Published in final edited form as:

J Biopharm Stat. 2019 ; 29(1): 82–97. doi:10.1080/10543406.2018.1489404.

Treatment effect on ordinal functional outcome using piecewise multistate Markov model with unobservable baseline: An application to the modified Rankin Scale

Christy Cassarly^{1,2,*}, Renee' H. Martin¹, Marc Chimowitz³, Edsel A. Peña⁴, Viswanathan Ramakrishnan¹, and Yuko Y. Palesch¹

¹Department of Public Health Sciences, Medical University of South Carolina, Charleston, SC

²Department of Otolaryngology–Head & Neck Surgery, Medical University of South Carolina, Charleston, SC

³Department of Neurology, Medical University of South Carolina, Charleston, SC

⁴Department of Statistics, University of South Carolina, Columbia, SC

Abstract

In clinical trials, longitudinally assessed ordinal outcomes are commonly dichotomized and only the final measure is used for primary analysis, partly for ease of clinical interpretation. Dichotomization of the ordinal scale and failure to utilize the repeated measures can reduce statistical power. Additionally, in certain emergent settings, the same measure cannot be assessed at baseline prior to treatment. For such a data set, a piecewise-constant multistate Markov model that incorporates a latent model for the unobserved baseline measure is proposed. These models can be useful in analyzing disease history data and are advantageous in clinical applications where a disease process naturally moves through increasing stages of severity. Two examples are provided using acute stroke clinical trials data. Conclusions drawn in this paper are consistent with those from the primary analysis for treatment effect in both of the motivating examples. Use of these models allows for a more refined examination of treatment effect and describes the movement between health states from baseline to follow-up visits which may provide more clinical insight into the treatment effect.

Keywords

longitudinal ordinal outcome; piecewise multistate models; panel data; modified Rankin Scale

1. INTRODUCTION

Outcomes on an ordinal scale is quite common in clinical trials (Roozenbeek et al., 2011). It is also common in these trials to analyze the data using a dichotomized version of the ordinal measure. For instance, in treatments of acute stroke, often the modified Rankin Scale (mRS) is used as the primary outcome. The mRS measures functional ability using a 7-point

*Corresponding author: Christy Cassarly, cassarly@musc.edu, Phone: (843) 792-4171.

ordinal scale (Table I). Although, mRS is measured on this 7-point scale, for testing primary hypotheses of interest, it is commonly dichotomized either by collapsing into {0,1} vs {2,3,4,5,6} or {0,1,2} vs {3,4,5,6}. Loss of information, when such ordinal outcomes are collapsed into a dichotomy, has been studied and shown to result in reduction of statistical power (Sankey & Weissfeld, 1998). In stroke trials, alternative analytic methods on the observed ordinal scale are gaining attention (The Optimising Analysis of Stroke Trials (OAST) Collaboration, 2007).

Another common approach in most longitudinal trials that measure mRS over time, is to use only the final measure made at 90 days, ignoring the earlier measurements because the outcome status at 90 days is considered to be the only measurement of clinical relevance. Moreover, typically, the ultimate goal of a treatment or therapeutic action is to achieve improvement in patient health compared to their baseline measure. To accurately quantify improvement, a measure of the outcome at baseline would be ideal. Conditions with sudden onset, such as traumatic brain injury, stroke and status epilepticus often require immediate attention and treatment, posing a challenge to assess baseline outcome measures for clinical trials that may also lack practical meaning based on patient status. Specific to acute stroke trials the mRS is not obtainable at baseline (Bruno et al., 2009). The current method of addressing this is to adjust for the severity of the condition (disease) at baseline.

In this manuscript a continuous-time non-homogeneous Markov process is proposed as an alternative to study the evolution of acute onset diseases. Of specific interest is exploration of potential differences in transition rates between two treatment groups. Using this method, it is possible to analyze treatment effects in the observed ordinal scale and incorporate data measured longitudinally. In addition, since much of the progression or recovery experienced post treatment by a patient suffering from an acute onset disease is expected to occur early, a method for predicting the baseline mRS state is proposed. This baseline mRS estimate may then be utilized in a model that more fully characterizes the evolution of disease over time.

The paper is organized as follows. Two motivating examples of large acute stroke therapy trials are described in Section 2. Homogeneous and piecewise-constant multistate Markov models (MSMs) are introduced. In Section 3, the baseline estimation procedure is described and demonstrated and piecewise-constant MSMs using the estimated baseline scores for the motivating examples are fit. Concluding remarks are presented in Section 4.

2. METHODS

Methods are developed and motivated through two phase III acute stroke therapy trials, namely the National Institute of Neurologic Disorders and Stroke tissue plasminogen activator (NINDS tPA) Stroke Study Part 2 (The National Institute of Neurological Disorders Stroke rt-PA Stroke Study Group, 1995) and the Albumin in Acute Stroke (ALIAS) Trial (Ginsberg et al., 2006).

The NINDS tPA Stroke Study Part 2 was designed to compare intravenous tPA versus placebo in subjects with acute ischemic stroke using a global test statistic (The National Institute of Neurological Disorders Stroke rt-PA Stroke Study Group, 1995). The global test

statistic simultaneously tested for treatment effect in four correlated outcomes (mRS, Barthel Index, Glasgow outcome scale and National Institutes of Health Stroke Scale). The Barthel Index is a simple index of independence that scores the ability of patients to care for themselves (Mahoney & Barthel, 1965). Patients that can perform all activities assessed with complete independence are given a score of 100. The Glasgow outcome scale is a global assessment of function that ranges from 1 indicating good recovery to 5, death (Jennett & Bond, 1975). The National Institutes of Health Stroke Scale (NIHSS) is a 42-point scale that measures neurologic deficit where 0 indicates normal function (Lyden et al., 1994).

A total of 624 patients were enrolled in NINDS tPA, with 312 in each treatment group. The primary analysis showed a significant global test score for the four dichotomized outcomes as well as for the dichotomized mRS alone at 90 days (Tilley et al., 1996). In addition to the primary outcome assessment at 90 days, the mRS was collected at 7-10 days, 180 days and 360 days from randomization for each subject.

