



Probing the Effective Treatment Thresholds for Alteplase in Acute Ischemic Stroke With Regression Discontinuity Designs

Andrew M. Naidech^{1*}, Patrick N. Lawlor^{2†}, Haolin Xu³, Gregg C. Fonarow⁴, Ying Xian^{3,5}, Eric E. Smith⁶, Lee Schwamm⁷, Roland Matsouaka^{8,9}, Shyam Prabhakaran¹, Ioana Marinescu¹⁰ and Konrad P. Kording¹¹

OPEN ACCESS

Edited by:

Guillaume Turc,
Centre Hospitalier
Sainte-Anne, France

Reviewed by:

Stefan Greisenegger,
Medical University of Vienna, Austria
Geoff Cohen,
University of Edinburgh,
United Kingdom

*Correspondence:

Andrew M. Naidech
a-naidech@northwestern.edu

†These authors share first authorship
The manuscript was written at
Northwestern; statistical analysis was
performed at Duke

Specialty section:

This article was submitted to
Stroke,
a section of the journal
Frontiers in Neurology

Received: 07 October 2019

Accepted: 24 July 2020

Published: 02 September 2020

Citation:

Naidech AM, Lawlor PN, Xu H, Fonarow GC, Xian Y, Smith EE, Schwamm L, Matsouaka R, Prabhakaran S, Marinescu I and Kording KP (2020) Probing the Effective Treatment Thresholds for Alteplase in Acute Ischemic Stroke With Regression Discontinuity Designs. *Front. Neurol.* 11:961. doi: 10.3389/fneur.2020.00961

¹ Department of Neurology, Northwestern University, Chicago, IL, United States, ² Division of Neurology, Children's Hospital of Philadelphia, Philadelphia, PA, United States, ³ Duke Clinical Research Institute, Duke University, Durham, NC, United States, ⁴ Division of Cardiology, University of California, Los Angeles, Los Angeles, CA, United States, ⁵ Department of Neurology, Duke University Medical Center, Durham, NC, United States, ⁶ Department of Neurology, University of Calgary, Calgary, AB, Canada, ⁷ Department of Neurology, Massachusetts General Hospital, Boston, MA, United States, ⁸ Department of Biostatistics and Bioinformatics, Duke University, Durham, NC, United States, ⁹ Program for Comparative Effectiveness Methodology, Duke Clinical Research Institute, Duke University, Durham, NC, United States, ¹⁰ School of Social Policy & Practice, University of Pennsylvania, Philadelphia, PA, United States, ¹¹ Departments of Neuroscience and Bioengineering, University of Pennsylvania, Philadelphia, PA, United States

Randomized Controlled Trials (RCTs) are considered the gold standard for measuring the efficacy of medical interventions. However, RCTs are expensive, and use a limited population. Techniques to estimate the effects of stroke interventions from observational data that minimize confounding would be useful. We used regression discontinuity design (RDD), a technique well-established in economics, on the Get With The Guidelines-Stroke (GWTG-Stroke) data set. RDD, based on regression, measures the occurrence of a discontinuity in an outcome (e.g., odds of home discharge) as a function of an intervention (e.g., alteplase) that becomes significantly more likely when crossing the threshold of a continuous variable that determines that intervention (e.g., time from symptom onset, since alteplase is only given if symptom onset is less than e.g., 3 h). The technique assumes that patients near either side of a threshold (e.g., 2.99 and 3.01 h from symptom onset) are indistinguishable other than the use of the treatment. We compared outcomes of patients whose estimated onset to treatment time fell on either side of the treatment threshold for three cohorts of patients in the GWTG-Stroke data set. This data set spanned three different treatment thresholds for alteplase (3 h, 2003–2007, $N = 1,869$; 3 h, 2009–2016, $N = 13,086$, and 4.5 h, 2009–2016, $N = 6,550$). Patient demographic characteristics were overall similar across the treatment thresholds. We did not find evidence of a discontinuity in clinical outcome at any treatment threshold attributable to alteplase. Potential reasons for failing to find an effect include violation of some RDD assumptions in clinical care, large sample sizes required, or already-well-chosen treatment threshold.

Keywords: alteplase, ischemic stroke, regression discontinuity design, causal inference, quasi-experiments

INTRODUCTION

Randomized controlled trials (RCTs) are considered the gold standard in clinical investigation because, ideally, RCTs remove both known and unknown imbalances in groups that could lead investigators to wrongly conclude a treatment is efficacious (1). Data from RCTs are generally required before regulatory approval is granted to market a drug treatment (e.g., alteplase for acute ischemic stroke) (2), and new interventions typically require data from RCTs on efficacy before they are widely accepted.

