10.1148/radiol.2021210383

# Appendix E1

## Statistical Analysis

Summary statistics were computed for demographics. General linear models were created to analyze structural connectivity and mutation status in all cognitively normal subjects. The following model was used:

$$M_{i,k} = \alpha_{MUT,k} + \alpha_{F,k} + \alpha_{ApoE,k} + b_{EYO,k}EYO + b_{Educ,k}Educ_i + b_{MMSE,k}MMSE_i + b_{NWM,k}NWM_{i,k} + b_{MUT*EYO,k}MUT*EYO_i + \varepsilon_k \quad \text{(E1)}$$

The outcome $M_i$ is the network metric (i.e., global efficiency) for subject *i* in network *k*. The term $\alpha_{MUT,k}$ denotes the additional effect of positive mutation status (M+) in network *k*. The term $\alpha_{F,k}$ denotes the additional effect of gender in females, in network *k*. The term $\alpha_{ApoE,k}$ denotes the additional effect of having the ApoE4 allele, in network *k*. The coefficients $b_{EYO,k}$, $b_{Educ,k}$, $b_{MMSE,k}$, and $b_{NWM,k}$ denote the effect of every additional unit of EYO, education (*Educ*), MMSE, and normalized total cerebral white matter volume (*NWM*) in network *k*. The coefficient $b_{MUT*EYO,k}$ denotes the interaction effect of positive mutation status (M+) and EYO in network k.

General linear models were then created for cognitively normal M+ subjects, which examined the relationship between white matter structural connectivity, amyloid burden and EYO. The following model was used:

$$M[MUT+]_{i,\mathrm{k}} = \alpha_{F,k} + \alpha_{ApoE,k} + b_{EYO,k}\mathrm{EYO} + b_{PIB,k}PIB_i + b_{Educ,k}Educ_i + b_{MMSE,k}MMSE_i + b_{NWM,k}NWM_i + \varepsilon_k \quad \text{(E2)}$$

The outcome $M[MUT+]_{I,k}$ is the network metric (i.e., global efficiency) for subject *i* in network *k*. The coefficients $b_{EYO,k}$ and $b_{PIB,k}$ denote the effect of every additional unit of EYO and SUVr calculated from the PiB amyloid PET images (PIB) of network *k*. The remainder of the terms are the same as in Equation 1 and described above.

## Correlation Network Analysis

Correlation network analysis is a data visualization technique in a form of a network in which nodes represent variables characterizing instances (subjects). Edges represent Spearman rank correlations existing between variables. Distance between nodes is defined as the inverse of the correlation: dist = 1/rho; small distance denotes large correlation. Fruchterman-Reingold force-directed graph technique is used to construct a layout in which strongly correlated concepts are next to each other and concepts that are strongly related to many other concepts are positioned in the center of the network. In this application, only edges denoting correlations with value rho ≥ 0.3 are presented. Four clinical cognitive measurements were included in the analysis: word recall, immediate; word recall, delayed; logical memory, immediate; and logical memory, delayed. Both EYO and age were included for M+ and M- subjects to demonstrate the

correlation between those variables. Correlation network graph visualization was performed using the *igraph* package in R.