

**The Michigan Stroke Transitions Trial (MISTT): Improving Care
Transitions for Acute Stroke Patients Through a Patient-centered
Home Based Case Management Program**

NCT02653170

June 17, 2019

Analysis of Patient Outcomes

Among randomized participants, descriptive statistics of demographic and clinical data, including frequencies and contingency tables, were constructed for categorical data, while means and standard deviations were generated for continuous variables.

The original study protocol specified that the analytical model for the intention-to-treat (ITT) analysis of the primary study outcomes would involve testing the differences in the mean change over time (90-day minus 7-day values or “delta”) in each of the 3 treatment groups using a one-way ANOVA.¹ This approach is referred to as a “complete case analysis,” since for an individual participant to be included, it requires that data be available from both time points; if data are missing for either time point the participant is dropped from the analysis.

Subsequently, after completing data collection, we faced the problem of missing data at both the 7-day and 90-day time points, which negatively impacted the “complete case analysis” sample size. Specifically, only 180 patients had data available for the primary outcomes at both time points. To increase the number of observations available for the ITT analysis, we therefore modified the statistical model from the ITT “complete case analysis” to an ITT “available case analysis” which relied on the following linear regression model approach.²

The primary outcomes (PROMIS-10 GPH and GMH subscales, and PAM) were analyzed as linear (continuous) variables using Proc Mixed (SAS V9.4, Cary, NC); 7-day and 90-day outcomes were analyzed as correlated continuous responses within patient. The statistical model for the mean response comprised of an intercept, two indicators for group-2 and group-3 (group-1 as referent), one indicator for time 90-day (7-day as referent), and two group-by-time interactions. Hence the model for the mean response is:

$$\beta_0 + \beta_1 (\text{Grp2}) + \beta_2 (\text{Grp3}) + \beta_3 (90d) + \beta_4 (\text{Grp2} * 90d) + \beta_5 (\text{Grp3} * 90d).$$

The change from 7-day to 90-day in mean response in group-1 is β_3 ; the corresponding change in group-2 is $\beta_3 + \beta_4$. Hence, the group-by-time interaction coefficient β_4 is the comparison of group-2 to group 1; β_5 is the analogous comparison of group-3 to group-1. The combined 2-degree freedom test of the group-by-time interactions (i.e., $\beta_4=\beta_5=0$) identifies whether the mean change over time differed between treatment groups. We refer to this interaction as a difference-in-differences (D-in-D) analysis or result.² Pair-wise contrasts (with 95% confidence intervals) were generated to compare treatment groups with each other.

This “available case analysis” approach has the advantage of increasing the number of individuals that contribute to the ITT analysis, since participants with data available from only one of the 2 time points can still be included. Thus, while only 180 patients (with 360 observations, one for each time point) could be included in the complete case analysis, 214 patients (with 434 observations) were included in the available case analysis. The net effect of using the available case analysis approach is that it minimized the amount of missing data and thereby reduces its potentially negative impact on the primary results. However, it does not eliminate the problem of missing data, so there is still the need to use a multiple imputation procedure which we did as part of a subsequent sensitivity analysis.

Following completion of main effects models for all the primary outcomes, we tested whether adjusting for potential confounding variables resulted in a meaningful change in the results (defined as at least a 10% change in the beta coefficient representing the effect size of each treatment group). We developed the following list of a priori potential confounders: age, race, sex, time between admission and returning home, study site, stroke severity, discharge modified Rankin score (mRS), discharge destination (home, acute rehabilitation, sub-acute rehabilitation), and consented caregiver included (yes, no). We also adjusted the analysis for any variables that were not sufficiently balanced across the 3 treatment groups following randomization. After testing for confounding effects, we determined whether each of these same variables acted as a modifier of treatment effect by testing covariate*intervention interaction terms.

Secondary patient outcomes included depression symptoms (PHQ-9), Neuro-QOL anxiety, 90-day unplanned hospital readmissions, 90-day stroke/TIA recurrence, and 90-day home time. The PHQ-9 and Neuro-QOL anxiety outcome measures were analyzed using the linear regression model described above. The binary measures of 90-day readmissions, and stroke/TIA were analyzed using logistic regression. Ninety-day home time was analyzed using a 90-day-inflated negative binomial regression model (NBRM) due to the large number of patients with a home time of 90-days.

