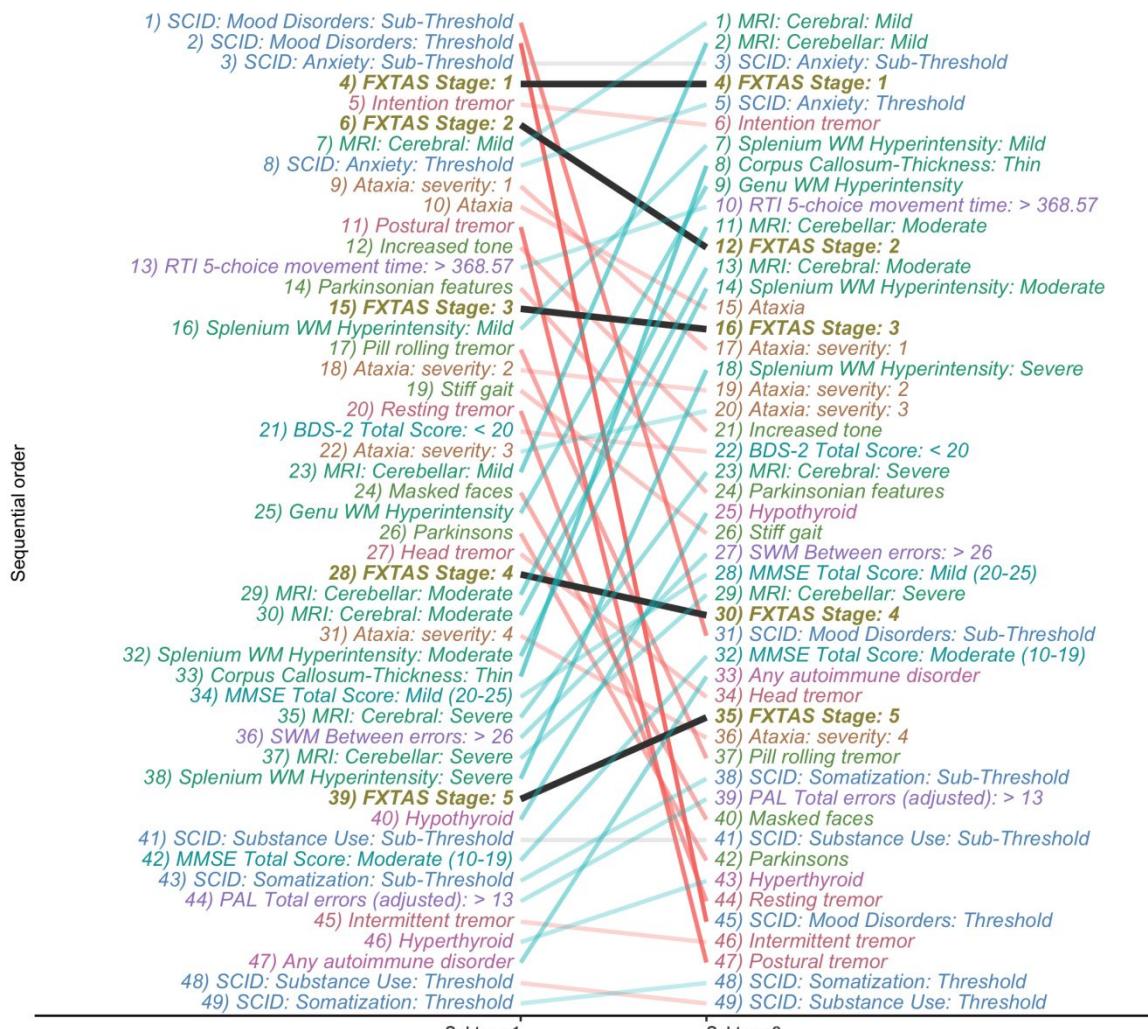
1
2
3
4 **Supplementary Fig. 4: Differences in event sequences between pairs of latent subtypes.**
5
6
7
8
9

10 Red lines indicate symptoms that moved to later positions between the left-hand subgroup
11 and the right-hand subgroup. Blue lines indicate symptoms that moved to earlier positions.
12 Line opacity levels indicate the number of positions changed (higher opacity represents
13 more positions changed).

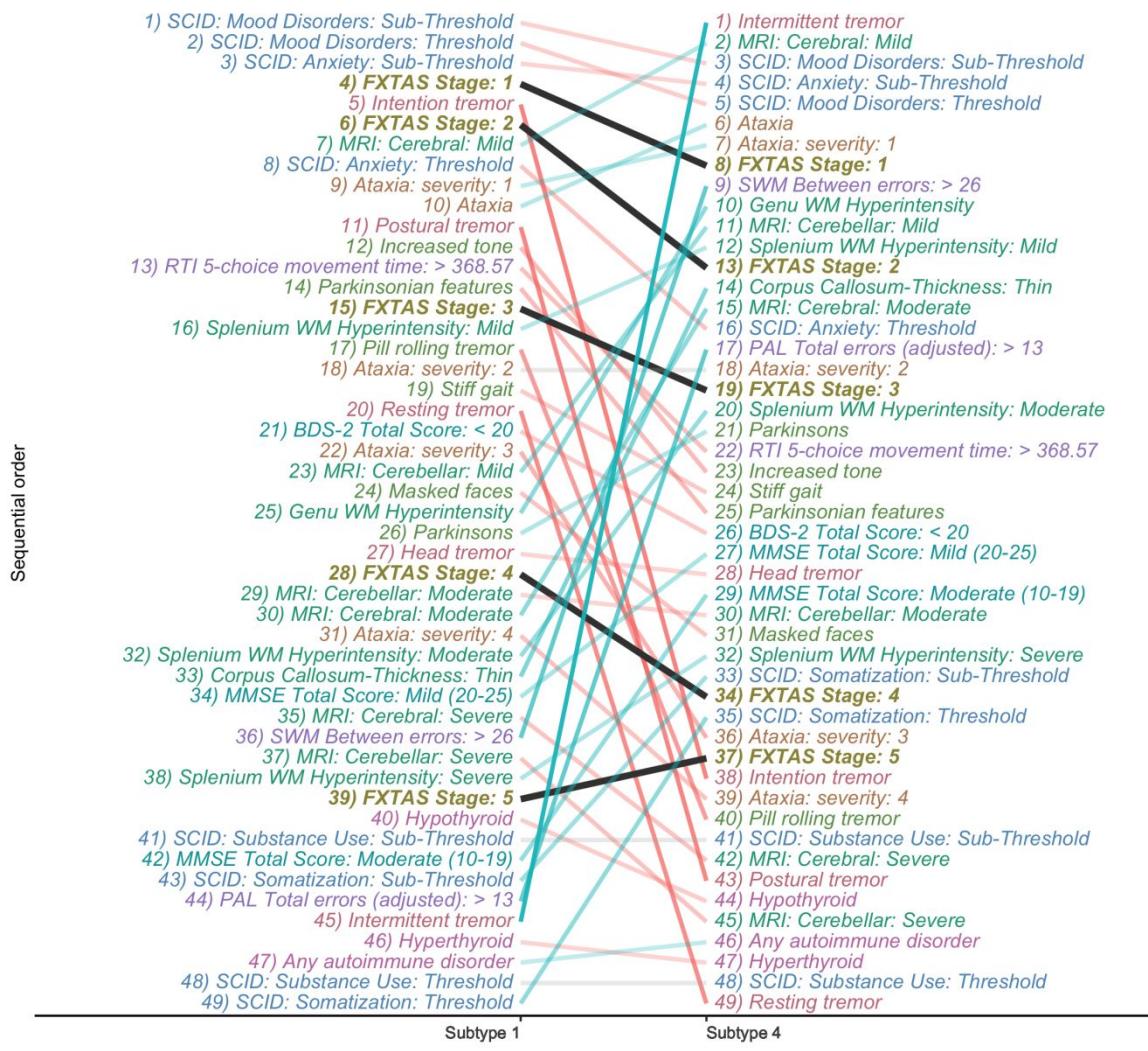
14 **(a) Subtype 1 compared to Subtype 2**
15



