Regression Discontinuity Evaluation of the Effect of Kidney Disease Classifications on Health Outcomes

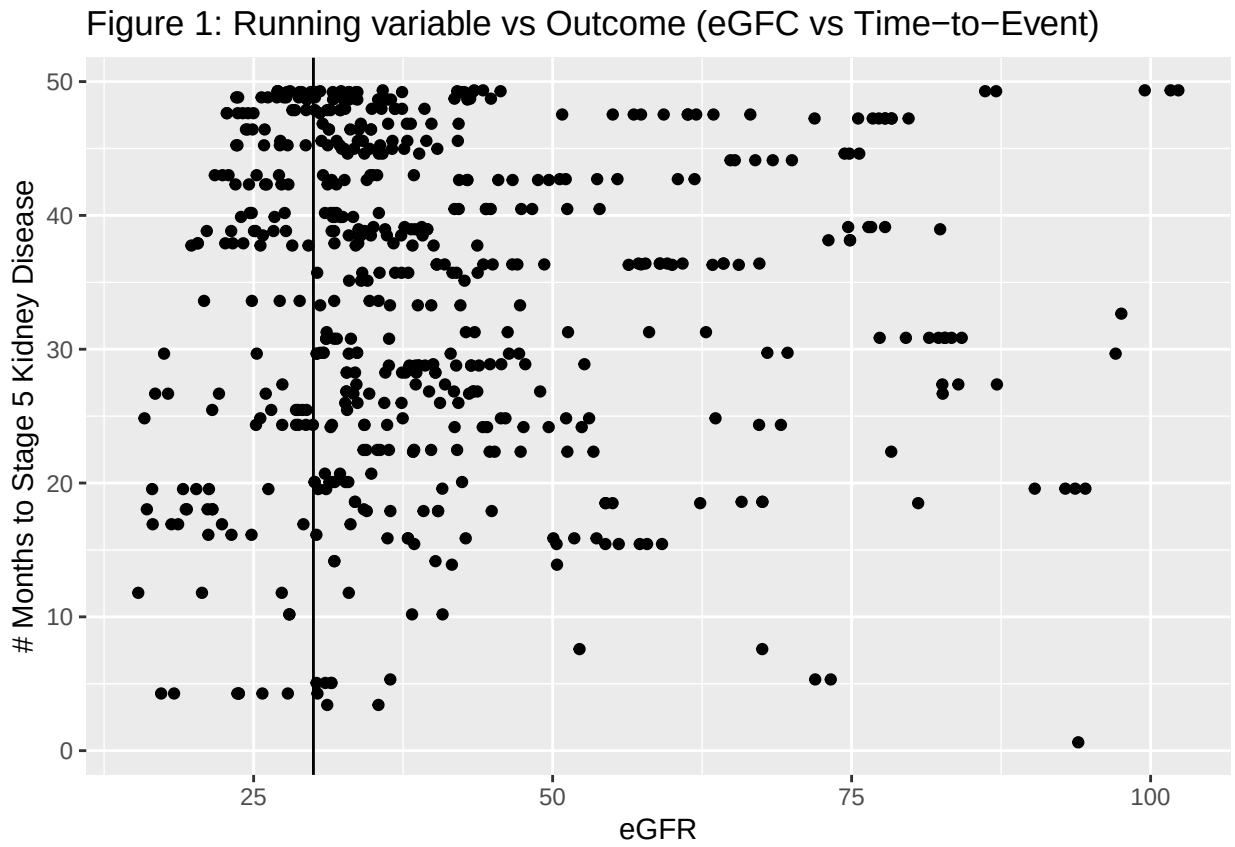
Robert Toto (RVT245)

5/12/2021

1 Introduction

The human kidney filters toxins, electrolytes, and proteins out of the bloodstream, constantly regulating the balance of the blood's molecular composition through filtering. If the kidneys fail to adequately balance the bloodstream through filtration, the patient is categorized into one of five stages of chronic kidney disease. Physicians use established eGFR (estimated glomerular filtration rate) cutoffs to sort patients into classification stages of kidney disease. The eGFR is measured in units of milliliters per minute per area of blood filtered by the kidneys (mL/min/1.73m2). At different stages, patients may jump to different treatments. For instance, at the higher stages (lower kidney function), patients begin to receive life-saving dialysis treatments, among other intensive forms of care. Using data on eGFR, I will run a Regression Discontinuity analysis to determine whether there exists a discontinuity in health outcomes on either side of the cutoff of 30 eGFR. An eGFR below 30 is considered Stage 4 kidney disease (the second most extreme stage on a scale of 1 to 5). When a patient falls below an eGFR of 30 and is diagnosed with Stage 4 kidney disease, her or she often begins to receive more intensive treatment. Patients who remain in Stage 3 kidney disease (eGFR above 30) are not moved to more intensive care as frequently. This Regression Discontinuity analysis will seek to determine whether patients on either side of this treatment cutoff (eGFR of 30) have discontinuous health outcomes as a result of differences in care. The outcome being measured is "time to event," which is the amount of months between when the eGFR measurement occurs and when the patient reaches a diagnosis of *Stage 5* (or end stage) kidney disease. The analysis therefore only applies to individuals in the data who ultimately reach Stage 5 kidney disease. The main concern of the paper is therefore whether being just above the Stage 4 eGFR cutoff of 30 (and therefore being less likely to receive more advanced care) causes patients to reach Stage 5 disease *sooner* than the patients just below the cutoff. The null hypothesis is that there is no difference in the amount of months to Stage 5 kidney disease between patients just below and just above the eGFR Stage 4 cutoff.