Therapies that have a treatment threshold lead to challenging problems about the choice of those thresholds. In the case of alteplase, only patients whose symptoms began before the threshold time (e.g., 3 h prior to presentation) are eligible for treatment. The time window curtails treatment in clinical practice, and off-label use beyond approved time windows introduces legal and ethical concerns. If a treatment is effective within a narrow time window (e.g., 3 h for alteplase), there is typically a desire to extend it further to increase the number of patients who might be treated. Yet, each new time window typically requires another RCT, with the associated time and expenses of planning and conducting the trial. In the case of alteplase, data from other clinical trials was utilized to propose additional RCTs. However, extending the time window for fibrinolytic treatment expressly carried an increased risk of intracranial hemorrhage, which was borne out in a RCT with an extended time window (3). The concern for symptomatic hemorrhage has guided the design and conduct of RCTs for ischemic stroke generally (4, 5). Publication of an RCT showing that treatment up to 4.5 h after symptom onset was efficacious required several more years (6). Clinicians are often hesitant to wait years for new RCTs, and may treat patients outside of rigorously applied clinical trial protocols (7). Each new therapy (e.g., endovascular therapy for large vessel occlusion) brings a similar invitation to extend the window as long as it is efficacious. Conversely, some RCTs of time-limited therapies are negative, leading to the testing of more stringent time windows in hopes of finding efficacy [e.g., shortening the window of recombinant Factor VII for intracerebral hemorrhage from 3 h of symptom onset (8) to 2.5 h (9)]. Methods to hasten the determination of effective time windows for treatments with a threshold are needed.

New analytic techniques may improve our ability to determine the optimal treatment window for time-limited treatments. Regression Discontinuity Design (RDD), well-validated in economics and epidemiology (10–13), could be particularly helpful for determining if treatment thresholds are correctly set. RDD uses *observational data* to examine whether patients just above and just below the treatment threshold have different outcomes. RDD depends on the assumption that patients within a small window on either side of a threshold are no different other than being eligible for a treatment. We hypothesized that RDD would be a useful technique to evaluate alteplase treatment thresholds using observational data (clinical data), and that we could compare the results with those from already-conducted

RCTs (6). RDD could eventually supplement RCTs in clinical decision making.

METHODS

We queried the Get With The Guidelines-Stroke (GWTG-Stroke) stroke data set, a long-standing, observational registry of patients with stroke. The methods of the GWTG-Stroke data set have been previously described in detail (14–16). To perform an RDD analysis, one requires for all patients the outcome variable of interest (e.g., home discharge), the continuous variable which determines the intervention (e.g., time from symptom onset until treatment decision), and the threshold value for this continuous variable that determines treatment administration [e.g., 3 h as shown in the original NINDS trial (2), or 4.5 h as in the ECASS III trial (6)]. If there is a difference (“discontinuity”) in outcomes between these two groups (that is not attributable to confounding variables), the difference in outcomes can be attributed to the treatment. To perform this analysis, a regression model is fit, including a parameter that estimates the magnitude of discontinuity in outcomes between patients on either side of the threshold.

We examined the following clinical outcomes for signs of a discontinuity: good discharge disposition, in-hospital mortality, length of hospital stay, ambulatory status at discharge, and modified Rankin Scale at discharge (see **Supplementary Table 1** for details). We also analyzed factors that might have influenced the use or effectiveness of alteplase by building more sophisticated, adjusted models. These characteristics included patient age, sex, race/ethnicity, relevant past medical history (e.g., prior ischemic stroke, coronary heart disease, etc., NIH Stroke Scale (NIHSS), as well as hospital characteristics such as setting (urban, rural), and annual volume of patients with ischemic stroke. These demographic variables allow us to more accurately estimate the effect of alteplase, but also to confirm the validity of our model. The unadjusted models do not include this additional information. Since patients are assumed to be similar on either side of the threshold, a discontinuity in a demographic variable at the threshold could be evidence that this assumption is violated. Thus, we checked for discontinuities in demographic factors as well as in outcomes.

All participating institutions were required to comply with local regulatory and privacy guidelines and, if required, to secure institutional review board approval. Because data were used primarily at the local site for quality improvement, sites were granted a waiver of informed consent under the common rule.

METHODS: RUNNING VARIABLE IMPUTATION

Our analysis required modification of the standard RDD setup due to different data recorded for treated and untreated patients. In a standard RDD, the running variable that determines the treatment is observed for both patients that are treated, as well as for patients that are not treated (e.g., a test score that determines a scholarship award). In our case, alteplase should

only be administered within the time window in question (e.g., before 3 h) starting with the time of symptoms onset. Thus, for patients who are treated, the sequence of events that must take place within this window are: (1) development of symptoms, (2) clinical imaging (CT scan) to determine the need for treatment, and (3) the treatment itself. A reasonable running variable for this situation is the sum of these three times, which we call *onset to treatment time*. The situation for patients who are not treated is different because there is no time of treatment with alteplase. To address this issue, for both treated and untreated patients, we imputed the time from clinical imaging to treatment. We imputed this value using the median from the patient's hospital in that particular year. We call this running variable *estimated onset to treatment time* (OTT_{est}). Multilevel modeling could be useful in such a scenario, but this is not yet routinely done in RDD analyses.

This method of estimating times for treated and untreated patients is reasonable. First, the time to administration of alteplase is known for each hospital each year, and reflects established protocols of stroke care. Second, the magnitude of the imputed value is about 20% of the total running variable; modest compared with the other two components. Third, the imputed *onset to treatment time* still reflects the most important clinical information that is used to determine alteplase administration (time from symptom development, and time of clinical imaging). We feel that the *onset to treatment time* and *estimated onset to treatment time* both contain essentially the same clinical information for treated and untreated patients, making it a reasonable running variable. Last, imputation of this value does not bias us toward finding an effect of alteplase, as would be the case if imputation introduced an artificial discontinuity in *estimated onset to treatment time*. Notably, the imputed time was used for both treated and untreated patients. Furthermore, *time from symptom development* and *time to imaging* for untreated patients are noisy measurements, and imputation of the *time from imaging to treatment* simply adds a fixed value to this. If anything, it introduces more noise, biasing us against finding a causal effect of alteplase.

