Olfactory Deficits and Donepezil Treatment in Cognitively Impaired Elderly

NCT01951118

Document Date: 03/05/2012

Pfeffer Functional Activities Questionnaire (FAQ) is a widely used 10-item instrument that takes 5 minutes to administer and focuses on instrumental, social and cognitive functioning[80]. We published the first report showing that informant-reported, but not self-reported, FAQ deficits strongly predicted conversion from MCI to AD[102]. We recently showed that specific FAQ items, particularly financial items and remembering appointments or to take medications, strongly distinguished patients with AD, MCI and healthy controls in baseline ADNI data[7]. The FAQ also showed improvement in an MCI trial using the AChel galantamine[50]. The FAQ may be less sensitive than the E-cog because it is short and does not tap into a wide range of functions. The FAQ is part of NACC and will be the secondary function measure. The informant is now required in all cases for study entry; some patients living alone without any available informant will not be eligible. The informant completes the FAQ[19,102] and the E-Cog[36].

Treatment Emergent Symptom Scale (TESS) is widely used to evaluate somatic side effects. For each item, a rating is made on a 3-point scale, with an additional rating on the likelihood that the medication caused the side effect. The total score with this safety assessment of possible AChel side effects will be an ancillary measure.

Apolipoprotein E genotype. In 501 elderly subjects, poor odor identification ability in combination with older age and the apolipoprotein E ε4 allele predicted greater global cognitive decline during follow-up.[76] This effect was not seen with a vocabulary test that used a multiple choice format similar to the odor identification test[76]. A similar association between the shorter 12-item CCSIT version of the UPSIT and the apolipoprotein E ε4 allele was reported in an epidemiological study[40]. Odor identification deficits characterize subjects with familial and genetic risk factors for AD[42] and predict episodic verbal memory decline over time in elderly subjects[42,101,112]. Therefore, we will evaluate family history in detail (NACC assessment) and apolipoprotein E genotype. Apolipoprotein-E genotype will be examined as a potential moderator of UPSIT prediction of donepezil response. Blood obtained will be stored at the Human Genetics Resources Core (HGRC) at Columbia University and annually sent in a batch to Prevention Genetics for identification of apolipoprotein E genotype. Using a standard protocol, DNA is amplified by the polymerase chase reaction (PCR). The genotypes are determined by the sizes of DNA fragments present. Recent evidence suggests that other genes, including clusterin, complement receptor 1 and phosphatidylinositol binding clathrin are associated with AD[91], but their effects are not as strong as apolipoprotein E and their associations with olfaction are unknown. Therefore, we will not study them and will not bank DNA because of the relatively small sample size for genetic analyses.

STATISTICS. General considerations and approach to multiple comparisons. In this proof-of-concept study, we do not consider it essential or even important to control for making one or more Type I errors between the two different study samples (MCI, AD) or between any of the cognitive or global clinical endpoints (SRT, ADAS-cog, CIBIC-plus). Testing this study's hypothesis (do changes in UPSIT predict outcome?) in these different samples with these different measures represents very different research questions, such that a positive finding from one sample or with one outcome cannot be claimed to support a positive result in a different sample or with a different outcome. However, for a given patient sample and a given outcome, our analysis will consider four parameters of interest, any one of which could be used as evidence supporting the scientific hypothesis that changes in UPSIT predict the given outcome in the given sample. This creates a four-dimensional parameter space and a corresponding family of hypotheses within which we will control the probability of making one or more Type I errors at 5%. Therefore, we will repeat the analyses described below for each study sample and for each endpoint, controlling the family-wise type I error rate in each repetition at the 0.05 level. Any positive findings will be viewed as hypothesis-generating for a subsequent, larger confirmatory study in the given sample and with the given cognitive or overall clinical endpoint.