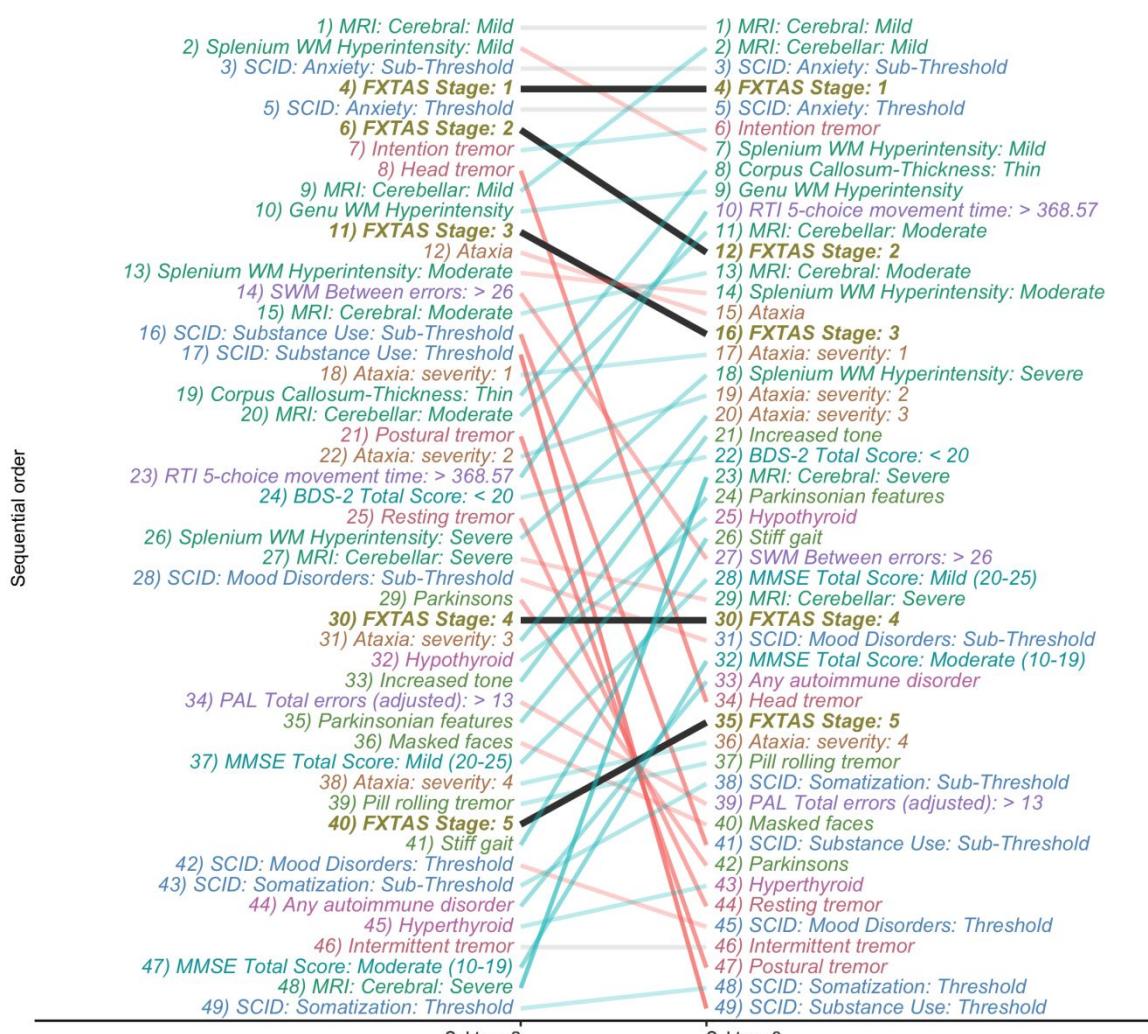
(b) Subtype 1 compared to Subtype 3



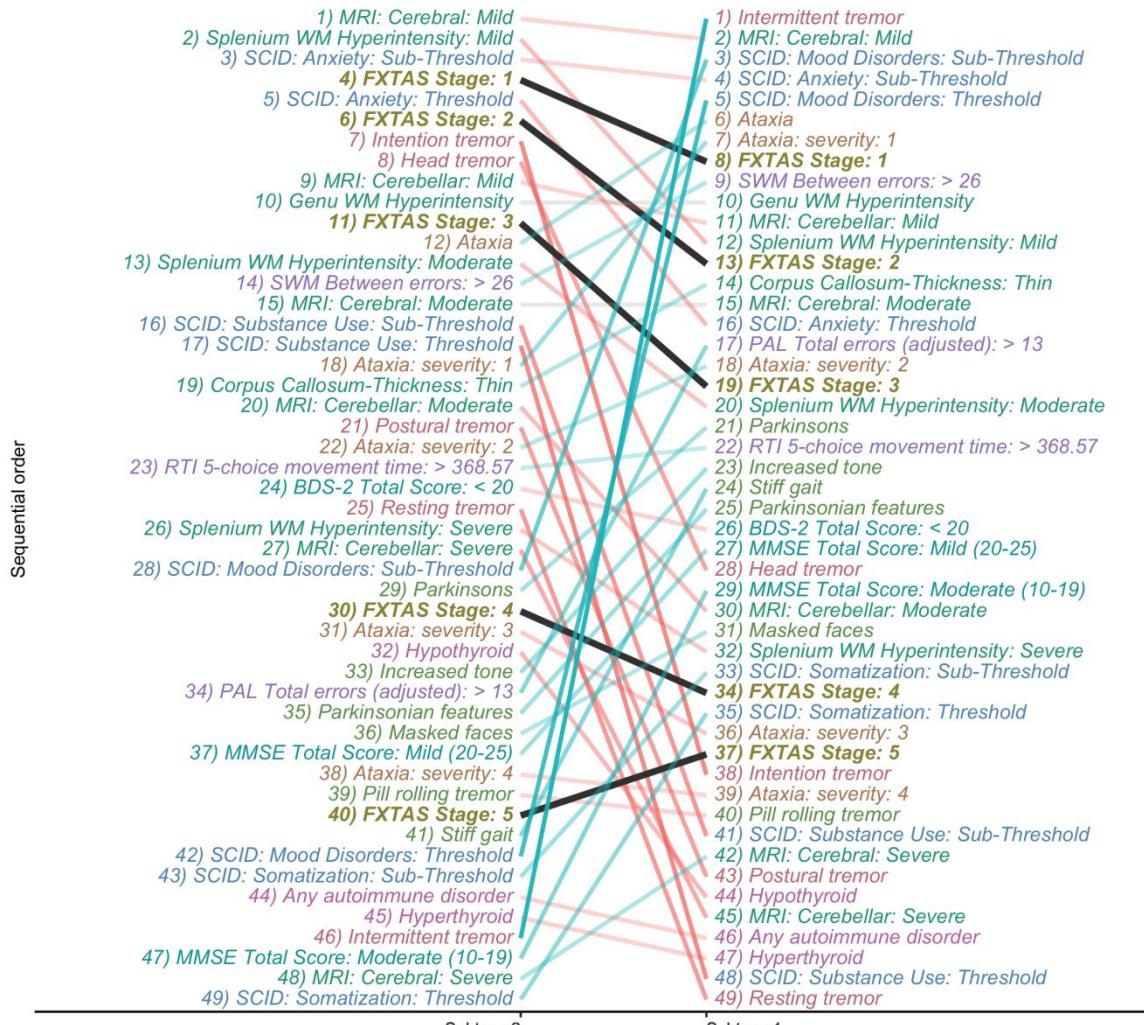
(c) Subtype 1 compared to Subtype 4



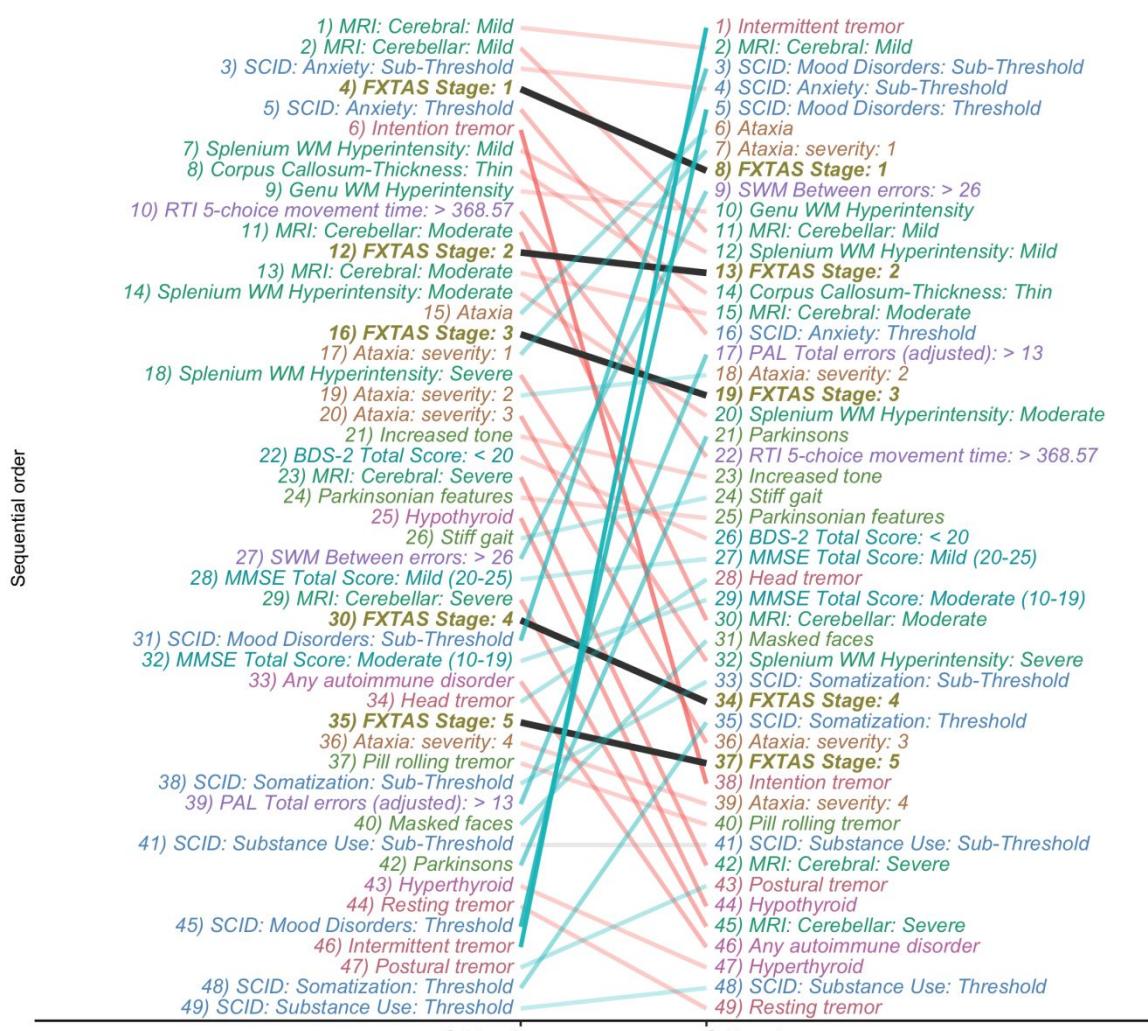
1
2
3
4 (d) Subtype 2 compared to Subtype 3
5
6



(e) Subtype 2 compared to Subtype 4



(f) Subtype 3 compared to Subtype 4



Analyses stratified by CGG repeats

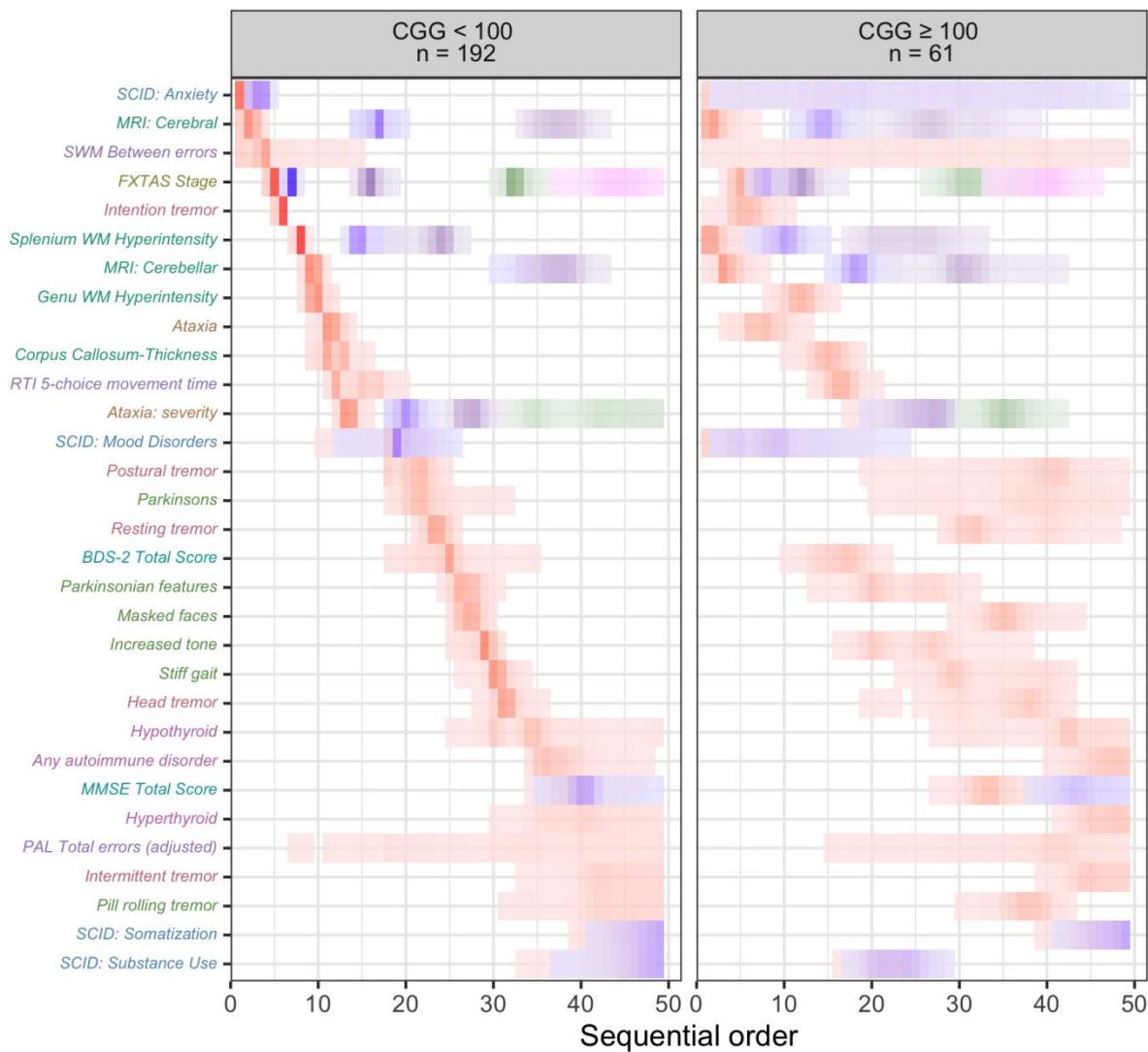
[Supplementary Fig. 5](#) shows the estimated sequences stratified by CGG repeat level. We did not find statistically significant evidence of a difference in event sequences between CGG < 100 and CGG ≥ 100 ($p = 0.176$), but some potential differences were noted ([Supplementary Fig. 5](#) and [Supplementary Fig. 6](#)). SWM between errors occurred prior to Stage 1 in participants with CGG repeats <100, whereas it occurred between stage 4 and 5 in those with CGG repeats ≥ 100 ([Supplementary Fig. 6](#)). Postural, resting, head, and intermittent tremor occurred earlier in those with CGG repeats <100 compared to those with CGG repeats ≥100. Several psychiatric disorders (as measured by SCID) occurred later in the event sequence in participants with CGG repeats <100 compared to those with CGG repeats ≥100.

The finding of only minimal differences between 55-99 vs 100-199 CGG repeats is surprising, since several studies have found that the higher the CGG repeats, the earlier the onset and the faster the progression of FXTAS.^{13,16,71} Perhaps the cut off of 100 is too high for this distinction to be made.

These differences were more evident among males than females (Supplementary material, [Comparing CGG levels stratified by sex](#)). When we fitted models for subgroup analyses by sex and CGG repeats level, the models for males showed the same patterns described above, except for head tremor (we found no difference in head tremor onset between CGG levels). [Supplementary Fig. 9](#) shows the corresponding stratified models and compares the estimates of the sequences for these two subgroups.

