SUPPLEMENTAL METHODS

NIH EXAMINER Subtests

Working Memory

Dot Counting: In this verbal working memory task, examinees are shown an array of green circles, blue circles, and blue squares and are asked to count the number of blue circles. The examinees are shown 2-7 trials and asked to count the number of blue circles on each. After all displays are shown, the examinee recalls the number of blue circles in each of the displays, in the correct order. Partial credit is given based on the number of correctly recalled totals.

Spatial N-Back: The NIH-EXAMINER N-back is a spatial version of the commonly used N-back working memory paradigm that requires flexible updating of information. Participants were shown a series of white squares in different positions on the screen. In the 1-back task, participants were asked to respond whenever the position of the square matched the position of the last square (1-back) that was presented. In the 2-back, participants responded if the square's position matched the position of the square shown two squares previously. The primary outcome was a discriminability index (d-prime) incorporating hits and false positives.

Cognitive Control:

Flanker: A row of five arrows was presented, and the examinee was instructed to focus on the center arrow. The examinee was then asked to indicate whether the center arrow pointed toward the right or left. There were two conditions interspersed during the task: either the remaining arrow pointed in the direction congruent or incongruent with the center arrow. The latter requires cognitive control and response inhibition. The outcome was a single variable combining accuracy and reaction time [1].

Dimensional Set-Shifting: This computerized set-shifting task was modeled after a previously published paradigm [2,3]. Examinees are asked to match a stimulus at the top of the screen with one of two stimuli in either corner at the bottom of the screen. The stimulus also indicates whether they are to match on color or shape. In task-homogenous block, examinees match either shape or color, whereas in the task-heterogeneous blocks, they must switch between matching shape and color. The outcome variable was again a composite of accuracy and reaction time [1].

Fluency:

Phonemic Fluency: Examinees were asked to quickly names as many words as possible beginning with a specific letter of the alphabet over two 60-second trials. They were not allowed to name people, places, numbers, or grammatical variants of previous responses.

Categorical Fluency: Examinees were asked to quickly name as many items as they could that belonged to a given category over two 60-second trials.

The outcomes for fluency were total correct words generated.

Genetic Testing

Samples were screened using targeted sequencing of a panel of genes previously implicated in neurodegenerative disorders. Exonic regions were captured using a custom designed library (SeqCap EZ

Choice Library, NimbleGen) and sequenced on an Illumina HiSeq4000 at the UCLA Neuroscience Genomics Core (http://www.semel.ucla.edu/ungc). Sequence reads were mapped to the GRCh37/hg19 reference genome, variants were joint-called with GATK [4], and annotated using ANNOVAR and the Ensembl Variant Effect Predictor tool [5,6]. The coding and exon-intron boundary regions of GRN and MAPT genes were screened for known (listed in the AD&FTD Mutation Database: http://www.molgen.ua.ac.be/ADMutations) or novel (likely) pathogenic variants (classified according to the American College of Medical Genetics and Genomics and the Association for Molecular Pathology guidelines) [7]. All pathogenic variants found were subsequently confirmed by Sanger sequencing. The presence of a pathological hexanucleotide repeat expansion in C9orf72 was detected using both fluorescent and repeat-primed PCR, as previously described [8].

Neuroimaging

The current analysis used the T1 weighted images, which were acquired as Magnetization Prepared Rapid Gradient Echo (MP-RAGE) images using the following parameters: 240x256x256 matrix; about 170 slices; voxel size = 1.05x1.05x1.25 mm^3; flip angle, TE and TR varied by vendor. Image processing is described in detail in the supplemental methods.

Image Processing

Before any prepossessing of the images, all T1-weighted images were visually inspected for quality control. Images with excessive motion or image artifact were excluded. The sample with neuroimaging that passed quality control included 80 mutation carriers with 140 visits. T1-weighted images undergone bias field correction using N3 algorithm, the segmentation was performed using SPM12 (Wellcome Trust Center for Neuroimaging, London, UK, http://www.fil.ion.ucl.ac.uk/spm) unified segmentation [9]. A group template was generated from segmented gray and white matter tissues and cerebrospinal fluid by non-linear registration template generation using Large Deformation Diffeomorphic Metric Mapping (LDDMM) framework [10]. Subjects' native space gray and white matter were normalized, modulated and smoothed in the group template. The applied smoothing used a Gaussian kernel with 10~mm full width half maximum. Every step of the transformation was carefully inspected from the native space to the group template. For statistical purposes, linear and non-linear transformations between the group template space and International Consortium of Brain Mapping (ICBM) [11] was applied. A standard parcellation atlas (Desikan et al., 2006) was transformed into ICBM space.