Figure 1 below shows the time-to-event (TTE)–or number of months to Stage 5 kidney disease–on the y-axis and eGFR on the x-axis. A low number of months implies a poor health outcomes becuase the patient entered end stage disease more quickly, and the vertical line at an eGFR of 30 shows the cutoff for a Stage 4 kidney disease diagnosis. Visually, there does seem to be poor health outcomes just to the right of the cutoff that could be prevented if the cutoff was slightly raised to capture marginally more people in the intensive treatment. However, there are also still very poor health outcomes just below the cutoff, so it is unclear from this visualization whether a discontinuity exists.



2 Data

The data is from a longitudinal study on kidney function across 505 adult patients with eGFR between 15 and 60 ml/min/1.73m2. For each patient, sex, age, UACR (urine albumin creatinine ration), and eGFR (estimated glomerular filtration rate) are recorded at a particular point in time. Each patient is subsequently tracked over time. The outcome of interest is the number of months between the initial recording of kidney function data and the point at which the patient reaches Stage 5 kidney disease. Patients never reaching this stage are excluded. The dataset is available from the Data Archiving and Networked Services (DANS) repository.

3 Density Tests

3.1 Histogram

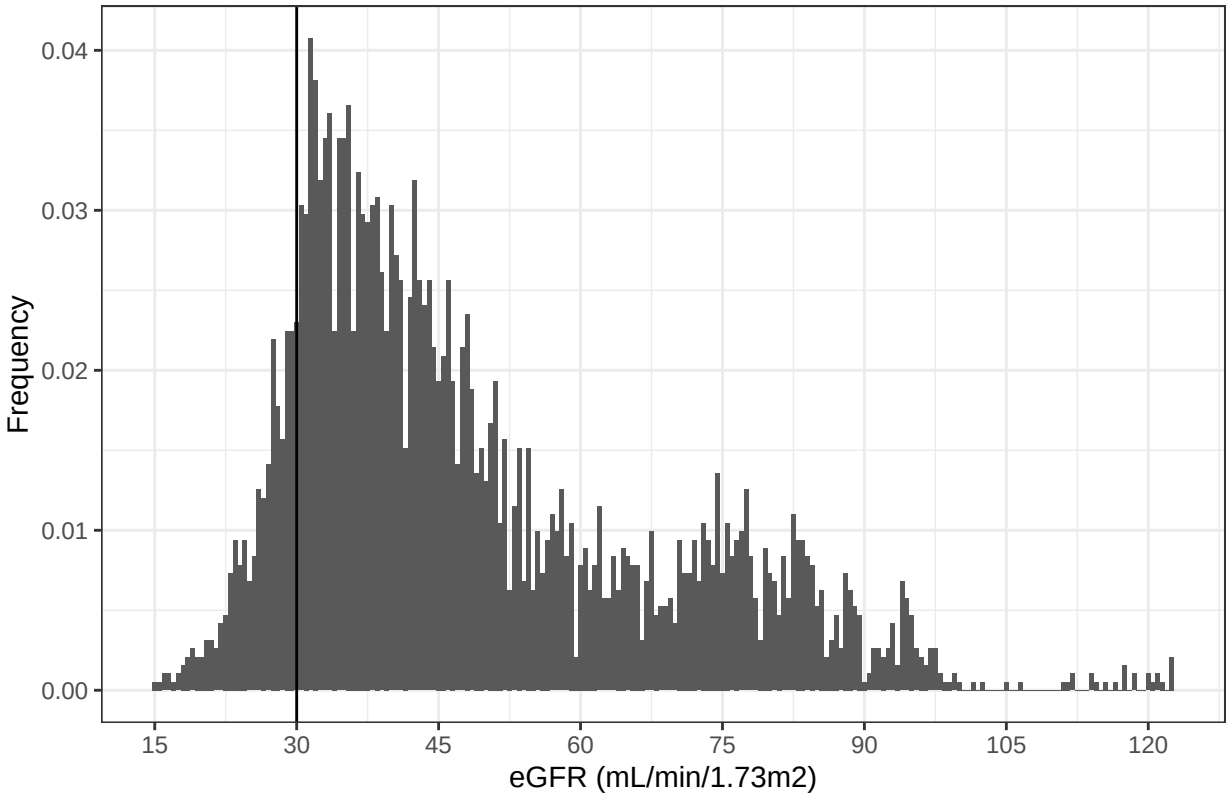
The histogram of eGFR measurements in Figure 2 shows no clear sign of bunching at the cutoff. The eGFR scores move smoothly across the cutoff, as they rise towards peak frequency just above 30. This is indirect evidence that individuals are not sorting themselves around the cutoff. This is expected since individuals cannot easily manipulate their kidneys' filtration capacities. Furthermore, even if it was physiologically possible to do so in a controlled manner, it could be be mortally dangerous for a person with kidney disease to try to lower their kidney function deliberately and therefore unlikely.

While patients are unable to manipulate their eGFR scores, it may be that medical staff (i.e. physicians and nurses) performing the eGFR measurement may round eGFR scores up or down near the cutoff. However this

is unlikely because the eGFR cutoff is not the sole determinant of whether a patient receives extra intensive treatment. A patient with an eGFR of 31 or 32 may still be given more intensive treatment (dialysis) by the physician if other health factors related to kidney failure (i.e. diabetes) indicate that more intensive treatment is practical even though the cutoff has not been surpassed. Therefore it is possible but also still unlikely that physicians may manipulate the cutoff.

