What Difference Does a Diagnosis Make? Evidence from Marginal Patients *

Mattan Alalouf, Sarah Miller, and Laura Wherry

June 2018

Abstract

This paper explores the impact of receiving a diagnosis of diabetes among patients with similar underlying health. We find that a marginally diagnosed patient spends $1,300 more on diabetes-related care and drugs during the first 7 years following diagnosis, but find no changes in self-reported health or health behaviors. We find that the diagnosis of preventable conditions such as diabetic retinopathy, neuropathy, kidney disease, ketoacidosis and diabetic coma increases following a diagnosis, but the most serious among these (ketoacidosis and diabetic coma) decrease in the diagnosed group over time. Using clinical data on health measures, we find some evidence that marginally diagnosed patients have lower BMI and blood pressure. Because a large fraction of diabetes patients have lab values close to the diagnostic threshold, our results imply that increasing the cutoff by 0.1 percentage points would reduce total spending on diabetes about $1.02 billion.

---

* This work benefited from helpful comments from seminar participants at the University of Michigan and UCLA. We would also like to thank Drs. Doug Bell, John Mafi, and Tannaz Moin for helpful conversations, as well as Amanda Do, Rob Follett, and Javier Sanz for their help in accessing and interpreting the UCLA medical records data. Laura Wherry benefitted from non-financial support from the California Center for Population Research at UCLA (CCPR), which receives core support (R24-HD041022) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). Research with UCLA medical records data was supported by NIH National Center for Advancing Translational Science (NCATS) UCLA CTSI Grant Number UL1TR001881.

I. Introduction

Since the late 1990s, the diagnostic criteria for several chronic diseases have been relaxed, creating millions of new patients who were considered healthy under previous diagnosis guidelines (Welch, Schwartz, and Woloshin 2011). For example, new guidelines on diagnosing hypertension outlined by the American College of Cardiology and the American Heart Association in 2017 are expected to increase the number of patients with hypertension by over 40%, resulting in an additional 31 million Americans recommended to receive treatment (Whelton et al. 2017; Mutner et al. 2017). These guideline redefinitions can have outsized impacts because patients diagnosed with an illness disproportionately have lab values that fall near the threshold for diagnosis.

This paper explores the impact of a diagnosis on such "marginally" sick patients by exploiting a cutoff in the diagnosis criteria for a common chronic illness, diabetes mellitus. The American Diabetes Association (ADA) recommends diagnosing diabetes in patients whose glycated hemoglobin (A1c) levels are at or above 6.5% (American Diabetes Association 2016). Patients with A1c values just below this threshold have similar underlying health as those with A1c values just above, but may receive very different types of care because they are not considered diabetic. We use claims data from 232,600 patients from 2009 to 2017 to examine the impact of receiving a diagnosis of diabetes on different types of spending, and matched survey data to examine how the diagnosis affects self-reported health and healthy behaviors. In addition, we examine data from electronic health records to measure the impact of a diagnosis on two clinical measures of health, BMI and blood pressure.

Implementing a regression discontinuity design, we find that those with A1c values falling right above the threshold are between 11 and 29 percentage points more likely to have a diabetes diagnosis appear on their health insurance claims in the first year following the test. This increase in diagnosis rate is associated with approximately $1,300 additional spending per year on recommended diabetes-related care, and this elevated rate of spending persists for the seven years over which we observe patients. Spending increases related to office-based care and laboratory tests are particularly large (representing in total approximately $1000 of the observed increase in spending), although we also detect statistically significant increases in spending on diabetes-related drugs and vision screening. We find no evidence of changes in self-reported

health or in healthy behaviors such as diet, exercise, or smoking following the initial lab test. The rate of diagnoses for complications related to diabetes—diabetic retinopathy, neuropathy, kidney disease, ketoacidosis and diabetic coma—initially increase, likely due to more intensive screening. However, we detect decreases in the rates of ketoacidosis among those marginally diagnosed with diabetes by year 6 following the initial lab test. Using electronic health records data, we find improvements in BMI and blood pressure that persist for at least three years following the initial lab test.

Marginally diagnosed patients are a relevant group to study because they represent a large fraction of those diagnosed with a disease (Welch, Schwartz, and Woloshin 2011). Most people have blood glucose levels near the normal range, and blood glucose levels in the extremes are rare. As a result, even though spending is relatively low among the "marginal" diabetics whose A1c fell just above the threshold, they have an outsized impact on total costs due to their large numbers. For example, in our data, over a quarter of all patients with A1c in the diabetic range had values within 0.1 percentage points of the diagnosis cutoff. Our spending results imply that moving the diagnosis threshold from 6.5 to 6.6 would reduce total diabetes-related spending by about $1.02 billion annually.

Our analysis builds on existing work in economics that evaluates the impact of diagnoses, and subsequent care, on health and economic outcomes. Analyses of infants born near the diagnosis threshold for "very low birthweight" (birthweights of less than 1500g) have documented effects of this diagnosis on health and human capital outcomes (Almond et al. 2010, Bharadwaj, Loken and Neilson 2010). Oster (forthcoming) documents that households in which someone was recently diagnosed with diabetes exhibit small reductions in their purchases of unhealthy food.

Recent work by Iizuka et al. (2017) uses a similar identification strategy to examine the use of health care around the lower "pre-diabetes" threshold using data from Japan. They find that patients who exceed this threshold have small but statistically significant increases in the probability of visiting a doctor for diabetes in the following year, but no improvement in any health measures. Our work builds on this paper in several ways. First, we focus on a higher cutoff, for diagnosis of actual diabetes (rather than pre-diabetes), which may be more relevant for treatment decisions and long-term health. Indeed, we observe increases in spending at the diabetes diagnosis threshold that are between 10 and 20 times larger in size than those observed

by Iizuka et al. (2017), likely due to the fact that there are relatively few interventions recommended for pre-diabetes other than counseling to improve diet and exercise. Second,, our context is the United States, which differs substantially from Japan both in terms of health policy and population health.

Our work is also related to the literature in medicine on heterogeneous effects of care using randomized controlled trials (e.g., Kent et al. 2016). In general, these papers find that interventions for diabetes are most effective for those with worse baseline health. Our analysis builds on this by incorporating cost estimates as well as measures of behavioral changes and self-assessed health for marginal patients, measures that are often excluded in medical studies.