The ALIAS Trial was designed to compare 25% human serum albumin and saline in patients with acute ischemic stroke (Ginsberg et al., 2006). Part 1 of the trial was suspended after enrolling 434 subjects due to safety concerns of albumin (Hill et al., 2011). Part 2 of the trial was slightly redesigned with unblinded safety analysis and enrolled 841 subjects. The analysis of both Parts 1 and 2, as well as the two combined, showed a lack of treatment effect on primary and secondary outcomes, including the dichotomized mRS at 90 days (Martin et al., 2016). In addition to the primary outcome assessment at 90 days, the mRS was also collected at 30 days, 180 days, 270 days and 360 days from randomization for each subject.

2.1 Multistate Markov models

MSMs in continuous-time has been used to model the course of many diseases (Jackson, 2011). The MSMs incorporate longitudinal ordinal data and can provide clinically relevant summary statistics to describe covariate effects, including sojourn times and transition rates. These models are advantageous in clinical applications, where a disease process naturally moves through increasing stages of severity (Jackson, 2011). Homogeneous continuous-time MSMs, where the transition rates are assumed to be constant over time, have been used to analyze various diseases (Aalen et al., 1997; Keiding et al., 2001; Marshall & Jones, 1995).

The assumption of homogeneity is not always realistic. With acute onset diseases, the rapid nature of onset and intervention likely characterize a process that changes quickly early on and tapers off after the initial acute recovery stage. The transition rates of the remainder of the longitudinal disease process, after the initial burst of rapid movement post-baseline, are likely to differ from those observed in the acute phase. In this case, a non-homogeneous model with piecewise constant intensity rates can be used.

Multistate Markov modeling requires that the Markov property holds for the observed data. Consider a system with a finite state-space $\Omega = \{1, 2, 3, \dots, I\}$, where I represents the number of states. Let $X(s+t)$ be a discrete random variable that indicates the state occupied by the system at time $s+t$. Let $F_{X(s)} = \{X(v): v \leq s\}$ which denotes all information pertaining to the

history of X up to time s (Chiang, 1980). The series of observations has the Markov property if the conditional distribution of $X(s+t)$ given $F_{X(s)}$, satisfies

$$P\{X(s+t) = j \mid F_{X(s)}, X(s) = i\} = P\{X(s+t) = j \mid X(s) = i\} = p_{ij}(s, s+t), i, j \in \Omega. \quad (1)$$

In other words, the present state depends only on the immediately preceding observation and not on the ones before it. In the context of clinical trials, the state of the system is the health state of one individual.

Though the ordinal outcome in acute therapy clinical trials is observed at discrete times, the disease process is continuous, where progression or recovery can occur at any time. Continuous-time MSMs can analyze this type of data, known as panel data, as long as the sampling times are considered to be noninformative (Jackson, 2011). In the clinical trial setting where the observation scheme is fixed in advance, this assumption is valid. If, however, a subject visited a clinic because of a change in symptoms and the outcome was collected, the sampling time would be informative and could bias inference. These continuous-time MSMs are flexible enough to model panel data with noninformative sampling times but also include exact time of death, which is commonly observed in clinical trials.

A common assumption of continuous-time MSMs is that of homogeneity, where transition probabilities remain constant over time. When homogeneity is assumed, the transition probabilities, $p_{ij}(t)$, are defined as

$$p_{ij}(t) = P\{X(s+t) = j \mid X(s) = i\} = P\{X(t) = j \mid X(0) = i\}. \quad (2)$$

Since this expression does not depend on s , the transition from state i to state j on a time interval of length t has the same probability at any time. The $p_{ij}(t)$ are elements of the transition probability matrix, $\mathbf{P}(t)$.

The movement of a subject between states is described by λ_{ij} , the transition intensities:

$$\lambda_{ij} = \lim_{\delta t \rightarrow 0} \frac{P(X(\delta t) = j \mid X(0) = i)}{\delta t}, \text{ for } i \neq j. \quad (3)$$

The intensities represent the instantaneous rate of moving from state i to state $j \neq i$ and form the generator matrix, Λ , whose rows sum to zero and the diagonal entries are $\lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}$. The transition probability matrix $\mathbf{P}(t)$ can be solved by taking a matrix exponential of Λ scaled by the time interval,

$$\mathbf{P}(t) = e^{t\Lambda} = \sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!} \quad (4)$$

where Λ^k is the k^{th} power of the generator matrix Λ .

Suppose X is observed over $t_0 < t_1 < t_2 < \dots < t_M$. Let $i_0, i_1, i_2, \dots, i_M$ be the observed states over these time points. Then, the associated likelihood function is

$$\begin{aligned} L(\Lambda | x) &= P(X_0 = i_0) \prod_{j=1}^M P(X_{t_j} = i_j | X_{t_{j-1}} = i_{j-1}) \\ &= P(X_0 = i_0) \prod_{j=1}^M P(X_{t_j - t_{j-1}} = i_j | X_0 = i_{j-1}) \quad (5) \\ &= P(X_0 = i_0) \prod_{j=1}^M P_{i_{j-1}i_j}(t_j - t_{j-1}). \end{aligned}$$

Using equations (4) in (5), the likelihood reduces to

$$L(\Lambda | x) = P_{i_0} \prod_{j=1}^M \left(\sum_{k=0}^{\infty} \frac{(t_j - t_{j-1})^k (\Lambda^k)_{i_{j-1}i_j}}{k!} \right), \quad (6)$$

where P_{i_0} is the initial probability that the process is at i_0 .

2.2 Piecewise-constant multistate Markov models

In the case of ischemic stroke occurrence and treatment, subjects can get better or worse very quickly. For this reason, the time homogeneity assumption is expected to fail for the first transition, from the estimated baseline to the first observed outcome. Therefore, assumption of homogeneity is relaxed and a non-homogeneous model is considered.

The MSM for panel data can be extended to accommodate piecewise-constant intensity matrices for the non-homogeneous case (Jackson, 2016). Here, the transition probability functions are dependent on s and the transition matrix function is $\mathbf{P}(s, t) = (p_{ij}(s, t))$. The transition intensity functions are now defined by

$$\lambda_{ij}(t) = \lim_{\delta t \rightarrow 0} \frac{P(X(t + \delta t) = j | X(t) = i)}{\delta t}, \quad i, j \in S, i \neq j. \quad (7)$$

The time-homogeneity assumption can be tested by using a likelihood ratio test to compare the time-homogenous model to piecewise-constant models with different cutoff times.