METHODS: SHARP RDD

To determine if there is a discontinuity in outcome at the desired treatment threshold (e.g., 3 h from symptom onset), we constructed regression models that allow clinical outcomes (e.g., favorable hospital disposition) to be explained by the time from symptom onset to treatment—whether the patient was treated with alteplase—and patient- and hospital-specific characteristics (Equation 1). In the simplest form of RDD, sharp RDD, all patients on the left of the threshold (e.g., before 3 h) would receive alteplase, and none of the patients on the right of the threshold (e.g., after 3 h) would receive it. Treatment with alteplase is thus modeled as binary, and depends only on *estimated onset to treatment time* and the treatment threshold. Its corresponding regression parameter (Equation 1, parameter b_3) is an estimate of the effect of treatment with alteplase. If there is a discontinuity in outcome at the treatment threshold (that is

not attributable to confounding factors), the parameter relating alteplase administration to clinical outcome will be significantly different from zero.

We fit a logistic regression model of the log odds of favorable clinical outcome (a dichotomous variable, see **Supplementary Table 1** for clinical outcomes examined), as a function of regressors given below, as well as patient- and hospital-specific characteristics (see **Supplementary Table 2**). NIH Stroke Scale was considered in a separate adjusted model due to a substantial amount of missing data (**Supplementary Table 7**). The adjusted equation for sharp RDD was:

$$\ln\left(\frac{P(Y)}{1-P(Y)}\right) = LO(Y) = b_0 + b_1(OTT_{est} - c) + b_2(OTT_{est} - c)I[OTT_{est} \leq c] + b_3I[OTT_{est} \leq c] + b_4X + \varepsilon \quad (1)$$

Y is the dichotomized clinical outcome, $P(Y)$ is the probability of that outcome, and $LO(Y)$ is the log odds of that outcome. OTT_{est} indicates *estimated onset to treatment time*; $I[OTT_{est} \leq c]$ is an indicator variable that equals 1 when the patient's onset to treatment time is less than the threshold time, c , and 0 otherwise; c has a value of 3 or 4.5 h depending on the cohort; X is a matrix containing the patient- and hospital-specific factors considered (see **Supplementary Table 2**); ε is an error term. The parameter b_3 is the one of greatest interest, as it models the effect of alteplase treatment on log odds of clinical outcome at the threshold. The meaning of b_3 can be seen by taking the difference in log odds of outcome of treated and untreated patients at the threshold (Equation 2); it is exactly the difference in log odds of outcome at the threshold. Exponentiating b_3 gives the odds ratio of good clinical outcome to poor clinical outcome at the threshold. Values of $\exp(b_3)$ greater than one (equivalent to $b_3 > 0$) should be interpreted as increasing the odds of the clinical outcome of interest.

$$LO(Y)|_{OTT=c_-} - LO(Y)|_{OTT=c_+} = b_3 \quad (2)$$

The parameter b_0 is an intercept term; b_1 gives the slope of log odds of Y as a function of OTT_{est} on the right side of the cutoff; b_2 is the parameter of the interaction term $(OTT_{est} - c)I[OTT_{est} \leq c]$, which effectively allows the slope to be different on the left and right sides of the threshold; b_4 is the set of weights corresponding to each of the patient- and hospital-specific variables contained in X . In RDD analyses, the intercept and slope parameters are not typically of primary interest for causal interpretation. One can still, for example, speculate about the meaning of the slope, with the caveat that patients become less similar as you get farther from the threshold, and so are less comparable. In contrast, the discontinuity parameter, b_3 , is interpreted as the causal effect of the treatment at the threshold.

METHODS: FUZZY RDD

In some cases, alteplase was administered outside of the treatment window, or not administered within the treatment window, meaning that the assumptions of sharp RDD do not

strictly apply. Thus, we also applied fuzzy RDD, a technique used to model imperfect treatment compliance. Fuzzy RDD is widely used in health policy research and elsewhere (17, 18).

We used a two-stage, instrumental variable approach. Stage one models alteplase treatment administration (i.e., compliance); stage two models the clinical outcome as a function of OTT_{est} and alteplase administration, similar to sharp RDD. Both stages also include patient- and hospital-specific characteristics, as these could differentially affect alteplase administration and clinical outcome.

We modeled the log odds of alteplase administration as a function of running variable, treatment window threshold, and patient- and hospital-specific factors. The equation for Stage 1 was:

$$\ln\left(\frac{P(\text{tPA})}{1 - P(\text{tPA})}\right) = LO(\text{tPA}) = b_0^{(1)} + b_1^{(1)}I[OTT_{est} \leq c] + b_2^{(1)}X + \varepsilon \quad (3)$$

Here, $P(\text{tPA})$ is the probability of alteplase administration, and $LO(\text{tPA})$ is the log odds of tPA administration. OTT_{est} indicates the imputed onset to treatment time. The parameters of this equation, $\{b_i^{(1)}\}$ have a superscript that denotes the stage of the model (first stage, here), and a subscript that uniquely identifies each term. The parameter $b_0^{(1)}$ is an intercept for the model of alteplase compliance; $b_1^{(1)}$ is a parameter that models alteplase compliance on the left side of the threshold; $b_2^{(1)}$ is a set of parameters that models alteplase administration as a function of patient- and hospital-specific factors.