Finally, our work builds on a literature in economics devoted to measuring the marginal impact of additional medical care. This work seeks to disentangle patient selection into care from the benefit of the care itself, usually with the goal of determining whether certain types of care reflect wasteful "flat of the curve" spending. In general, the literature here is mixed, with some studies showing large benefits of additional spending associated with certain types of care (e.g., Doyle, Graves, Gruber and Kleiner 2015; Almond et al. 2010) and other papers showing limited or no benefit associated with additional spending (e.g., Frakes 2013). A limitation in the existing literature is that much of it focuses on care for acute conditions like heart attacks, high-risk deliveries, or at-risk newborns. This paper expands upon this by looking at disease care and management for a common chronic illness that affects much of the population and by looking over a longer time horizon to understand the dynamic nature of these types of health investments. Understanding the marginal impact of this type of care is particularly relevant given that treatment for chronic illnesses represent a substantial fraction of total medical spending (e.g., Gerteis et al. 2014) and that over 100 million individuals in the United States are diagnosed with diabetes or pre-diabetes (Center for Disease Control 2017).

II. Background

Many chronic illnesses are diagnosed by comparing the outcome of a laboratory test to a "cutoff" value specified in guidelines created by health or medical associations. To diagnose type 2 diabetes, physicians primarily rely on a measure of blood sugar called glycated hemoglobin A1c,

which is collected via a blood draw. A1c measures the exposure of hemoglobin cells to plasma glucose (blood sugar) over the past three months. Normal A1c levels range from 4.0 to 5.6 percent. High A1c levels, or amounts of glycated hemoglobin, indicate poor control of glucose, a marker of diabetes.

Although diabetes may be diagnosed using alternative tests of single or episodic glucose levels, A1c tests are considered the preferred diagnostic tool since they measure chronic glycemic exposure.[1] The International Expert Committee recommended A1c as a superior method to diagnose diabetes in a 2009 consensus report, noting that it is convenient (does not require the patient to fast) and more stable (less fluctuation within and between days).[2] Using data from cross-sectional epidemiological studies, this committee observed that the rate of any retinopathy, a long-term complication of diabetes, "increases substantially at A1c values starting between 6.0 and 7.0%." The committee also noted that when A1c levels were examined in 0.5% increments, the incidence of diabetes-specific "moderate" retinopathy noticeably increases at A1c levels $\geq$6.5%.[3] They concluded that this provides a strong justification for assigning an A1c cutoff value of $\geq$6.5% for the diagnosis of diabetes, although they note that "this cut point should not be construed as an absolute dividing line between normal glycemia and diabetes." The committee also emphasizes that while there is evidence for a continuum of risk for the development of diabetes based on A1c levels, "there does not appear to be a specific level at which risk for diabetes clearly begins" (International Expert Committee 2009).

Concordant with the IEC guidelines, the American Diabetes Association subsequently affirmed the recommendation of a diagnosis of diabetes if a patient's A1c is greater than or equal to 6.5% (American Diabetes Association 2010).[4] Once diagnosed, recommended care includes a

[1] Other tests that can be used to screen for diabetes are measurement of fasting plasma glucose (FPG) and a two-hour plasma glucose during an oral glucose tolerance test (OGTT). The OGTT is not commonly used other than among pregnant women to screen for gestational diabetes.
[2] The IEC was convened in 2008 to consider the diagnosis of diabetes among nonpregnant individuals. Members were appointed by the American Diabetes Association, the European Association for the Study of Diabetes, and the International Diabetes Federation.
[3] The analysis included data on approximately 28,000 individuals from nine countries. The IEC also considered specificity and sensitivity and relied on receiver operating characteristic curve analysis when determining the optimal cutpoint.
[4] Beginning in 2011, the ADA guidelines specify a value of 6.5% or higher on two separate tests unless the patient exhibits "unequivocal" hyperglycemia. However, in practice, we do not find that patients are receiving multiple tests.

comprehensive medical evaluation to form a care management plan, annual diagnostic tests to detect comorbid conditions or complications (urinary albumin, lipid panel, glomerular filtration rate), A1c monitoring at least twice per year, antidiabetic medication, statin therapy (for those age 40 and above),[5] annual assessment for diabetic retinopathy by an ophthalmologist, annual comprehensive foot evaluation, as well as flu, pneumonia, and hepatitis B vaccines (ADA 2017). In addition, the ADA recommends that diabetics receive more aggressive treatment for comorbid conditions; e.g., diabetics with blood pressure above 140/90 mmHg should have prompt initiation of drug therapy at the "maximum tolerated dose" to achieve blood pressure goals, while patients at high risk for kidney disease require additional monitoring and treatment.

Patients whose A1c falls below 6.5, but at or above 5.7 are considered "prediabetic" with impaired glucose tolerance and at a higher risk of developing diabetes in the future. As such, they also receive some care aimed at improving health; specifically, the ADA recommends that they have their A1c tested annually and that they are counseled on making lifestyle interventions such as losing weight or increasing exercise.

Because individuals with A1c values falling just below or above the 6.5 threshold have similar underlying health, we can use the cutoff in diabetes classification to evaluate the impact of receiving a diagnosis on the health care utilization and associated costs, health, and behaviors for patients at the margin of being diagnosed. Understanding the role of diagnosis for a marginal patient is particularly interesting because of the underlying distribution of A1c in the population of those who are tested. In our data, comprised of claims from a large health insurer, among patients whose A1c level meets the diagnostic threshold for diabetes, over 25 percent have a blood test between 6.5 and 6.6, inclusive. Although these patients are the healthiest within the diabetic group, they have the potential to have an outsized impact on total spending because of their large numbers. In addition, most discussions about changing treatment guidelines involves adjusting diagnosis criteria around the "margins" of existing rules.

III. Data

A. OptumInsight Claims and Health Risk Assessment Data

[5] This guideline was recommended in 2013 by the American College of Cardiology and the American Heart Association and confirmed by the American Diabetes Association at the end of 2014.