The effect of a covariate, specifically treatment, is incorporated into the model as transition intensity functions (Pérez-Ocón et al., 2001). Let z be a vector of observed covariates then

$$\lambda_{ij}(t; z) = \lambda_{ij}(t) \exp(z' \beta_{ij}(t)), \quad \lambda_{ij}(t) > 0 \quad (8)$$

where $\beta_{ij}(t)$ is the parameter vector associated with the covariate vector z in the transition between states i and j in time t . The transitional rates are represented by $\lambda_{ij}(t; z)$ at time t for the subjects with vector z .

2.3 Latent baseline estimation

To consider the full evolution of ischemic stroke over time, an estimated baseline functional outcome is needed because baseline mRS is not obtainable in the acute setting. While functional outcome is not available at baseline, many other measures that are correlated with functionality are available. An estimation procedure using baseline characteristics known to be associated with functional outcome was developed.

As a preliminary step, to summarize information from numerous baseline measurements considered clinically relevant for functional outcome in ischemic stroke patients, data reduction was performed using a Principal Components Analysis (PCA). The items included in the PCA were age at randomization, baseline glucose, time from symptom onset to randomization, the Alberta Stroke Program Early CT score (ASPECTS), and NIHSS. Age (Knoflach et al., 2012) and time from symptom onset to thrombolysis (Newby et al., 1996) are two well established predictors of outcome. In the PCA, time from symptom onset to randomization was used as a proxy for time from symptom onset to treatment. Another baseline characteristic associated with poor outcome is “stress” hyperglycemia (Capes et al., 2001). This hyperglycemia can be quantified using acute post-stroke glucose levels. The ASPECTS is also a strong prognostic indicator of outcome (Hill et al., 2014), which is a 10-point quantitative topographic CT scan score, where a normal scan receives a score of 10 (Pexman et al., 2001). For each defined region of the brain, a point is subtracted if there is evidence of ischemic change. The NIHSS is commonly used to measure baseline stroke severity. The individual items of the scale are presented in Appendix A. Because ASPECTS is calculated from a CT image and NIHSS is a neurologic exam, these two assessments can be done at baseline when the mRS cannot. Baseline NIHSS is known to strongly predict outcome in acute stroke therapy trials (Adams et al., 1999). Although total score is typically used for indicating stroke severity, each item of the scale was used in the PCA individually in order to more efficiently assess the contribution of each facet of the scale.

After reducing the data to fewer PCA’s sextiles (six categories because the seventh category, namely mRS = 6 corresponds to death) based on the joint distributions of the PCA’s will be used to define the baseline states of the individuals. Then the MSM likelihood ratio tests will be used to compare treatment effects.

3. RESULTS

Detailed descriptions of the PCA for each trial are presented in Appendix B. For both NINDS tPA and ALIAS, most of the variability was adequately explained by the first two components. Thus, component scores were calculated for Components 1 and 2 for each trial. Larger values on the component scores were expected to be associated with worse functional outcome. To assign the baseline mRS state, each of the component scores were divided into sextiles with equal probability. In Figure 1, the joint distribution of the two discretized scores are shown were used to assign values of mRS = 0 to mRS = 5 with equal probability

(subjects cannot be dead at baseline, so no one was assigned an mRS = 6). There is uncertainty in the assignment of baseline states which needs to be accounted for in the hypothesis tests comparing treatments.

3.1 The longitudinal data

In the NINDS tPA trial, several subjects either did not have available CT scans or the scans were not of sufficient quality to obtain ASPECTS (16/624). Scores for the second component could not be calculated and subsequently the baseline mRS could not be estimated. These subjects were not excluded from the analysis unless they died in the first week of follow-up, leaving them with only one observed mRS score (5/624). Therefore, the total number of subjects included in the MSM was 619, where 308 were randomized to receive tPA and 311 to receive placebo.

In the ALIAS Trial, a small number of subjects withdrew consent or were lost-to follow-up prior to the 30 day visit and had only one observation available, the estimated baseline mRS (17/1275). Excluding these subjects, a total of 1258 were included in the MSM, where 628 were randomized to receive albumin and 630 to receive placebo. In Appendix C, baseline characteristics of the 17 excluded subjects are summarized alongside those of the subjects that were not excluded. No notable differences exist between the groups and thus the exclusion of these subjects should be inconsequential in the analysis.

3.2 Development of MSM model for the NINDS tPA study and ALIAS

In order to analyze panel data with continuous-time Markov chains it is important to consider which transitions can realistically occur in continuous time. When states represent severity it is assumed that in order for a subject to transition from one state to another non-adjacent state they also transition through the intermediate states. Thus, a reduced transition intensity matrix should be estimated, where non-adjacent state transitions are fixed to equal zero. The exception is when a state represents death. The reality is that a subject can die from any state.

In practice, if there is not enough information from the data, on a certain transition rate, more transition intensities may need to be set to zero (Jackson, 2011). State tables display counts of the pairs of transitions between states in successive observation times and summarize them in frequency tables of previous state against current state. These state tables can be used to identify counts that are too few to model.

The state table of all aggregate transitions from the NINDS tPA data (baseline to 360 days) is displayed in Table II. Even though it is possible for subjects to die from any state, it is highly unlikely to occur from states 1, 2, 3 or 4. This is not surprising as subjects are only observed over the course of one year. The relatively healthy subjects have a low risk of death. Therefore more constraints are required for this model. When the table is stratified by time for the piecewise model and again by treatment, the frequency of death from the healthier states (1, 2, 3 and 4) is too small to estimate the transition intensities. As a result, the original model was modified to no longer allow death from any state. Constraints were imposed such that death was only allowable from states 5 and 6.

Similarly, when the state table was examined for ALIAS, small counts were observed for death from states 1, 2 and 3. The reduced allowable transition matrix for ALIAS fixed the intensities from these states to death to equal 0, only allowing the intensities from 4, 5 and 6 to death to be estimated.

3.3 Analysis of the NINDS tPA study and ALIAS

The entries of the transition intensity matrices were estimated by applying the maximum-likelihood method and accounting for the two constant intervals partitioned at time $t = .333$ (representative of 7-10 days on a month-long interval) in the NINDS tPA trial and time $t = 1$ (representative of 30 days) in ALIAS. Next, the MSM likelihood ratio test was used to compare treatment effects. In this likelihood ratio test, the uncertainty of the estimated baseline states had to be considered. This was achieved through bootstrap approach, using which an empirical distribution was derived to determine the p-values. The steps used in this non-parametric bootstrap approach were as follows:

1. Sample with replacement from the original dataset 1,000 times. Each sample is of the same size as the original (1258 for ALIAS and 619 for NINDS t-PA).
2. Use the baseline estimation procedure from the original data on each of the 1,000 bootstrapped samples (fixing the number of significant components as well as the variables used in each of the component score calculation).
3. Fit the piecewise MSM to each resample and obtain the test statistic.
4. Compare the original test statistic to the new bootstrap distribution, made up of the 1,000 test statistics from the bootstrap samples.
5. The bootstrap p-value was calculated by finding the proportion of bootstrap samples in which the test statistics is larger than or equal to the one calculated from the original sample.