Given that the eGFR is not the sole sorting variable for the treatment, but still a very strong one, this is not quite a sharp RDD. Since there is no data on whether a patient was classified as Stage 4 or not, we will assume that an eGFR below 30 has caused the patient to be classified as such (moved to the treatment group) while remembering that this classifier is partly fuzzy, given physician discretion to override the eGFR sorting mechanism. Overall, most individuals below the cutoff are Stage 4, and most above are only Stage 3, and physicians use the eGFR cutoff as a global standard, so it is still an extremely strong running variable, with the cutoff playing a major role in diagnosis and extra care.

Figure 2: Histogram of Running Variable (eGFR)



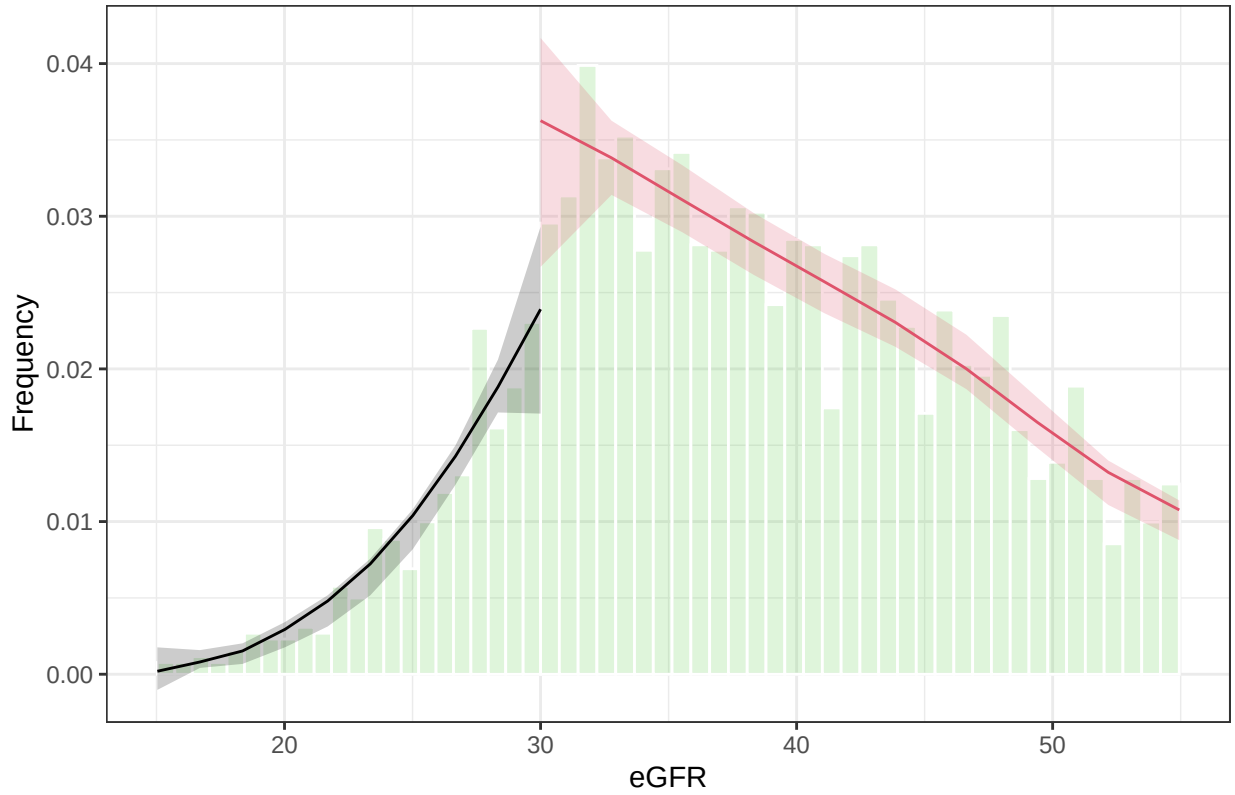
3.2 McCrary Density Test

The McCrary Density test in Figure 3 and related output below show that there is a greater density of individuals just above the cutoff, compared to just below, implying a potential violation of the smoothness assumption due to manipulation (sorting above the cutoff). This indicates that there may actually be bunching on the right (higher) side of the eGFR cutoff. This would indicate that medical staff may be rounding eGFR measurements *upward* to the less extreme kidney disease stage (Stage 3 above 30). This is highly dubious, however, because physicians and nurses are more likely to take an excess of caution with such low kidney function. If medical staff are sorting on eGFR, they are likely to round *downward*, not up out of caution so that patients who may need more intensive care will receive it, and not the inverse as the density test seems to show. In other words, nurses may see an eGFR of 30.5 and report is as 30 so the patient qualifies as Stage 4 and could receive advanced care more immediately. A nurse or physician seeing

an eGFR of 29.5 would not round the patient up into the less extreme Stage 3 classification, as doing so would be medically unethical and potentially dangerous.

Given the small sample size of this dataset (505 patients), there may not be sufficient size to show high-resolution smoothness where it may actually otherwise occur. The McCrary test may be picking up noise in this small sample set, and a larger sample is needed to better understand possible manipulation. Furthermore, the fact that, visually, the frequency of eGFR scores spikes just above the cutoff in a possibly-natural fashion may imply that the test is picking up normal bunching near the peak of the distribution. The p-value of the McCrary Density test is 0.03, which is significant at a typical 5% cutoff. Until a larger sample is available to run this analysis (nearly all are under lock-and-key of national research organizations), the smoothness assumption is not fully satisfied. The intuitively-backwards results of the McCrary bunching test in this small sample sheds further uncertainty on what it truly going on with the running variable at the cutoff.