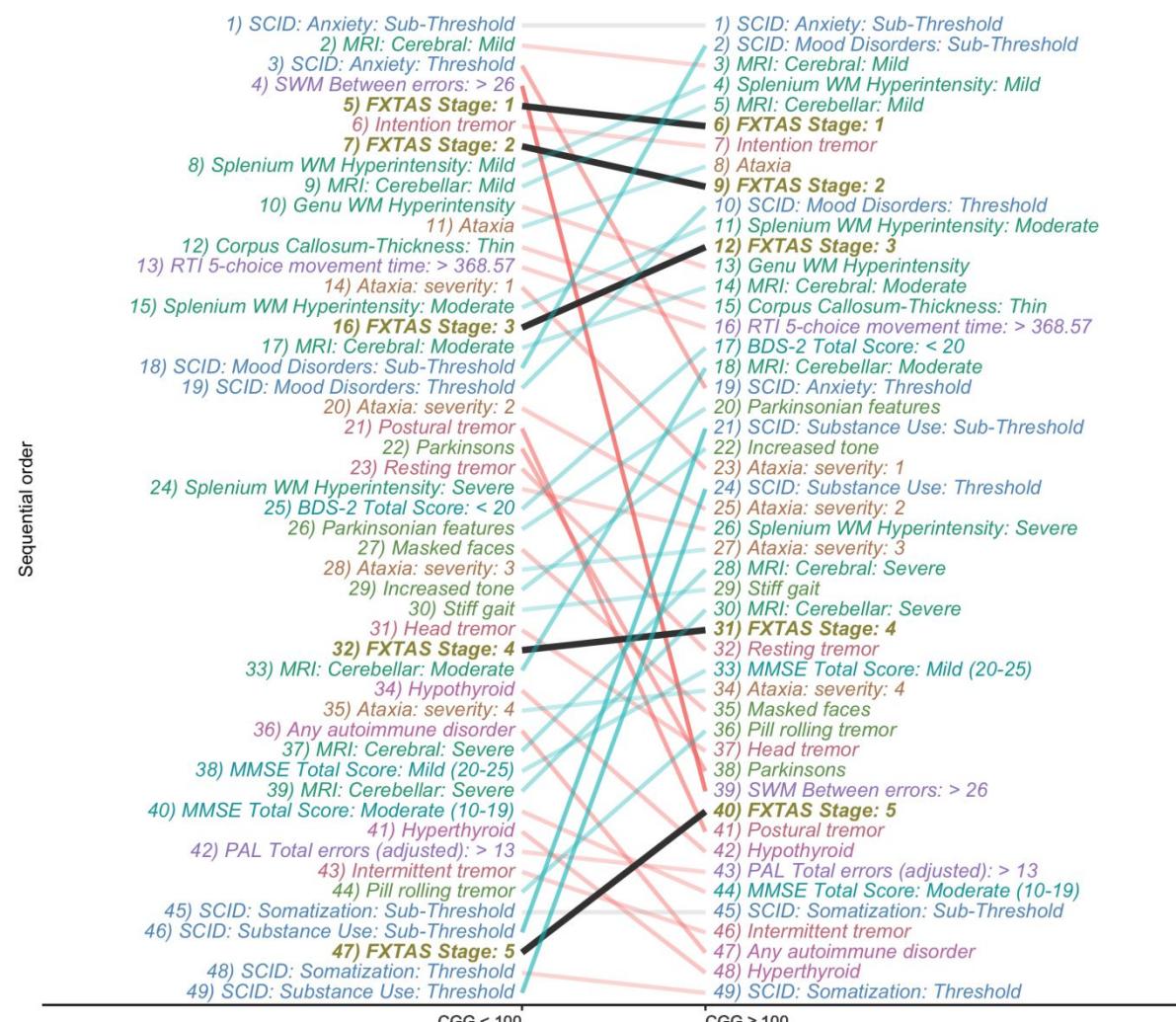
We did not find statistically significant evidence of a difference between CGG <100 and CGG ≥100 among females ($p = 0.596$). [Supplementary Fig. 10](#) shows the estimated stratified models and compares the estimates. Females had later onset of SWM Between Errors for CGG repeats ≥100 and later onset of intermittent tremors for CGG repeats <100, but did not show evidence of the other patterns observed for males and for the models without sex-stratification. Limited sample size in females for subgroup analyses was likely an issue; the available data included only 25 females with CGG ≥100.

Supplementary Fig. 5: Event sequences stratified by CGG repeats (<100 vs 100+). The different colors (red, blue, purple, green, magenta) indicate the ordinal levels of symptom progression. Color gradient intensity represents the likelihood of sequence position. The brighter the color, the more likely that the corresponding symptom event occurs in that position in the sequence.



1
2
3
4 **Supplementary Fig. 6: Event sequences stratified by CGG repeats (<100 vs 100+).**

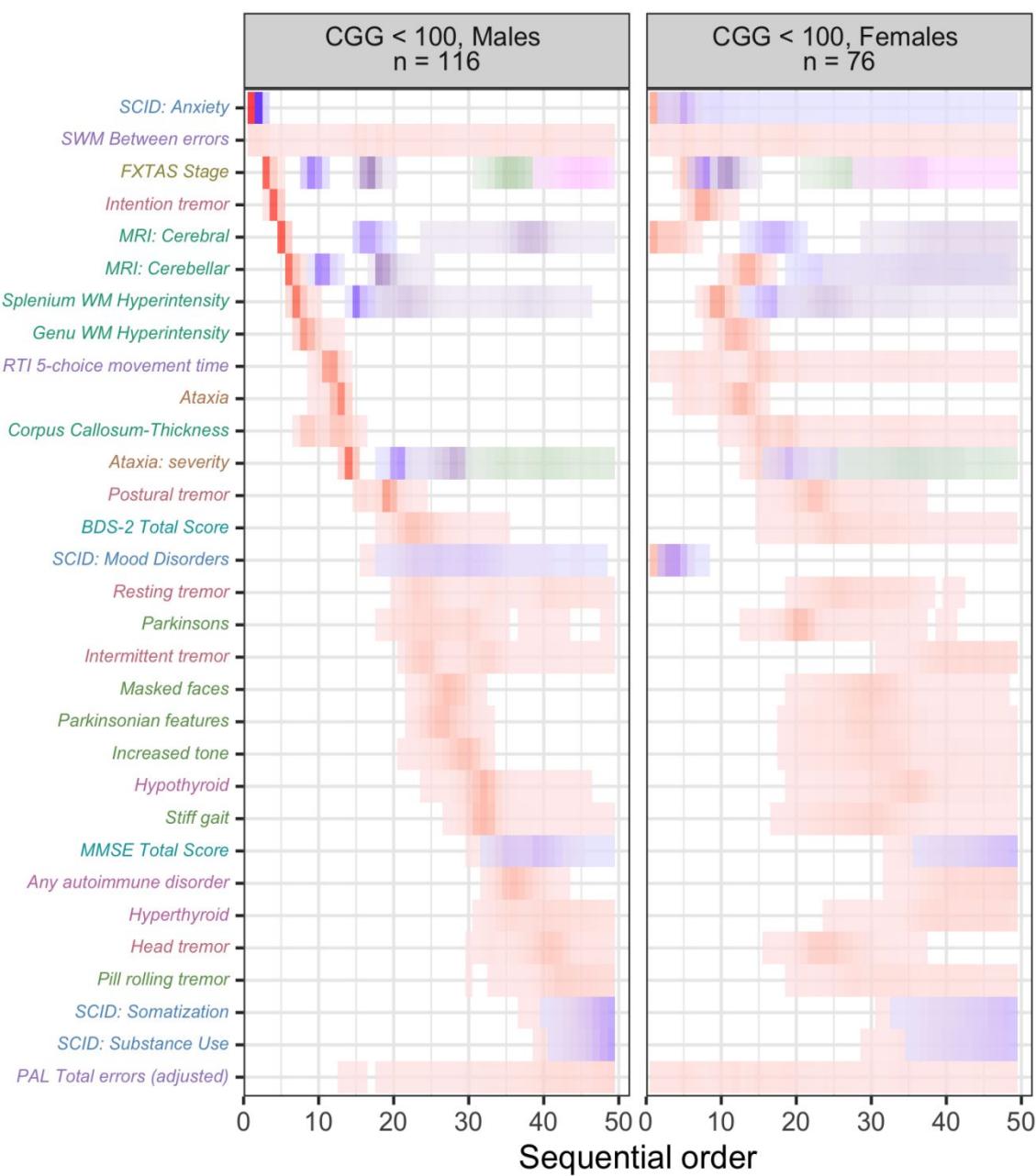
5 Differences in event sequences between CGG repeat sizes. Red lines indicate symptoms
6 that moved to later positions between the left-hand subgroup and the right-hand subgroup.
7 Blue lines indicate symptoms that moved to earlier positions. Line opacity levels indicate
8 the number of positions changed (higher opacity represents more positions changed).
9
10