Analysis was performed using the *msm* package in R using the Broyden-Fletcher-Goldfarb-Shanno method (Jackson, 2011).

For the NINDS tPA trial, using a likelihood ratio test to compare the piecewise model including treatment ($-2LL = 6251$) to one without treatment ($-2LL = 6282$), the model with treatment was preferred, indicating a statistically significant treatment effect on the transition rates ($p = 0.002$, $df = 12$). It is interesting to note that in models excluding estimated baseline mRS, the comparison of a model with treatment ($-2LL = 4033$) adjusting for baseline NIHSS to one without treatment ($-2LL = 4053$) failed to detect a statistically significant treatment effect ($p = 0.053$, $df = 12$).

In ALIAS, using a likelihood ratio test to compare the piecewise model including treatment ($-2LL = 12730$) to one without ($-2LL = 12745$), the reduced model was preferred, indicating no significant treatment effect on the transition rates ($p = 0.29$, $df = 13$), confirming the results from the primary analysis of the trial (Ginsberg et al., 2013).

For NINDS tPA, after comparing the empirical distribution to the observed test statistic of 31.18, the p-value from the bootstrap procedure was 0.04 which is larger than the observed

p-value of 0.002 but still indicative of a significant treatment effect. For ALIAS, the p-value calculated from the bootstrap procedure, where the test statistic from the analysis of the original sample was 15.25, was 0.49 which is also larger than the observed p-value of 0.29 but results in the same conclusion that there is no a significant treatment effect.

In order to assess the general goodness of fit of MSMs for each trial, we used a graphical comparison of the observed versus expected number of patients in each state at various time points (Jackson, 2016). In these plots the number of individuals in each state at each time was well estimated, indicating a good fit.

Table III shows the transition intensities estimated in the piecewise-constant model for the NINDS tPA Stroke Study. Transition rates differ between the placebo and tPA groups. The differences between groups can also be presented using hazard ratios (Table IV). A hazard ratio greater than one indicates that the rate of transition is higher in the treatment group.

The only statistically significant hazard ratio was from state 2 to state 3 ($HR = 0.51$). This can be interpreted as the most significant impact of tPA is to reduce the hazard of transitioning from $mRS = 1$ to $mRS = 2$. The other hazard ratios, although not statistically significant, suggest a trend of tPA reducing the hazard of negative transitions.

4. SUMMARY AND DISCUSSION

Dichotomization of ordinal outcomes is common but results in a loss of information and can reduce statistical power. Some patients with severe disability at baseline may never have the potential to achieve success as defined by the dichotomy. Thus, the prognostic heterogeneity of subjects does not allow for potential equal contribution to the estimation of treatment effect for all subjects with a dichotomized outcome (Price et al., 2013).

A number of alternative methods for ordinal outcome data have received attention in recent years (Cassarly et al., 2017a). Linear regression and analysis of variance have been suggested where the ordinal outcome is treated as a continuous variable. Summary statistics from these models do not have straightforward interpretations because non-integer values from ordinal scales do not have a clear meaning.

A number of ordinal analyses have also been suggested. Ordinal logistic regression, under the assumption of proportional odds, assumes an identical effect of the predictors for each cumulative probability (Agresti, 1999). If the assumption fails, this analysis could mask important effects at one end of the ordinal outcome (DeSantis et al., 2014). The partial proportional odds model relaxes this assumption and includes a term that allows the odds ratios to increase proportional to the outcome scale (DeSantis et al., 2014). Alternatively, the cumulative logit model allows for the calculation of odds ratios for each adjacent category of response in relation to covariates and does not require the proportional odds assumption (Agresti, 1999). The sliding dichotomy is another alternative method for the analysis of ordinal outcomes. It allows for the definition of success to vary based on patient-specific baseline prognostic variables while maintaining a dichotomized outcome. While the sliding dichotomy allows for baseline severity adjusted outcome, it still ignores any non-successful transitions (Saver, 2007). The Cochran-Mantel Haenszel (CMH) shift test can also be used

to analyze the distribution of ordinal data (Savitz et al., 2007). It is not feasible for large scale clinical trials with non-simple randomization schemes because it can only accommodate a limited number of covariates.

An approach to transform the mRS into a patient-centered outcome measure was recently proposed (Chaisinanunkul et al., 2015). This method is easy to implement and can provide greater statistical power but it is based on only two populations and may not be representative of patients in other locations. None of the previously mentioned methods utilize the repeated measures even though outcome is collected over time. In fact, a literature search for repeated measures analysis of acute stroke trial data only returned four articles where a generalized estimating equations approach was used for repeated measures analysis (Cassarly et al., 2017a; Cassarly et al., 2017b; Feng et al., 2011; Palesch et al., 2015).

The results presented in this manuscript are the first to estimate a missing baseline ordinal outcome for use in a MSM. In the case of ischemic stroke occurrence and treatment much of the progression or recovery experienced by a patient is expected to occur early. Functional outcome measures are not suitable at baseline and as a result, functional changes over time from baseline cannot be measured. Therefore, latent estimation the functional baseline was warranted, allowing for inclusion of an informative transition from baseline to first follow-up to be included in a longitudinal model.

Using the proposed method, we were no longer restricted to modeling the functional outcome of acute stroke patients beginning with the first follow-up visit. The method predicted the unobservable baseline state for use in a model that could more fully characterize the evolution of disease starting almost immediately after stroke (at randomization into the trial). After a stroke subjects can get better or worse quickly but we do not expect rapid movement through states after the acute event. Because of this initial rapid movement through states, a piecewise-constant MSM was chosen to best accommodate the stroke process.

Using the longitudinal data, including the estimated baseline, this work showed that there are differences in the rate of transitions between the treatment and placebo groups in the NINDS tPA trial, confirming the results of the primary analysis while allowing for examination of the effect on all adjacent-state transitions. The ability to estimate the intensities for all adjacent-state transitions allowed for examination of the most significant effect of treatment. Specifically, it was determined that the most significant impact of tPA is reduction the hazard of transitioning from mRS = 1 to mRS = 2. In addition, the conclusion of no treatment effect in the ALIAS Trial data was also consistent with the primary analysis from that trial.

