Published in final edited form as:

Ophthalmic Epidemiol. 2018 August ; 25(4): 315–322. doi:10.1080/09286586.2018.1469156.

Regression discontinuity and randomized controlled trial estimates: An application to the Mycotic Ulcer Treatment Trials

Catherine E. Oldenburg¹, N. Venkatesh Prajna², Tiruvengada Krishnan³, Revathi Rajaraman⁴, Muthiah Srinivasan², Kathryn J Ray¹, Kieran S. O'Brien¹, M. Maria Glymour⁵, Travis C. Porco^{1,5,6}, Nisha R Acharya^{1,5,6}, Jennifer Rose-Nussbaumer^{1,6}, and Thomas M. Lietman^{1,5,6}

¹Francis I. Proctor Foundation, University of California, San Francisco

²Aravind Eye Care System, Madurai, India

³Aravind Eye Care System, Pondicherry, India

⁴Aravind Eye Care System, Coimbatore, India

⁵Department of Epidemiology and Biostatistics, University of California, San Francisco

⁶Department of Ophthalmology, University of California, San Francisco

Abstract

Background—We compare results from regression discontinuity (RD) analysis to primary results of a randomized controlled trial (RCT) utilizing data from two contemporaneous RCTs for treatment of fungal corneal ulcers.

Methods—Patients were enrolled in the Mycotic Ulcer Treatment Trials I and II (MUTT I & MUTT II) based on baseline visual acuity: patients with acuity $\leq 20/400$ (logMAR 1.3) enrolled in MUTT I, and $>20/400$ in MUTT II. MUTT I investigated the effect of topical natamycin versus voriconazole on best spectacle-corrected visual acuity (BSCVA). MUTT II investigated the effect of topical voriconazole plus placebo versus topical voriconazole plus oral voriconazole. We compared the RD estimate (natamycin arm of MUTT I [N=162] versus placebo arm of MUTT II [N=54]) to the RCT estimate from MUTT I (topical natamycin [N=162] versus topical voriconazole [N=161]).

Results—In the RD, patients receiving natamycin had mean improvement of 4-lines of visual acuity at 3 months (logMAR -0.39 , 95% CI: -0.61 , -0.17) compared to topical voriconazole plus placebo, and 2-lines in the RCT (logMAR -0.18 , 95% CI: -0.30 , -0.05) compared to topical voriconazole.

Conclusions—The RD and RCT estimates were similar, although the RD design overestimated effects compared to the RCT.

Corresponding Author: Catherine E. Oldenburg, ScD, MPH, Francis I. Proctor Foundation, University of California, San Francisco, 513 Parnassus Ave, San Francisco, CA, catherine.oldenburg@ucsf.edu, Phone: 415-502-8843.

Conflicts of Interest. None to report.

Meeting Presentation. Results presented in part at the 38th Annual Conference of the Society for Clinical Trials, Liverpool, United Kingdom, May 9, 2017.

Keywords

regression discontinuity; fungal keratitis; randomized controlled trial; causal inference

INTRODUCTION

The regression discontinuity (RD) design has recently received greater attention in the epidemiologic literature for the estimation of causal effects in observational data. RD can be used in scenarios where a continuously measured variable is used to determine treatment eligibility.^{1–5} For example, clinical decision rules are based on continuous biomarkers such as prostate specific antigen, which is used to establish eligibility for further workup for prostate cancer⁶, or CD4 count to determine eligibility for antiretroviral therapy among individuals living with HIV.^{3,7} In such scenarios, interventions are assigned based on a threshold rule: whether a patient falls above or below the threshold determines their treatment eligibility.

Regression discontinuity designs have been described as the observational epidemiologic method that most closely resembles a randomized controlled trial,⁸ as randomized treatment assignment can be conceptualized as a threshold assignment process.² However, direct comparisons between RD designs and randomized controlled trials (RCTs) are difficult, primarily because the RD design estimates a local effect at the threshold, and the intention-to-treat (ITT) effect from the RCT is a global effect.^{8,9} This makes true validation of the RD design against the gold standard RCT challenging. However, useful insights into the differences between each design and the estimates they yield can be obtained in scenarios where both RD and RCT analyses are possible.