To assess the impact of a diagnosis on health care utilization and spending, self-reported health and subsequent blood sugar, we use the OptumInsight (Optum) database from 2009 to 2017.[6] Optum provides claims data from a large health insurance company; over our sample period, 61 million patients are included in these data. The data include all claims associated with these patients that are paid for by the health insurance company. The claims data include procedure and diagnosis codes, as well as the date the procedure occurred and the cost of the claim (including the amount paid by the patient and by the insurance company); for pharmacy claims, the data include the national drug code, quantity dispensed, and cost. Additionally, the data contain some limited information on demographic characteristics of the patient, such as age, race, and gender. Optum data is primarily composed of private or employer-sponsored health insurance plans, but includes some Medicare Advantage plans as well (see Table 1).

When an A1c test is used for diagnosis, the provider sends a blood sample to a lab for evaluation. In the Optum data, we observe the lab test results for laboratories that are affiliated with Optum. We do not observe test results for un-affiliated laboratories, although we do see spending associated with these laboratories. In total, we observe lab results for 80% of all labs taken by patients within our dataset. In nearly all cases, lab test results are rounded to the nearest 0.1 percent; for the small number of lab results that are reported with greater precision, we round to the nearest 0.1 percent so the data are comparable across all labs reported.

The data include unique patient identifiers that allow us to track patient spending over time. We are able to track patient spending if patients change insurance plans as long as the insurance plan is offered by an insurance company associated with Optum. If patients drop their insurance coverage, or switch to an insurance plan offered by a different company, we are no longer able to observe them in the data. However, Optum provides records of patients' enrollment, so we can distinguish patients who dropped coverage from patients who had no claims in a given year. We do detect a small increase in attrition associated with receiving an A1c test exceeding the threshold; in Section VI we examine this differential attrition and evaluate how it affects our estimates.

[6] Although OptumInsight data are available for earlier years, we use 2009 as our first initial A1c test year because this was the first year that A1c was recommended for the diagnosis of diabetes.

A small subset of patients completes a Health Risk Assessment (HRA) survey in any given year. This survey asks a wide range of questions about self-assessed health and behaviors and is typically administered by the insurance company at the behest of an employer. As such, patients who take the HRA are not representative of patients in the Optum database; however, selection into taking an HRA is not associated with an A1c value above the threshold (see Section VI). Therefore, our estimates provide an internally valid measure of these self-assessed outcomes in the HRA subsample.

A full list of questions considered from the HRA is found in Appendix Section 1. In order to avoid conducting an excessive number of statistical tests, which might generate a large number of false positives, and in order to improve our statistical power, we construct four summary indices for the following question domains: self-assessed health, physical and emotional well-being, healthy diet and exercise behavior, and smoking and drinking behavior. We do this by constructing a standardized z-score (subtracting the mean and dividing by the standard deviation) for each component and taking the average. This applies equal weight to each index component and is similar to the approach used in Finkelstein et al. (2012) and Kling, Liebman, and Katz (2003). Positive values of the index indicate better outcomes (i.e., better health or lower rates of unhealthy behavior).

To construct our sample, we pull all adult patients who have an A1c test with observable results over our sample period. From this group, we drop individuals who had a diagnosis of diabetes on their claim at any time before their first observed A1c test in order to eliminate patients who are receiving the test in order to monitor, rather than diagnose, diabetes. We also exclude individuals with a pregnancy-related diagnosis during the year prior to their first A1c test in order to drop cases associated with gestational diabetes. We further restrict to patients who are in the sample at least 12 months prior to the initial lab test and 12 months following the lab test in order to be able to observe their spending before and after. This leaves us with 232,586 patients. Appendix Figure 2 provides a diagram of our sample definition with patient counts for each exclusion criterion.

Our primary outcome measures look at spending for all types of care recommended for diabetics by the American Diabetes Association. This includes doctor office visits and consultations related to diabetes; spending on diabetes-controlling drugs such as metformin, as well as statins,

and ACE inhibitors; lab tests for relevant conditions; and spending related to vision testing. We examine total spending on recommended care as well as each category individually. In addition, we evaluate spending on endocrinology specialty care and indicators for diagnoses of five conditions associated with uncontrolled diabetes: diabetic retinopathy, neuropathy, kidney disease, diabetic coma, and ketoacidosis.

Table 1 presents descriptive statistics for these patients in the left panel. Of these patients, 54% are female, average age is 49.7, and about 25 percent are enrolled in a Medicare advantage plan (with the remaining on private employer-sponsored insurance). Women are slightly over-represented relative to men. This is likely because screening is recommended for women with a previous diagnosis of gestational diabetes.

B. Electronic Health Records Data

The Optum data allow us to see detailed information on utilization and costs, but relatively little information on actual health measures. To investigate the impact of a diagnosis on health measures, we use electronic health records (EHR) data for patients receiving care in the University of California, Los Angeles health system. There are approximately 4.9 million distinct patient records and these data include clinical health measures recorded any time a patient visits a UCLA medical provider. They also include some information on utilization of services (number of encounters with the UCLA system), diagnoses, procedures, and prescription drugs, although they do not include cost measures.

Similar to the Optum data, we constructed our analytic sample using UCLA data available from 2009 to 2016.[7] Like the Optum data, the UCLA data allow us to follow patients over time. In the UCLA data, we are able to track a patient as long as they continue to have encounters within the UCLA health system, which includes two large hospitals and all clinics. We limit our sample to patients with a primary care provider (PCP) at UCLA in order to increase the likelihood that we observe their ongoing care. We explore differential attrition in Section VI below.

[7] Technically we only had data for encounters occurring through November 30, 2016. While we had data prior to 2009, we limited our analytic sample to those patients who had their first test in 2009 or later. Please see appendix for additional details on sample construction for the EHR data.

We construct an analytic sample using similar criteria to those used with the Optum data (see appendix for additional details). Our outcomes of interest are two measures that are collected for a large percentage of patients: body mass index and blood pressure.[8] Although the UCLA data allow us to observe encounters for patients with A1c tests dating back to 2009, the clinical health measures are only available starting in March 2, 2013, when UCLA implemented its current EHR program CareConnect, which uses Epic Systems technology. We present results for individuals with A1c tests conducted in 2013 and later, but provide additional analysis using all individuals with A1c tests conducted in 2009 and later in the appendix. The results are similar between the two samples.