In the second stage, we modeled the log odds of clinical outcome as a function of running variable, treatment window threshold, and patient- and hospital- specific factors. The equation for Stage 2 was:

$$\ln\left(\frac{P(Y)}{1 - P(Y)}\right) = b_0^{(2)} + b_1^{(2)}(OTT_{est} - c) + b_2^{(2)}(OTT_{est} - c)LO(\text{tPA}) + b_3^{(2)}LO(\text{tPA}) + b_4^{(2)}X + \varepsilon \quad (4)$$

Here, again, Y refers to the dichotomized clinical outcome and $P(Y)$ to the probability of the clinical outcome. $LO(\text{tPA})$ refers to the predicted log odds of alteplase administration, from the first stage (Equation 3). The parameters of this equation, $\{b_i^{(2)}\}$ have a superscript that denotes the stage of the model (second stage, here), and a subscript that uniquely identifies each term. The parameter $b_3^{(2)}$ is again of greatest interest, as it models the effect of alteplase treatment on log odds of clinical outcome. The meaning of $b_3^{(2)}$ can be seen by taking the difference in log odds of outcome of treated and untreated patients at the threshold (Equation 5). Unlike sharp RDD, the difference in log odds depends on $b_3^{(2)}$ as well as $b_1^{(1)}$. This makes sense because $b_1^{(1)}$ is the additional log odds of alteplase administration for patients with OTT_{est} within the treatment window. In other words, outcomes in a fuzzy RDD model depend both on the treatment effect size

and the compliance. Values of $b_3^{(2)} > 0$ should be interpreted as increasing the odds of the clinical outcome of interest.

$$LO(Y)|_{OTT=c_-} - LO(Y)|_{OTT=c_+} = b_3^{(2)}b_1^{(1)} \quad (5)$$

The parameter $b_0^{(2)}$ is an intercept term; $b_1^{(2)}$ gives the slope of log odds of Y as a function of OTT_{est} independent of alteplase administration (analogous to b_1 in Equation 1); $b_2^{(2)}$ gives the slope of log odds of Y as a function of $(OTT_{est} - c)LO(\text{tPA})$ (analogous to b_2 in Equation 1); b_4 is the set of weights corresponding to each of the patient- and hospital-specific variables contained in X .

Another important issue with instrumental variable models is the “strength” of the instrument. If there were a weak or non-existent relationship between *estimated onset to treatment time* and *odds of tPA administration*, there would be no point in examining whether *estimated onset to treatment time* affects clinical outcomes, since our proposed model is that *estimated onset to treatment time* affects *odds of alteplase administration* which in turn may affect *clinical outcome*. **Supplementary Table 3** shows percentage of patients receiving alteplase on each side of the threshold in each cohort. In all cohorts there is meaningful difference in odds of alteplase administration. Thus, we have a relatively strong instrument to examine the effect of alteplase on clinical outcomes.

Analyses were performed using the SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) and R version 3.4.4 (R Core Team, 2018, Vienna, Austria). All P -values are 2-sided tests and were considered statistically significant at <0.05 . Duke Clinical Research Institute (DCRI) served as the data analysis center.

RESULTS

We analyzed three separate cohorts within the GWTG-Stroke registry (**Table 1**). Cohort A includes patients prior to 2008, when 3 h was the accepted treatment threshold (**Table 2**). Calendar year 2008 was not modeled, as this was the year that ECASS III reported alteplase was effective up to 4.5 h after symptom onset (6), which represented a new treatment threshold, and practice shifted. In 2009 and afterwards, treatment thresholds of both 3 and 4.5 h were observed for ongoing study. Cohort B includes patients after 2009 who were treated according to the 4.5 h threshold (**Table 3**). Cohort C includes patients after 2009 who were treated according to the 3 h threshold (**Table 4**). We included only patients with *estimated onset to treatment times* within 20 min of the treatment threshold, with valid imaging, and valid imaging times. We excluded inter-hospital transfers, stroke occurring after hospital arrival, contraindications to alteplase, and enrollment in clinical trials of alteplase. After these restrictions, there were 1,869 patients for analysis in Cohort A, 6,550 patients in Cohort B, and 13,086 patients in Cohort C. Cohorts on either side of the treatment threshold were generally well-matched, although there were some differences (**Supplementary Tables 4–6**; and see below).

The central result of our study was that, in terms of clinical outcomes (e.g., odds of good discharge disposition), there was no strong evidence of a discontinuity at either the 3 or 4.5 h

TABLE 1 | Cohort Demographics.