Here, we describe the regression discontinuity method and its application in ophthalmology, with an illustrative example from the Mycotic Ulcer Treatment Trials (MUTT I and MUTT II).^{10,11} The MUTT trials were randomized controlled trials designed to evaluate treatment strategies for fungal keratitis (or fungal corneal ulcers), a potentially blinding infection of the cornea. MUTT I and MUTT II were run concurrently, with MUTT I enrolling patients with better baseline visual acuity to compare topical natamycin versus topical voriconazole, and MUTT II enrolling patients with worse visual acuity to compare topical voriconazole plus oral voriconazole versus topical voriconazole plus placebo. By utilizing data from two concurrent RCTs, we are able to further compare the estimates derived from the RD compared to the original RCT result for the comparison of voriconazole versus natamycin for the treatment of fungal corneal ulcers.

MATERIALS AND METHODS**Regression Discontinuity Methods**

RD designs take advantage of scenarios in which treatment is assigned by a threshold rule. In clinical applications, this often refers to a measured continuous variable (the “assignment variable”) that determines treatment status based on a certain cutoff (the “threshold”). Previous clinical applications of the RD design have utilized threshold rules such as CD4

count determining antiretroviral therapy initiation for people living with HIV,¹² prostate-specific antigen for determining more extensive workup for prostate cancer,⁶ and age cohort for determining HPV vaccination.¹³ Examples in ophthalmology could include visual acuity, intraocular pressure and glaucoma treatment initiation, or corneal thickness and indication for LASIK. Under such threshold rules, patients who present immediately above and below the threshold are expected to be exchangeable with respect to baseline characteristics, similar to the exchangeability that is expected to arise in a randomized controlled trial. When the probability of receiving treatment given that an individual is below the threshold does not equal the probability of receiving treatment given that an individual is above the threshold, a discontinuity in treatment assignment exists. This discontinuity can be utilized to estimate causal effects in a narrow window around the threshold.

“Sharp regression discontinuity” refers to situations in which the threshold assigns treatment perfectly. In this case, the difference in the mean outcome can be calculated above and below the threshold, and the effect estimate is equal to the average causal effect in the population around the threshold. Alternatively, regression discontinuity designs can be used when the threshold rule imperfectly assigns treatment, where not all patients receive the treatment they were assigned to based on the rule or adherence to a treatment regimen is not 100%. In this case, the difference in means is equivalent to the intention-to-treat effect in a randomized controlled trial, which is the effect of randomization arm on outcomes, for the population close to the threshold.²

Assumptions for Causal Inference with Regression Discontinuity

There are two primary identifying assumptions for causal effect estimation with RD.¹⁴ The central assumption of the RD design is that patients immediately above and immediately below the threshold are exchangeable, or more flexibly that any non-exchangeability can be fully accounted for by modeling the association between the assignment variable and the outcome. Under this exchangeability assumption, causal effects can be estimated as the difference in outcomes above and below the threshold. For all RD scenarios, causal effects are only valid if the exchangeability, or continuity, assumption is not violated.^{13,15} Although the assumption of exchangeability cannot be empirically verified, there are several tests that can help assess the plausibility of the assumption. Covariate balance tests can be performed to demonstrate balance of baseline covariates above and below the threshold, baseline characteristics can be plotted to ensure there is no discontinuity in baseline characteristics at the threshold, and baseline characteristics can be included in the regression discontinuity model as sensitivity analyses to ensure that the model is robust to the inclusion of baseline characteristics.

Secondly, the assignment variable must be continuous, and the threshold value must be known. When the threshold rule is well characterized, the consistency assumption that is required for causal inference in any observational study will be met.¹⁶ This assumption can be assessed by verifying that the variable has been measured and reported continuously and via knowledge that threshold is known and that outcomes were measured for subjects independently of the threshold. Continuity of the assignment variable can be assessed by plotting a histogram of baseline values to see if there is bunching at the threshold. Bunching

at the threshold would indicate that there may have been manipulation of values of the assignment variable that could jeopardize causal inference.

