Social Science & Medicine 314 (2022) 115420

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Does health information affect lifestyle behaviours? The impact of a diabetes diagnosis

Alessio Gaggero [a], Joan Gil [b], Dolores Jiménez-Rubio [c,*], Eugenio Zucchelli [d,e,f]

[a] Department of Quantitative Methods, University of Granada, Granada, Spain
[b] Department of Economics and BEAT, Universitat de Barcelona, Spain
[c] Department of Applied Economics, University of Granada, Spain
[d] Madrid Institute for Advanced Study (MIAS) and Department of Economic Analysis, Universidad Autónoma de Madrid (UAM), Spain
[e] Lancaster University, UK
[f] IZA, Germany

ARTICLE INFO

JEL classification:
JEL: C21
I10
I12

Keywords:
Diabetes
Health information
Lifestyle behaviours
Body mass index
Regression discontinuity design
Administrative data

ABSTRACT

Despite an increasing interest in the effect of health information on health-behaviours, evidence on the causal impact of a diagnosis on lifestyle factors is still mixed and does not often account for long-term effects. We explore the role of health information in individual health-related decisions by identifying the causal impact of a type-2 diabetes diagnosis on body mass index (BMI) and lifestyle behaviours. We employ a fuzzy regression discontinuity design (RDD) exploiting the exogenous cut-off value in the diagnosis of type-2 diabetes provided by a biomarker (glycated haemoglobin) drawn from unique administrative longitudinal data from Spain. We find that following a type-2 diabetes diagnosis individuals appear to reduce their weight in the short-term. Differently from previous studies, we also provide evidence of statistically significant long-term impacts of a type-2 diabetes diagnosis on BMI up to three years from the diagnosis. We do not find perceivable effects of a type-2 diabetes diagnosis on quitting smoking or drinking. Overall, health information appears to have a sustained causal impact on weight reduction, a key lifestyle and risk factor among individuals with type-2 diabetes.

All the authors contributed to the design of the study, econometric analysis and drafting of the manuscript, and all of them reviewed and agreed on the final draft.

1. Introduction

The role of health information is at the core of economic models of health investment. Conventional theoretical models assume that individuals have complete information and perfect knowledge of their health capital and make their investment decisions based on corresponding costs and benefits (Grossman, 1972; Becker and Murphy, 1988; Cawley and Ruhm, 2011). Yet, these assumptions have been questioned by a number of empirical studies, including experimental contributions (Bhargava et al., 2017; Kettlewell, 2020) and the latest literature on the relationship between biased health perceptions and risky health-behaviours (Arni et al., 2021). The ability to process health information rationally becomes especially relevant among individuals with chronic conditions. This is because the effectiveness of most health treatments for patients affected by chronic conditions, such as diabetes or hypertension, largely depends on an individual's ability to change health-behaviours as a response to new information acquired through screening tests and diagnoses.

Accordingly, an important question is whether health information would *causally* affect the health-behaviours of individuals with chronic conditions. Although this has been explored in the fields of medicine and economics, evidence appears to be overall inconclusive. More specifically, medical studies tend to focus on either the effects of portable devices (Patel et al., 2015; Jo et al., 2019) or specific cancer diagnoses (Jazieh et al., 2006; Burris et al., 2015) on behavioural changes, providing mixed results often based on standard statistical associations. The economics literature has traditionally examined the impact of different types of broad health information campaigns (Brown and Schrader, 1990; Chern et al., 1995; Kim and Chern, 1999; Roosen et al., 2009; Allais et al., 2015; Fichera and von Hinke, 2020), and only more recently has attempted to identity the causal impact of a *diagnosis* on lifestyle behaviours among patients with different chronic conditions

* Corresponding author.
E-mail address: dolores@ugr.es (D. Jiménez-Rubio).

https://doi.org/10.1016/j.socscimed.2022.115420
Received 25 April 2022; Received in revised form 21 September 2022; Accepted 28 September 2022
Available online 17 October 2022

(Zhao et al., 2013; Kim et al., 2019; Gaggero, 2020; Iizuka et al., 2021). Findings from such literature appear to suggest overall limited impacts, including positive short-term effects only among specific sub-groups of individuals (e.g. older individuals, Gaggero et al., 2020; individuals at high risk of diabetes when information is combined with an invitation for a second visit, Kim et al., 2019) or no perceivable behavioural changes (e.g. Alalouf et al., 2019). In addition, these studies do not often consider potential long-term effects. Hence, whether a diagnosis might have a causal impact on lifestyle behaviours remains an open empirical question.

The main objective of this paper is to explore the role of health information in affecting individual health-related decisions by identifying the causal impact of a type-2 diabetes mellitus (T2DM) diagnosis on weight loss and lifestyle behaviours such as smoking and alcohol consumption. We focus on T2DM as it is among the most widespread chronic conditions worldwide, currently affecting 462 million individuals (6.28% of the world's population); it is the ninth leading cause of death globally; and its burden of disease is projected to raise at an increasingly faster pace in both developing and developed countries (Khan et al., 2020). We employ a fuzzy regression discontinuity design (RDD) on comprehensive longitudinal administrative data from Spain and exploit the exogenous cut-off of an established biomarker used for T2DM diagnoses (glycated haemoglobin, HbA1c). We estimate the impact of a T2DM diagnosis on clinically measured BMI, quitting smoking and drinking, both short-term (within one year from the diagnosis) and long-term (up to three years after the diagnosis).