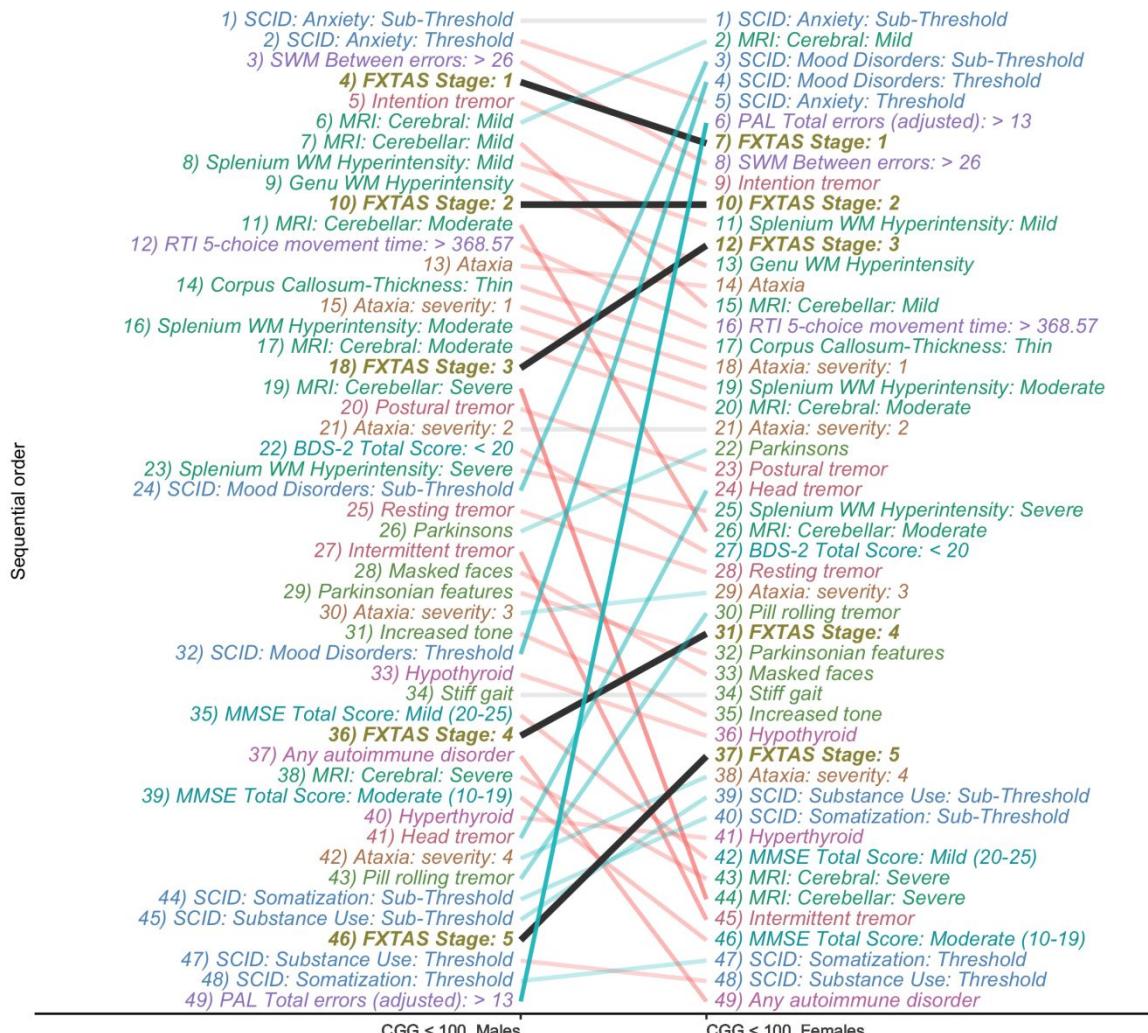
Comparing sexes stratified by CGG level

Supplementary Fig. 7: Event sequences stratified by sex, for CGG repeats <100. **(a)** The different colors (red, blue, purple, green, magenta) indicate the ordinal levels of symptom progression. Color gradient intensity represents the likelihood of sequence position. The brighter the color, the more likely that the corresponding symptom event occurs in that position in the sequence. **(b)** Differences in event sequences by sex, for CGG repeats <100.

(a)

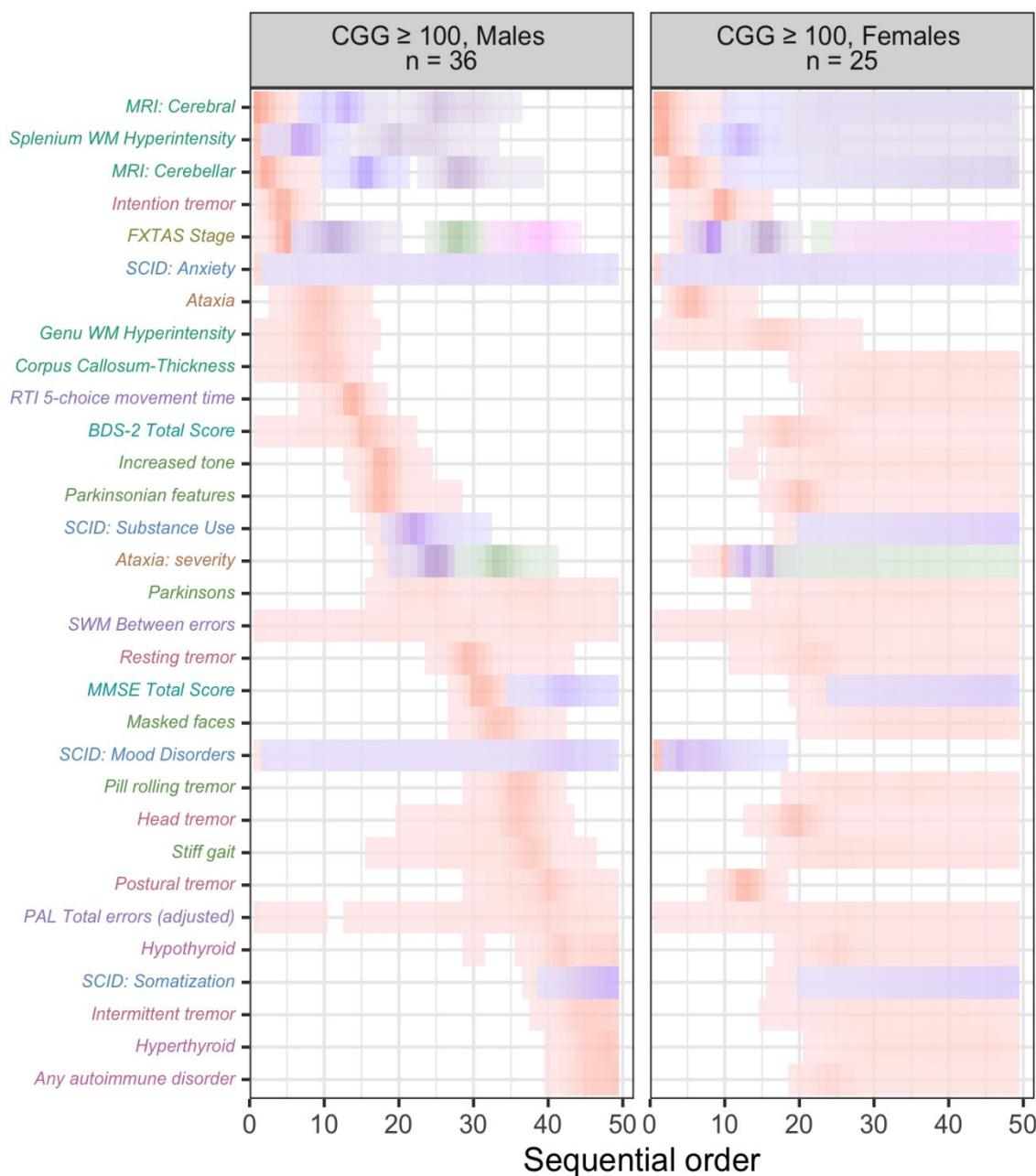


(b)