Generalizability and Interpretation of Effects

Without additional assumptions, the effect estimated from the RD model is only directly generalizable to individuals who present in a narrow window around the threshold (the “bandwidth”). This holds true even if information is utilized from participants who present further from the threshold. Therefore, the estimated effect is interpreted as the mean difference in outcome for patients above versus participants below the threshold (or vice versa), among participants who have baseline values that are close to the threshold. To interpret the estimate as an average treatment effect for the entire population, the functional form of the potential outcomes must be known and they must be constant across the entire population. However, this is a strong and untestable assumption, as the local effect estimated by the model must be extrapolated to unobserved regions of the assignment variable.⁹ If there is effect modification of the effect size by the assignment variable, the assumption will not hold. Interpretation of any effect in regression discontinuity must carefully consider generalizability and the degree to which the effect is generalizable across values of the assignment variable.

In practice, regression discontinuity is subject to a bias-variance tradeoff. The assumptions for causal inference with regression discontinuity are most likely to hold for individuals who are closest to the threshold value. However, in most applications there are very few observations in narrow ranges of the bandwidth, which limits statistical power. Information can be borrowed from individuals in a wider bandwidth to improve power, however as the bandwidth widens correctly modeling the functional form of the assignment variable becomes increasingly important to maintain exchangeability.

Application to the Mycotic Ulcer Treatment Trials

The Mycotic Ulcer Treatment Trials were two double-masked randomized controlled trials designed to evaluate treatment strategies for fungal keratitis. Methods and primary results from each trial have previously been reported.^{10,11} In brief, MUTT I compared topical natamycin (TN_{MUTTI}) to topical voriconazole (TV_{MUTTI}) among patients with a baseline visual acuity of 20/400 and better. MUTT II enrolled patients with baseline visual acuity of worse than 20/400, and compared oral voriconazole versus placebo TVP_{MUTTII}) in conjunction with topical voriconazole. Of note, visual acuity of 20/20 is considered perfect vision and acuity worse than 20/200 in the better eye is considered legal blindness in the United States. Following the release of the results from MUTT I, topical natamycin was added to the treatment regimen for all patients in MUTT II. Patients in MUTT I were enrolled in three clinical sites in India (Aravind Eye Hospitals, Madurai, Pondicherry, and Coimbatore). MUTT II enrolled patients in the three clinical sites in India as well as two additional sites in Nepal. The present analysis was restricted to patients enrolled in India prior to the addition of topical natamycin in MUTT II to treatment regimens. Institutional review board approval was obtained from the University of California, San Francisco and the Aravind Eye Care System. Written informed consent was obtained from all participants.

With the exception of the baseline visual acuity requirement, enrollment criteria were similar between the two studies. Due to the use of oral voriconazole, patients with known liver disease or who weighed less than 40 kg were excluded in MUTT II. Patients were followed on an identical schedule between the two studies, and topical antifungal agents were applied according to an identical protocol (every hour while the patient was awake for the first week, then every 2 hours while awake until 3 weeks from enrollment). All patients in both studies were examined at baseline, every 3 days until re-epithelialization, and at 3 weeks and 3 months from enrollment. A similar proportion of patients were excluded from the studies because they did not meet the enrollment criteria at screening (56.3% in MUTT I, 56.4% in MUTT II), and a similar proportion were lost to follow-up by three months (13.0% in MUTT I, 13.8% in MUTT II).

Corneal scrapings for smear and culture were obtained after visual acuity assessment and slit lamp examination to determine causative organisms. Fungal smears were considered to be positive when fungal elements were seen under low-power magnification and reduced light, and fungal cultures were considered to be positive with growth on any two media or moderate to heavy growth on one media.