We find a statistically significant reduction in BMI following a T2DM diagnosis. Specifically, our estimates show that individuals diagnosed with T2DM exhibit a decrease of around 1 point on the BMI scale and this effect appears to be larger among patients with pre-existing comorbidities, those who remain active in the labour market and those who live alone. Differently from the majority of previous studies, we also provide evidence of statistically significant long-term impacts of a T2DM diagnosis on weight loss up to three years since the diagnosis. However, we do not detect any statistically significant causal impacts of a diagnosis on either quitting smoking or drinking. Our results appear to be consistent throughout alternative parametric and non-parametric specifications with varying bandwidths and further robustness checks.

This paper contributes to the literature in several important ways. First, our findings are produced using uniquely suited panel data, including a reliable biomarker used as the running variable in our RDD approach, and several clinically measured anthropometric and health-related variables. Importantly, such data allows for the possibility of identifying long-term effects. Secondly, we focus our empirical exploration on individuals affected by T2DM, one the most prevalent conditions globally. We do so by using a large sample of individuals drawn from Spain, a European country with one of the highest rates of metabolic diseases (Rojo-Martínez et al., 2020). Thirdly, and differently from the large majority of previous studies on the impact of a T2DM diagnosis employing a sharp RDD, our fuzzy RDD approach explicitly accounts for the possibility that a T2DM diagnosis may not be exclusively based on the value of a single biomarker crossing a predetermined threshold, but on a broader set of patients' characteristics. That is, our approach allows to more realistically consider that physicians could inform their diagnoses using a range of information, including family history and the presence of further health conditions. Fourthly, whereas recent studies in the area (e.g. Kim et al., 2019; Gaggero, 2020; Iizuka et al., 2021) tend to use an objective measure of blood sugar (Fasting Plasma Glucose, FPG), such measure is sensitive to short-term lifestyle changes and tend to underestimate diabetes prevalence. In this study, we employ a more reliable biomarker (glycated haemoglobin, HbA1c) which does not suffer from these limitations (American Diabetes Association, 2020). Finally, differently from previous studies on data from the US (Alalouf et al., 2019); Japan (Iizuka et al., 2021); and Korea (Kim et al., 2019), our analysis allows for an exploration of the causal impact of a diabetes diagnosis on BMI and lifestyle factors within a universal health care system where access to health care is entirely free of charge and a detailed follow-up programme for individuals diagnosed with diabetes is offered to all patients.

Overall, this study provides novel causal evidence on the short- and long-term impacts of a T2DM diagnosis by employing a fuzzy RDD on rich longitudinal administrative data. Accordingly, this paper contributes directly to the growing economics literature by providing one of the first empirical explorations of both short- and long-term causal impacts of health information via a diabetes diagnosis on behavioural changes among people with chronic conditions.

2. Relevant literature

Recent medical studies find that the diagnosis of different types of cancers may affect positively the lifestyle (as measured by smoking, drinking, dietary changes and exercise) of a minority of older male individuals (Jazieh et al., 2006) and that a significant proportion of patients diagnosed with lung or head/neck cancer continue smoking (Burris et al., 2015). Yet, such medical evidence appears to mainly focus on cancer diagnoses and is often based on limited cross-sectional samples as well as simple statistical correlations.[1]

The economics literature on health information has mainly concerned its effect on the demand for food. Previous studies explored the effects of either public health information (Brown and Schrader, 1990; Chern et al., 1995; Kim and Chern, 1999; Roosen et al., 2009) or nutritional labels (e.g. Allais et al., 2015; Fichera and von Hinke, 2020) on the demand for fats, oils and the nutritional composition of shopping baskets, with more recent works exploiting casual inference via policy experiments and structural models (e.g. Allais et al., 2015; Fichera and von Hinke, 2020). However, an emerging stream of economic studies have started exploring the causal impact of a *diagnosis* on lifestyle behaviours (Zhao et al., 2013; Kim et al., 2019; Gaggero, 2020; Iizuka et al., 2021; Dai et al., 2022).

Zhao et al. (2013) employ a RDD on data from the China Health and Nutrition Survey (CHNS) to investigate the impact of a hypertension diagnosis on several health outcomes. They find that individuals respond to a hypertension diagnosis by substantially decreasing their fat intake with the effect being especially relevant among wealthier individuals. Dai et al. (2022) recently revisited the effect of a hypertension diagnosis on the same data and their results suggest that individuals may respond differently also depending on the information provided to diagnose their condition (systolic vs diastolic blood pressure). As for diabetes, Kim et al. (2019) use a RDD on data drawn from the Korea's National Health Screening Program on type-2 diabetes mellitus (T2DM); hyperlipidaemia (high cholesterol); and obesity diagnoses. Their findings suggest overall limited effects of disease-related risk information if provided in isolation. However, they also find that individuals classified as high risk for diabetes who are prompted for a second visit exhibit reductions in waist circumference in the short-run.[2] Exploiting a RDD on a sample of individuals aged 50 years and above, Gaggero (2020)

[1] The broader medical literature also includes studies on the effects of workplace wellness programs on health such as Song and Baicker (2019). They exploit a randomised controlled trial (based on nutrition, physical activity, stress reduction and related topics implemented by registered dietitians at the treatment worksites) and find that while the intervention appeared to have improved self-reported health behaviours, it did not have an impact on clinical measures of health nor health care use.