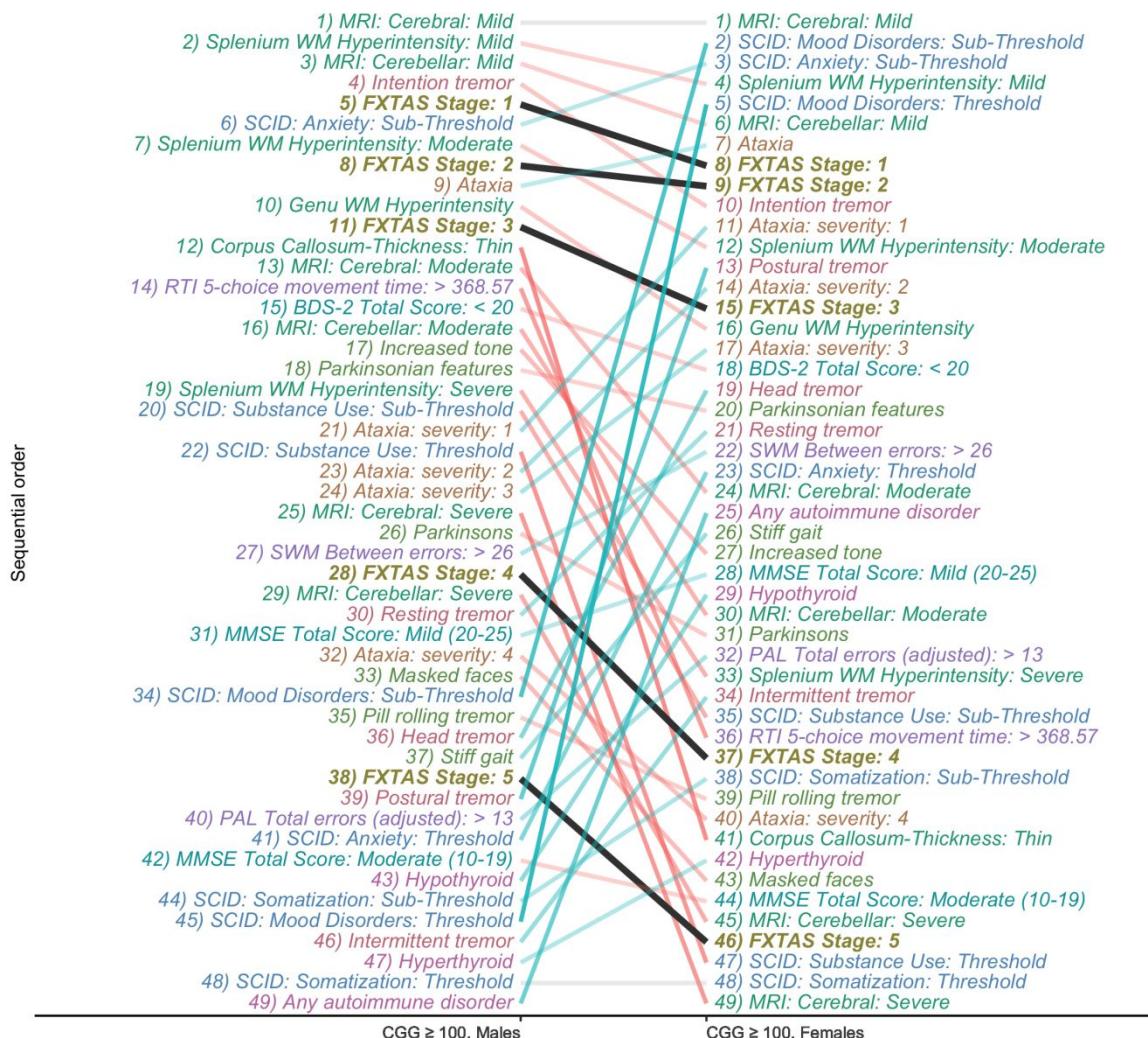


Supplementary Fig. 8: **Event sequences stratified by sex, for CGG repeats ≥ 100 .** (a) The different colors (red, blue, purple, green, magenta) indicate the ordinal levels of symptom progression. Color gradient intensity represents the likelihood of sequence position. The brighter the color, the more likely that the corresponding symptom event occurs in that position in the sequence. (b) Differences in event sequences by sex, for CGG repeats ≥ 100 .

(a)



(b)