The primary outcome for MUTT I was best spectacle-corrected visual acuity (BSCVA) at three months from enrollment, which was a secondary outcome for MUTT II. The primary outcome for MUTT II was corneal perforation and/or the need for a penetrating keratoplasty (TPK; modeled as a composite outcome), a secondary outcome for MUTT I. These outcomes were measured according to an identical protocol. The primary analysis for each trial considered corneal ulcers from all organisms, and a pre-specified subgroup analysis considered effects in ulcers due to *Fusarium* and non-*Fusarium* species. Visual acuity was analyzed as the logarithm of the minimal angle of resolution (logMAR). On the logMAR scale, in this study visual acuity ranges from -0.30 to 2.0, and higher values indicate worse visual acuity. LogMAR visual acuity can be converted to the Snellen visual acuity chart. A 0.10 difference in logMAR is equivalent to one line of visual acuity on a standard Snellen chart (e.g., the difference in 20/20 and 20/25 is approximately 0.10 logMAR).

Regression Discontinuity Methods for the MUTT Trials

In the present application, we utilized the continuous enrollment criterion of baseline visual acuity for enrollment in either MUTT I or MUTT II as a threshold-based assignment rule. We utilized this rule to replicate MUTT I using a regression discontinuity design, using the natamycin arm from the MUTT I trial ($TN_{\text{MUTT I}}$) and the topical voriconazole plus placebo arm from MUTT II ($TVP_{\text{MUTT II}}$). This scenario represents a sharp or deterministic regression discontinuity design, where all individuals at or below the 20/400 cutoff received natamycin, and all those above 20/400 received voriconazole plus oral placebo. Although examining a narrow bandwidth around the cutoff is typical in RD studies, due to the limited sample size in the current application, we chose a bandwidth of the entire range of data for all RD models.

Sample code for the analyses described below is included in the Supplemental Appendix. For the outcome of three-month BSCVA, a linear regression model was fit with terms for being above or below the threshold (20/400; logMAR 1.3), and a term for baseline visual

acuity allowing for a differential slope above and below the threshold. Perforation/TPK was modeled with logistic regression including a term for above or below the threshold and baseline visual acuity, similar to the BSCVA model. Sensitivity analyses including a squared term for baseline visual acuity above and below the threshold were run to assess potential bias in specification of the functional form of BSCVA.

We calculated the effect in MUTT I of TN_{MUTTI} versus TV_{MUTTI} (the primary RCT contrast) at the threshold by creating a continuous term for baseline visual acuity centered at the threshold (20/400) in MUTT I. We then fit a linear regression model predicting 3-month BSCVA in MUTT I with terms for study arm (TN_{MUTTI} versus TV_{MUTTI}), centered baseline visual acuity, and the interaction between centered baseline acuity and study arm. The main effect of study arm was interpreted as the RCT effect at the 20/400 threshold.

To test whether effects in the RD model could be explained by differences between the two studies, an additional analysis was run comparing patients in the voriconazole arm of MUTT I (TV_{MUTTI}) to the placebo arm in MUTT II (TVP_{MUTTI}) (e.g., an RD analysis comparing topical voriconazole versus topical voriconazole plus placebo). Effect modification in the RCT was assessed by running the original model for natamycin versus voriconazole in MUTT I with an interaction term for baseline visual acuity by treatment arm. All analyses were run in Stata 14.1 (StataCorp, College Station, TX).

RESULTS

RD analyses included 216 patients in the TN_{MUTTI} ($N=162$) or the TVP_{MUTTI} prior to the addition of natamycin ($N=54$). Descriptive characteristics for these two groups are listed in Table 1. Demographic characteristics were similar between the two groups, with a median age of 48 years (interquartile range [IQR] 39 to 58) among individuals with baseline acuity <20/400 and 51.5 years (IQR 45 to 65) among individuals with baseline acuity ≥20/400. As expected, ulcers with baseline acuity <20/400 had better clinical characteristics, including smaller infiltrate/scar sizes and shorter duration of symptoms, compared to those with worse visual acuity. There was no evidence of bunching at the threshold with respect to the distribution of baseline BSCVA (Figure 2).