Sensitivity Analysis to Account for Missing Data

To assess the potential impact of missing data on primary patient outcomes, the analyses were repeated after missing data were generated using a multiple imputation procedure that generated 25 iterations of the fully imputed data. The linear regression results from the imputed data were then compared to the original analysis to determine if the presence of missing data had introduced any bias with respect to the statistical inferences drawn regarding overall treatment effect of the intervention (this approach can be regarded as a form of sensitivity analysis). To do this, we first applied the multiple imputation procedure 'Proc MI' (SAS software (SAS/ACCESS® 9.4) which is based on the missing at random (MAR) assumption where, using a Markov chain Monte Carlo (MCMC) method, missing outcome data are imputed using variables that are known or suspected to be associated with missing data. The Proc MI procedure imputed missing values for the primary outcomes (PROMIS-10 GPH, PROMIS-10 GMH, and PAM) each separately for the two different time points (7-day and 90-day), as well as the secondary outcomes of PHQ-9 and NeuroQOL anxiety. The imputation model included the treatment group (randomization arm) variable and the following 24 auxiliary variables: age, sex, race, stroke type (ischemic vs. hemorrhagic), past medical history/comorbidities (including history of stroke, TIA, myocardial infarction, coronary artery disease, atrial fibrillation, hypertension, hyperlipidemia, diabetes mellitus, depression, and heavy alcohol consumption), living alone prior to stroke event, having a consented caregiver, caregiver living with patient, relationship of caregiver to patient, severity of stroke (NIH Stroke Scale [NIHSS]/ Glasgow Coma Scale [GCS]), modified Rankin Scale (mRS) at discharge, score on 6-item cognitive screening test, study site, hospital discharge destination (home, acute rehabilitation, sub-acute rehabilitation), and the interval between admission and discharge to home (number of days). None of the auxiliary variables had any missing values. These auxiliary variables were chosen

on the basis that they were statistically significantly associated with missing outcomes data, or directly described the reason for missing-ness (i.e., health-related reason, refused, or no-contact). Trace plots (across 2,500 iterations) and autocorrelation plots were used to monitor the convergence behavior of the imputation algorithm. To allow for replication of the results, a random seed number was used for all the imputations. The final imputed data set contained 13,250 observations that resulted from 25 imputations of the original 265 cases. An identical MI approach was used to impute missing values for two of the secondary patient outcomes, including PHQ-9 and Neuro-QOL anxiety.

After generating the imputed data set, the same linear regression model used in the available case analysis (generated using SAS PROC MIXED) was applied to each imputed data set. The SAS PROC MIANALYZE procedure was used to summarize these results across the 25 imputed datasets, taking into account the uncertainty in imputation due to missing data. PROC MIANALYZE generates valid statistical inferences under a missing at random (MAR) assumption by combining the estimated model coefficients and associated covariance matrices across the multiple imputed datasets. As with the available case analysis, the quantities of substantive interest were the mean responses between the two time points and among the three study arms. In the full factorial model (main effects for the treatment groups and time, and their interactions), these quantities are the combined 2-degree freedom test of the group-by-time interactions (i.e., $\beta_4=\beta_5=0$), which identifies whether the mean change over time differed between treatment groups.

Because the MIANALYZE procedure is unable to generate parameter estimates and standard errors for interaction terms, the mean models were parameterized using parameters that directly represented the cell means for each of the 6 combinations of treatment group and time (i.e., Trt_group1= Usual Care at 7-day, Trt_group2= Usual Care at 90-day, Trt_group3= SWCM at 7-day, Trt_group4= SWCM at 90-day, Trt_group5= SWCM+MISTT website at 7-day and Trt_group6= SWCM+MISTT website at 90-day). From this parameterization, the group-by-time interaction effects in the mean responses were expressed as contrasts of the 6 cell means and evaluated using the TEST statement of PROC MIANALYZE with MULT option (a joint numerator 2-df test). The mean differences across the two time points for any two of the three study arms were compared using a t- distribution with df that take into account the uncertainty due to imputations.

Analysis for Caregiver Outcomes

Primary caregiver outcomes included the Bakas Caregiver Outcome Scale (BCOS) and PHQ-9 (depression), which were analyzed using the same linear regression model with group-by-time interactions as described above.

Secondary caregiver outcomes included PROMIS informational support, PROMIS emotional support, and unhealthy days. The support measures were also analyzed using the linear regression model, while the number of unhealthy days were analyzed using a zero inflated negative binomial model to account for the large number of caregivers who reported no unhealthy days.

We did not apply MI to caregiver outcomes as we lacked equivalent information on caregiver characteristics or details for why they did not participate, which could be used as auxiliary variables in the MI step.

REFERENCES

1. Reeves MJ, Hughes AK, Woodward AT, et al. Improving transitions in acute stroke patients discharged to home: the Michigan Stroke Transitions Trial (MISTT) protocol. *BMC Neurol* 2017;17:115.
2. Dimick JB, Ryan AM. Methods for evaluating changes in health care policy: the difference-in-differences approach. *Jama* 2014;312:2401-2.