[2] Though not based on a formal diagnosis, Cook (2019) also exploits a RDD to estimate the role of personalised weight information on weight change and finds only a moderate impact concentrated among the "very overweight" with higher income.

provides evidence of a short-term reduction in BMI and waist circumference following a T2DM diagnosis. Finally, Alalouf et al. (2019) and Izuka et al. (2021) are, to the best of our knowledge, the only two causal RDD studies that have attempted to examine short vs long-term impacts of a diagnosis. Alalouf et al. (2019) employ US data including information on HbA1c and find that while a T2DM diagnosis does not appear to have either short- or long-run statistically significant effects on BMI or lifestyle behaviours (apart from short-term improvements in blood sugar), it seems to increase diabetes-related spending. Iizuka et al. (2021) exploit Japanese data and show that following a pre-diabetes diagnosis individuals appear to consume more (preventive) care and improve health outcomes, while a full diabetes diagnosis affects blood sugar levels only among high-risk individuals with hypertension and elevated baseline blood sugar. Yet, a full diabetes diagnosis does not seem to causally change lifestyle behaviours either in the short- or long-run. Hence, results on the causal impact of a diagnosis on lifestyle behaviours appear to be mixed and there is currently no evidence of statistically significant long-term impacts.[3]

3. Data and institutional setting

3.1. Data

We make use of rich individual-level longitudinal data drawn from administrative records of patients followed over seven consecutive years (2004–2010) in six primary care centres and two hospitals in the municipality of Badalona (north-east of Barcelona, Spain).[4] This dataset includes patients aged 16+ who had at least one contact with the health care system between January 1, 2004 and December 31, 2010 and were assigned to one of such health care centres.[5] The dataset includes information about each patient's clinical measurements of height and weight; chronic conditions and other diagnosed diseases, according to the International Classification of Primary Care, second edition (ICPC-2); lifestyles (smoking and drinking); date of admission and discharge; health care use (number of both general practitioners and specialists visits, hospital length of stay); and reason for visit.[6] Moreover, the dataset contains patients' age; gender; employment status (active versus inactive); as well as immigration status (defined by the place of birth, i.e. European Union, EU, versus non-EU) and current residence. While the overall sample includes a total of 123,453 patients, given the purpose of our analysis we restrict our estimating sample initially to 13,971 individuals with at least one HbA1c measurement per year over the period 2004–2010. Importantly, since virtually all individuals in the sample who are at risk of or are diagnosed with type-2 diabetes present at least one biomarker measurement per year, this is unlikely to bias our estimates.[7]

3.2. Key variables

Our analysis focuses on the impact of a T2DM diagnosis on modifiable lifestyle factors (alcohol and tobacco consumption, weight), while accounting for several relevant health-related and sociodemographic observed variables.[8] Our key treatment variable is a binary indicator based on a physician's diagnosis of T2DM reported using the corresponding ICPC-2 code. We couple this information with the one provided by a biomarker, glycated haemoglobin (HbA1c), which provides an accurate measure of glucose concentration over the previous 8–12 weeks (Goldstein et al., 2004; IEC, 2009; Lyons and Basu, 2012). HbA1c measurements are endorsed by the International Expert Committee (IEC) and the American Diabetes Association (ADA) as they appear to be more reliable if compared to other measures of blood sugar such as the ones based on Fasting Plasma Glucose (FPG) (IEC, 20; ADA, 2010; 2020; Ho-Pham et al., 2017). For instance, FPG tests have a limited time validity; are sensitive to short-term lifestyle changes and stress; and have been found to underestimate the prevalence of diabetes.[9]

Our outcomes of interest are body mass index (BMI) and two lifestyle behaviours whose improvement is monitored as part of the standard treatment for T2DM prescribed by physicians: smoking and alcohol consumption.[10] Importantly, our data includes clinically measured BMI, thus avoiding potential issues related to measurement error and reporting bias. Smoking and alcohol consumption are binary variables defining whether an individual is currently a smoker and drinks alcohol, respectively. These are also derived from patients' clinical records and are taken from patients' interviews done by physicians. Since we can only observe smoking and drinking at the extensive margin, our interest is on the impact of a T2DM diagnosis on quitting smoking and drinking.

[3] The broader correlational literature on the effects of a T2DM diagnosis on health-behaviours also appears to be generally inconclusive. For instance, Chong et al. (2017) and Oster (2018) suggest a weak association between a T2DM diagnosis and lifestyle as well as dietary changes (calorie purchases), respectively, whereas Seuring et al. (2020) find some contradicting results among women (decreased BMI but also increased hypertension and reduced physical activity). Relevant to this paper, in one of the very few correlational studies exploring long-term effects, Slade (2012) suggests that the association between a T2DM diagnosis and changes in health-behaviours tend to disappear over time with evidence of recidivism (i.e. weight gain and decreased physical exercise) two years after a diagnosis.

