Trajectory of Recovery in the Elderly (TORIE)
PI: Dr. Joshua Mincer
NCT02275026
Document Date: 10-31-2013

protocols and forms will be compiled into an operations manual that will be revised as necessary. The manual will include a question-by-question written script for interviewing purposes. Neuropsychological assessment will be conducted by specially trained personnel with experience in psychological assessment. Training will be conducted, as part of the overall investigators' training meeting, by Drs. Sewell. Training seminars will be repeated every six months and as needed, to ensure standardization of procedures.

**Data Reliability**: Diagnostic, case ascertainment and methodologic consistency will be assured by implementing at least every six month training sessions for all the staff. An initial training meeting before study enrollment, and monthly conferences thereafter will ensure reliability and consistency in the use of diagnostic criteria and written standard operating procedures. All data will be independently audited to confirm consistency among participant records, study data sheets, and the main database. Additionally, we will use standard statistical and graphical methods to identify outlying data and confirm their validity

Data will be maintained in a custom-designed relational database (Research Electronic Data Capture REDCap). The REDCap Consortium is composed of 549 active institutional partners from CTSA, GCRC, RCMI and other institutions in 50 countries. The consortium supports a secure web application (REDCap) designed exclusively to support data capture for research studies.

Although the current proposal is not a clinical trial, we will incorporate a **Data and Safety Monitoring Plan** with a **Data and Safety Monitoring Board** (DSMB) as described in the section E.

### *10. Data Analysis*

*All analyses will use a type I error of 5% (two-sided) to test for statistical significance, and will be performed using SAS v9.3 (SAS Institute Inc., Cary, NC). The fit of all models to the data will be examined using standard approaches, such as examination of residuals and the proportional hazards assumption for Cox regression. All baseline and outcome data on participants enrolled in this study will be summarized. Categorical variables will be described in terms of counts and percentages, and quantitative variables in terms of medians and ranges or means and standard deviations, as appropriate.*

***10.1 Specific Aim I:*** *The primary goal is to test whether age is associated with the time of return to baseline cognitive function, assessed by PQRS (primary) and NIH toolbox (secondary), following general anesthesia.*

*Scales of the PQRS are performed prior to anesthesia, and then following anesthesia at defined time intervals (see Section 4). We will focus on the recovery of cognitive function (the primary endpoints as described in 5.0) as related to age and controlling for other participant characteristics using Cox discrete time regression.[74] We expect that everyone will return to baseline cognitive function by 30 days, most early in the course of testing. Thus, if later time points are missing most will have been observed before loss to follow-up. For intermittent missingness before returning to baseline, we will truncate the observations at the time of missing observations. In a subsequent sensitivity analysis, we will assume that they are missing at random and use last value carried forward (worse-case scenario). We will also explore multiple imputation if >5% of the outcome data are missing prior to returning to baseline function.*

*For the secondary NIH Toolbox outcomes, we will calculate six primary standard Adjusted Scale Scores, as described in the Scoring Manual for Toolbox measures described in 5.1. We will again use discrete time Cox regression with covariate adjustment; however, the granularity of the timing of measurements is less than the primary outcome of PQRS. For both the primary and secondary outcomes, we will use the Hochberg test to adjust for multiple comparisons.[75]*

*To study the time course of the PQRS and Toolbox measures, we will use generalized linear mixed models (GLMMs),[76] which account for the covariance structure of repeated observations within participant, and we will test whether a random intercepts and slopes for participants best fit the data. GLMMs are statistical models for data with correlations or non-constant variability and where the response is not necessarily normally distributed. GLMMs assume that missing data are missing at random; thus, participants with partial data contribute to the model estimation. We will fit the appropriate distribution (e.g. dichotomous, count, ordinal, normal) for each outcome using Proc GLIMMIX in SAS v9.3.*

***10.2 Specific Aim II:*** *The measures for fMRI analyses generally fall into three groups: resting, task-driven, and rCBF.*