Figure 3: McCrary Density Plot




```
##
## Manipulation testing using local polynomial density estimation.
##
## Number of obs =       3828
## Model =               unrestricted
## Kernel =              triangular
## BW method =           estimated
## VCE method =          jackknife
##
## c = 30                Left of c           Right of c
## Number of obs         442                 3386
## Eff. Number of obs    371                 1015
## Order est. (p)        2                   2
## Order bias (q)        3                   3
```

```
## BW est. (h)           7.001               8.315
##
## Method                T                   P > |T|
## Robust                2.1455              0.0319
##
##
## P-values of binomial tests (H0: p=0.5).
##
## Window Length / 2          <c     >=c    P>|T|
## 0.128                       5      15    0.0414
## 0.256                      20      25    0.5515
## 0.385                      31      41    0.2888
## 0.513                      36      55    0.0586
## 0.641                      56      75    0.1155
## 0.769                      63      83    0.1155
## 0.898                      72      95    0.0884
## 1.026                      88     112    0.1036
## 1.154                      99     132    0.0350
## 1.282                     106     145    0.0163
```

4 Covariate Balance

4.1 RD Estimation for Prtreatment Covariates: Male & Age

To check for covariate balance, Table 1 shows the local average treatment effect (LATE) estimates for the two pretreatment characteristics (Age and Male) at the running variable cutoff (eGFR = 30). These estimates were calculated using a triangular kernel and bandwidth of 3 eGFR points around the cutoff (i.e. 27 to 33). "Honest" standard errors were separately estimated using the "RDHonest" package. The LATE estimate for the Male covariate at the cutoff is -0.189 with a large standard deviation of 0.256. This estimate is neither statistically significant nor statistically different from zero, given the large standard error and the estimate's nearness to zero. Therefore, the sex of patients (represented by the Male covariate) is balanced at the cutoff. The LATE estimate for the Age covariate is -14.238 with an honest standard error of 5.476. This indicates that Age may not be balanced at the cutoff and that as a patient passes the cutoff, their age has a discontinuous relation to eGFR. To explore this further, I provide linear and quadratic plots of these pretreatment covariates below, modeled separately on either side of the cutoff to visualize possible discontinuities at an eGFR of 30.

```
##
## Table 1: LATE of Pretreatment Characteristics at Cutoff
## ========================
##             Male    Age
## ------------------------
## eGFR       -0.189 -14.238
## Honest_SE 0.256   5.476
## ------------------------
```

4.2 Pretreatmant Covariate Discontinuity Models: Linear & Quadratic

As expected from the LATE estimate of Male at the cutoff, there is not convincing evidence of an observable discontinuous change in the Male covariate at the cutoff. In Figure 4, the linear model shows a slight increase in the Male covariate at the cutoff, but this is insignificant, given that this gap is covered by the

confidence intervals shown in gray. The quadratic model in Figure 5 shows a slightly larger gap for Male at the discontinuity of the same sign, but also overwhelmed by the confidence intervals . These models provide further indirect evidence that the sex of patients is balanced at the cutoff.

Less expected are the results for the Age covariate. The LATE estimate for Age in Table 1 implied a possibly significant discontinuity in Age at the cutoff. However, Figure 4 shows that Age is essentially perfectly continuous across the cutoff when modeled linearly. The quadratic model for Age in Figure 5 shows a small but insignificant gap, similar to that of Male but of reversed sign. These models provide indirect evidence that Age is also balanced at the cutoff, but with less confidence. Given the significant results in Table 1, it is possible that Age is slightly unbalanced at the cutoff, and this possibility needs to be considered when interpreting the ultimate results. The implication of an unbalanced Age covariate is that Age is interacting with eGFR at the cutoff, thereby explaining some of the discontinuity found in the LATE for the outcome variable. As individuals age, their kidney function naturally decreases, however there is no common sense connection between this natural decrease and the cutoff. It is reasonable to assume that Age is, more likely than not, balanced at the cutoff. That said, the LATE for the outcome variable at the cutoff could be different when comparing very young and very old patients. The sample size of this dataset is too small to detect meaningful differences between age groups. However, further analyses could account for any imbalance in age by measuring the LATE estimate for the outcome (months to Stage 5 kidney disease) by separate age groups. For the purpose of this analysis, we will use the provided evidence as mild confidence that Male and Age are both invariant to changes in the treatment assignment (cutoff). It is therefore possible to be mildly confident in the regression discontinuity design's ability to provide unbiased estimates on the number of months to Stage 5 kidney disease caused by the eGFR cutoff. Stronger confidence would require a larger dataset.