In the MSM of the NINDS tPA data, where the estimated baseline was not included and the model adjusted for baseline NIHSS instead, as is done in most other types of analysis of the mRS, the effect of treatment was only marginally significant. Thus, it seems as though inclusion of the estimated baseline mRS improved the ability to detect a treatment effect. It is hypothesized that the inclusion of the latently estimated baseline mRS improves the

model because of the acute nature of ischemic stroke therapy and the expected early recovery and disease progression directly following treatment.

The MSM can incorporate longitudinal ordinal data and provide clinically interpretable summary statistics to describe covariate effects on all transition rates and sojourn times. Estimation of transition rates can describe treatment effect in a much finer gradient than modeling collapsed ordinal scale allowing for a more comprehensive understanding of differences in the effect of treatment. The MSM also allows for specific hypotheses to be tested--for example a likelihood ratio test could be used to test whether the effect of treatment is the same for all forward transitions and backward transitions (Jackson, 2016).

Another benefit of these models is the potentially decreased sample size needed to detect a treatment effect. A previous simulation study indicated that the power of MSMs applied to acute onset clinical trial data was significantly increased when the number of follow-up visits was increased (Cassarly et al., 2017a). Future trials could collect the ordinal outcome more frequently over the course of follow-up, increasing the power to detect differences using this modeling technique.

A limitation of MSMs is that they are computationally intensive, especially when using bootstrapping to obtain the bootstrap empirical distribution. In addition, use of these models requires a priori decisions about the transitions that can realistically occur, which may be a data driven decision. Interpretation of the full model could potentially be overwhelming, as there are many parameters that describe the effect of one covariate; however, the model also allows for testing whether the effect of a covariate is the same for certain transitions, which could reduce the number of parameters. The flexibility to estimate the full or reduced model allows a number of clinical questions to be answered using one approach.

Future directions could include more complex methods for baseline estimation. For example, a Bayesian PCA could be used in alternative baseline estimation procedure. This method could potentially address the uncertainty of the assigned baseline scores via extraction of the posterior distributions for the component scores (Smycka & Keil, 2015). In addition, the methods presented for estimation could be extended to cases where more than two components were used in the estimation procedure.

ACKNOWLEDGEMENTS:

Supported by grant U01-NS061861-05 from the National Institute of Neurological Disorders and Stroke.

Appendix A: modified NIHSS summary

In this appendix, Table A.1 is presented, which shows the 15 items of the NIHSS. The form for recording the data contains detailed instructions for use of the scale.

Table A.1.

The Modified National Institutes of Health Stroke Scale Summary.

1A	Level of Consciousness	0 = Alert
----	------------------------	-----------

		1 = Not alert, obtunded 3 = Unresponsive
1B	LOC Questions	0 = Answers both correctly 1 = Answers one correctly 2 = Answers neither correctly
1C	LOC Commands	0 = Performs both tasks correctly 1 = Performs one task correctly 2 = Performs neither task
2	Best Gaze	0 = Normal 1 = Partial gaze palsy 2 = Total gaze palsy
3	Visual	0 = No visual loss 1 = Partial hemianopia 2 = Complete hemianopia 3 = Bilateral hemianopia
4	Facial Palsy	0 = Normal 1 = Minor paralysis 2 = Partial paralysis 3 = Complete paralysis
5	Motor Arm a. Left b. Right	0 = No drift 1 = Drift before 10 seconds 2 = Falls before 10 seconds 3 = No effort against gravity 4 = No movement
6	Motor Leg c. Left d. Right	0 = No drift 1 = Drift before 10 seconds 2 = Falls before 10 seconds 3 = No effort against gravity 4 = No movement
7	Limb Ataxia	0 = Absent 1 = Present in one limb 2 = Present in two limbs
8	Sensory	0 = Normal 1 = Mild to moderate sensory loss 2 = Severe to total sensory loss
9	Best Language	0 = No aphasia, normal 1 = Mild to moderate aphasia 2 = Severe aphasia 3 = Mute or global aphasia
10	Dysarthria	0 = Normal 1 = Mild to moderate 2 = Severe
11	Extinction and Inattention (Neglect)	0 = No abnormality 1 = Mild 2 = Severe

Appendix B: data reduction using principal component analysis

When applied to a set of variables, PCA can group correlated variables into a smaller set of independent composite variables (principal components). The resulting linear combinations of variables account for as much variability in the data as possible and can be used to calculate component scores. A number of methods exist for computation of component scores. Non-refined methods are simple, easy to compute and easy to interpret, while refined methods are more complex and exact (DiStefano et al., 2009). Refined computation methods are generally less stable across samples; hence, a non-refined method is implemented for the motivating data sets. Of the non-refined methods, the summation of standardized variables is

preferred when the standard deviations of the raw data vary widely, as was found in the NINDS tPA and ALIAS data.

PCA was used to group measures of severity (individual items of the NIHSS) and other baseline variables known to be associated with functional ability (age, baseline glucose, time from stroke onset to randomization and ASPECTS score) into components and calculate component scores. For the analysis using PCA, variables were reformatted so that the direction of effect was consistent across all candidate variable and standardized scales were used. Component scores from PCA are more intuitive if the expected relationship with the variables and outcome is in the same direction. For example, increased age increases the risk of a negative outcome. All of the variables included in the PCA are positively correlated with bad outcome except ASPECTS. A smaller ASPECTS is predictive of negative outcome. For interpretation purposes, the direction of the ASPECTS scale was reformatted such that 10 represented the worst and 0 normal.

In the process of determining the number of components to retain from the PCA analysis, the eigenvalues of the correlation matrix were visually assessed using scree plots (Figure B. 1). From this illustration it can be seen that the line flattens after the second component in the scree plot for both NINDS tPA (a) and ALIAS (b) and thus the first two components were found to adequately explain most of the variability. In order to assess how strongly the components and variables are related, the component loadings for retained components were examined. Component loadings can be interpreted as the correlation between an observed variable and a component. A general rule of thumb deems components with loadings of at least $|.4|$ highly explanatory (Stevens, 2009).

Figure B.1.

Scree plot of eigenvalues from PCA of baseline variables from: (a) NINDS tPA and (b) ALIAS.

