Supplemental Online Content

Stamm B, Royan R, Giurcanu M, Messe SR, Jauch EC, Prabhakaran S. Door-in-door-out times for interhospital transfer of patients with stroke. *JAMA*. doi: 10.1001/jama.2023.12739

This supplemental material has been provided by the authors to give readers additional information about their work.

e Methods

To analyze the time from door-in-door-out (DIDO), we first log-transformed the response (log-DIDO) and then fitted a generalized estimating equation (GEE) model for log-DIDO as the response vs hospital and patient level variables (as categorical predictors) using the overall dataset as well as stratified by stroke group. The categorical predictors were selected based on the literature review and *a priori* expert knowledge about the variables that were known to be associated with the DIDO time as well as on their availability in the dataset. To fit the GEE models for the log-DIDO, we specified an exchangeable (aka equicorrelation) working correlation structure to incorporate the within-hospital (cluster) correlation in the response variable. The advantage of the GEE compared to other methods (e.g., maximum likelihood for generalized linear mixed effects models) is that it is robust to the misspecification of the within-cluster correlation structure.[1] Specifically, the GEE estimates are consistent and efficient if the correlation structure is correctly specified and consistent (although potentially no longer efficient) if the correlation structure is misspecified.

After we fitted the GEE models for log-DIDO, we back-transformed the coefficient estimates using the exponential transformation. Under the normality assumption for the distribution of log-DIDO (equivalently, under the logNormal assumption for the distribution of DIDO time), the back-transformed contrasts of the parameter estimates were used to estimate differences in the median DIDO time between the groups of categorical predictors and the baseline group on the original time scale. The standard error estimates of the difference in the median DIDO time were calculated based on the delta method. The normality assumption of the log-DIDO was scrutinized using a histogram plot for the studentized residuals from the model fit with a superimposed density of the standard normal distribution. We did not find evidence of serious deviations from the normal distribution of the residuals of the GEE models for log-DIDO. In order to assess the predictive power of the models, we also calculated the coefficient of determination (R^2) for each GEE model and found that the predictors explain between 10% to 15% of the variability in log-DIDO, showing that a good portion in the variation of the response was unexplained by the GEE models. We also found that the equicorrelation parameter estimates of the GEE models range between 0.18 and 0.24, showing that the GEE models capture the cluster correlation in log-DIDO unexplained by the set of predictors.

To analyze the risk of DIDO >120 minutes, we fitted a GEE model for the binary outcome DIDO > 120 minutes as the response vs. hospital and patient level variables (as categorical predictors) using the logit link to the overall dataset and stratified by stroke group. We reported the odds ratios of DIDO >120 minutes for the groups of categorical variables vs. the baseline group and the differences in the absolute risks of DIDO >120 minutes between the groups of categorical predictors and the baseline group. The standard errors for the differences in the absolute risks of binary DIDO categories were calculated using the delta method. We also found that the equicorrelation parameter estimates on the logit scale of the GEE models range between 0.13 and 0.17, showing that the models capture the cluster correlation in the risk of DIDO >120 minutes on the logit scale unexplained by the set of predictors.

To account for missing data, we also conducted two imputation methods for all the GEE models. In the first imputation method, we imputed the missing values with the most prevalent category among the groups of each categorical variable. In the second imputation method, we used the multiple imputations by chained equations algorithm[2] to generate multiple imputed datasets (S=10). The GEE models were fit for each imputed data set and the estimates, their standard errors, and the p-values of the z-tests were pooled over the multiple imputations.

Finally, we derived additional referring hospital-level variables of interest and ran exploratory GEE models to evaluate the effects on DIDO time adjusting for the predictors included in our main analyses. These variables included: 1) stroke transfer volume, 2) telestroke capability, 3) Medicaid/self-pay proportion, and 4) tenecteplase (TNK) utilization. These hospital-level variables were not present in the original GWTG-Stroke registry but were hypothesized to likely affect DIDO time, warranting this additional exploratory analysis. A telestroke-capable hospital was defined as a hospital where one or more patients from our sample had a telestroke performed over the time period of our study. A TNK-utilizing hospital was defined as a hospital in which one or more patients from our sample had tenecteplase (TNK) administered over the time period of our study. Stroke transfer volume per transferring hospital was defined as a count variable for the absolute number of patients in our final sample per each

hospital. A distribution of this count variable for our dataset was created, and the results were distributed into the following quartiles: 0-50, 51-100, 101-150, and >150 transfers. The proportion of patients with Medicaid/self-pay insurance per hospital was derived by calculating the overall median number of Medicaid or self-pay patients per transferring hospital. There is broad variability in the definitions used to define safety-net hospitals.[3] We thus created a distribution of the Medicaid/self-pay proportion variable for our dataset, and the results were distributed into the following quartiles: 0-5%, 6-10%, 11-15%, and >15%. In these exploratory GEE models, we also included one additional patient-level variable: last known well (LKW) to ED arrival time. This variable was calculated from the respective parent variables (time of LKW and ED arrival time) in the GWTG-Stroke registry. We categorized elapsed time from LKW, as follows: < 3 hours, 3-4.5 hours, 4.5-24 hours, >24h/unknown, similar to an approach previously taken in a GWTG-Stroke publication.[4] There were 15,990 missing values (14.7%), which crossed the pre-specified threshold of 10% missingness used to exclude variables from the main analyses. Thus, this covariate was additionally included in the exploratory GEE models, with results available in Supplemental Table 4.