Figure 4: Linear Discontinuity for Pretreatment Characteristics

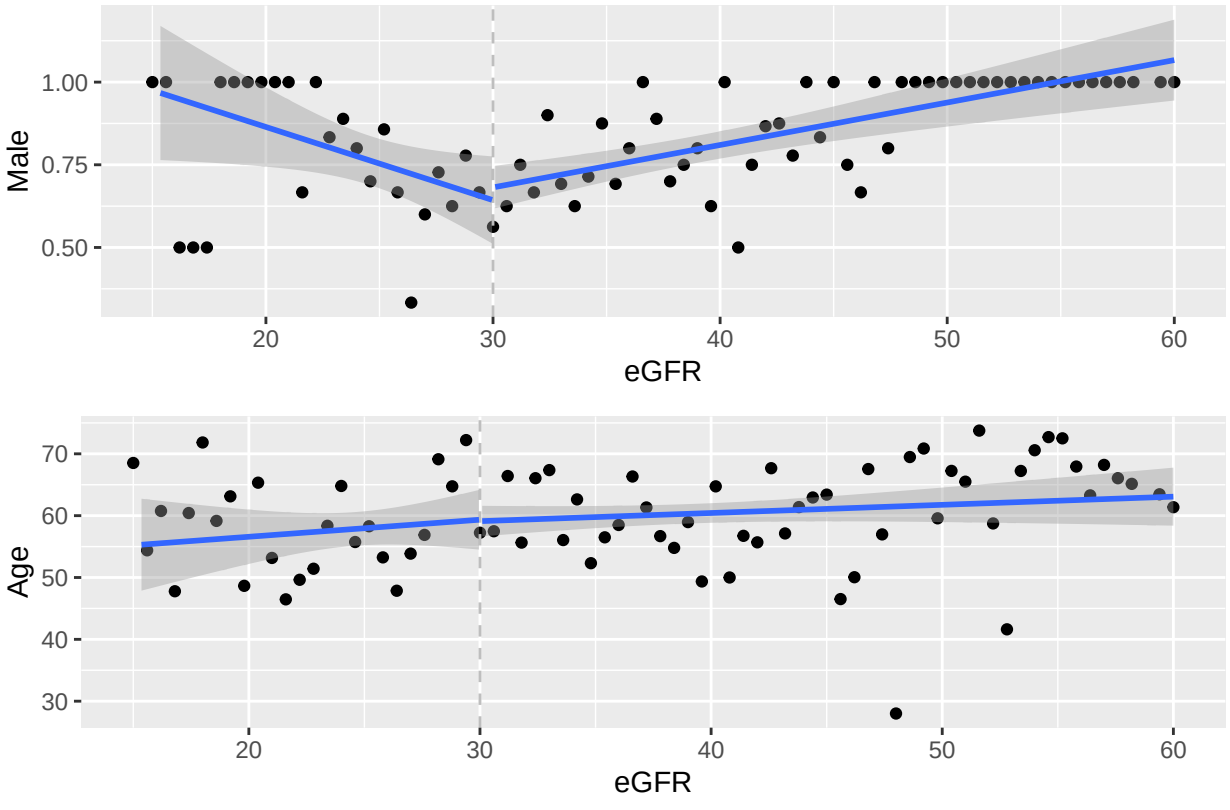
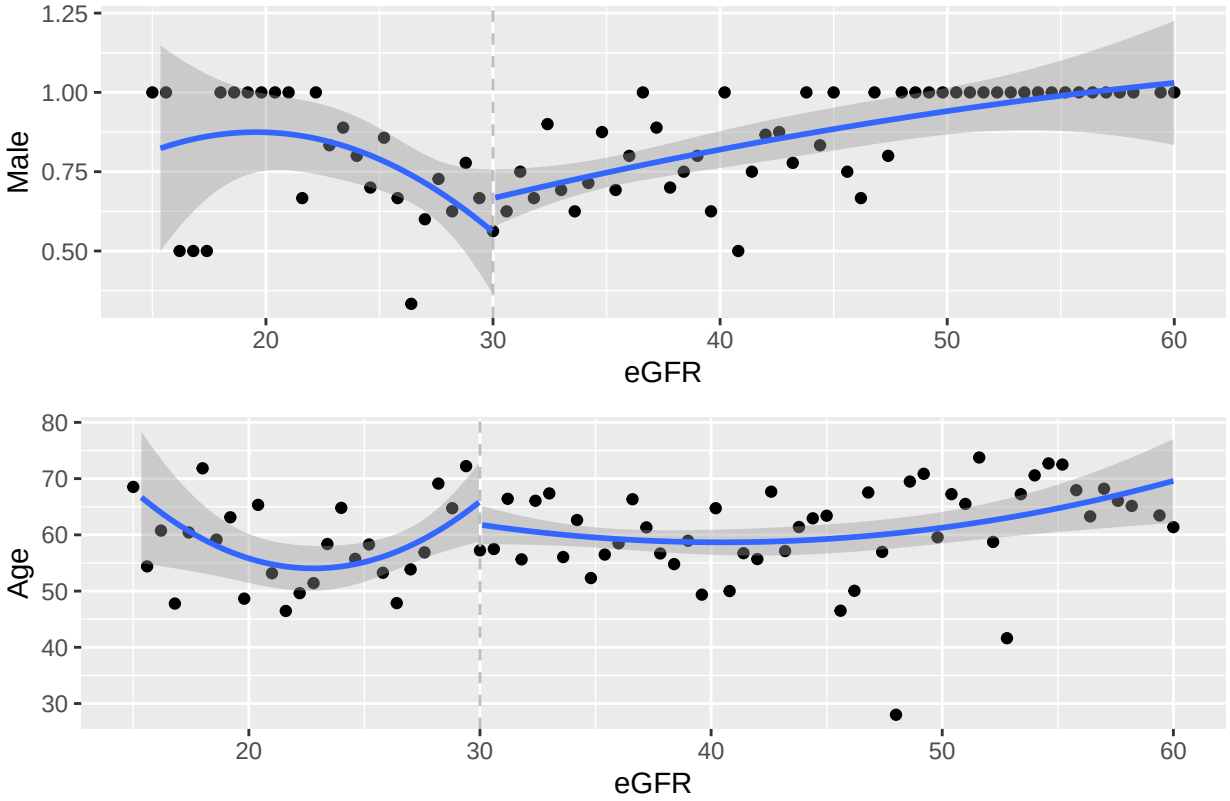


Figure 5: Quadratic Discontinuity for Pretreatment Characteristics



5 Regression Discontinuity Estimation

```
##
## Table 2: LATE for # Months to Stage 5 (outcome)
## ================================================================
##                    eGFR  Linear_Interact Linear_Quadradic_Interact.
## ----------------------------------------------------------------
## Months_to_Stage5 -8.207     -8.207                -8.207
## Honest_SE         6.167      6.167                 6.167
## ----------------------------------------------------------------
```

Using a bandwidth of 3 eGFR points on either side of the cutoff (27 to 33), the Local Average Treatment effect for the number of months until a patient reaches Stage 5 kidney disease (outcome variable) is -8.207 months, with an honest standard error of 6.17 months (Table 2). This implies that an individual just above the eGFR cutoff progresses to Stage 5 kidney disease 8.2 months sooner than an individual just below the cutoff. This would support the alternative hypothesis of discontinuity in outcomes resulting from the arbitrary cutoff. The medical implication is that individuals just above the cutoff have *worse* health outcomes, measured by speed of disease progression, than those just below.