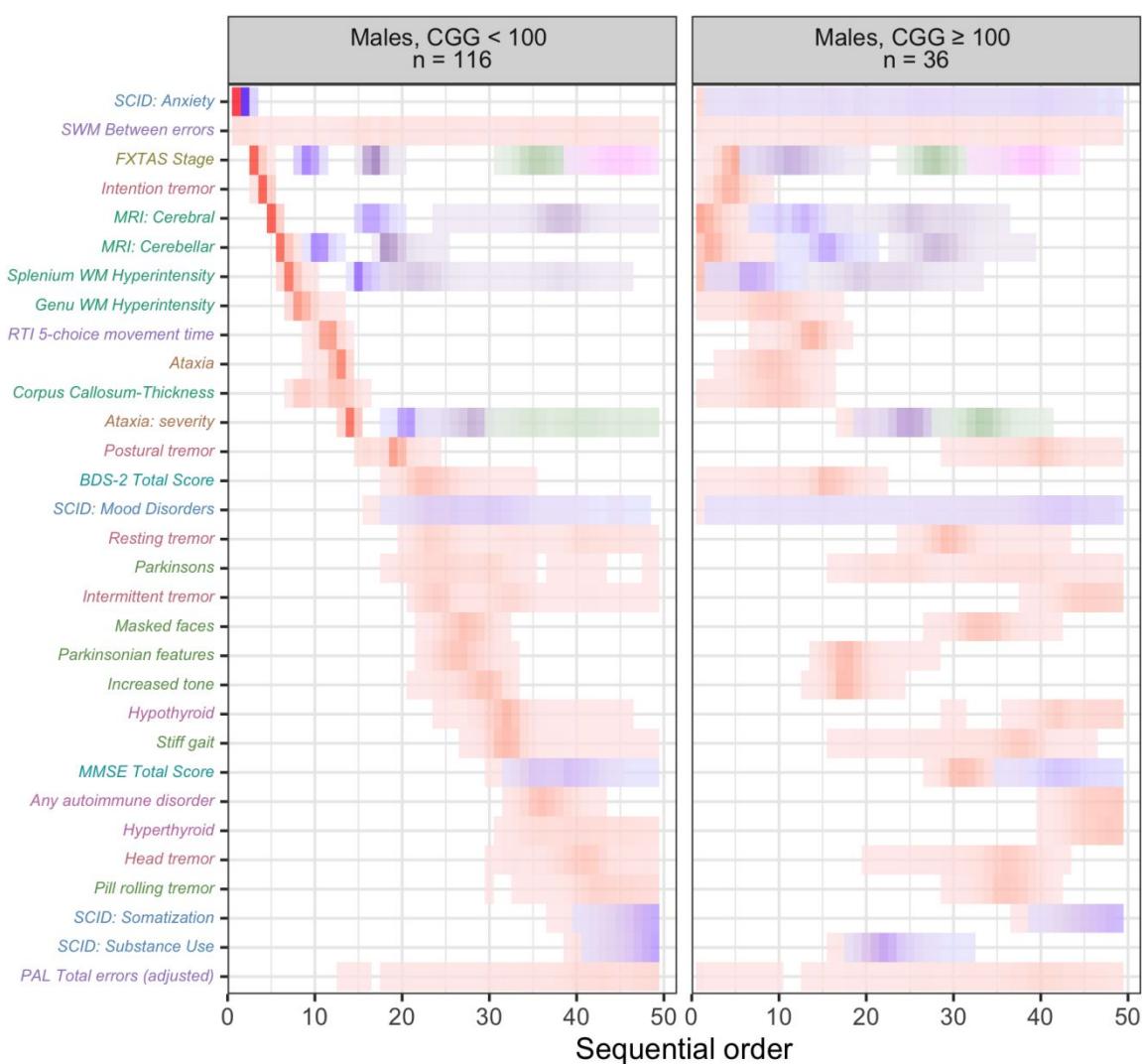


Comparing CGG levels stratified by sex

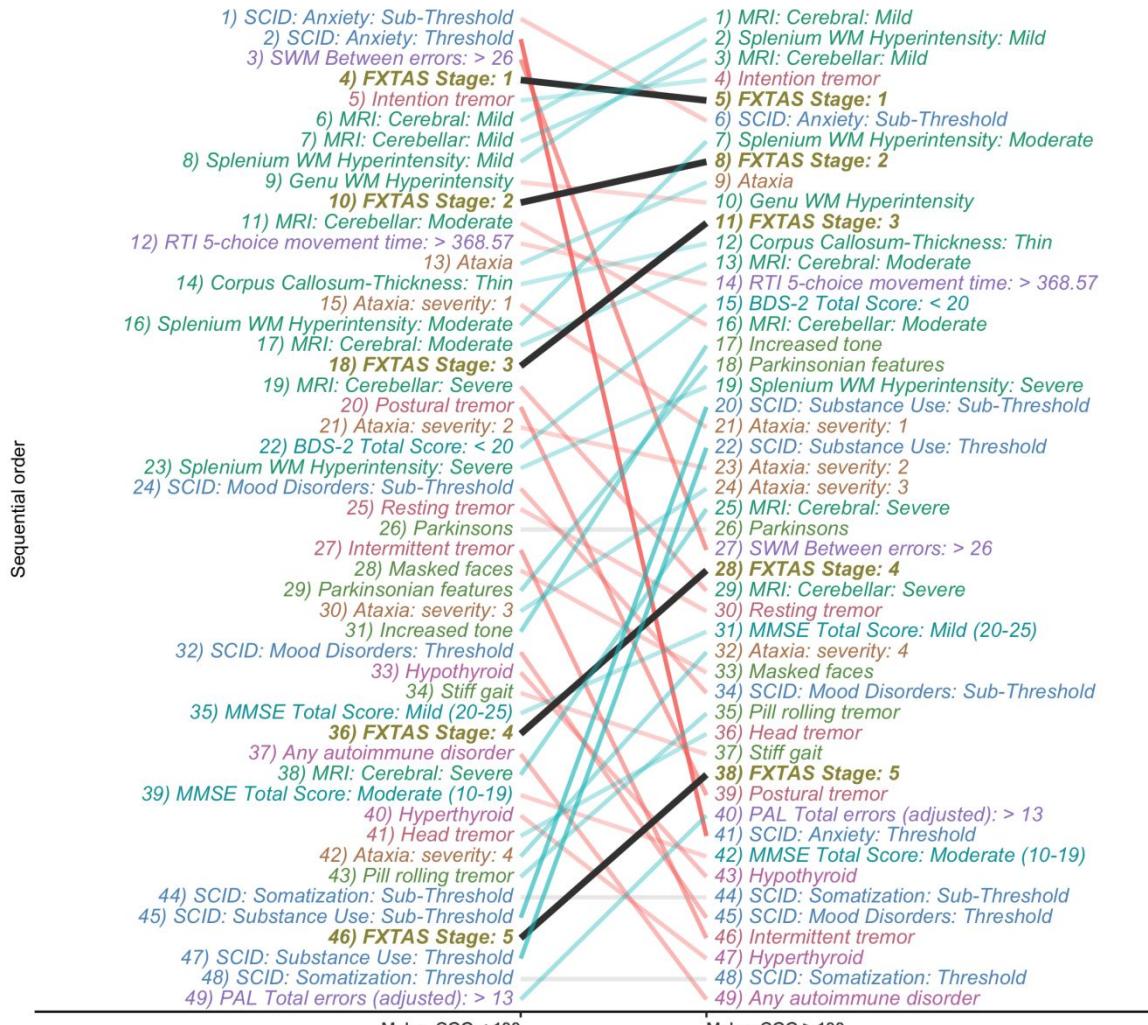
Supplementary Fig. 9: Event sequences stratified by CGG repeats, among males. (a) Estimated event sequences, stratified by CGG repeats (<100 vs 100+). The different colors (red, blue, purple, green, magenta) indicate the ordinal levels of symptom progression. Color gradient intensity represents the likelihood of sequence position. The brighter the color, the more likely that the corresponding symptom event occurs in that position in the sequence. (b) Positional differences in estimated event sequence between CGG repeat levels. Red lines indicate symptoms that moved to later positions between the left-hand subgroup and the right-hand subgroup. Blue lines indicate symptoms that moved to earlier

positions. Line opacity levels indicate the number of positions changed (higher opacity represents more positions changed).

(a)



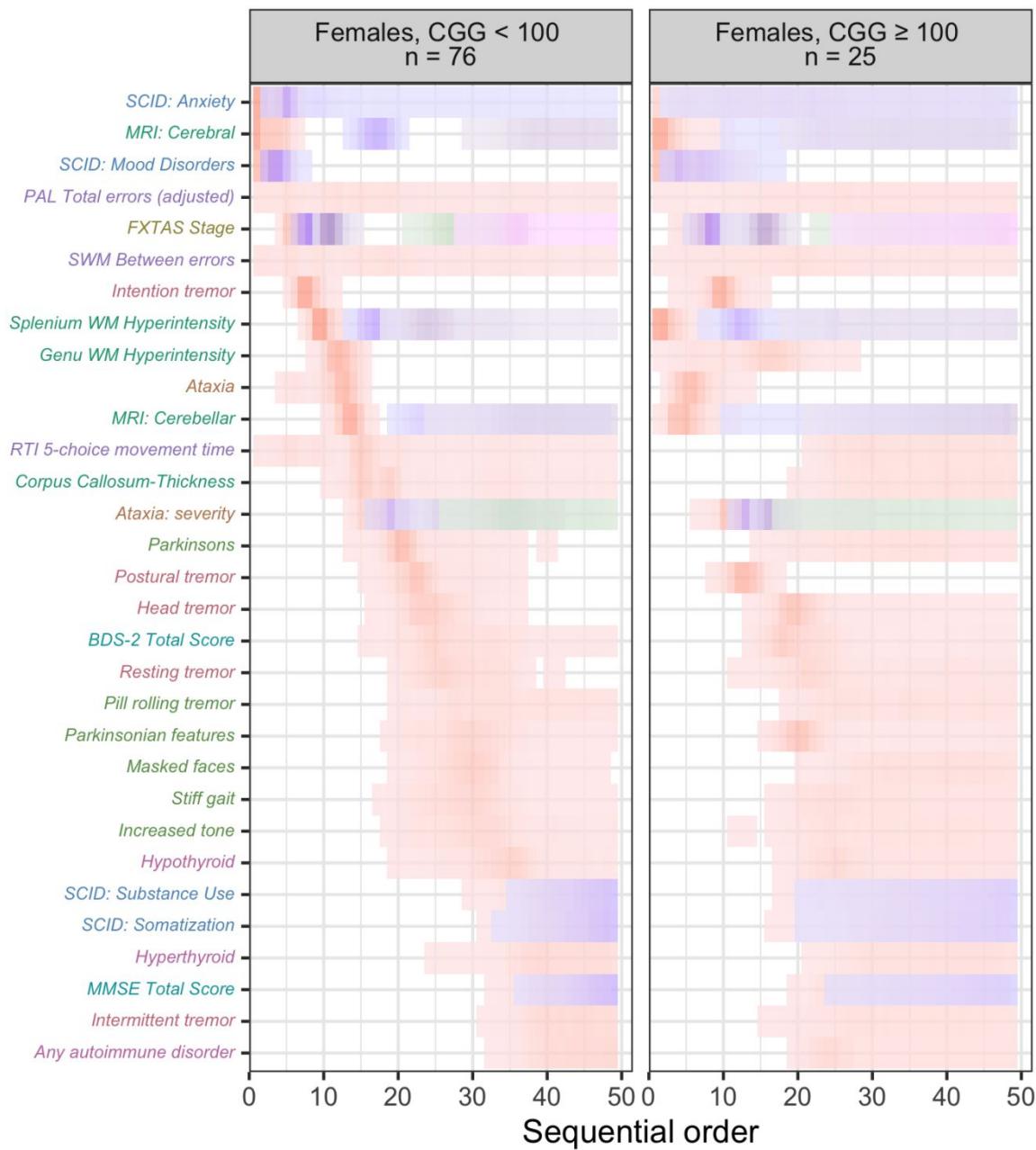
(b)