The data analysis was conducted on the AHA precision medicine website (https://precision.heart.org/) using SAS version 9.4 (SAS Institute Inc) for preliminary data processing and RStudio for additional data processing and data analysis.[5] The R package geepack[6] was used to fit the GEE models, the R package mice were used for the multiple imputations method,[2] the R package table one was used to generate summary tables,[7] the R package car was used to calculate the SE using the delta method,[8] and the R package quantreg was used to estimate the quantiles of the DIDO time stratified by state.[9]

eResults

In GEE logistic regression models (Supplemental Table 1), the following were significantly associated with DIDO time >120 minutes in the overall cohort: age 70-79 years (aOR 1.06, 95% CI 1.01-1.11), age ≥80 years (aOR 1.22, 95% CI 1.16-1.28); female sex (aOR 1.12, 95% CI 1.08-1.16); Black (aOR 1.15, 95% CI, 1.09-1.21) and Hispanic ethnicity (aOR 1.16, 95% CI 1.07-1.25); diabetes (aOR 1.16, 95% CI, 1.12-1.21); prior stroke (aOR 1.19, 95% CI, 1.14-1.24); peripheral vascular disease (aOR 1.15, 95% CI, 1.02-1.30); smoking (aOR 1.13, 95% CI, 1.08-1.19); AIS-other stroke type (aOR 2.05, 95% CI, 1.91-2.19); EMS no notification (aOR 1.07, 95% CI, 1.01-1.14); after hours (aOR 1.07, 95% CI, 1.03-1.11); during pandemic (aOR 1.26, 95% CI, 1.21-1.32), urban location (aOR 1.13, 95% CI, 1.00-1.28), daily hospital census 100-199 (aOR 1.22, 95% CI, 1.07-1.39), daily census ≥200 (aOR 1.35, 95% CI, 1.12-1.63), and annual thrombolysis volume 30-126 (aOR 1.27, 95% CI, 1.02-1.57). In GEE logistic regression models, the following were significantly associated with DIDO time <120 minutes in the overall cohort: AIS-EVT eligible stroke type (aOR 0.87, 95% CI, 0.80-0.95); NIHSS 2-4 (aOR 0.54, 95% CI, 0.50-0.58); NIHSS 5-12 (aOR 0.33, 95% CI, 0.31-0.36); and NIHSS >12 (aOR 0.25, 95% CI, 0.23-0.26). Results for the subgroups are depicted in eTable 1.

The secondary analyses, including both the rule-based and multiple imputation GEE models, did not substantially differ from the primary model findings. The effect size of the association between Black race, Hispanic ethnicity, and DIDO times was somewhat attenuated. For instance, in the rule-based imputed overall model, Black race was still associated with a significantly longer DIDO time (+7.23 min, as compared to 8.2 min increase in the original model), whereas the higher DIDO time associated with Hispanic ethnicity became nonsignificant. Full results are seen in eTable 2 and 3. Additionally, in the rule-based imputed model (eTable 2), the following insurance variables were significantly associated with decreased DIDO times: private/VA/Champus/other insurance type (-4.47 minutes, 95% CI -6.89, -2.04) and self-pay/no insurance (-11.36 minutes, 95% CI -16.66, -6.05). In the multiple imputations model (eTable 3), the following insurance variables were significantly associated with decreased DIDO times: private/VA/Champus/other insurance type (-3.29 minutes, 95% CI -6.08, -0.50) and self-pay/no insurance (-6.25 minutes, 95% CI -11.08, -1.42).

The exploratory GEE model with post-hoc created variables (eTable 4) overall yielded similar results to those from the primary analysis models. From the additionally included covariates, the following were significantly associated with increased DIDO time: LKW to arrival time 3-4.5 hours (+11.13 minutes, 95% CI 8.11-14.15); LKW to arrival time 4.5-24 hours (+22.49 minutes, 95% CI 20.15-24.82); LKW to arrival time >24 hours / unknown (+59.69 minutes, 95% CI 56.50-62.88) compared to LKW 0-3 hours as the reference group; and telestroke site (+5.31 minutes, 95% CI 0.51-10.12). The following were associated with significantly decreased DIDO time: 51-100 stroke transfers over study time period (-9.24 minutes, 95% CI -14.44, -4.05); 101-150 stroke transfers (-16.29 minutes, 95% CI -22.35, -10.23); and >150 stroke transfers (-20.52 minutes, 95% CI -27.79, -13.26) compared to 0-50 stroke transfers as the reference group.

e References:

1. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics*. 1988;44(4):1049-1060.

2. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw*. 2011;45:1-67.

3. Hefner JL, Hogan TH, Opoku-Agyeman W, Menachemi N. Defining safety net hospitals in the health services research literature: a systematic review and critical appraisal. *BMC Health Serv Res*. 2021;21(1):278.

4. Tong D, Reeves MJ, Hernandez AF, et al. Times from symptom onset to hospital arrival in the Get with the Guidelines--Stroke Program 2002 to 2009: temporal trends and implications. *Stroke*. 2012;43(7):1912-1917.

5. R Core Team. *R: A Language and Environment for Statistical Computing*. Accessed May 8, 2023. https://www.r-project.org/

6. Højsgaard S, Halekoh U, Yan J. The R Package geepack for Generalized Estimating Equations. *J Stat Softw*. 2006;15:1-11.

7. Yoshida K. Create "Table 1" to Describe Baseline Characteristics with or without Propensity Score Weights [R package tableone version 0.13.2]. Published online April 15, 2022. Accessed May 8, 2023. https://cran.r-project.org/web/packages/tableone/index.html

8. Fox J, Weisberg S. *An R Companion to Applied Regression*. SAGE Publications; 2011.

9. Koenker R. Quantile Regression [R package quantreg version 5.95]. Published online April 8, 2023. Accessed May 8, 2023. https://CRAN.R-project.org/package=quantreg