Descriptive statistics from the UCLA data are presented in the third column of Table 1. The individuals in the UCLA data are relatively similar to those in the Optum data, although there are slightly more Asians represented and fewer blacks. Age at first encounter is also somewhat higher, at 47.5 years.

IV. Empirical Strategy

We use a regression discontinuity (RD) design framework to analyze the impact of a diagnosis of diabetes across patients with similar underlying health. To implement this, we use local linear regression to model the trend in A1c on either side of the cutoff, using the Imbens and Kalyanaramen (2012) optimal bandwidth selection procedure. In the appendix, we present results that assess the sensitivity of our results to other bandwidth sizes (see Appendix Table 2).

We estimate two versions of this model. The first version is a "reduced form" model that looks at the change in the outcome at the cutoff. The second version presents a "fuzzy" regression discontinuity estimate of the effect of receiving a diagnosis of diabetes using the diagnostic cutoff as an instrument. If all patients with A1c values below the cutoff were not diagnosed with diabetes, and all patients above the cutoff did receive a diagnosis, these two estimates would be identical. However, the change in the prevalence of a diagnosis does not change from 0 to 1 at

[8] We used BMI reported in the EHR data when available. In other cases, we calculate BMI ourselves using weight and height measures when collected on the same day and at the same encounter. Patients occasionally had multiple measurements taken on the same day during an encounter. In these cases, we used the average value for each measurement. We also dropped outlier BMI values that fell in the first and 99th percentiles of the distribution.

the 6.5 threshold, and the fuzzy regression discontinuity design scales the reduced form estimate by the observed change in diagnosis prevalence.

One concern about the validity of our design is whether or not physicians can manipulate the test score in some way as to generate selection across the threshold. However, due to the nature of the test, this is unlikely in our setting. The patient's A1c reflects the plasma glucose levels of the patient during the past 3 months, making it difficult to manipulate. In addition, the test values are reported to Optum or recorded in the patient's medical record directly from the lab, making physician or patient manipulation impossible.

The first panel of Figure 1 presents the sample size in the Optum data. There are three noticeable features. First, the density appears to be smooth across the threshold, indicating there is no bunching on one side of the threshold or the other (we test for this formally in Appendix Table 3). Second, given the shape of the distribution of A1c values within the population that takes the test, a substantial fraction of patients whose A1c is at or above the threshold fall very close to the threshold. Indeed, 15% of patients with A1c at or above 6.5 in the Optum data have an A1c of 6.5; 25% have an A1c of 6.5 or 6.6. We see similar patterns in the EHR data, presented in the second panel of Figure 1. Finally, the UCLA sample appears healthier overall, with more of the distribution falling in the normal A1c range. This could reflect differences in the underlying health of the UCLA population relative to the national population captured in the Optum database, or it could reflect more aggressive A1c testing within the UCLA health system. While these differences do not affect the internal validity of our estimates, they could affect the types of populations our results are applicable to.[9]

V. Results

Short Run Effects

[9] We do not observe spending in the EHR data, making it difficult to directly compare the Optum and EHR results. However, in the appendix, we do report changes in the number of observed encounters and the number of labs related to diabetes. See Appendix Table X.

We first analyze the impact of having an A1c value exceeding the 6.5 threshold on diagnosis and use of medical care in the first 12 months following the initial lab test. The first panel of Figure 2 plots the rates of diagnosis within A1c values for the year following the initial test. We see a large and discontinuous increase in diagnosis rates at the cutoff value of 6.5. At the 6.5 value, diagnosis rates jump by approximately 11.4 percentage points. The second panel of Figure 2 shows average spending on all recommended diabetes care. Here, we also observe a clear discontinuous increase in spending at the diagnosis threshold. The subsequent panels show the results for spending on specific types of diabetes related care: office-based care (consultations and evaluation), lab test spending, and drug spending for diabetes-controlling drugs. We observe noticeable jumps in office-based and lab spending, but no noticeable jump for lab spending in the first year. Finally, we also examine spending for statins, ACE inhibitors, and endocrinology specialty care; all appear to increase at the cutoff.

Table 2 presents the local linear regression estimates associated with the outcomes in Figure 2. We see a statistically significant increase in the presence of a diagnosis of about 11.4 percentage points (row 1). The next row shows the effect of crossing the diagnosis threshold on spending on recommended care. At the cutoff of A1c equal to 6.5, we observe a discontinuous increase in spending of about $201. We scale this increase in spending by the fraction of the sample who gain a diagnosis of diabetes at the cutoff using a fuzzy RDD estimate in column 2. This estimate shows that a diagnosis of diabetes is associated with approximately $1,200 of additional spending on recommended care.[10] When we break this spending down into subcategories, we see that the spending increases in the first year are split between office consultations and evaluations, which experiences an increase of about $70, and spending related to lab work, which experiences an increase of about $91. We also see significant increases in spending related to vision, which is to be expected given that diabetics are recommended to have at least one vision check-up per year. In addition, we estimate marginally significant increases in spending on statins and on endocrinology specialty care. While we observe a positive estimate on spending on hypertension-controlling drugs, it is not statistically significant.

[10] Note that because the optimal bandwidth selection procedure chooses different bandwidths for each outcome, the first stage differs slightly across outcomes; in all cases, it is highly statistically significant.

Figure 3 shows the effect of receiving an A1c value above the diabetes diagnosis cutoff on the incidence of preventable complications associated with diabetes: diabetic retinopathy (eye damage), diabetic neuropathy (nerve damage), diabetic kidney disease, ketoacidosis and diabetic coma. The incidence of all of these complications appears to increase substantially at the A1c cutoff. This is not necessarily surprising given that the ADA guidelines stipulate that much of the recommended care for diabetics involves screening for these conditions. The associated estimates are reported in Table 3. For all outcomes except ketoacidosis, we find statistically significant increases in the diagnosis of these complications.

The results from the Optum claims data demonstrate substantial increases in spending associated with the cutoff A1c value and indicate that patients with otherwise similar underlying health are treated very different depending on where their A1c falls relative to the diagnosis guidelines cutoff. Next, we evaluate whether these additional interventions affect clinical measures of health reported on the UCLA electronic health record.