The loadings for the first two components are displayed in Table B.1. Based on the criteria outlined above, baseline glucose, time from onset to randomization, and individual NIHSS items Q1A and Q07 were not included in the component score calculation for NINDS tPA. Age, baseline glucose, time from onset to randomization, and individual NIHSS items Q07 and Q10 were not included in the component score calculation for ALIAS.

Table B.1.

Factor loadings on components based on PCA.

	NINDS tPA Components		ALIAS Components	
	1	2	1	2
Age	0.15	-0.03	0.16	0.13
Baseline glucose	0.10	-0.04	0.06	-0.02
Stroke onset to randomization	-0.08	-0.06	0.03	-0.04
NIHSS item				
Q1A – Level of Consciousness (LOC)	0.38	0.39	0.32	0.42
Q1B – LOC Questions	0.78	-0.18	0.80	-0.01
Q1C – LOC Commands	0.75	-0.01	0.78	0.14
Q02 – Best Gaze	0.37	0.66	0.20	0.71
Q03 – Visual	0.36	0.61	0.25	0.64
Q04 – Facial Palsy	0.22	0.50	0.01	0.48
Q5A – Motor Arm Left	-0.39	0.77	-0.63	0.58
Q5B – Motor Arm Right	0.84	-0.20	0.82	-0.01
Q6A – Motor Leg Left	-0.25	0.74	-0.53	0.56
Q6B – Motor Leg Right	0.80	-0.11	0.76	0.03
Q07 – Limb Ataxia	-0.20	-0.21	-0.15	-0.24
Q08 – Sensory	0.20	0.60	-0.02	0.54
Q09 – Best Language	0.86	-0.19	0.87	-0.03
Q10 – Dysarthria	0.60	0.04	0.35	0.22
Q11 – Extinction and Inattention	0.15	0.70	-0.09	0.68
ASPECTS	0.15	0.49	0.15	0.48

Each of the variables deemed important with the cutoff of $>|.4|$ on the first two components were standardized to range [0, 1]. The standardized values of variables that loaded high on each component were summed to calculate the component scores as follows:

- NINDS tPA Component Score 1 =
 $Q1B + Q1C + Q5B + Q6B + Q09 + Q10$
- NINDS tPA Component Score 2 =
 $Q02 + Q03 + Q04 + Q5A + Q6A + Q08 + Q11 + ASPECTS$
- ALIAS Component Score 1 =
 $Q1B + Q1C + Q5B + Q6B + Q09$
- ALIAS Component Score 2 =
 $Q1A + Q02 + Q03 + Q04 + Q5A + Q6A + Q08 + Q11 + ASPECTS.$

These component scores were then used in the main manuscript to assign baseline mRS state.

Appendix C: baseline characteristics of ALIAS subjects included and excluded from MSM analysis

In this appendix, Table C.1 is presented with baseline characteristics of the 17 ALIAS subjects excluded from the analysis because of withdrawn consent or lost-to-follow-up prior to the 30 day visit in combination with all mRS scores missing post-baseline.

Table C.1.

Baseline Characteristics of the Subjects.

Characteristic		Included (N = 1258)	Excluded (N = 17)
Age [mean (SD)]		66.1 (13.6)	62.6 (15.3)
Male sex [n (%)]		674 (53.6)	10 (58.8)
Race [n (%)]			
	White	969 (77.0)	13 (76.5)
	Black	197 (15.7)	3 (17.7)
	Asian	60 (4.8)	0 (0.0)
	American Indian/Alaska Native/First Nations People	7 (0.6)	0 (0.0)
	Native Hawaiian or Pacific islander	3 (0.2)	0 (0.0)
	Multiple, Other, or Unknown	1 (0.1)	0 (0.0)
Ethnic group [n (%)]			
	Non-Hispanic/Latino	1147 (91.2)	15 (88.2)
	Hispanic/Latino	60 (4.8)	2 (11.8)
	Unknown	51 (4.1)	0 (0.0)
Medical history[N (%)]			
	Hypertension	911 (72.4)	10 (58.8)
	Atrial fibrillation	257 (20.4)	2 (11.8)
	Past congestive heart failure	55 (4.4)	1 (5.9)
	Past myocardial infarction	155 (12.3)	5 (29.4)
	Past stroke	238 (18.9)	6 (35.3)
	Past transient ischaemic attack	157 (12.5)	2 (11.8)
	Diabetes mellitus	261 (20.8)	4 (23.5)
	Hyperlipidemia	554 (44.0)	8 (47.1)
	Peripheral vascular disease	75 (6.0)	1 (5.9)
Baseline NIHSS score [median, (inter-quartile range)]		11 (8 - 17)	9 (7 - 13)
Baseline ASPECTS score			
Baseline ASPECTS > 7 [n/N, (%)]		932/1238 (75.3)	15/16 (93.8)

Characteristic		Included (N = 1258)	Excluded (N = 17)
Clinical findings [mean (SD)]	Systolic blood pressure, mm Hg	156.9 (29.0)	157.3 (31.3)
	Plasma glucose, mmol/L	7.4 (3.2)	7.9 (3.3)
	Creatinine, μ mol/L	90.2 (25.9)	92.7 (27.4)