To explain further, if an individual has an eGFR of 29 and is therefore classified into Stage 4 kidney disease, this patient begins receiving more intensive or advanced care (e.g. dialysis, daily monitoring, and prescription drugs). This extra care slows or prevents the progression to Stage *5* kidney disease (eGFR below 15) for most patients. Therefore, this group is expected to have a higher outcome value (more months elapsed before Stage 5 diagnosis). Conversely, an individual with an eGFR of 31 is not diagnosed with Stage 4 kidney disease (the patient remains in Stage 3) and therefore would receive the extra care as immediately.

The lack of extra care would make the individual more likely to reach Stage 5 kidney disease sooner because even though they eventually fall below the eGFR cutoff of 30, they spent a greater length of time without crucial extra care than the individual with an eGFR of 29 and therefore could suffer more kidney damage while waiting. If accurate, the LATE estimate of -8.207 would reject the null hypothesis of no discontinuous effect at the cutoff. The honest standard error of 6.17 is relatively high but does not entirely overwhelm the negative sign of the LATE. Even at one standard deviation below the estimate, a patient just above the cutoff reaches Stage 5 roughly 2 months sooner–a non-trivial span of time for a patient with Stage 3 kidney disease.

Therefore, the LATE gives evidence of discontinuous health outcomes as a result of the cutoff. This is attenuated by the inability to fully satisfy the smoothness assumption and pretreatment balance. Should the LATE sign truly be negative, however, it could be inferred that individuals just above the cutoff reach a more dangerous diagnosis (Stage 5) sooner than those just below the cutoff. The negative LATE is important because it implies that starting advanced care earlier for individuals near an eGFR of 30 gives them about 8 additional months of life before reaching Stage 5 kidney disease. The most important implication is for the patients who *never* reach Stage 5 kidney disease as a result of receiving the extra care. While this dataset only provides patients who ultimately experienced Stage 5 disease, this result implies that by providing advanced care sooner, many patients can recover kidney function and never progress to Stage 5. The question is whether this data holds for a much larger sample of patients with kidney disease.

While eGFR cutoffs are necessary to classify patients into stages of progressive disease, knowing that the cutoff may cause significant discontinuous health outcomes could lead medical staff to rely less on the cutoff when assigning extra care. Medical staff of course do not rely solely on eGFR for sending patients to more advanced treatment regimens, but knowledge of a discontinuity could downweight reliance on the cutoff for assigning extra care and thereby reduce the amount of patients progressing to Stage 5 kidney disease more quickly than necessary. This is crucial becuase recovery from Stage 5 is extremely difficult, and slowing the progression to this stage is essential to saving lives. The purpose of this analysis is therefore to provide physicians with information on patient outcomes (as it relates to the cutoffs used in medical practice) to hopefully alter physician behavior in a way that causes this discontinuity to disappear.

Note that Table 2 also shows that when the cutoff is interacted with the binary Stage 4 covariate linearly and quadratically, the LATE is unchanged. Across all three models, there is roughly an 8-month negative discontinuity in outcomes with about a 6-month honest standard error.

6 Discontinuity Models (Linear, Quadratic, and Loess)

The discontinuity found in the regression discontinuity estimateion in Table 2 is modeled visually in three ways: linear, quadratic, and loess. These models are shown below with ensuing discussion.

Figure 6: eGFR and Time to Stage 5 (Linear)