Notation and statistical analysis model. Let Y_{it} denote the value for subject i at time index t for a given outcome measure for patients in a given study sample. Here i = 1,...,N and the time index t takes values t=0 for baseline, t=1 for the measurement at week 26, and t=2 for the measurement at week 52. For t =1 and 2, define the change-score $\Delta Y_{it} = Y_{it} - Y_{i0}$. Let U_{it} denote the UPSIT value for subject i at time index t, where now t=0 corresponds to the pre-atropine challenge measurement, t=1 to the post-challenge measurement, and t=2 to the week-eight measurement. Define the change-score $\Delta U_{it} = U_{it} - U_{i0}$. Thus ΔU_{i1} is the change (typically a decrease) in the UPSIT due to the atropine challenge and ΔU_{i2} is the change (typically an increase) from baseline to week eight. The statistical model we will use is a bivariate multiple regression model: $\Delta Y = \beta' X + E$, where ΔY is the 2 x 1 response vector $(\Delta Y_{i1}, \Delta Y_{i2})'$, β is a 7 x 2 matrix of multivariate regression coefficients, X is the 7 x 1 vector of explanatory factors (ΔU_{i1}, ΔU_{i2}, age, sex=female, education ≥ 8 yrs, baseline MMSE, and an intercept dummy), and E is a 2 x 1 vector of errors terms which we will assume has a bivariate normal distribution with mean 0 and 2 x 2 variance-covariance matrix Σ for each given X. [If there is an unexpected departure from normality in any outcome measure, as evidenced by visual inspection of residual histograms and a formal Kolmogorov-Smirnov goodness of fit statistic at p<.10, we will use a

normalizing transformation for ΔY in an attempt to restore the normality of the residuals.] The first two rows of β are our parameters of interest: slope coefficient β_{rs} is the expected change in mean ΔY_{is} per unit change in ΔU_{ir} (for r,s=1,2). They will be estimated by ordinary least squares (OLS) with 95% confidence intervals. We will test the global null hypothesis of no ability of UPSIT to predict the cognitive or clinical outcome by testing H_0: $\beta_{11}=\beta_{12}=\beta_{21}=\beta_{22}=0$, at level alpha=0.05, using an F test with 4 and N–14 degrees of freedom. Simultaneous 95% confidence intervals for each coefficient, and for linear contrasts between coefficients will be prepared using the standard union-intersection principle for multivariate analysis.

Secondary analyses. (i) All measures other than the hypothesized predictor and outcome variables will be evaluated in secondary analyses, e.g., E-cog and FAQ as functional outcomes. Pearson correlation coefficients will be used to examine the associations between these measures and UPSIT, other cognitive measures, and continuous baseline demographic variables. Student's t test will compare differences for dichotomous variables (e.g., sex). The statistical analytic approach will be similar for testing the primary hypotheses. (ii) In another secondary analysis we will attempt to confirm the findings of the British study[108] using their choice of dichotomization at 4 or more points change in week eight UPSIT for predicting outcome. We will not use that (or any other) dichotomization for our primary analysis in order to retain full power. We will, however, be able to evaluate whether the four-point change in 8-week UPSIT cut-point is better or worse than other cut-points in predicting outcomes. For this we will use ROC methods for predicting a good clinical outcome, such as a 4-point improvement in SRT or ADAS-cog, or a 2-point improvement in CIBIC-plus. (iii) In other secondary analyses we will determine the robustness of our findings by preparing completers-only analyses (see Handling of Missing Data below) and comparing the results to the primary analyses.

Exploratory analyses. We will not assume any time-evolution curve for the responses at 26 and 52 weeks. The response-curve is unlikely to be linear over the year of follow-up, as evidence shows a plateau or a small decline in cognitive and global measures after an initial positive response to donepezil. In exploratory analyses we will check if a simple parametric model in time with subject-specific parameters can parsimoniously fit the response data and if the findings relating to UPSIT persist. We will check for interactions between the slope parameters β_{11}, β_{12}, β_{21}, β_{22} and the other covariates of age, sex, education, and baseline MMSE, each at the alpha = 0.10 level given reduced power to detect interactions compared with main effects. For exploratory hypothesis with baseline UPSIT as a significant predictor in both the MCI and AD samples, we will conduct separate simple regression analyses with covariates as appropriate. We will also explore the relative utility of the baseline UPSIT compared to the change scores by entering them in the same regression model and examining which has a larger standardized coefficient (i.e., the regression coefficient times the ratio of standard deviations of the predictor and dependent variables). For the exploratory hypothesis on prediction of conversion to AD (MCI sample only), logistic regression analyses will be conducted. For variables that may impact outcome, e.g., incident depression, stroke, in secondary analyses we will check for the effect of any such potential confounders as individual terms in a bivariate multiple regression model. Apolipoprotein-E genotype will be examined as a potential moderator of UPSIT prediction of donepezil response.