[4] The municipality of Badalona is part of the province of Barcelona and currently includes 223,006 residents (2021, Statistical Institute of Catalonia - Idescat). According to data on the municipality of Badalona drawn from Idescat (https://www.idescat.cat) for 2010 (i.e. the final year used for this analysis) and compared to census data from the same year from the National Statistical Institute (www.ine.es), sociodemographic characteristics of individuals residing in the municipality of Badalona appear to reflect quite closely national averages in terms of age and gender with slight differences in activity rate (65,5% vs 60,2% in Badalona and Spain, respectively) and proportion of immigrants (around 15% vs 12,2%, Badalona and Spain, correspondently). In addition, our sample also appears to broadly resemble key observables of individuals residing in the municipality of Badalona together with a lower proportion of immigrants and a higher proportion of employed individuals (around 9.5%. vs 15% and 71% vs 65.5% in our sample vs the whole municipality, respectively – note that this comparison is based on average values for the whole initial sample vs Idescat data for the reference year 2010). However, we should keep in mind that some differences might be expected as the administrative data used here, as any other administrative data based on health records, are primarily intended to collect information on individuals using the health care service.

[5] We exclude from the analysis patients transferred or moved to other centres and those from other areas or regions. However, since movements across centres are rare events, this is unlikely to have a significant impact on our results.

[6] Although our paper focuses on lifestyle factors and behaviours, we also explored the impact of a type-2 diabetes diagnosis on health care utilisation employing the same fuzzy regression discontinuity design. Results suggest that a type-2 diabetes diagnosis does not appear to have a statistically significant impact on health care use. Estimates are available upon request.

[7] This is because within the Spanish health care system the HbA1c test is used as the main tool to diagnose patients at risk of developing T2DM as well as to monitor patients already diagnosed with T2DM. This test is also administered as part of routine health checks to all individuals presenting relevant risk factors or symptoms of hyperglycaemia.

[8] In this case weight (reductions) would be proxying a series of behaviours related to eating and exercising.

[9] Importantly, although HbA1c was not universally considered the main tool to diagnose diabetes until 2009 (IEC, 2009), the Spanish national health care system was already routinely using this test as one of the key measures to diagnose diabetes, including during the years used for this analysis. Also relevant to this analysis, it is highly unlikely for patients to manipulate HbA1c scores and this is important in the light of our RDD strategy where HbA1c is running variable (see section 4 as well).

[10] BMI is defined as weight divided by the square of height.

3.3. Institutional setting

Our data are drawn from medical records of General Practitioners (GP) practices and hospitals based in Spain, an EU country with a universal health care system free at the point of delivery except for pharmaceuticals, which may require co-payment (Bernal et al., 2018). We exploit the fact that current national health guidelines recommend annual check-ups to adults of any age with at least one risk factor for diabetes in addition to three-year full examinations to all individuals who are 45 years old or older (Mata et al., 2013). In addition and following the national strategy for the management of diabetes (Spanish Ministry of Health, Social Services and Equality, 2012), diagnosed patients are recommended a series of annual follow up visits where weight, height and HbA1c measurements are collected. This normally happens during an initial visit as well as further visits at 6 and 12 months from the diagnosis. Lifestyle changes might also be monitored through further follow up visits (Mata et al., 2013).

In this setting, physicians follow national medical guidelines for managing patients with T2DM where the threshold value of HbA1c $\geq$ 6.5 percent is used as the main criteria for a T2DM diagnosis.[11] However, it is worth noting that realistically medical doctors may not base their diagnosis exclusively on HbA1c values and could look at broader patients' characteristics, including family history around T2DM and whether they present other metabolic conditions, such as hypertension or dyslipidaemia (i.e. an excessive amount of lipids in blood, including triglycerides and cholesterol). Ultimately, this may imply that some patients with several metabolic conditions and a value of HbA1c just below 6.5 percent may be diagnosed with T2DM and thus may also be recommended to start changing some of their health-related behaviours. We account for this in our empirical approach by both employing a fuzzy rather than a sharp RDD, allowing the discontinuity driven by a T2DM diagnosis not to depend only on the value of our running variable (HbA1c), and including GP fixed effects.[12] We also check whether patients might be changing their behaviours before the actual diagnosis by considering different (and lower) HbA1c cut-offs. Note that upon a T2DM diagnosis, patients of the Spanish health care system are normally recommended to follow a non-pharmacological treatment consisting of educational training sessions for diabetes self-management aimed at improving their lifestyle through dietary changes and regular exercise.

4. Econometric methods

4.1. Regression discontinuity design

We exploit the fact that the probability of being diagnosed with T2DM changes discontinuously as a function of the value of the biomarker HbA1c, our running variable, being above the cut-off of 6.5, via a regression discontinuity design (RDD). The main idea behind a RDD is that individuals just above or below the pre-identified cut-off point are otherwise identical but in the probability of being diagnosed, and hence selection bias around the cut-off should disappear (Thistlethwaite and Campbell, 1960; Hahn et al., 2001; Imbens and Lemieux, 2008; Lee and Lemieux, 2010).