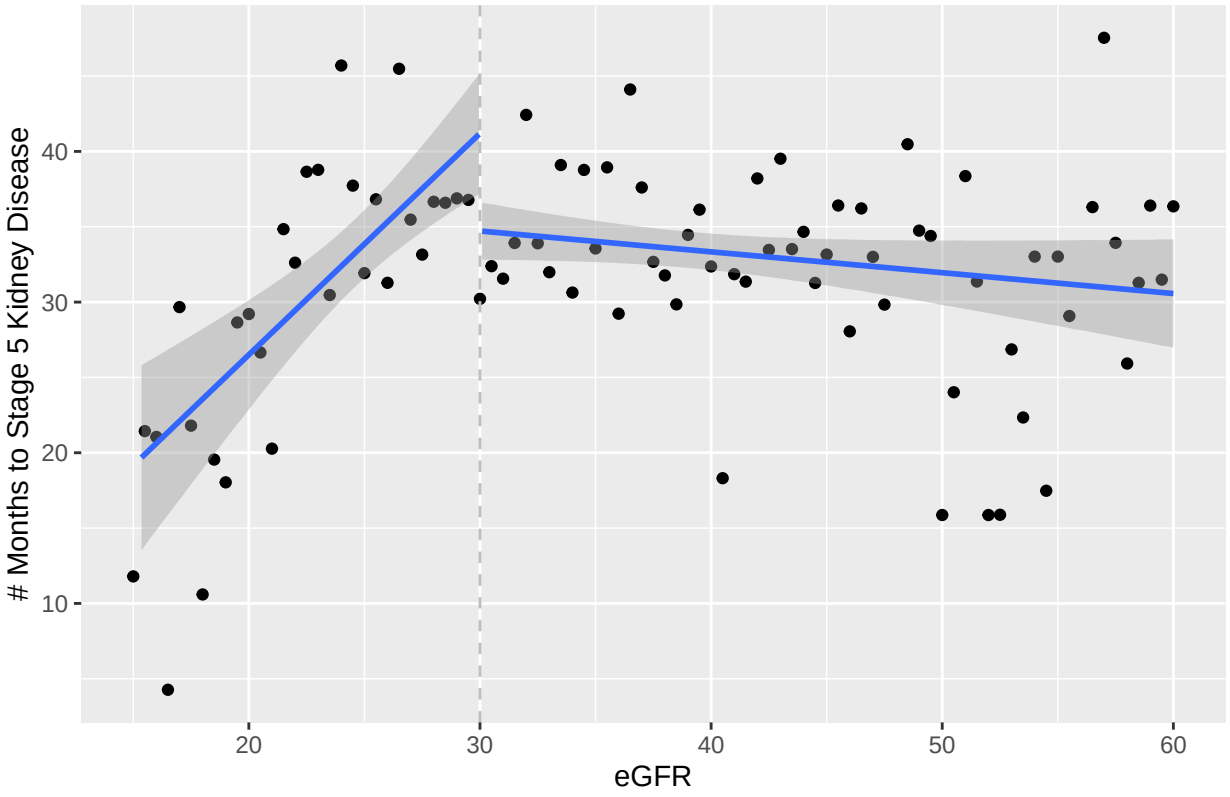


Figure 7: eGFR and Time to Stage 5 (Quadratic)

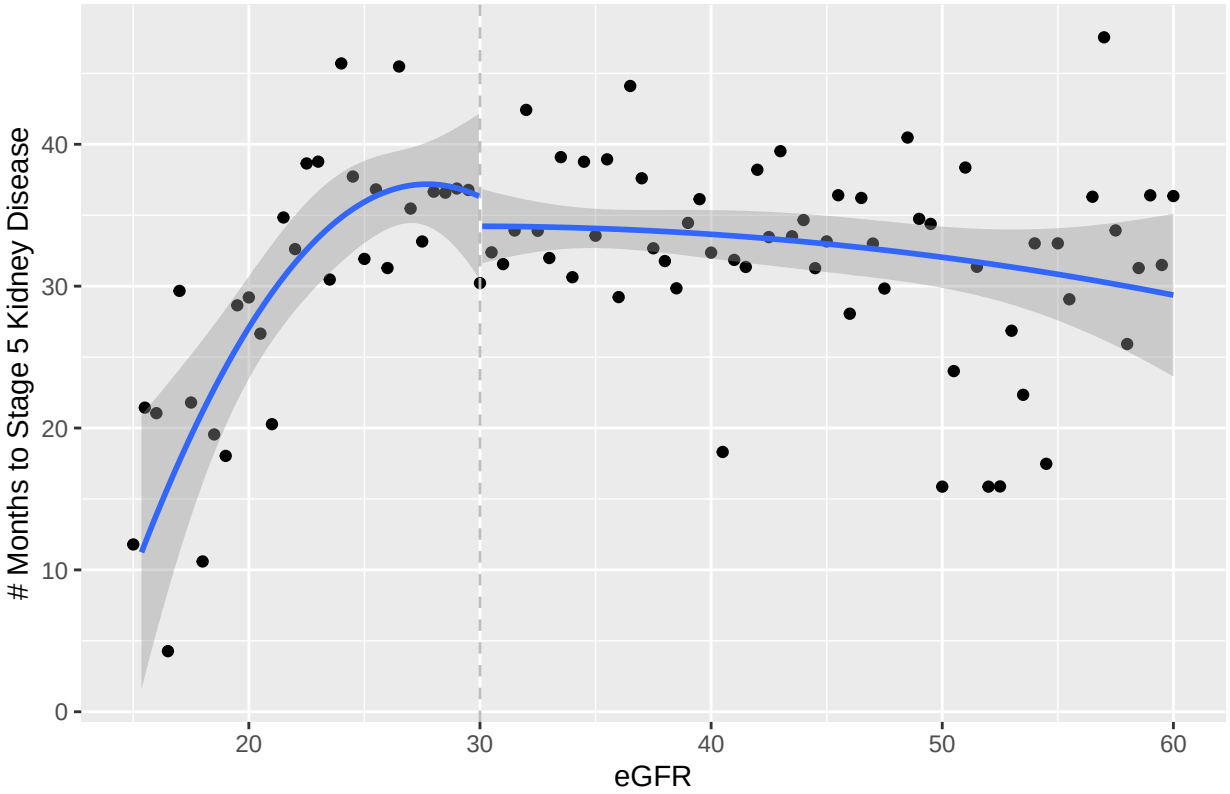
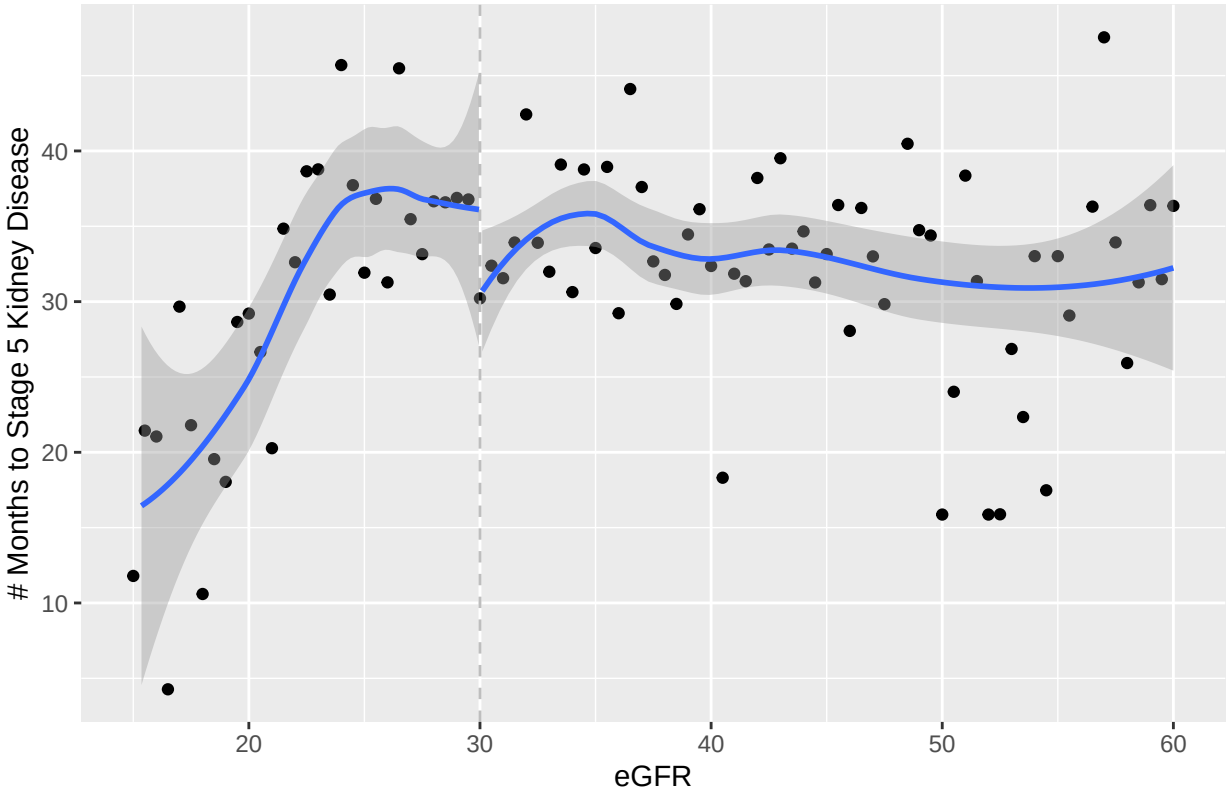