We examine the effect of diagnosis on BMI, systolic, and diastolic blood pressure. The figures associated with these outcomes are presented in Figure 4. There is no visible change in these outcomes at the cutoff for BMI and systolic blood pressure, although diastolic blood pressure appears to be lower for those whose A1c values fall above 6.5. Table 4 presents the regression discontinuity estimates for these results. In the first row, we replicate the increase in diagnosis of diabetes at the cutoff A1c value, and see that it increases by 29 percentage points. In the subsequent rows, we show that diabetes diagnosis is associated with lower BMI and diastolic blood pressure, although there is no effect on systolic blood pressure.

Finally, we examine how the diagnosis of diabetes is associated with measures of self-reported health in the Health Risk Assessment. We do not find any change in self reported health or healthy behaviors associated with diagnosis of diabetes. Our point estimates indicate that those diagnosed with diabetes report worse subjective health.

Long Run Effects

The initial effects of diabetes diagnosis may not persist if, for example, those with A1c values directly below the cutoff gain a diagnosis relatively soon after the initial test. We examine how persistent the spending and health effects are by re-estimating our fuzzy RDD model using

outcomes from years 2 to 7 following the initial test. These estimates demonstrate how a diagnosis in year 1 affects spending in subsequent years.

Figure 5 shows how spending evolves over time, in years 2-7 following the initial lab test. We see fairly stable impacts of a diagnosis in year 1 on spending on recommended care in these subsequent years, with little evidence that the gap in spending between those above and below the threshold closes. Although the estimates become less precise due to the reduction in sample, we see consistently positive effects for all years, and statistically significant effects for the first 5 years. Our estimated effects on spending related to lab and office-based care are similarly stable over time. Even five years after the initial diagnosis, we observe that those with initial A1c values just above 6.5 have considerably higher spending in these categories. Spending on diabetes-controlling drugs appears to increase over time, with much higher spending observed in subsequent years. In contrast, spending on endocrinology specialist care, statins, and ACE inhibitors do not appear to be consistently higher in years 2 through 7 after the initial test.

Overall, our results using claims data show large increases in spending and use of care around the diabetes diagnosis threshold. These spending increases stay relatively stable or, in the case of drug spending, grow larger over time.

We next examine the incidence of preventable complications related to diabetes in Figure 6. Here, we see a consistently higher increase in the incidence of diagnosis for all subsequent years for retinopathy, neuropathy, and diabetic kidney disease. However, for the more serious and life-threatening complications, we see a decrease in the incidence by year six or seven. For ketoacidosis, the decrease in the incidence in year 6 is statistically significant. This pattern indicates that there may be prevention of serious complications associated with the spending on recommended care.

We also use the EHR data to evaluate how persistent the improvements in clinical health are. For these measures, we observe only 3 years after the initial test. Effects are similar across all three years. Finally, we look at the impact of a diagnosis on the self-reported outcomes in the health risk assessment survey. We continue to find no statistically significant effect on these outcomes in subsequent years, but our confidence intervals are large.

VI. Specification Checks and Additional Analyses

A. Robustness to Choice of Regression Specification and Bandwidth

We assess the robustness of our results to using several alternative bandwidths, reported in Appendix Section 1. Our results are very similar both quantitatively and qualitatively when alternative bandwidths are used, although larger bandwidths tend to result in somewhat larger estimates.

We also assess whether our results are robust to a more parametric approach which uses a quadratic model, estimated on either side of the bandwidth, to estimate the RDD effect. These results are reported in Appendix Section 1. Again, we find quite similar results using this approach.

B. Checking for Discontinuities in Baseline Characteristics

As an additional check, we evaluate whether demographic characteristics (age, gender, and, available in the EHR sample only, race) or insurance characteristics (Medicare versus non-Medicare) change discontinuously at the cutoff. If we found discontinuities in these seemingly unrelated characteristics, it would be a sign that our empirical strategy was potentially compromised.

We operationalize this by making each of these demographic and insurance characteristics our dependent variable. The results are reported in Table 6. We do not find statistically significant changes in any of the demographic characteristics we consider, although we do find a statistically significant reduction in the probability a patient has Medicare coverage in the Optum sample. To evaluate whether this differential enrollment in Medicare affects our results, we have re-estimated our spending models to include this covariate. The inclusion of this Medicare indicator has little effect on our results, suggesting that the spending effects are not driven by differential Medicare enrollment. Taken together our analysis suggests that there are not systematic changes in patient characteristics at the diabetes diagnosis threshold that are confounding our analysis.

C. Sample Attrition

a. Testing for Discontinuities in Sample Inclusion by Year

Both our Optum and UCLA samples are comprised only of patients who are associated with these organizations. If diagnosis with diabetes causes patients to, for example, switch out of Optum-associated plans or out of the UCLA health system, this could affect the interpretation of our outcomes for later years. Similarly, we may not observe patients in certain subsequent years based on when the patient's initial A1c test occurs. For example, if a patient has her initial test in 2016, we are only to observe 1 post-test year (2017), so that patient will not be available for regressions that estimate the effects in year 2 and beyond. In Appendix Section 2, we test for whether individuals with A1c values just above the threshold are more likely to leave the sample. We find no significant effects on the probability of attrition for either the UCLA or the Optum sample.

VIII. Conclusion

The distribution of health in the population is such that many individuals diagnosed with a chronic illness are near the margin of being diagnosed. This paper explores the impact of a diagnosis on such "marginal" patients. We find that individuals whose blood glucose levels are just above the threshold for diabetes diagnosis have significantly higher spending than similar patients whose blood glucose levels fall just below this cutoff. The increase in spending is persistent: even after 5 years, those whose initial test put them directly above the diagnosis cutoff have spending that is significantly higher. A back-of-the-envelope calculation suggests that increasing the threshold at which diabetes is diagnosed by just 0.1 percentage point would reduce total diabetes spending by over $1 billion per year.