*Change scores will be calculated for each of 20-25 components (characterizing resting state and activity driven brain networks) related to fMRI imaging at each time point (Section 4). These changes (expressed in standardized units) associated with age will be analyzed using GLMMS adjusted for participant characteristics as described for Aim I. We will use the False Discovery Rate test to adjust for multiple comparisons.*[77]

*For the ASL analysis, an additional 5 measures for rCBF (global and right/left frontal and temporal) will be generated for each of three time points and analyzed with GLMMs corrected for multiple comparisons as described above.*

*fMRI components will be entered as covariates in the above-mentioned Cox discrete-time models for the primary outcomes (PQRS and NIH Toolbox scales) to assess their association with the primary outcomes and with age. Similarly, the fMRI components will be entered into the GLMMs described in 10.1 to estimate their adjusted associations. We will use the False Discovery Rate test to adjust for multiple comparisons.*[77]

***10.3 Sample Size, Power, and Effect Size Estimation:***

*Sample and power for secondary outcomes were calculated with PASS12 (PASS12, Kayville, Utah, USA).* ***SA1****: A Cox regression*[78] *on the time to return to baseline cognitive performance requires 72 participants to detect a HR of 1.03 per year or age (β=0.033), assuming a standard deviation for age of 11.25 years, 80% power, a Type I error of 0.05, and adjusting for other characteristics expected to have a generalized* $R^2$ *of 0.2.*[7980]

*The NIH toolbox yields six primary standardized scaled scores (mean=100, SD=15). Although there will be fewer discrete time measurements for Cox regression, the effect-size calculation is similar. Assuming 72 participants with a standard deviation for age of 11.25 years, an adjusted Type 1 error rate of 0.05/6=0.0083, and other covariates having a generalized* $R^2$ *of 0.2 with age, a HR of 1.04 (β=0.041) can be detected with 80% power.*[80]

***SA2****: A sample size of 72 maintains 80% power for 30 MRI components to detect a change of 0.04 in each component testing for variation across the age spectrum where the assumed standard deviation of age is 11.25 years, the SD of the normalized MRI component outcomes is 1.0, and the two-sided Type I error of 0.0017 to account for multiplicity of the 30 tests (Bonferroni correction). We recognize that in fMRI, multiple comparisons across massive numbers of voxels make it extremely difficult to formulate valid assumptions regarding variability. However, we are using summary measures for these components and have used a Bonferroni adjustment for power calculation.*

*We expect missing data to be low as the total time under observation is 30 days and all participants are expected to have returned to baseline cognitive function by 30 days. Preliminary data had 75% return to baseline at T+15 minutes and all by 7 days. Thus, outcomes are expected to be captured on the majority of subjects even if they do not attend all follow-up visits. However, to insure sufficient outcomes we will increase by 5.5% to 76 thus adding a participant to each age decade.*

## 11. Scientific Limitations and Rationale for Design Decisions

*The results will be a first step and cannot completely elucidate age-related differences in subjects and conditions associated with clinical indications for surgery. The study focuses on the main effect of anesthesia on cognitive recovery but cannot address the possibility that administration of anesthesia interacts in some way with what happens due to the surgery itself. There are multiple anesthetics in current use. This proposal includes a fairly standard combination of anesthetic agents. Additional agents cannot be accommodated in this model. Safety is a primary concern. The current design is focused on safety as opposed to results at any cost.*

There are multiple alternative means of assessing the central nervous system apart from cognitive testing and MRI based imaging. The investigators considered using PET scanning for amyloid *and EEG as an assessment of cerebral function.* While each of these presents an advantage, neither could be accommodated in the current design. PET scanning involves further risk of radiation exposure. *EEG in the context of MRI is not available in our institution at this time (we are working on it). EEG as a primary alternative is possible. The choice of MRI represents a decision by the investigators as to the most likely scientific benefit.*