Variable	Cohort A (2.67–3.3 h, <2009)	Cohort B (4.16–4.83 h, >2009)	Cohort C (2.67–3.3 h, >2009)
N	1,869	6,550	13,086
Age, years	73 (61–82)	73 (61–83)	72 (61.8)
Race, White	77.7%	71.5%	71.9%
Black	12.2%	15.5%	14.3%
Asian	1.8%	2.4%	7.0%
Other	3.9%	3.5%	3.8%
NIH Stroke Scale	9 (5–15)	5 (2–11)	8 (4–14)
Women	48.7%	50.2%	50.4%
Atrial fibrillation	22.1%	19.1%	19.7%
Prior Stroke or TIA	27.8%	31.1%	27.6%
Coronary artery disease	29.4%	26.2%	25.1%
Diabetes mellitus	28.7%	31.7%	29.7%
Historical hypertension	72.6%	76.3%	74.2%
Tobacco Use	19.8%	17.5%	17.2%
Dyslipidemia	35.9%	45.3%	43.9%
Primary Stroke Center	44.8%	46.4%	45.9%
Academic Hospital	66.6%	61.8%	63.1%

Data are N, %, or mean (25–75%ile).

TABLE 2 | RDD results from Cohort A (treatment threshold of 3 h, prior to 2008).

Outcomes	Coefficients	Unadjusted		Adjusted		NIHSS Adjusted	
		Discontinuity, exp(b_3)	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)
Sharp RDD							
Good discharge disposition	Before vs. after	0.91 (0.63–1.32)	0.621	0.88 (0.57–1.36)	0.570	0.78 (0.44–1.38)	0.394
Discharged home	Before vs. after	0.89 (0.63–1.24)	0.480	0.86 (0.60–1.24)	0.424	0.91 (0.53–1.54)	0.718
In-hospital mortality	Before vs. after	0.86 (0.46–1.61)	0.644	0.87 (0.46–1.64)	0.669	0.69 (0.29–1.62)	0.390
LOS ≤4 days	Before vs. after	0.71 (0.51–1.00)	0.052	0.70 (0.49–1.00)	0.048	0.85 (0.54–1.33)	0.469
Ambulate independently at discharge	Before vs. after	0.90 (0.63–1.28)	0.550	0.85 (0.58–1.26)	0.429	0.90 (0.52–1.55)	0.696
mRS 0–1 vs. 2–6	Before vs. after	–	–	–	–	–	–
Fuzzy RDD							
		Estimate	P-value	Estimate	P-value	Estimate	P-value
Stage 1	$b_1^{(1)}$						
tPA	Before vs. after cut-off	1.055	<.0001	1.113	<.0001	1.177	<.0001
Stage 2	$b_3^{(2)}$						
Good disposition	tPA Yes/No	–0.366	0.621	–0.764	0.357	–1.879	0.074
Discharged home	tPA Yes/No	–0.468	0.480	–0.668	0.362	–0.867	0.398
In-hospital mortality	tPA Yes/No	–0.574	0.644	–0.085	0.949	0.328	0.851
LOS ≤4 days	tPA Yes/No	–1.312	0.052	–1.460	0.035	–1.021	0.230
Able to ambulate independently at discharge	tPA Yes/No	–0.417	0.550	–1.100	0.158	–1.133	0.289
mRS 0–1 vs. 2–6	tPA Yes/No	–	–	–	–	–	–

threshold. This was true using both sharp and fuzzy RDD, suggesting that the result does not depend strongly on the formulation of the model. This was also true using both simple models without patient- and hospital-specific factors and NIH stroke scale, as well as models with those factors. This suggests that the result does not depend strongly on model complexity, and is not related to changes in patient characteristics across the alteplase administration threshold.

We did detect some weak effects with this analysis, however. In Cohort A, the odds of length of stay <4 days was significantly

different across the threshold in both the sharp and fuzzy adjusted models (Table 2). But the direction of the effect was opposite what would be expected; patients treated with alteplase were less likely to have a length of stay <4 days than were the untreated patients. In Cohorts B and C, there were multiple outcomes with significant differences across the threshold in the NIHSS-adjusted models (Tables 3, 4). Some of these effects, again, had the opposite direction that would be expected with alteplase treatment (negative values of $b_3^{(2)}$). Furthermore, there was a significant amount of missing NIHSS

TABLE 3 | RDD results from Cohort B (treatment threshold of 4.5 h, after 2009).

Outcomes	Coefficients	Unadjusted		Adjusted		NIHSS Adjusted	
		Discontinuity, exp(b_3)	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)
Sharp RDD							
Good discharge disposition	Before vs. after	0.95 (0.76–1.18)	0.627	0.92 (0.72–1.17)	0.504	0.86 (0.64–1.15)	0.300
Discharged home	Before vs. after	1.13 (0.93–1.37)	0.230	1.11 (0.90–1.38)	0.312	1.20 (0.92–1.55)	0.171
In-hospital mortality	Before vs. after	1.09 (0.70–1.71)	0.701	1.12 (0.71–1.77)	0.637	1.27 (0.71–2.28)	0.417
LOS \leq 4 days	Before vs. after	1.21 (1.00–1.47)	0.056	1.20 (0.98–1.46)	0.081	1.33 (1.05–1.67)	0.017
Able to ambulate independently at discharge	Before vs. after	0.97 (0.80–1.19)	0.790	0.95 (0.77–1.17)	0.615	0.92 (0.72–1.18)	0.498
mRS 0–1 vs. 2–6	Before vs. after	1.26 (0.85–1.88)	0.254	1.29 (0.84–2.00)	0.250	1.42 (0.89–2.28)	0.144
Fuzzy RDD							
Stage 1		$b_1^{(1)}$					
tPA	Before vs. after cutoff	0.977	<.0001	1.013	<.0001	1.043	<.0001
Stage 2		$b_3^{(2)}$					
Good discharge disposition	tPA Yes/No	–0.304	0.627	–0.517	0.408	–0.452	0.482
Discharged home	tPA Yes/No	0.653	0.230	0.482	0.353	1.292	0.028
In-hospital mortality	tPA Yes/No	0.484	0.701	0.071	0.951	0.652	0.579
LOS \leq 4 days	tPA Yes/No	1.051	0.056	0.859	0.073	1.072	0.031
Able to ambulate independently at discharge	tPA Yes/No	–0.150	0.790	–0.543	0.304	–0.124	0.827
mRS 0–1 vs. 2–6	tPA Yes/No	1.276	0.254	0.265	0.791	1.037	0.291