In general, there are two types of RDD: sharp versus fuzzy designs (Trochim, 1984). A sharp design assumes that treatment assignment is a deterministic function of the running variable, whereas in a fuzzy design treatment assignment is assumed to be a stochastic function of the running variable. In our case, we would use the sharp design if the T2DM diagnosis would depend solely on the HbA1c being above the pre-determined threshold value of 6.5. However, because a T2DM diagnosis is ultimately at the discretion of the physician, and hence may depend also on other patients' characteristics, in this case a fuzzy design is deemed to be more appropriate.

Formally, let $h_0 = 6.5$ be the cut-off of interest and $h_{i,t} = (HbA1c_{i,t} - h_0)$ the normalised running variable at time t. The corresponding indicator variable $H_{i,t} = 1(h_{i,t} \geq 0)$ defines the treatment assignment rule: it takes the value of 1 for patients with HbA1c scores greater than or equal to the predetermined cut-off level. Finally, let $D_{i,t}$ be a binary indicator taking the value of 1 for individuals diagnosed with T2DM.

A first parameter of interest is the intention-to-treat (ITT) effect, denoted by π, which identifies the effect of being assigned to the treatment ($H_{i,t}$) on the outcome of interest around the cut-off:

$$\pi = E\left[Y_{i,t} \middle| \boldsymbol{X}_{i,t}, h_{i,t} = h_0^+\right] - E\left[Y_{i,t} \middle| \boldsymbol{X}_{i,t}, h_{i,t} = h_0^-\right], \tag{1}$$

where h_0^+ and h_0^- represent, respectively, patients with values of the (normalised) running variable, $h_{i,t}$, just above and below the cut-off point, h_0.

Given the fuzzy nature of our RDD, we rescale the ITT effect (π) by the impact of treatment assignment rule, $H_{i,t}$, on the probability of T2DM diagnosis, $D_{i,t}$, around the cut-off as follows (i.e. the first stage):

$$\beta = \frac{E\left[Y_{i,t} \middle| \boldsymbol{X}_{i,t}, h_{i,t} = h_0^+\right] - E\left[Y_{i,t} \middle| \boldsymbol{X}_{i,t}, h_{i,t} = h_0^-\right]}{E\left[D_{i,t} \middle| \boldsymbol{X}_{i,t}, h_{i,t} = h_0^+\right] - E\left[D_{i,t} \middle| \boldsymbol{X}_{i,t}, h_{i,t} = h_0^-\right]} \tag{2}$$

Here, β is the main parameter of interest which, under the additional assumption of monotonicity, represents the local average treatment effect (LATE) of a T2DM diagnosis among compliers around the cut-off. These are the patients whose T2DM diagnosis is actually affected by the HbA1c value being above or below the cut-off (Hahn et al., 2001). Note that difference between ITT and LATE estimates might be of interest in this case. This is because it would help identifying the impact of the information provided by an actual diabetes diagnosis on BMI and lifestyle behaviours versus the one of simply having an HbA1c score beyond the cut-off value.

In practice, following Hahn et al. (2001) we estimate Equation (2) via a two-stage least-squares (2SLS) estimation using $H_{i,t}$ as an instrument for $D_{i,t}$, as follows:

$$\Delta Y = Y_{i,t+1} - Y_{i,t} = g\left(h_{i,t}\right) + \beta D_{i,t} + \boldsymbol{X}_{i,t}^{'}\gamma + \varepsilon_{i,t}, \tag{3}$$

$$D_{i,t} = f\left(h_{i,t}\right) + \rho H_{i,t} + \lambda H_{i,t} * h_{i,t} + \boldsymbol{X}_{i,t}^{'}\Omega + \varepsilon_{i,t}. \tag{4}$$

As shown in Equation (4), following the recommendations of Gelman and Imbens (2018), our baseline specification is estimated via a linear model including an interaction term between the running variable and the treatment assignment indicator, $H_{i,t}$. This interaction term allows the linear function to vary at the two sides of the cut-off (effectively permitting the function to have different slopes before and after the cut-off). As a robustness check, we estimate both equations including

[11] Note that blood sugar levels ranging between 5.7% $\leq$ HbA1c $\leq$ 6.4% would define a prediabetes condition, i.e. blood glucose levels would be high, but not high enough to prompt a diabetes diagnosis.

[12] Note that mathematically this is analogous to controlling for initial body weight and initial smoking or drinking.

higher order polynomials (between 2 and 4) of the running variable that are also allowed to vary around the cut-off via interactions with $H_{i,t}$. Finally, we also estimate equation (3) non-parametrically using a local randomization approach (Cattaneo et al., 2020). While the parametric approaches described above make use of the full sample, this non-parametric approach focuses on individuals in an arbitrarily small neighbourhood around the cut-off.

As presented in Equation (3), our outcome variables are *changes* in BMI and lifestyle behaviours between time $t + 1$ and t. Accordingly, changes in BMI were coded as $BMI_{t+1} - BMI_t$, while changes in smoking and drinking behaviour were coded as binary variables, taking the value of 1 for those individuals quitting smoking or drinking between time t and $t + 1$.[13] $\boldsymbol{X}_{i,t}$ is a set of pre-diagnosis controls. These include patients' demographic characteristics, such as age and gender; a set of dummy variables to account for employment status (active vs non-active); marital status (living vs not living alone); and immigration status (we include this as there is evidence suggesting systematic differences in lifestyle behaviours among the immigrant populations, e.g. Carrasco-Garrido et al., 2007). We also account for pre-existing health conditions (e.g. hypertension and dyslipidaemia) and time elapsed from the diagnosis.