Missing data and handling of dropouts. All due diligence will be done to minimize the occurrence of losses to follow-up, as the best first-line defense of study validity. Close-out due to death from any cause is *not* a loss to follow-up as we will use worst-possible scores for patients who die. Dropout due to withdrawal of consent, moving away, and other forms of loss to follow-up will be handled as follows. Patients who drop out prior to their week eight visit will be omitted from the analysis, as one of the two key predictor variables will be unobservable as will every outcome change-score. In each study, we expect 5% attrition before the week eight visit. Other patients dropping out after week eight (estimated additional 10%, i.e., total 15% dropout in MCI, and an additional 15% or total 20% dropout in AD) will be included in the analysis by using their last observations carried forward. While this may underestimate the change in cognitive and clinical outcomes at 26 and 52 weeks, these measures are not expected to change much after the point of dropout, especially if the patient discontinues donepezil treatment. Patients switched from donepezil to another AChEI during the 52 weeks will be considered as completers assuming they remain in the study. All patients who discontinue donepezil, whether or not they switch to another AChEI, will still be followed for the entire 52 weeks (whenever possible) for the primary intent-to-treat analyses.

Power and sample size considerations**.** In the multivariate regression model the ordinary least squares estimate of the vectorized regression coefficient $\beta^* = \beta^V$, has variance-covariance matrix $\Sigma \otimes (X'X)^{-1}$ where $\otimes$ is the Kronecker product, and the estimates of the four regression coefficients of interest, $\beta = (\beta_{11}, \beta_{21}, \beta_{12}, \beta_{22})'$, has variance-covariance matrix $\Sigma \otimes D$, where D is the 2 x 2 submatrix of $(X'X)^{-1}$ corresponding to $(\Delta U_{i1}, \Delta U_{i2})$. For purposes of the power calculation we assume D will be approximately equal to

$$\tilde{D} = (N-1)^{-1}\left\{Cov\begin{bmatrix}\Delta U_1\\ \Delta U_2\end{bmatrix}\right\} = \{(N-1)(1-\rho_U{}^2)\}^{-1}\begin{bmatrix}S_1^{-2} & -\rho_U S_1^{-1}S_2^{-1}\\ -\rho_U S_1^{-1}S_2^{-1} & S_2^{-2}\end{bmatrix},$$

where S_j^2 is the variance of ΔU_{ij} (j=1,2) and ρ_U is the correlation between ΔU_{i1} and ΔU_{i2}. The approximation consists of ignoring the effects of the other covariates and assuming that the distribution of (ΔU_{i1}, ΔU_{i2}) we will observe in the study will have variances and correlation given by S_1^2, S_2^2, and ρ_U. Then the quadratic form $\hat{\beta}'(\hat{\Sigma}\otimes D)^{-1}\hat{\beta}$ has a non-central F distribution on four and N–14 degrees of freedom with non-centrality parameter $\beta'(\Sigma\otimes D)^{-1}\beta$, and the probability that a non-central F variable with non-centrality parameter $\beta'(\Sigma\otimes\tilde{D})^{-1}\beta$ will exceed the critical value cutting off 5% in the upper tail of the F distribution on four and N–14 degrees of freedom will approximate the power of the F test. Since Cov(ΔY) = β'Cov(ΔU) β + Σ, we could estimate Σ by Cov(ΔY) - β'Cov(ΔU)β with assumptions about the marginal variances and covariance of ΔY and the same about ΔU. However, to be conservative, we will simply assume that Σ is given by the marginal parameters listed below and not reduce them by the assumed regression effects. We assume the following standard deviations and within-variable correlation parameters (between times t=1 and t=2).