Figure 8: eGFR and Time to Stage 5 (Loess)



The graphical results modeling outcome discontinuity provide varying support for the RDD finding in Table 2. The linear model in Figure 6 most strongly supports the finding of a discontinuity close to 8 months between patients just above and just below the cutoff. Figure 6 shows that individuals just above the cutoff are expected to reach Stage 5 kidney disease within 35 months while individuals just below the cutoff are not expected to reach Stage 5 kidney disease for 42 months. This is a difference of about -7 months, which is very close to the LATE estimates in Table 2 of about -8.2 months. Furthermore, the confidence intervals around these estimates do not overlap, providing good evidence that at least some discontinuous gap exists in this data sample. This confidence is of course attenuated by the low sample size and can only be believed by scaling up this analysis to a very large sample. The preliminary evidence of a possible discontinuity, and the medical implications of its possible veracity, make exigent the exploitation of a larger sample size.

Figures 7 and 8, which show quadratic and loess models, respectively, provide weaker evidence of discontinuity, with the quadratic model (Figure 7) being the weakest. The quadratic model shows a very small gap in the estimates around the cutoff of only about -3 months (34 months just above the cutoff and 37 months just below the cutoff). Despite this smaller gap, which effectively disappears due to the overlapping confidence intervals, the estimate remains negative, as indicated in Table 2. The loess model in Figure 8 shows a larger gap of about -5 months, but the confidence intervals near the cutoff are also very large.

These findings provide some support for a negative discontinuity, with the linear model providing the strongest support. Under the linear model, and with the support of the LATE results in Table 2, it can be cautiously asserted that falling just below the eGFR cutoff at the time of measurement slows the patient's progression to more dangerous Stage 5 kidney disease by 7 to 8 months (assuming smoothness and covariate balance in a large sample). Since falling below the cutoff is strongly associated with receiving advanced care (as a result of Stage 4 classification below eGFR of 30), this discontinuity is presumably the result of some patients receiving advanced care sooner than patients *with effectively the same kidney function status* as a result of the arbitrary eGFR cutoff.

As stated, further analysis must be conducted on a larger sample size. The ideal dataset for such an analysis

is the *Chronic Renal Insufficiency Cohort* (CRIC), which contains kidney function and demographic data for tens of thousands or patients across the world. The present analysis is limited by its low sample size and suffers from a possible violation of the smoothness assumption (see Density Test discussion in Section 3.2), slight imbalance in the Age pretreatment covariate, and visual evidence of discontinuity in only one of three models (i.e. the linear model in Figure 6). The LATE estimate in this analysis, however, does provide mild evidence of a possible discontinuity. While the estimate has a large standard error, the true sign is likely negative, supporting the alternative hypothesis of a discontinuity. Should a larger study bear out these results with greater confidence, awareness of such discontinuity among physicians could lead to a closing of this outcome gap. Patients within a certain range above the cutoff may be more often assigned advanced care than before, extending the amount of months before Stage 5 to a period similar to those just under the cutoff. In this way, knowledge of the discontinuity could improve patient care strategies, removing the discontinuity and leading to better health outcomes.