Additionally, our econometric specifications included time-, health area- and GP fixed-effects (FE). In particular, the inclusion of GP FE allows us controlling for systematic differences across physicians that might affect both a T2DM diagnosis and health outcomes. For example, some GPs may be stricter and diagnose patients with T2DM as soon as the HbA1c approaches the pre-determined cut-off, while others might only formally diagnose T2DM when HbA1c values reach higher levels. Similarly, some physicians may be more cautious about the use of drugs and more prone to recommend lifestyle changes, whereas others may rely mainly on pharmaceutical treatments for the management of T2DM. Our empirical strategy enables accounting for such differences. We cluster standard errors on the running variable based on the recommendation of Lee and Card (2008).[14]

4.2. Validity

RDD methods such as the one employed in this study rely on the standard local continuity assumption typical of all RDD approaches. That is, we are assuming that patients with a HbA1c value just above and below the cut-off should be identical, in terms of both pre-determined observables and unobservables. One indirect way to test for this assumption is to show that the distribution of the observed baseline covariates does not change discontinuously around the cut-off point. Accordingly, Fig. 1 displays a set of local polynomial smoothing (LPS) regressions for pre-diagnosis outcomes of interest and covariates: *(a)* pre-diagnosis BMI *(b)* pre-diagnosis smoking behaviour (c) pre-diagnosis drinking behaviour *(d)* age; *(e)* gender; *(f)* employment status; *(g)* immigration status; and other potential medical conditions, such as *(h)* hypertension, and *(i)* dyslipidaemia.[15] Interestingly, Fig. 1 does not appear to show any significant discontinuity at the cut-off for all variables apart from hypertension, suggesting that the continuity assumption should generally hold in this case. The slight discontinuity around hypertension might simply suggest that patients with hypertension may have already been recommended by their physicians to change their lifestyle even before a potential T2DM diagnosis. This might be expected given the sample of individuals we are focusing on. However, robustness checks show that our main results hold in the absence and presence of individuals with hypertension.[16]

An additional concern for the validity of our identification strategy relates to the potential manipulation of the running variable. McCrary (2008) shows that if individuals can systematically manipulate the value of the running variable in order to (or not to) receive the treatment, then the continuity assumptions will not hold. In our case, this would imply that patients could successfully change their HbA1c level just before the blood test. However, here it seems implausible to assume that patients would be able to manipulate their HbA1c test values, and hence influence a T2DM diagnosis, since the HbA1c value captures the average glucose concentration over the previous 8–12 weeks. Thus, in order to significantly change the value of the biomarker, individuals should have started to substantially change their lifestyle 8–12 weeks prior to the test. Although such behaviour seems unlikely, in Fig. 2 we investigate this issue by examining whether the density function of our running variable is smoothly distributed around the cut-off. As expected, Fig. 2 reveals no evidence of a discontinuity at the (normalised) cut-off.[17]

5. Results

5.1. Descriptive statistics

Table 1 displays summary statistics of the variables used in our empirical analysis. In terms of our main outcomes of interest, the table shows that individuals in the estimating sample present an average BMI of around 30 (widely considered as the standard threshold for obesity); that 18 percent of them are current smokers, while only 3 percent report drinking alcohol.[18] The table also reports that 67 percent of individuals in the sample have been diagnosed with T2DM together with an average of 3.1 years since the onset of the disease. The mean HbA1c value for the patients in our sample is of 6.6 percent, just above the threshold used to diagnose T2DM. With respect to sociodemographic characteristics, Table 1 shows an average age of 65; that 52 percent of the sample consists of women; and that 87, 27, and 2 percent of the individuals in the sample are not living alone; still active in the labour market; and are immigrants (born outside the European Union), respectively. Additionally, the Table reports that 59, 53 and 18 percent of the patients are also diagnosed with hypertension, dyslipidaemia and depression, correspondingly.

5.2. Main results

Fig. 3 presents evidence of the causal (ITT) impact of a T2DM

[13] Following Kolesár and Rothe (2018), we also estimate our models using Eicker-Huber-White (EHW) heteroskedasticity-robust standard errors. These are recommended when the number of support points around the cut-off is sufficiently large and are based on a smaller bandwidth. Results are very similar and available upon request.

[14] Similar graphs can be obtained for the full set of covariates and are available upon request.

[15] In the Appendix, we test the consistency of our results when excluding, respectively, patients with hypertension and dyslipidaemia. See results in Table A5.

[16] A formal McCrary test for the manipulation of the assignment variable confirms this finding, with an estimated log difference in height of $\widehat{\theta} = -0.071$ (s.e. 0.0157), failing to reject the null hypothesis of no discontinuity. Similarly, following Cattaneo et al. (2018) we also run a robust bias-corrected manipulation test based on an unrestricted inference with a local-polynomial of order 2, triangular Kernel, a Jackknife standard error estimator and a data-driven bandwidth selection. A manipulation test score of T = - 0.4614 (p-value of 0.6445) was obtained, indicating no evidence of systematic manipulation of the running variable.