We do observe small improvements in clinical measures of health for those who are just above the threshold of diagnosis. We also find significant changes in the incidence of one serious, preventable condition associated with uncontrolled diabetes, ketoacidosis, which materializes after 6 years of treatment. These results indicate that there are some health benefits of this treatment even among the relatively healthy diabetics whose A1c is close to the cutoff. We do not find any changes in self-reported health or self-reported healthy behaviors.

Overall, results suggest the effects of the increased spending on relatively healthy diabetes patients are modest, and would indicate that future revisions of diagnosis guidelines for chronic diseases should exercise caution when relaxing these guidelines to include healthier patients.

References

Almond, Douglas, Joseph J. Doyle Jr., Amanda E. Kowalski, and Heidi Williams. (2010). Estimating Marginal Returns to Medical Care: Evidence from At-Risk Newborns. *Quarterly Journal of Economics*, 125(2): 591-634.

American College of Cardiology and the American Heart Association Panel of Experts. 2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk. *Circulation*

American Diabetes Association. 2010. "Diagnosis and Classification of Diabetes Mellitus." *Diabetes Care* 33 Suppl 1: S62-S69.

Bharadwaj, Prashant, Katherine V. Loken, and Christopher Neilson. (2013). Early Life Health Interventions and Academic Achievement. *American Economic Review*, 103(5): 1862-91.

Center for Disease Control and Prevention. 2017. National Diabetes Statistics Report. Retrieved from https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf on July 12, 2018.

Doyle, J., J. Graves, J.Gruber and S. Kleiner. 2015. Measuring Returns to Hospital Care: Evidence from Ambulance Referral Patterns. *Journal of Political Economy* 123(1): 170-214.

Finkelstein, A., M Gentzkow and H. Williams. 2016. "Sources of Geographic Variation in Healthcare." *Quarterly Journal of Economics* 131(4):1681-1726.

Frakes, M. 2013. "The Impact of Medical Liability Standards on Regional Variations in Physician Behavior: Evidence from the Adoption of National-Standard Rules." *American Economic Review* 103 (1):257-276.

Gerteis J, Izrael D, Deitz D, LeRoy L, Ricciardi R, Miller T, Basu J. Multiple Chronic Conditions Chartbook. AHRQ Publications No, Q14-0038. Rockville, MD: Agency for Healthcare Research and Quality; 2014

Iizuka, T, K Nishiyama, B Chen, K Eggleston. 2017. "Is preventive care worth the cost? Evidence from mandatory checkups in Japan." NBER Working Paper Series No. 23413.

The International Expert Committee. 2009. "International Expert Committee Report on the Role of the A1C Assay in the Diagnosis of Diabetes." *Diabetes Care* 32(7): 1327-1334.

Kent, D, J Nelson, I Dahabreh, P Rothwell, D Altman and R Haward. 2016. “Risk and treatment effect heterogeneity: re-analysis of individual participant data from 32 large clinical trials.” *International Journal of Ipeidemiology* 2075-2088.

Welch, HG, L. Schwarz, S. Woloshin. Overdiagnosed: Making People Sick in the Pursuit of Health. 2011, Beacon Press Books.

Table 1. Demographic Statistics, Optum and UCLA EHR Data

	Optum	UCLA EHR	Health Risk Assessment
Female	0.541	0.612	0.581
Age	49.65	47.15	43.45
White	N/A	0.654	N/A
Black	N/A	0.048	N/A
Hispanic	N/A	0.111	N/A
Asian	N/A	0.105	N/A
Medicare	0.251	0.181	0.004

Note: Authors' calculations from Optum claims and UCLA EHR data.

Table 2. Regression Discontinuity Design Estimates of Effects of Diabetes Diagnosis in First 12 Months Following A1c Test (Optum Results)

	Baseline Mean (Std Dev)	Reduced Form Estimate	Fuzzy RD Estimate of Effect of Diagnosis
Diabetes Diagnosis	0.171	0.114***	N/A
		(0.017)	
Spending on Recommended Care	345.50 (727.14)	201.376***	1,215.441***
		(22.434)	(125.914)
Spending on Office-Based Care Related to Diabetes	25.73 (124.58)	70.402***	522.381***
		(7.027)	(50.515)
Spending on Diabetes-Related Labs	86.60 (374.13)	91.196***	537.849***
		(13.093)	(71.400)
Spending on Diabetes-Controlling Drugs	5.60 (125.06)	5.598*	24.003*
		(3.208)	(13.707)
Spending on Diabetes-Related Vision and Hearing Services	2.32 (40.06)	5.401***	24.755***
		(1.317)	(5.976)
Spending on Statins	143.12 (432.94)	19.860*	114.035*
		(10.179)	(58.416)
Spending on ACE Inhibitors	90.40 (294.79)	7.275	26.468
		(5.314)	(19.314)
Endocrinology Specialist Care Spending	20.17 (282.66)	28.952*	103.832*
		(15.751)	(56.488)

Notes: Baseline means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009).

Table 4: Regression Discontinuity Design Estimates of Effects of Diabetes Diagnosis on Subsequent Diagnoses of "Preventable" Conditions

	Baseline Mean	Reduced Form Estimate	Fuzzy RD Estimate of Effect of Diagnosis
Diabetic Retinopathy	0.002	0.004**	0.026**
		(0.002)	(0.012)
Ketoacidosis	0.0005	0.001	0.007
		(0.001)	(0.004)
Diabetic Coma	0.0005	0.002**	0.009**
		(0.001)	(0.004)
Diabetic Neuropathy	0.007	0.011***	0.064***
		(0.003)	(0.019)
Kidney Disease	0.006	0.017***	0.088***
		(0.003)	(0.015)

Notes: Baseline means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009).