MCI sample parameters	UPSIT change score	SRT change score	ADAS-cog change score	CIBIC-plus change score
Standard deviation (t=1)	3.75 (S_1)	2.0 (σ_1)	2.0 (σ_1)	1.0 (σ_1)
Standard deviation (t=2)	3.75 (S_2)	2.0 (σ_2)	2.0 (σ_2)	1.0 (σ_2)
Correlation (t=1, t=2)	0.80 (ρ_U)	0.85 (ρ_Y)	0.85 (ρ_Y)	0.70 (ρ_Y)
AD sample parameters				
Standard deviation (t=1)	3.75 (S_1)	3.0 (σ_1)	3.0 (σ_1)	1.0 (σ_1)
Standard deviation (t=2)	3.75 (S_2)	3.0 (σ_2)	3.0 (σ_2)	1.0 (σ_2)
Correlation (t=1, t=2)	0.80 (ρ_U)	0.85 (ρ_Y)	0.85 (ρ_Y)	0.70 (ρ_Y)

It would be clinically meaningful if a four-point change in UPSIT either at the atropine challenge test or the week eight observation were to predict a four-point change in either the SRT or the ADAS-cog endpoint, or a one-point change in CIBIC-plus, at week 26. Similarly, the scientific hypothesis implies that no change in UPSIT predicts no change in the cognitive or clinical endpoint, and this would imply that the slope coefficients $\beta_{12} = \beta_{22} = 1.0$ for SRT and ADAS-cog and $\beta_{12} = \beta_{22} = 0.25$ for CIBIC-plus. However, this assumes perfect correlation between UPSIT and the outcome measures, which we do not expect. A set of smaller regression coefficients are expected, and the next table presents the magnitudes of regression coefficients that can be detected with 90% power with a proposed sample size of N=100 in the AD sample. Regarding effect sizes at 52 weeks, we assume equal effects or, alternatively, half the effect size, compared to 26 weeks.

Effect sizes (θ) detectable at 90% (with alpha=.05 two-tailed).

Primary Hypothesis Test	SRT change score	ADAS-cog** change score	CIBIC-plus change-score***
$\beta_{11} = \beta_{12} = \beta_{21} = \beta_{22} = \theta$ (equal effect size at 26 & 52 weeks and for both UPSIT scores)	0.165	–0.165	0.053
$\beta_{11} = \beta_{12} = \theta$ with $\beta_{21} = \beta_{22} = 0$ (equal effect size for the atropine challenge measure at 26 and 52 weeks, but no effect of UPSIT at 8 weeks)*	0.312	–0.312	0.100
$\beta_{11} = 2\beta_{12} = \beta_{21} = 2\beta_{22} = \theta$ (effect size at week 52 equal to half that at week 26; same for both UPSIT measures)	0.143	–0.143	0.055
$\beta_{11} = 2\beta_{12} = \theta$ with $\beta_{21} = \beta_{22} = 0$ (effect size at week 52 equal to half that at week 26 for atropine challenge, but no effect of UPSIT at 8 weeks)*	0.271	–0.271	0.105
Test of H_0: $\beta_{11} = \beta_{12}$			
$\lvert\beta_{11} - \beta_{12}\rvert = \theta$ (difference of UPSIT effect between weeks 26 and 52)	0.241	0.241	0.114

* Note that because of the assumed symmetries, results assuming effects for UPSIT at atropine challenge and no effect for UPSIT at week 8 are the same as for the reverse. **The signs are reversed for ADAS-cog since ADAS-cog is an inverse scale. *** The non-monotonicities in the table reflect the effect of the different correlation structure assumed for SRT or ADAS-cog versus CIBIC-plus.

The table shows power is at least 90% to declare a slope coefficient 0.312 or greater significant at the 5% level. We have greater power for the MCI sample since the standard deviations are smaller for SRT and ADAS-cog.