[17] The low prevalence of smoking and drinking in our sample might not be entirely surprising as these are all individuals who have been seeking medical attention at some point during the period considered in this study. In addition, a large proportion of them were diagnosed with medical conditions. Hence, it is likely that these patients have been previously advised to change their lifestyle behaviours.

[18] Figure A1 in the Appendix plots the probability of being diagnosed with T2DM as a function of the (normalised) running variable, also corresponding to the results of the first stage estimation.

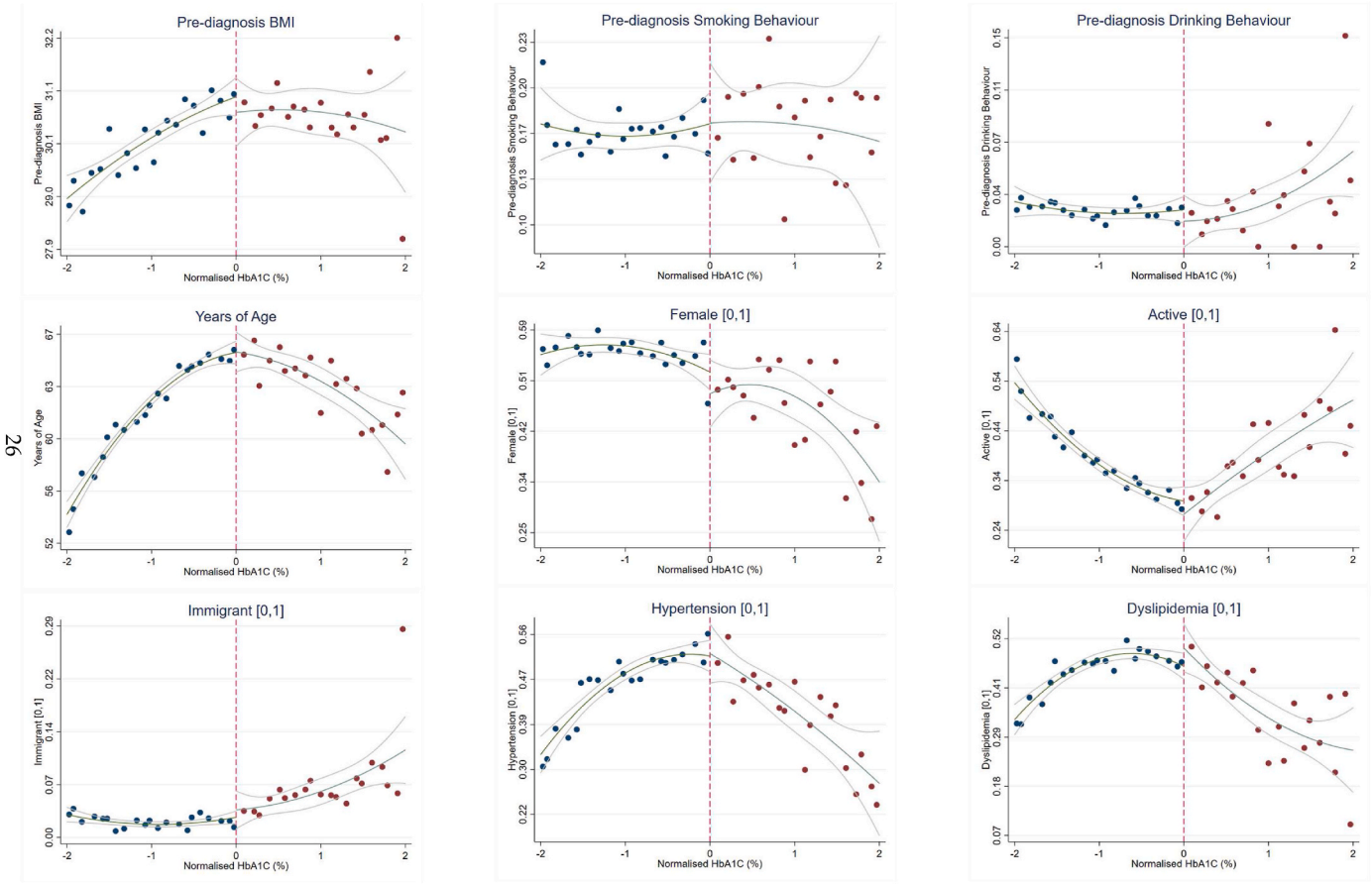

Fig. 1. Continuity test
Note: The Figure shows local polynomial estimates of a number of covariates as a function of the running variable.