REFERENCES

- Aalen OO, Farewell V, De Angelis D, Day N, & Gill O (1997). A Markov model for HIV disease progression including the effect of HIV diagnosis and treatment: application to AIDS prediction in England and Wales. *Stat Med*, 16(19), 2191–2210. [PubMed: 9330428]
- Adams HP, Davis PH, Leira EC, Chang K-C, Bendixen BH, Clarke WR, Woolson RF, & Hansen MD (1999). Baseline NIH Stroke Scale score strongly predicts outcome after stroke: A report of the Trial of Org 10172 in Acute Stroke Treatment (TOAST). *Neurology*, 53(1), 126. [PubMed: 10408548]
- Agresti A (1999). Modelling ordered categorical data: recent advances and future challenges. *Stat Med*, 18(17-18), 2191–2207. [PubMed: 10474133]
- Bruno A, Saha C, & Williams LS (2009). Percent Change on the National Institutes of Health Stroke Scale: A Useful Acute Stroke Outcome Measure. *J Stroke Cerebrovasc Dis*, 18(1), 56–59. [PubMed: 19110146]
- Capes SE, Hunt D, Malmberg K, Pathak P, & Gerstein HC (2001). Stress Hyperglycemia and Prognosis of Stroke in Nondiabetic and Diabetic Patients. *A Systematic Overview*, 32(10), 2426–2432.
- Cassarly C, Martin RH, Chimowitz M, Peña EA, Ramakrishnan V, & Palesch YY (2017a). Assessing Type I error and power of multistate Markov models for panel data—A simulation study. *Communications in Statistics - Simulation and Computation*, 46(9), 7040–7061. [PubMed: 29225407]
- Cassarly C, Martin RH, Chimowitz M, Peña EA, Ramakrishnan V, & Palesch YY (2017b). Comparison of multistate Markov modeling with contemporary outcomes in a reanalysis of the NINDS tissue plasminogen activator for acute ischemic stroke treatment trial. *PLOS ONE*, 12(10), e0187050. [PubMed: 29073195]
- Chaisinankul N, Adeoye O, Lewis RJ, Grotta JC, Broderick J, Jovin TG, Nogueira RG, Elm JJ, Graves T, Berry S, Lees KR, Barreto AD, Saver JL, Furlan A, Baxter B, Lutsep HL, Ribo M, Jansen O, Gupta R, & Pereira VM (2015). Adopting a Patient-Centered Approach to Primary Outcome Analysis of Acute Stroke Trials Using a Utility-Weighted Modified Rankin Scale. *Stroke*, 46(8), 2238–2243. [PubMed: 26138130]
- Chiang CL (1980). *An Introduction to Stochastic Processes and Their Applications*. New York: Robert E. Krieger Publishing Co.
- DeSantis SM, Lazaridis C, Palesch Y, & Ramakrishnan R (2014). Regression Analysis of Ordinal Stroke Clinical Trial Outcomes: An Application to the NINDS t-PA Trial. *Int J Stroke*, 9(2), 226–231. [PubMed: 23803174]
- Feng W, Vasquez G, Suri MFK, Lakshminarayan K, & Qureshi AI (2011). Repeated-Measures Analysis of the National Institute of Neurological Disorders and Stroke rt-PA Stroke Trial. *J Stroke Cerebrovasc Dis*, 20(3), 241–246. [PubMed: 20621509]
- Ginsberg MD, Palesch YY, & Hill MD (2006). The ALIAS (ALbumin In Acute Stroke) Phase III randomized multicentre clinical trial: design and progress report. *Biochemical Society Transactions*, 34(6), 1323–1326. [PubMed: 17073812]
- Ginsberg MD, Palesch YY, Hill MD, Martin RH, Moy CS, Barsan WG, Waldman BD, Tamariz D, & Ryckborst KJ (2013). High-dose albumin treatment for acute ischaemic stroke (ALIAS) part 2: a

- randomised, double-blind, phase 3, placebo-controlled trial. *The Lancet Neurology*, 12(11), 1049–1058. [PubMed: 24076337]
- Hill MD, Demchuk AM, Goyal M, Jovin TG, Foster LD, Tomsick TA, von Kummer R, Yeatts SD, Palesch YY, & Broderick JP (2014). Alberta Stroke Program Early Computed Tomography Score to Select Patients for Endovascular Treatment. *Interventional Management of Stroke (IMS)-III Trial*, 45(2), 444–449.
- Hill MD, Martin RH, Palesch YY, Tamariz D, Waldman BD, Ryckborst KJ, Moy CS, Barsan WG, & Ginsberg MD (2011). The Albumin in Acute Stroke Part 1 Trial. *An Exploratory Efficacy Analysis*, 42(6), 1621–1625.
- Jackson C (2011). Multi-State Models for Panel Data: The msm Package for R. *J Stat Softw*, 38(8), 28.
- Jackson C (2016). *Multi-state modelling with R: the msm package*. Cambridge, U.K.: MRC Biostatistics Unit.
- Jennett B, & Bond M (1975). Assessment of outcome after severe brain damage. *Lancet*, 305, 480–484.
- Keiding N, Klein JP, & Horowitz MM (2001). Multi-state models and outcome prediction in bone marrow transplantation. *Statistics in medicine*, 20(12), 1871–1885. [PubMed: 11406847]
- Knoflach M, Matosevic B, Rücker M, Furtner M, Mair A, Wille G, Zangerle A, Werner P, Ferrari J, Schmidauer C, Seyfang L, Kiechl S, & Willeit J (2012). Functional recovery after ischemic stroke —A matter of age. *Data from the Austrian Stroke Unit Registry*, 78(4), 279–285.
- Lyden P, Brott T, Tilley B, Welch KMA, Mascha EJ, Levine S, Haley EC, Grotta J, Marler J, & Group, N. T. S. S. (1994). Improved reliability of the NIH Stroke Scale using video training. *Stroke*, 25, 2220–2226. [PubMed: 7974549]
- Mahoney FI, & Barthel DW (1965). Functional evaluation: the Barthel index. *Md State Med J*, 14, 61–65.
- Marshall G, & Jones RH (1995). Multi-state models and diabetic retinopathy. *Statistics in medicine*, 14(18), 1975–1983. [PubMed: 8677398]
- Martin RH, Yeatts SD, Hill MD, Moy CS, Ginsberg MD, & Palesch YY (2016). ALIAS (Albumin in Acute Ischemic Stroke) Trials. *Analysis of the Combined Data From Parts 1 and 2*, 47(9), 2355–2359.
- Newby LK, Rutsch WR, Califf RM, Simoons ML, Aylward PE, Armstrong PW, Woodlief LH, Lee KL, Topol EJ, & Van de Werf F (1996). Time from symptom onset to treatment and outcomes after thrombolytic therapy. *Journal of the American College of Cardiology*, 27(7), 1646–1655. [PubMed: 8636549]
- Palesch YY, Yeatts SD, Tomsick TA, Foster LD, Demchuk AM, Khatri P, Hill MD, Jauch EC, Jovin TG, Yan B, von Kummer R, Molina CA, Goyal M, Schonewille WJ, Mazighi M, Engelter ST, Anderson C, Spilker J, Carrozzella J, Ryckborst KJ, Janis LS, Simpson A, Simpson KN, & Broderick JP (2015). Twelve-Month Clinical and Quality-of-Life Outcomes in the Interventional Management of Stroke III Trial. *Stroke*, 46(5), 1321–1327. [PubMed: 25858239]
- Pérez-Ocón R, Ruiz-Castro JE, & Gámiz-Pérez ML (2001). Non-homogeneous Markov models in the analysis of survival after breast cancer. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 50(1), 111–124.
- Pexman JHW, Barber PA, Hill MD, Sevvick RJ, Demchuk AM, Hudon ME, Hu WY, & Buchan AM (2001). Use of the Alberta Stroke Program Early CT Score (ASPECTS) for Assessing CT Scans in Patients with Acute Stroke. *American Journal of Neuroradiology*, 22(8), 1534–1542. [PubMed: 11559501]
- Price M, Hertzberg V, & Wright DW (2013). Does the sliding dichotomy result in higher powered clinical trials for stroke and traumatic brain injury research? *Clinical Trials*, 10(6), 924–934. [PubMed: 23027647]
- Roozenbeek B, Lingsma HF, Perel P, Edwards P, Roberts I, Murray GD, Maas AIR, & Steyerberg EW (2011). The added value of ordinal analysis in clinical trials: an example in traumatic brain injury. *Critical Care*, 15(3).
- Sankey SS, & Weissfeld LA (1998). A study of the effect of dichotomizing ordinal data upon modeling. *Communications in Statistics - Simulation and Computation*, 27(4), 871–887.