TABLE 4 | RDD results from Cohort C (treatment threshold of 3 h, after 2009).

Outcomes	Coefficients	Unadjusted		Adjusted		NIHSS Adjusted	
		Discontinuity, exp(b_3)	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)
Sharp RDD							
Good discharge disposition	Before vs. after	1.00 (0.86–1.17)	0.986	1.07 (0.90–1.26)	0.442	0.98 (0.81–1.19)	0.850
Discharged home	Before vs. after	1.00 (0.88–1.15)	0.948	1.06 (0.91–1.24)	0.431	1.04 (0.87–1.23)	0.691
In-hospital mortality	Before vs. after	1.07 (0.78–1.45)	0.681	1.04 (0.77–1.40)	0.795	1.22 (0.87–1.71)	0.246
LOS \leq 4 days	Before vs. after	1.04 (0.90–1.19)	0.601	1.06 (0.91–1.22)	0.462	1.02 (0.87–1.19)	0.830
Able to ambulate independently at discharge	Before vs. after	0.92 (0.81–1.06)	0.252	0.96 (0.83–1.11)	0.588	0.91 (0.78–1.06)	0.237
mRS 0–1 vs. 2–6	Before vs. after	0.96 (0.74–1.24)	0.739	0.98 (0.75–1.29)	0.902	0.95 (0.71–1.28)	0.747
Fuzzy RDD							
Stage 1		$b_1^{(1)}$					
tPA	Before vs. after	0.723	<.0001	0.756	<.0001	0.858	<.0001
Stage 2		$b_3^{(2)}$					
Good discharge disposition	tPA Yes/No	0.012	0.986	–0.266	0.687	–0.864	0.125
Discharged home	tPA Yes/No	0.039	0.948	–0.204	0.736	–1.193	0.019
In-hospital mortality	tPA Yes/No	0.557	0.681	1.166	0.327	1.830	0.153
LOS \leq 4 days	tPA Yes/No	0.315	0.601	0.260	0.647	–1.067	0.021
Able to ambulate independently at discharge	tPA Yes/No	–0.674	0.252	–0.952	0.107	–1.534	0.001
mRS 0–1 vs. 2–6	tPA Yes/No	–0.376	0.739	–1.232	0.260	–2.267	0.009

data (Supplementary Table 7), and likely as a result, there were significant differences in initial NIHSS score across the thresholds (see below). Overall, we feel that these results are more likely explained by the fact that we tested many outcomes and would expect some to be positive by chance, and by biased subgroups. We thus do not feel justified in claiming any strong effects of alteplase administration at the treatment thresholds of 3 or 4.5 h.

We also checked for confounding by examining whether patient- and hospital-specific quantities (e.g., demographic factors) had discontinuities at the treatment thresholds. For an

RDD to be valid, there should not be discontinuities in these quantities at the threshold, since any putative discontinuity in outcome could be attributed to the discontinuity in these factors rather than to the treatment itself. Thus, this analysis would be more important if there were significant differences in clinical outcome, which we did not find. Nevertheless, we tested many patient- and hospital-specific quantities (Supplementary Tables 4–6). Cohorts B and C included significantly more patients than Cohort A, and had more significant differences in patient- and hospital-specific

quantities across the threshold. In cohorts B and C, initial NIHSS was higher in post-threshold patients than pre-threshold patients. NIHSS was missing in a larger percentage of pre-threshold patients in Cohorts A and C. This likely means that NIHSS score measurement was not random, and that the NIHSS-possessing subgroups were not as comparable on either side of the treatment thresholds (see paragraph above). This casts some doubt on the results of the NIHSS-adjusted models. Ambulatory status was different across the threshold in Cohorts B and C. Overall, there were some patient- and hospital-specific quantities that were imbalanced across the threshold. Some of this may be due to true imbalance, some due to chance, and perhaps some due to differences in data collection on each side of the threshold, which is likely not as consistent as in an RCT. Again, the lack of differences in outcome make this analysis less crucial. Experimental techniques short of experimental randomization do have their limitations, but it is also important to keep in mind that RCTs can also have such imbalance that does not necessarily threaten causal validity.

DISCUSSION

Using RDD in a nationwide data set, we found no convincing evidence of a discontinuity in the effectiveness of alteplase on clinical outcomes after ischemic stroke around the 3 and 4.5 h treatment thresholds. This result is not necessarily inconsistent with an RCT (6) showing positive, but modest, effects in the 3 to 4.5 h window. These data suggest that using RDD to extend treatment windows for alteplase may be challenging.