In the overall RD sample, there was an approximately 4-line improvement in BSCVA among individuals in the TN_{MUTTI} arm compared to TVP_{MUTTI} arm at the 20/400 threshold (logMAR −0.39, 95% CI −0.61 to −0.17; Table 2; Figure 3). This was similar to the MUTT I RCT trial results, which found an approximately 2-line improvement in BSCVA among individuals randomized to natamycin (logMAR −0.18, 95% CI −0.30 to −0.05). Results were consistent with the estimated effect at the threshold in the RCT (−0.18, 95% CI −0.41 to 0.05). In the RD, individuals in the TN_{MUTTI} arm had 0.31 (95% CI 0.12 to 0.77; Table 2; Figure 4) times the odds of perforation/TPK compared to the TVP_{MUTTI} arm. In the RCT, individuals randomized to TN_{MUTTI} had 0.42 (95% CI 0.22 to 0.80) times the odds of perforation/TPK compared to TV_{MUTTI} . The effect at the threshold in the RCT was slightly weaker (OR 0.63, 95% CI 0.25 to 1.57).

In *Fusarium* ulcers, the effect of natamycin was larger in both the RD and the RCT (Table 3). In the RD, individuals with *Fusarium* ulcers had more than 9 lines (logMAR 0.96, 95% CI -1.30 to -0.62) of acuity improvement in TN_{MUTTI} compared to TVP_{MUTTI} . In the RCT, individuals with *Fusarium* ulcers randomized to TN_{MUTTI} had approximately 4 lines (logMAR -0.41, 95% CI -0.30 to -0.05) of improvement in BSCVA compared to TV_{MUTTI} , an effect that was similar at the threshold (logMAR -0.37, 95% CI -0.80 to 0.05). *Fusarium* ulcers assigned to TN_{MUTTI} in the RD (OR 0.04, 95% CI 0.002 to 0.62), RCT (OR 0.06, 95% CI 0.01 to 0.28), and RCT at the threshold (OR 0.12, 95% CI 0.02 to 0.76) also had reduced odds of perforation/TPK compared of TVP_{MUTTI} . There was no difference between TN_{MUTTI} and TVP_{MUTTI} in either BSCVA or perforation/TPK in non-*Fusarium* ulcers.

There was no statistically significant difference between the TV_{MUTTI} and TVP_{MUTTI} (Supplemental Table 1). There was no evidence of effect modification of the effect of TN_{MUTTI} versus TV_{MUTTI} by baseline visual acuity in the RCT (Supplemental Table 2), suggesting that results were not necessarily different at different values of visual acuity.

DISCUSSION

Here, we demonstrate that the quasi-experimental regression discontinuity approach yielded similar estimates to those derived from an RCT investigating an identical comparison, although the RD estimates were consistently larger than the RCT estimates, including the estimated RCT effect at the 20/400 threshold. Given the additional assumptions required for generalizability of RD estimates, the RCT outcomes are best suited for guiding clinical practice if available, although RD estimates can provide important additional information such as the effect of cutoff values for health policy. In the current application, it is possible that fungal ulcers with worse baseline visual acuity have a larger response to treatment with natamycin than do those with better baseline visual acuity, which could partially explain this discrepancy. Previous work from MUTT I demonstrated that *Fusarium* isolates were the least susceptible to voriconazole, and thus may require a higher concentration of voriconazole to effectively treat.¹⁷ It is possible that this is particularly true for patients with worse visual acuity at baseline. There was no statistically significant evidence of effect modification for either visual acuity or perforation/TPK. The slopes for visual acuity outcomes above and below the threshold were similar, however the slope for probability of perforation was considerably steeper, suggesting that there may be differential effects at worse values of baseline visual acuity. Analyses of effect modification could be underpowered, and thus generalizability of the RD effect should not be done for baseline visual acuity values far from the threshold.

Although there was no statistically significant difference in the TV_{MUTTI} and TVP_{MUTTI} , mean visual acuity was approximately 2 lines better in TV_{MUTTI} . If participants were randomized within the same study and placebo had no effect, we would not expect a difference between these two groups. This two-line difference could arise for several reasons: differences in the two studies resulting in bias, random chance, or an effect of the placebo itself. To obtain a “bias-minimized” estimate for our contrast of interest (TN_{MUTTI} versus TN_{MUTTI}), we subtracted the two-line estimate from our calculated estimate of

approximately 4-lines difference, and obtained a difference of 2-lines (logMAR 0.18), equivalent to our RCT estimate. However, it is possible that these results are not due to bias or chance, but are due to the addition of the placebo itself in the TVP_{MUTTI} arm. It is possible that knowledge that there was a 50% chance of getting an additional active therapy changed patients' expectations and affected clinical outcomes. If outcomes are improved due to the use of the oral placebo itself, this would constitute a violation of the continuity assumption, because patients above and below the threshold would no longer be exchangeable. It is therefore possible that differences in the RD and RCT estimates are due in part to the use of oral placebo. However, given that visual acuity and perforation are clinical outcomes, it is unlikely that the use of placebo would explain the entire difference.¹⁸