Table 5: Regression Discontinuity Design Estimates of Effects of Diabetes Diagnosis in First 12 Months Following A1c Test (UCLA EHR Results)

	Baseline Mean (Std. Dev)	Reduced Form Estimate	Fuzzy RD Estimate of Effect of Diagnosis
Diabetes Diagnosis	0.080	0.292***	--
		(0.041)	
BMI	28.61 (5.55)	-1.094**	-3.303**
		(0.481)	(1.522)
Normal BMI	0.212	0.045	0.126
		(0.030)	(0.087)
Overweight	0.339	0.026	0.058
		(0.037)	(0.090)
Obese	0.421	-0.128***	-0.406**
		(0.045)	(0.158)
Diastolic blood pressure	76.13 (8.31)	-2.056***	-4.357**
		(0.768)	(1.975)
Systolic blood pressure	129.34 (14.29)	-1.197	-2.769
		(0.976)	(2.011)
High blood pressure	0.570	-0.082**	-0.118
		(0.041)	(0.082)

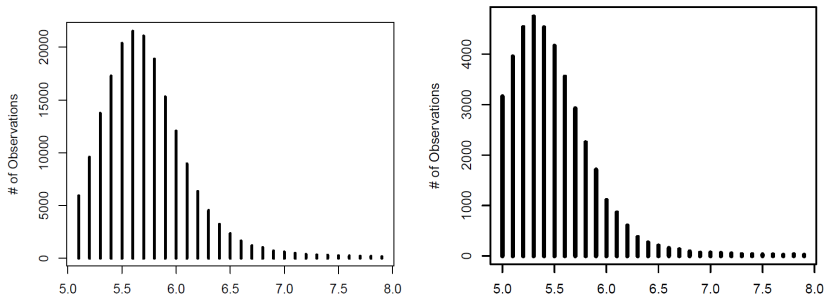
Notes: Baseline means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009).

Table 6: Regression Discontinuity Design Estimates of Effects of Diabetes Diagnosis in First 12 Months Following A1c Test (Health Risk Assessment Results)

	Reduced Form Estimate	Fuzzy RD Estimate of Effect of Diagnosis
Diet and Exercise	-0.013	-0.067
	(0.051)	(0.270)
Physical and Emotional Well-Being	-0.043	-0.306
	(0.055)	(0.282)
Self Reported Health Very Good or Excellent	-0.092	-0.593
	(0.082)	(0.549)
Healthy Smoking and Drinking Choices	-0.066	-0.219
	(0.061)	(0.287)

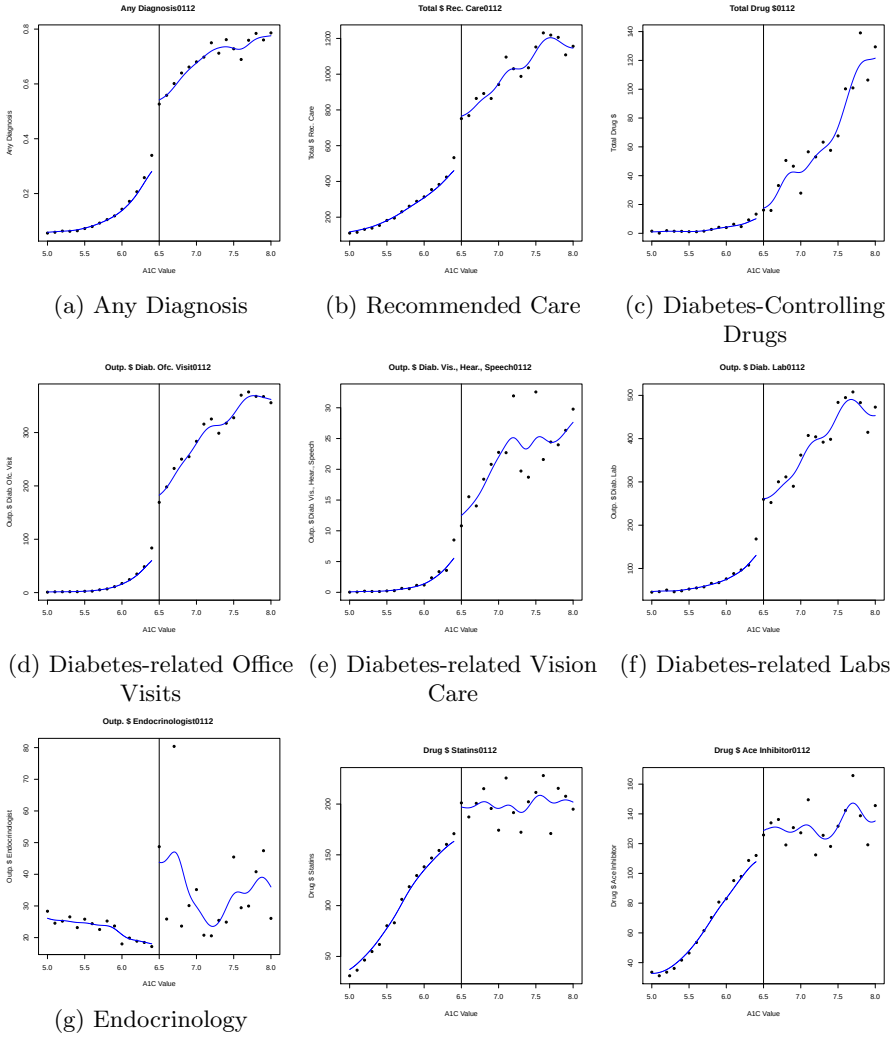
Notes: Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009).

Figure 1: Sample size by A1c Value, Optum and EHR



Source: Authors' calculations from the Optum claims data (left panel) and UCLA EHR data (right panel). Horizontal axis indicates each value of A1c. Vertical axis indicates sample size within each A1c bin.

Figure 2: Effect of Diabetes Diagnosis on Spending in 12 Months Following Initial Lab Test



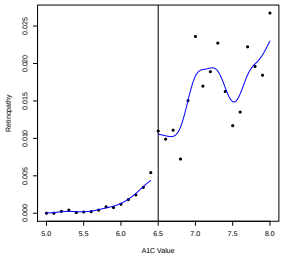
(a) Any Diagnosis (b) Recommended Care (c) Diabetes-Controlling Drugs

(d) Diabetes-related Office Visits (e) Diabetes-related Vision Care (f) Diabetes-related Labs

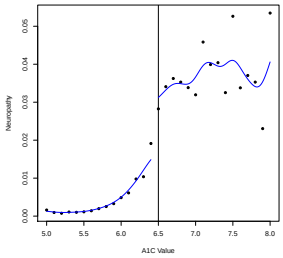
(g) Endocrinology Specialist Care (h) Statins (i) ACE Inhibitors

Source: Authors' calculations from the Optum claims data. Horizontal axis indicates the year relative to the initial lab test. Vertical axis indicates size of the regression discontinuity effect estimated using local linear regression. 95 percent confidence intervals displayed.