There are several potential explanations for not detecting an effect of alteplase at these thresholds. First, our analyses sought to detect a difference during the hospitalization or at the time of hospital discharge. Alteplase leads to improved outcome at 3 months, not to improved discharge disposition, so functional outcomes [e.g., the modified Rankin Scale (2, 6) months later] or health-related quality of life (19–21), might be more likely to show effects of alteplase treatment. Next, our study may have been underpowered. We chose a window of 20 min on either side of the treatment threshold, which led to smaller cohorts. These potential limitations may have attenuated our ability to detect an effect using RDD. Next, our continuous variable, *estimated onset to treatment time*, was imputed using hospital-specific characteristics, which added some noise near the threshold, further reducing power. Lastly, treatment thresholds chosen by expert clinicians may be accurate guesses of the optimal treatment threshold, and so further extensions of time windows for treatment may have small marginal benefits (22, 23). For example, if patients just outside the treatment threshold could benefit from alteplase, one would expect the difference in outcomes at the threshold to be larger. Thresholds chosen by experts may also incorporate other information that affects outcome, such as the increased risk of intracerebral hemorrhage due to alteplase, which increases with an expanded time window from symptom onset to treatment (3). These potential limitations may have attenuated our ability to detect an effect using RDD. Future research using RDD will need to carefully choose

questions where there are substantial existing data that closely approximate the clinical trial question of interest.

In summary, our RDD analysis did not find a discontinuity in disposition around 3 or 4.5 h alteplase treatment thresholds in acute ischemic stroke. Future investigations with RDD might leverage large data sets with outcome measures sensitive to the intervention, that can be culled from the electronic health record, potentially combining multiple hospitals and institutions.

DATA AVAILABILITY STATEMENT

The datasets generated for this study will not be made publicly available. Data are maintained by the Get With The Guidelines (GWTG) consortium at the Duke Clinical Research Institute, where the analysis was performed. Authors should contact GWTG with data requests.

ETHICS STATEMENT

The DUHS IRB has determined the specific components in this study to be in compliance with all applicable Health Insurance Portability and Accountability Act (HIPAA) regulations. The DUHS IRB waived the need for formal ethical approval for this study. Written informed consent was not required as per local legislation.

AUTHOR CONTRIBUTIONS

AN introduced the concept with KK and co-wrote the manuscript. PL co-wrote the manuscript. HX, YX, and RM performed and oversaw statistical analysis and edited the manuscript. GF, ES, LS, and RM revised the manuscript for intellectual content. SP developed the concept with AN and KK and revised the manuscript. IM revised the manuscript for intellectual content and oversaw the statistical analysis. KK developed the concept and revised the manuscript for intellectual content. All authors contributed to the article and approved the submitted version.

FUNDING

The Get With The Guidelines[®]-Stroke (GWTG-Stroke) program is provided by the American Heart Association/American Stroke Association. GWTG-Stroke was sponsored, in part, by Novartis, Boehringer Ingelheim Lilly, Novo Nordisk, Sanofi, AstraZeneca and Bayer. The authors declare that this study received funding from Boehringer-Ingelheim and Merck. The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fneur.2020.00961/full#supplementary-material>