When treatment effects vary with the assignment variable, the intention-to-treat RD and RCT methods estimate different effects. The RD estimate is a local effect. In this case, this means the estimate is only generalizable to individuals near the 20/400 threshold. Conversely, the RCT effect is a global effect, and represents an average in the entire study population (in this case, individuals with baseline visual acuity up to 20/400), though this average effect may differ from the effect at any given value of the assignment variable. These local and global effects are the same when the treatment effect is constant across values of the assignment variable.⁹ The plausibility of the RCT and RD results being comparable for the BSCVA outcome in this setting is supported by several pieces of information. First, we estimated the effect of the RCT at the threshold, and found results that were broadly similar to the global RCT effect. Consistent with this, there was no evidence of effect modification in the RCT of the effect of natamycin versus voriconazole by baseline visual acuity. Second, visual inspection of Figure 3 suggests that the slopes above and below the threshold are similar for visual acuity, although the slopes were different for the probability of perforation. Evidence from this study suggests that effects are not heterogeneous across the assignment variable for visual acuity, and therefore the RD and RCT approaches are estimating approximately the same quantity.

The results of this study must be considered in the context of several limitations. As with any RD analysis, there may be bias in effect estimates if the functional form of baseline visual acuity was misspecified. Sensitivity analyses including a squared term for baseline visual acuity did not substantially change estimates. Due to the relatively small sample size, we were not able to assess very narrow bandwidths around the threshold in this study, although assessing robustness to different bandwidths is generally recommended in RD designs.³ The small sample size also precluded subgroup analysis of specific non-*Fusarium* organisms. Furthermore, because the RD and RCT estimate different effects (a local effect at the threshold in the RD and a global average effect in the RCT), it is not possible to determine whether differences in effect sizes are due to differential effects by baseline acuity or bias in the RD due to differences in the trials, although we did not see evidence of effect modification by baseline acuity, particularly for visual acuity outcomes.

In summary, this study illustrated the use of RD in an ophthalmology application, and suggested broadly similar effects of natamycin versus voriconazole on clinical outcomes in fungal keratitis when outcomes were analyzed using both an RD and an RCT design, although the RD was an overestimate of the RCT estimate. Although the RD effect is

immediately generalizable only at 20/400 without additional assumptions whereas the RCT is a global effect, the consistency of results between the RD and RCT designs was illustrated through assessment of effects in *Fusarium* and non-*Fusarium* ulcers. The results of this study support the value of RD designs for the estimation of effects in non-randomized settings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial Support. The Mycotic Ulcer Treatment Trials were supported by the National Eye Institute (Grant U10EY018573 to TML and NRA). CEO was supported in part by the National Institute of Mental Health R25MH083620. JRN was supported by the National Eye Institute K23EY025025.