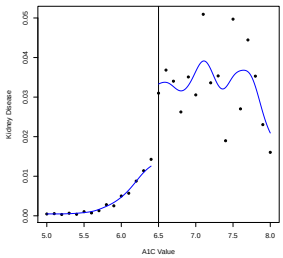
Figure 3: Effect of Diabetes Diagnosis on Diagnosis of Complications Related to Diabetes



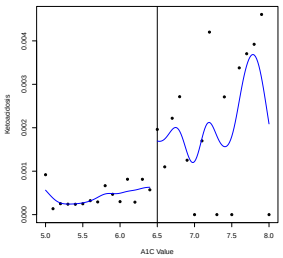
(a) Diabetic Retinopathy



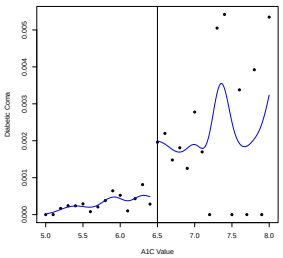
(b) Diabetic Neuropathy



(c) Kidney Disease



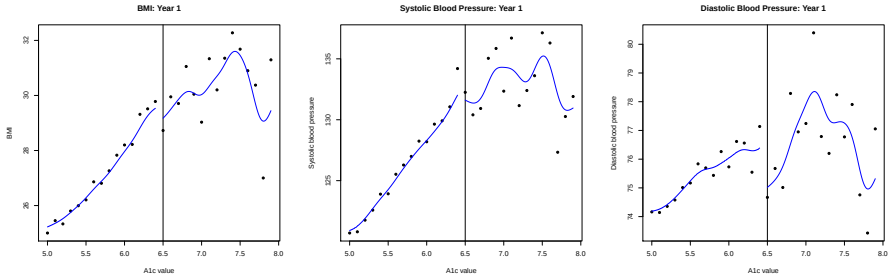
(d) Ketoacidosis



(e) Diabetic Coma

Source: Authors' calculations from the Optum claims data. Horizontal axis indicates the year relative to the initial lab test. Vertical axis indicates size of the regression discontinuity effect estimated using local linear regression. 95 percent confidence intervals displayed.

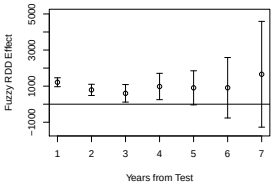
Figure 4: Effect of Diabetes Diagnosis on Clinical Measures of Health, 12 Months After Initial Diagnosis



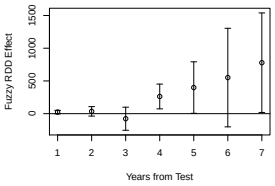
(a) BMI (b) Systolic BP (c) Diastolic BP

Source: Authors' calculations from the Optum claims data. Horizontal axis indicates the year relative to the initial lab test. Vertical axis indicates size of the regression discontinuity effect estimated using local linear regression. 95 percent confidence intervals displayed.

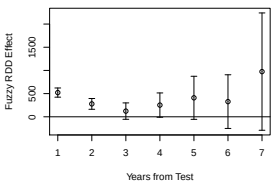
Figure 5: Effect of Diabetes Diagnosis on Spending by Year After Initial Test



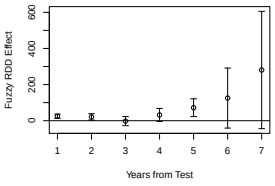
(a) Recommended Care



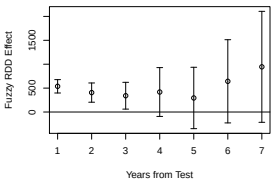
(b) Diabetes-Controlling Drugs



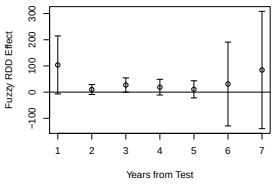
(c) Diabetes-related Office Visits



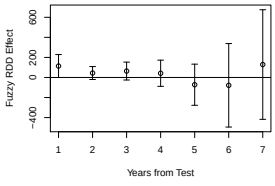
(d) Diabetes-related Vision Care



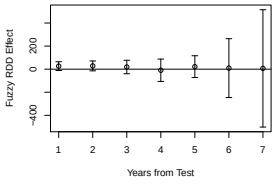
(e) Diabetes-related Labs



(e) Endocrinology Specialist Care



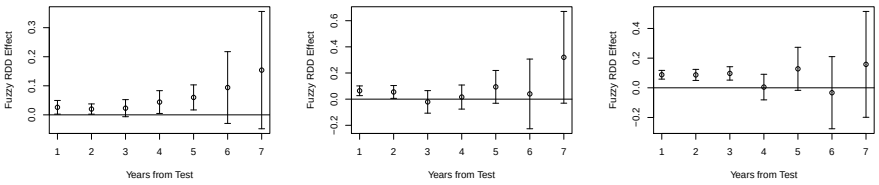
(f) Statins



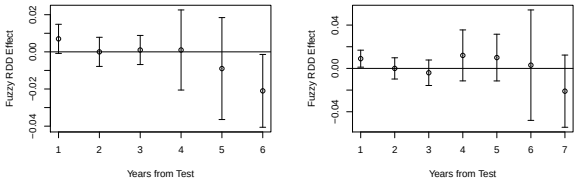
(g) ACE Inhibitors

Source: Authors' calculations from the Optum claims data. Horizontal axis indicates the year relative to the initial lab test. Vertical axis indicates size of the regression discontinuity effect estimated using local linear regression. 95 percent confidence intervals displayed.

Figure 6: Effect of Diabetes Diagnosis on Incidence of Complications by Year After Initial Test



(a) Diabetic Retinopathy (b) Diabetic Neuropathy (c) Kidney Disease



(d) Ketoacidosis (e) Diabetic Coma

Source: Authors' calculations from the Optum claims data. Horizontal axis indicates the year relative to the initial lab test. Vertical axis indicates size of the regression discontinuity effect estimated using local linear regression. 95 percent confidence intervals displayed.