- Saver JL (2007). Novel End Point Analytic Techniques and Interpreting Shifts Across the Entire Range of Outcome Scales in Acute Stroke Trials. *Stroke*, 38(11), 3055–3062. [PubMed: 17916765]
- Savitz SI, Lew R, Bluhmki E, Hacke W, & Fisher M (2007). Shift Analysis Versus Dichotomization of the Modified Rankin Scale Outcome Scores in the NINDS and ECASS-II Trials. *Stroke*, 38(12), 3205–3212. [PubMed: 17975102]
- Smycka J, & Keil P (2015). Bayesian PCA. Retrieved from <http://www.petrkeil.com/?p=2316>
- The National Institute of Neurological Disorders Stroke rt-PA Stroke Study Group. (1995). Tissue Plasminogen Activator for Acute Ischemic Stroke. *N Engl J Med*, 333(24), 1581–1588. [PubMed: 7477192]
- The Optimising Analysis of Stroke Trials (OAST) Collaboration. (2007). Can We Improve the Statistical Analysis of Stroke Trials? Statistical Reanalysis of Functional Outcomes in Stroke Trials. *Stroke*, 38(6), 1911–1915. [PubMed: 17463316]
- Tilley BC, Marler J, Geller NL, Lu M, Legler J, Brott T, Lyden P, & Grotta J (1996). Use of a Global Test for Multiple Outcomes in Stroke Trials With Application to the National Institute of Neurological Disorders and Stroke t-PA Stroke Trial. *Stroke*, 27(11), 2136–2142. [PubMed: 8898828]

Figure 1.
Baseline mRS score from summed standardized component scores for: (a) NINDS tPA and
(b) ALIAS

Table I.

Modified Rankin Scale.

Score	Description
0	No symptoms at all
1	No significant disability despite symptoms; able to carry out all usual duties and activities
2	Slight disability; unable to carry out all previous activities but able to look after own affairs without assistance
3	Moderate disability requiring some help, but able to walk without assistance
4	Moderately severe disability; unable to walk without assistance and unable to attend to own bodily needs without assistance
5	Severe disability; bedridden, incontinent, and requiring constant nursing care and attention
6	Dead

Table II.

NINDS tPA state table.

	To (state j)							
	1	2	3	4	5	6	7	Total
From (state i)								
1	206	52	9	3	5	0	2	277
2	87	245	39	12	23	7	8	421
3	23	73	85	35	32	14	10	272
4	13	41	62	148	52	25	10	351
5	11	26	27	84	187	53	41	429
6	4	8	7	25	83	142	83	352
Total	344	445	229	307	382	241	154	2102

Table III.

Estimated transitions rates from piecewise-constant MSM for NINDS tPA.

Transition	Placebo		tPA	
	0 t 0.333	t > 0.333	0 t 0.333	t > 0.333
1→2	10.83 (2.36, 49.71)	0.09 (0.06, 0.14)	7.91 (1.89, 33.06)	0.07 (0.05, 0.09)
2→3	19.30 (6.07, 61.37)	0.09 (0.06, 0.15)	9.75 (3.15, 30.22)	0.05 (0.03, 0.07)
3→4	12.84 (6.53, 25.28)	0.19 (0.12, 0.32)	16.59 (7.09, 38.86)	0.25 (0.14, 0.44)
4→5	24.03 (10.50, 54.98)	0.12 (0.08, 0.19)	16.07 (6.71, 38.50)	0.08 (0.05, 0.13)
5→6	2.97 (1.92, 4.59)	0.05 (0.04, 0.10)	2.63 (1.70, 4.07)	0.05 (0.03, 0.09)
5→7	0.48 (0.27, 0.86)	0.05 (0.04, 0.07)	0.46 (0.26, 0.83)	0.05 (0.03, 0.07)
6→7	0.50 (0.28, 0.88)	0.12 (0.09, 0.18)	0.60 (0.35, 1.02)	0.15 (0.10, 0.21)
2→1	7.53 (1.64, 34.55)	0.08 (0.05, 0.11)	7.54 (1.85, 30.72)	0.08 (0.06, 0.10)
3→2	18.06 (5.49, 59.44)	0.23 (0.16, 0.32)	20.59 (6.43, 65.91)	0.26 (0.18, 0.36)
4→3	14.15 (6.74, 29.71)	0.22 (0.15, 0.32)	14.06 (6.09, 32.46)	0.22 (0.15, 0.33)
5→4	6.29 (2.59, 15.26)	0.20 (0.15, 0.27)	6.86 (2.86, 16.46)	0.22 (0.16, 0.30)
6→5	3.08 (2.01, 4.72)	0.23 (0.17, 0.32)	2.46 (1.61, 3.77)	0.18 (0.13, 0.26)

Table IV.

Hazard ratios from piecewise-constant MSM for NINDS tPA.

Transition	Hazard ratio (95% CI)
1→2	0.73 (0.43, 1.23)
2→3	0.51 (0.28, 0.90)
3→4	1.29 (0.66, 2.52)
4→5	0.67 (0.38, 1.17)
5→6	0.88 (0.53, 1.48)
5→7	0.96 (0.55, 1.66)
6→7	1.20 (0.75, 1.92)
2→1	1.00 (0.62, 1.62)
3→2	1.14 (0.71, 1.83)
4→3	0.99 (0.59, 1.68)
5→4	1.09 (0.72, 1.65)
6→5	0.80 (0.52, 1.23)