References

1. Oldenburg CE, Seage GR, Tanser F, et al. Antiretroviral therapy and mortality in rural South Africa: A comparison of causal modeling approaches. *American Journal of Epidemiology*. 2018 Forthcoming.
2. Oldenburg CE, Moscoe E, Barnighausen T. Regression Discontinuity for Causal Effect Estimation in Epidemiology. *Current Epidemiology Reports*. 2016; 3(3):1–9. DOI: 10.1007/s40471-016-0080-x
3. Bor J, Moscoe E, Mutevedzi P, Newell M-L, Barnighausen T. Regression Discontinuity Designs in Epidemiology. *Epidemiology*. 2014; 25(5):729–737. DOI: 10.1097/EDE.0000000000000138 [PubMed: 25061922]
4. Moscoe E, Bor J, Barnighausen T. Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: a review of current and best practice. *Journal of Clinical Epidemiology*. 2015; 68(2):132–143. DOI: 10.1016/j.jclinepi.2014.06.021
5. Vandenbroucke JP, le Cessie S. Regression discontinuity design: Let's give it a try to evaluate medical and public health interventions. *Epidemiology*. 2014; 25(5):738–741. DOI: 10.1097/EDE.0000000000000145 [PubMed: 25076150]
6. Shoag J, Halpern J, Eisner B, et al. Efficacy of prostate-specific antigen screening: Use of the regression discontinuity in the PLCO cancer screening trial. *JAMA Oncology*. 2015; 1(7):984–986. [PubMed: 26291583]
7. Bor J, Fox MP, Venkataramani A, Tanser F, Pillay D, Barnighausen T. Treatment Eligibility and Retention in Clinical HIV Care: Regression-Discontinuity Evidence From South Africa. 2017 Vol Forthcoming.
8. Labrecque JA, Kaufman JS. Commentary. *Epidemiology*. 2016; 27(4):500–502. DOI: 10.1097/EDE.0000000000000485 [PubMed: 27031041]
9. Bor J, Moscoe E, Barnighausen T. Three approaches to causal inference in regression discontinuity designs. *Epidemiology*. 2015; 26(2):e28.doi: 10.1097/EDE.0000000000000231
10. Prajna NV, Krishnan T, Mascarenhas J, et al. The Mycotic Ulcer Treatment Trial: A Randomized Trial Comparing Natamycin vs Voriconazole. *JAMA Ophthalmol*. 2013; 131(4):422.doi: 10.1001/jamaophthalmol.2013.1497 [PubMed: 23710492]
11. Prajna NV, Krishnan T, Rajaraman R, et al. The Mycotic Ulcer Treatment Trial II: a Randomized Trial Comparing Oral Voriconazole Versus Placebo. 2016:1–9. Vol In press.
12. Oldenburg CE, Bor J, Harling G, et al. Impact of early antiretroviral therapy eligibility on HIV acquisition. *AIDS*. 2018; 32:635–643. DOI: 10.1097/QAD.0000000000001737 [PubMed: 29334546]
13. Smith LM, Strumpf EC, Kaufman JS, Lofters A, Schwandt M, Levesque LE. The Early Benefits of Human Papillomavirus Vaccination on Cervical Dysplasia and Anogenital Warts. *Pediatrics*. 2015; 135(5):e1131–e1140. DOI: 10.1542/peds.2014-2961 [PubMed: 25917991]

14. Barnighausen T, Oldenburg C, Tugwell P, et al. Quasi-experimental study designs series – Paper 7: assessing the assumptions. *Journal of Clinical Epidemiology*. Mar.2017 :1–29. DOI: 10.1016/j.jclinepi.2017.02.017
15. Smith LM, Kaufman JS, Strumpf EC, Levesque LE. Effect of human papillomavirus (HPV) vaccination on clinical indicators of sexual behaviour among adolescent girls: the Ontario Grade 8 HPV Vaccine Cohort Study. *CMAJ*. 2015; 187(2):E74–E81. DOI: 10.1503/cmaj [PubMed: 25487660]
16. Hernan MA, Taubman SL. Does obesity shorten life? The importance of well-defined interventions to answer causal questions. *Int J Obes Relat Metab Disord*. 2008; 32:S8–S14. DOI: 10.1038/ijo.2008.82
17. Lalitha P, Sun CQ, Prajna NV, et al. In Vitro Susceptibility of Filamentous Fungal Isolates From a Corneal Ulcer Clinical Trial. *American Journal of Ophthalmology*. 2014; 157(2):318–326. DOI: 10.1016/j.ajo.2013.10.004 [PubMed: 24439440]
18. Kaptchuk TJ, Miller FG. Placebo Effects in Medicine. *N Engl J Med*. 2015; 373(1):8–9. DOI: 10.1056/NEJMp1506446 [PubMed: 26132938]

Figure 1.
Flow diagram of the Mycotic Ulcer Treatment Trials I & II for randomized controlled trial and regression discontinuity comparisons.