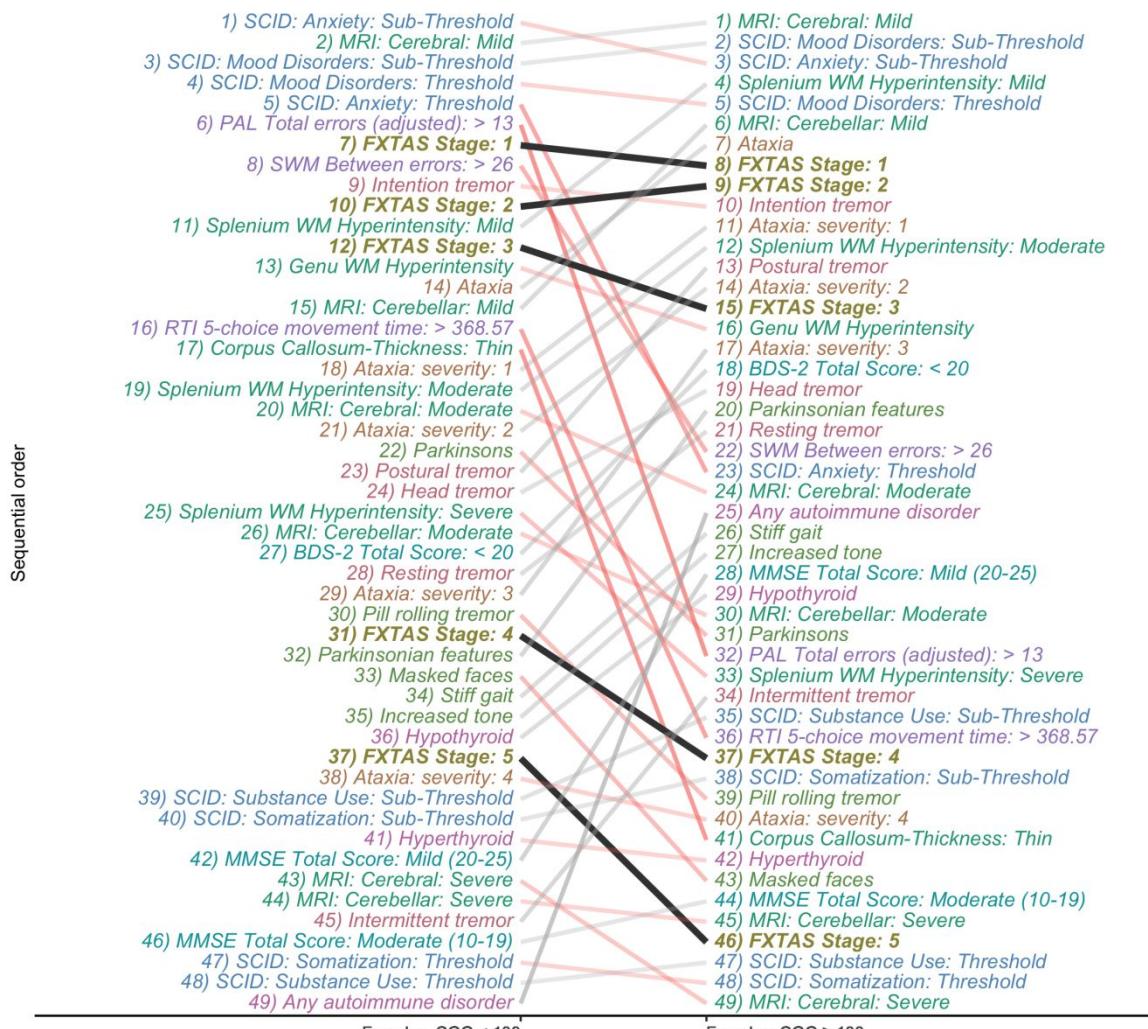


1
2
3
4 Supplementary Fig. 10: **Event sequences stratified by CGG repeats, among females.** **(a)**
5 Estimated event sequences, stratified by CGG repeats (<100 vs 100+). The different colors
6 (red, blue, purple, green, magenta) indicate the ordinal levels of symptom progression.
7 Color gradient intensity represents the likelihood of sequence position. The brighter the
8 color, the more likely that the corresponding symptom event occurs in that position in the
9 sequence. **(b)** Positional differences in estimated event sequence between CGG repeats.
10 Red lines indicate symptoms that moved to later positions between the left-hand subgroup
11 and the right-hand subgroup. Blue lines indicate symptoms that moved to earlier positions.
12 Line opacity levels indicate the number of positions changed (higher opacity represents
13 more positions changed).

14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

(a)



1
2
3
4
5
(b)

1
2
3
4 **STROBE** statement: Reporting guidelines checklist for cohort, case-control and cross-sectional studies
5

SECTION	ITEM NUMBER	CHECKLIST ITEM	REPORTED ON PAGE NUMBER:
TITLE AND ABSTRACT			
	1a	Indicate the study's design with a commonly used term in the title or the abstract	1
	1b	Provide in the abstract an informative and balanced summary of what was done and what was found	1
INTRODUCTION			
Background and objectives	2	Explain the scientific background and rationale for the investigation being reported	3
	3	State specific objectives, including any pre-specified hypotheses	4
METHODS			
Study design	4	Present key elements of study design early in the paper	4
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4
Participants	6a	Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants	4
	6b	Cohort study—For matched studies, give matching criteria and number of exposed and unexposed Case-control study—For matched studies, give matching criteria and the number of controls per case Variables	5
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5

SECTION	ITEM NUMBER	CHECKLIST ITEM	REPORTED ON PAGE NUMBER:
Data sources/measurements	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group.	6
Bias	9	Describe any efforts to address potential sources of bias.	N/A
Study size	10	Explain how the study size was arrived at	4
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why .	5-6
Statistical methods	12a	Describe all statistical methods, including those used to control for confounding	6-8
	12b	Describe any methods used to examine subgroups and interactions	7
	12c	Explain how missing data were addressed	7
	12d	Cohort study—If applicable, explain how loss to follow-up was addressed Case-control study—If applicable, explain how matching of cases and controls was addressed Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy	N/A
	12e	Describe any sensitivity analyses	N/A
RESULTS			
Participants	13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	8
	13b	Give reasons for non-participation at each stage	N/A
	13c	Consider use of a flow diagram	N/A
Descriptive Data	14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8, 23
	14b	Indicate number of participants with missing data for each variable of interest	29-33
	14c	Cohort study—Summarise follow-up time (eg, average and total amount)	N/A
Outcome Data	15*	Cohort study—Report numbers of outcome events or summary measures over time	29-33

SECTION	ITEM NUMBER	CHECKLIST ITEM	REPORTED ON PAGE NUMBER:
		Case-control study—Report numbers in each exposure category, or summary measures of exposure Cross-sectional study—Report numbers of outcome events or summary measures	
Main Results	16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g. 95% confidence interval). Make clear which confounders were adjusted for and why they were included	8-11
	16b	Report category boundaries when continuous variables were categorized	5-6
Is this a repeat of #11? Not sure what they're getting at	16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
	16d	Report results of any adjustments for multiple comparisons	N/A
Other Analyses	17a	Report other analyses done—e.g. analyses of subgroups and interactions, and sensitivity analyses	9-11
	17b	If numerous genetic exposures (genetic variants) were examined, summarize results from all analyses undertaken	N/A
	17c	If detailed results are available elsewhere, state how they can be accessed	15
DISCUSSION			
Key Results	18	Summarise key results with reference to study objectives	11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	13-14
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	11-13
Generalisability	21	Discuss the generalisability (external validity) of the study results Other information	14
FUNDING			

SECTION	ITEM NUMBER	CHECKLIST ITEM	REPORTED ON PAGE NUMBER:
	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	14

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

For Review Only