Statistical analyses.

Effect Sizes

Effect sizes for each measure were evaluated by calculating sample sizes (per arm) required to detect 25% and 40% reductions in decline [12], using 10,000-fold bootstrapping as described in supplemental methods. Sample size estimates were calculated using annualized change scores and the standard deviation of this annualized change. Annualized change was defined as the difference in performance between participants' first and second visit, divided by the time interval between these visits. We chose to use a two-time point calculation of change, rather than using parameters from the LME model, to remain consistent with standard clinical trial methodology. Each sample size estimate was obtained using the *boot* package from 'R', and 95% confidence intervals were estimated using a 10,000-fold bootstrapping procedure.

Follow-up Analyses

In follow up analyses, we assessed the associations between the NIH-EXAMINER Executive Composite score, FTLD-CDR-SB, and volume of the 4 lobar volumes (see supplemental methods for details). Following Neuhaus et al. [13,14], we decomposed the Executive Composite score into within-(time-varying) and between-subject (time-invariant) components to directly relate purely within-subject change in Executive Composite with changes in brain volume and FTLD-CDR-SB and to avoid estimation bias resulting from incorrectly assuming common within- and between-subject effects. We first calculated a time-invariant mean score across visits for each person. We then subtracted each participant's mean from his/her EXAMINER composite at each time point to estimate a mean-centered, within-person metric of change. Both were entered as predictors into LME models with FTLD-CDR-SB or neuroimaging ROIs as the outcomes. Covariates included gender, baseline age and education, as well as their interactions with time. Total intracranial volume and its interaction with time were entered in the models with neuroimaging outcomes. We also analyzed the association between the Executive Composite and lobar volumes in the non-carrier controls to further validate the association between this measure and neural tissue.

[1] Weintraub S, Dikmen SS, Heaton RK, Tulsky DS, Zelazo PD, Bauer PJ, et al. Cognition assessment using the NIH Toolbox. Neurology 2013;80:S54–64. doi:10.1212/WNL.0b013e3182872ded.

[2] Monsell S. Task switching. Trends Cogn Sci 2003;7:134–40.

[3] Kray J, Lindenberger U. Adult age differences in task switching. Psychol Aging 2000;15:126–47.

[4] McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:1297–303. doi:10.1101/gr.107524.110.

[5] Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164–e164. doi:10.1093/nar/gkq603.

[6] McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome Biol 2016;17:122. doi:10.1186/s13059-016-0974-4.

[7] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405–23. doi:10.1038/gim.2015.30.

[8] DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 2011;72:245–56. doi:10.1016/j.neuron.2011.09.011.

[9] Ashburner J, Friston KJ. Unified segmentation. Neuroimage 2005;26:839–51. doi:10.1016/j.neuroimage.2005.02.018.

[10] Ashburner J, Friston KJ. Diffeomorphic registration using geodesic shooting and Gauss-

Newton optimisation. Neuroimage 2011;55:954–67. doi:10.1016/j.neuroimage.2010.12.049.
[11] Mazziotta JC, Toga AW, Evans A, Fox P, Lancaster J. A probabilistic atlas of the human brain: theory and rationale for its development. The International Consortium for Brain Mapping (ICBM). Neuroimage 1995;2:89–101.
[12] Sakpal TV. Sample size estimation in clinical trial. Perspect Clin Res 2010;1:67–9.
[13] Neuhaus JM, Kalbfleisch JD. Between- and within-cluster covariate effects in the analysis of clustered data. Biometrics 1998;54:638–45.
[14] Neuhaus JM, McCulloch CE. Separating between- and within-Cluster Covariate Effects by Using Conditional and Partitioning Methods. J R Stat Soc Ser B (Statistical Methodol 2006;68:859–72.