REFERENCES

- McAlister FA, Laupacis A, Wells GA, Sackett DL. Users' guides to the medical literature: XIX. Applying clinical trial results B. Guidelines for determining whether a drug is exerting (more than) a class effect. *JAMA*. (1999) 282:1371–7. doi: 10.1001/jama.282.14.1371
- The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med*. (1995) 333:1581–7. doi: 10.1056/NEJM199512143332401
- Clark WM, Wissman S, Albers GW, Jhamandas JH, Madden KP, Hamilton S. Recombinant tissue-type plasminogen activator (Alteplase) for ischemic stroke 3 to 5 hours after symptom onset. The ATLANTIS study: a randomized controlled trial. Alteplase thrombolysis for acute noninterventional therapy in ischemic stroke. *JAMA*. (1999) 282:2019–26. doi: 10.1001/jama.282.21.2019
- Adeoye O, Sucharew H, Khoury J, Vagal A, Schmit PA, Ewing I, et al. Combined approach to lysis utilizing eptifibatid and recombinant tissue-type plasminogen activator in acute ischemic stroke-full dose regimen stroke trial. *Stroke*. (2015) 46:2529–33. doi: 10.1161/STROKEAHA.115.010260
- Pancioli AM, Broderick J, Brott T, Tomsick T, Khoury J, Bean J, et al. The combined approach to lysis utilizing eptifibatid and rt-PA in acute ischemic stroke: the CLEAR stroke trial. *Stroke*. (2008) 39:3268–76. doi: 10.1161/STROKEAHA.108.517656
- Hacke W, Kaste M, Bluhmki E, Brozman M, Dávalos A, Guidetti D, et al. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. *N Engl J Med*. (2008) 359:1317–29. doi: 10.1056/NEJMoa0804656
- Katzan IL, Furlan AJ, Lloyd LE, Frank JI, Harper DL, Hinchey JA, et al. Use of tissue-type plasminogen activator for acute ischemic stroke: the Cleveland area experience. *JAMA*. (2000) 283:1151–8. doi: 10.1001/jama.283.9.1151
- Mayer SA, Brun NC, Begtrup K, Broderick J, Davis S, Diringer MN, et al. Efficacy and safety of recombinant activated factor VII for acute intracerebral hemorrhage. *N Engl J Med*. (2008) 358:2127–37. doi: 10.1056/NEJMoa0707534
- Broderick JB, Diringer MN, Hill MD, Brun NC, Mayer SA, Steiner T, et al. Determinants of intracerebral hemorrhage growth: an exploratory analysis. *Stroke*. (2007) 38:1072–5. doi: 10.1161/01.STR.0000258078.35316.30
- van Leeuwen N, Lingsma HF, de Craen AJ, Nieboer D, Mooijaart SP, Richard E, et al. Regression discontinuity design: simulation and application in two cardiovascular trials with continuous outcomes. *Epidemiology*. (2016) 27:503–11. doi: 10.1097/EDE.0000000000000486
- Bor J, Moscoe E, Bärnighausen T. Three approaches to causal inference in regression discontinuity designs. *Epidemiology*. (2015) 26:e28–30. doi: 10.1097/EDE.0000000000000256
- Moscoe E, Bor J, Bärnighausen T. Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: a review of current and best practice. *J Clin Epidemiol*. (2015) 68:122–33. doi: 10.1016/j.jclinepi.2014.06.021
- Bor J, Moscoe E, Mutevedzi P, Newell ML, Bärnighausen T. Regression discontinuity designs in epidemiology: causal inference without randomized trials. *Epidemiology*. (2014) 25:729–37. doi: 10.1097/EDE.0000000000000138
- Fonarow GC, Zhao X, Smith EE, Saver JL, Reeves MJ, Bhatt DL, et al. Door-to-needle times for tissue plasminogen activator administration and clinical outcomes in acute ischemic stroke before and after a quality improvement initiative. *JAMA*. (2014) 311:1632–40. doi: 10.1001/jama.2014.3203
- Ellrodt AG, Fonarow GC, Schwamm LH, Albert N, Bhatt DL, Cannon CP, et al. Synthesizing lessons learned from get with the guidelines: the value of disease-based registries in improving quality and outcomes. *Circulation*. (2013) 128:2447–60. doi: 10.1161/01.cir.0000435779.48007.5c
- Reeves MJ, Smith E, Fonarow G, Hernandez A, Pan W, Schwamm LH, et al. Off-hour admission and in-hospital stroke case fatality in the get with the guidelines-stroke program. *Stroke*. (2009) 40:569–76. doi: 10.1161/STROKEAHA.108.519355
- Müller T, Shaikh M. Your retirement and my health behavior: evidence on retirement externalities from a fuzzy regression discontinuity design. *J Health Econ*. (2017) 57:45–59. doi: 10.1016/j.jhealeco.2017.10.005
- Miller G, Pinto D, Vera-Hernández M. Risk protection, service use, and health outcomes under Colombia's health insurance program for the poor. *Am Econ J Appl Econ*. (2013) 5:61–91. doi: 10.1257/app.5.4.61
- Naidech AM, Beaumont JL, Berman M, Francis B, Liotta E, Maas MB, et al. Dichotomous “good outcome” indicates mobility more than cognitive or social quality of life. *Crit Care Med*. (2015) 43:1654–9. doi: 10.1097/CCM.0000000000001082
- Katzan IL, Fan Y, Uchino K, Griffith SD. The PROMIS physical function scale: a promising scale for use in patients with ischemic stroke. *Neurology*. (2016) 86:1801–7. doi: 10.1212/WNL.0000000000002652
- Baumhauer JF. Patient-reported outcomes - are they living up to their potential? *N Engl J Med*. (2017) 377:6–9. doi: 10.1056/NEJMp1702978
- Marler JR, Tilley BC, Lu M, Brott TG, Lyden PC, Grotta JC, et al. Early stroke treatment associated with better outcome: the NINDS rt-PA stroke study. *Neurology*. (2000) 55:1649–55. doi: 10.1212/WNL.55.11.1649
- Marinescu IE, Triantafyllou S, Kording K. *Regression Discontinuity Threshold Optimization*. (2019). Available online at SSRN: <https://ssrn.com/abstract=3333334>

Conflict of Interest: GF: research support from PCORI and NIH, and an employee of University of California which holds a patent on an endovascular device for stroke. AN: research support from NIH and AHRQ. SP: research support from PCORI and NIH. KK: research support from NIH. LS: reports being the principal investigator of an investigator-initiated study of extended-window intravenous thrombolysis funded by the National Institutes of Neurological Disorders and Stroke (clinicaltrials.gov/show/NCT01282242) for which Genentech provided alteplase free of charge to Massachusetts General Hospital as well as supplemental per-patient payments to participating sites; serving as chair of the AHA/ASA GWTG stroke clinical work group and Quality Oversight Committees, serving as a stroke systems consultant to the Massachusetts Department of Public Health; and serving as a scientific consultant regarding trial design and conduct to Penumbra (data and safety monitoring committee, Separator 3D and MIND trials) and Medtronic (Victory AF and Stroke AF trials).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Naidech, Lawlor, Xu, Fonarow, Xian, Smith, Schwamm, Matsouka, Prabhakaran, Marinescu and Kording. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.