Figure 2.
Histogram of baseline best spectacle-corrected (BSCVA) in logMAR. The red dotted line indicates the threshold (20/40; logMAR 1.3).

Figure 3.
Average best spectacle-corrected visual acuity (BSCVA) at three months binned for every 0.10 logMAR (one line) of visual acuity. The solid red line indicates the threshold (20/400; logMAR 1.3).

Figure 4. Probability of perforation and/or requiring therapeutic penetrating keratoplasty binned for every 0.10 logMAR (one line) of visual acuity. The solid red line indicates the threshold (20/400; logMAR 1.3).

Table 1

Descriptive characteristics in the study sample

	Natamycin (<20/400) ¹	Voriconazole (<20/400) ²	Voriconazole (>20/400) ³	Difference in Natamycin vs Voriconazole at threshold ⁴
Age (median, IQR)	48 (39 to 58)	45 (38 to 55)	51.5 (45 to 65)	2.42 (−3.15 to 7.99)
Female sex	73 (45.1%)	67 (41.6%)	20 (37.0%)	2.31 (0.89 to 6.03)
Occupation				
Agriculture	80 (49.4%)	76 (47.2%)	26 (48.2%)	0.89 (0.36 to 2.19)
Non-agriculture	82 (50.6%)	85 (52.8%)	28 (51.9%)	
Duration of symptoms (median, IQR), days	5 (3 to 9)	5 (3 to 10)	8.5 (6 to 14)	−1.60 (−4.80 to 1.59)
Baseline BSCVA	0.66 (0.38 to 0.92)	0.64 (0.38 to 0.76)	1.70 (1.30 to 1.80)	n/a ⁵
Baseline infiltrate/scar size	3.09 (2.45 to 3.99)	3.19 (2.65 to 3.97)	5.41 (4.69 to 6.69)	−1.55 (−2.09 to −1.00)

¹TNMUTTI: topical natamycin in MUTT I;²TMUTTI: topical voriconazole in MUTT I;³TVPMTTII: topical voriconazole plus oral placebo in MUTT II;⁴Estimated in a regression discontinuity model comparing TNMUTTI and TVPMTTII with a term for each baseline characteristic, odds ratios for dichotomous variables and difference in means for continuous variables;⁵Assignment variable, not estimable in regression discontinuity model

Table 2
Primary results for regression discontinuity and randomized controlled trial (N=216)

	Regression Discontinuity	Randomized Controlled Trial, average ¹	Randomized Controlled Trial, at threshold ²
3-month BSCVA (mean logMAR difference, 95% CI)	-0.39 (-0.61 to -0.17)	-0.18 (-0.30 to -0.05)	-0.18 (-0.41 to 0.05)
Perforation/TPK (OR, 95% CI)	0.31 (0.12 to 0.77)	0.42 (0.22 to 0.80)	0.63 (0.25 to 1.57)

¹ Overall intention-to-treat estimate from RCT;

² Estimated as main effect in a model with an interaction term between baseline visual acuity centered at the threshold and treatment arm

Table 3Subgroup analyses for *Fusarium* (N=83) and non-*Fusarium* (N=133)

	Regression Discontinuity	Randomized Controlled Trial, average ¹	Randomized Controlled Trial, at threshold ²
3-month BSCVA			
<i>Fusarium</i>	-0.96 (-1.30 to -0.62)	-0.41 (-0.61 to -0.20)	-0.37 (-0.80 to 0.05)
Non- <i>Fusarium</i>	-0.19 (-0.47 to 0.10)	-0.02 (-0.17 to 0.13)	-0.05 (-0.33 to 0.24)
Perforation/TPK			
<i>Fusarium</i>	0.04 (0.002 to 0.62)	0.06 (0.01 to 0.28)	0.12 (0.02 to 0.76)
Non- <i>Fusarium</i>	0.62 (0.20 to 1.89)	1.08 (0.48 to 2.43)	1.58 (0.48 to 5.22)

¹ Overall intention-to-treat estimate from RCT;² Estimated as main effect in a model with an interaction term between baseline visual acuity centered at the threshold and treatment arm