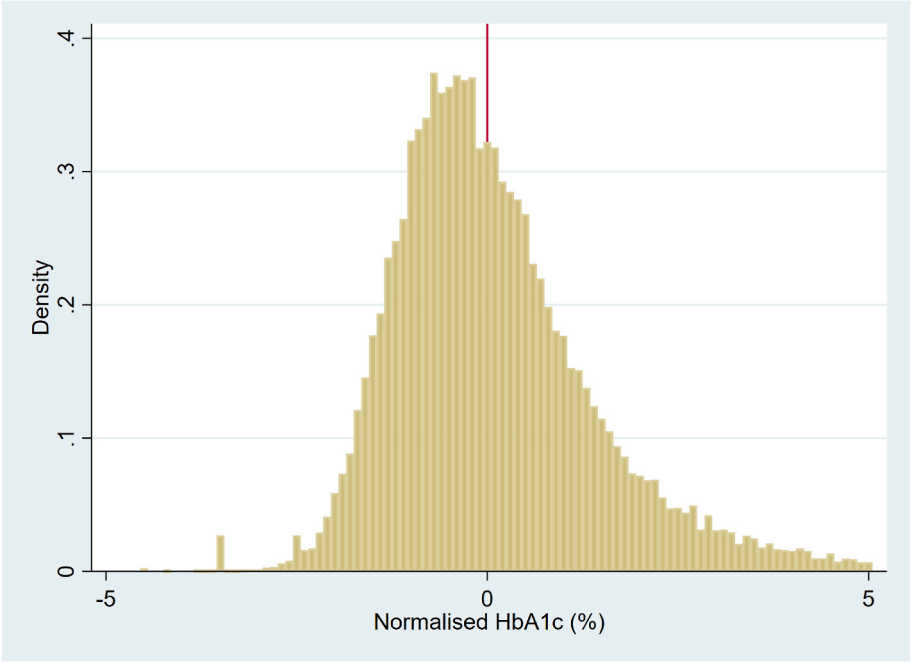

Fig. 2. Density of the running variable
Note: The Figure shows evidence of no manipulation of the running variable. Bin size = 0.1. The bin size has been selected by means of the McCrary test Stata routine, i.e. DCdensity.

Table 1
Summary statistics.

	Mean	S.D.	Min	Max	Obs
Lifestyle Behaviours:					
Body Mass Index	30.29	5.05	14	68	27,920
Smoking [0,1]	0.18	0.38	0	1	39,688
Drinking [0,1]	0.03	0.17	0	1	39,688
T2DM Variables:					
T2DM Diagnosis [0,1]	0.67	0.47	0	1	39,688
Onset of T2DM	3.13	3.72	0	39	34,741
HbA1C (%)	6.60	1.43	0	20	39,994
Demographics:					
Age	65.10	12.62	16	106	39,688
Female [0,1]	0.52	0.50	0	1	39,688
Not Living Alone [0,1]	0.87	0.33	0	1	39,688
Active [0,1]	0.27	0.44	0	1	39,594
Immigrant [0,1]	0.02	0.13	0	1	39,688
Other Conditions:					
Hypertension [0,1]	0.59	0.49	0	1	39,688
Dyslipedimia [0,1]	0.53	0.50	0	1	39,688
Depression [0,1]	0.18	0.38	0	1	39,688
Years:					
2004	0.11	0.31	0	1	39,994
2005	0.12	0.32	0	1	39,994
2006	0.13	0.34	0	1	39,994
2007	0.14	0.35	0	1	39,994
2008	0.15	0.36	0	1	39,994
2009	0.17	0.37	0	1	39,994
2010	0.18	0.39	0	1	39,994
Observations	39,994				

Note: The Table reports summary statistics of the main variables of interest.

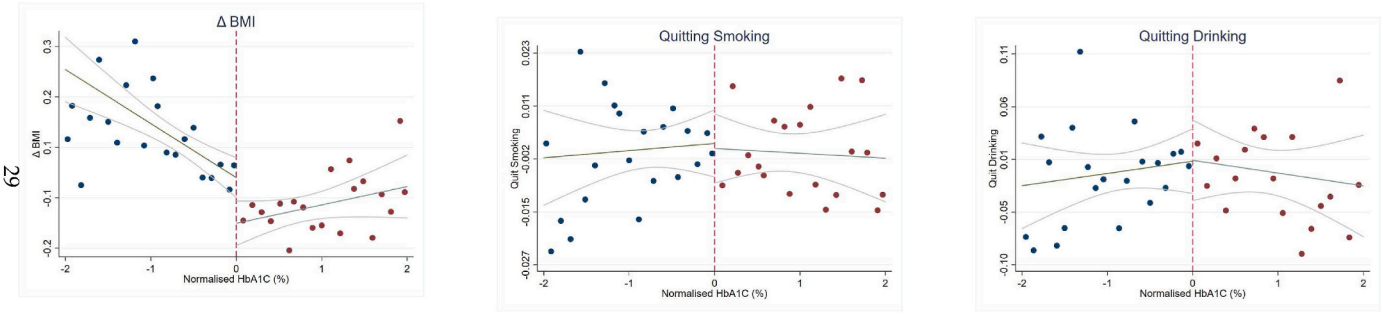

Fig. 3. Graphical Evidence
Note: The figure shows local polynomial estimates of the outcomes of interest at time $t + 1$, as a function of the running variable at time t.

diagnosis by means of a series of RDD plots of the outcomes of interest as a function of the (normalised) HbA1c, our running variable. Interestingly, the Figure exhibits a statistically significant discontinuity at the normalised threshold for the BMI outcome. Specifically, the plots appear to imply that patients with a HbA1c value just above the cut-off show a lower BMI than their counterparts. However, the Figure does not display any significant discontinuity with respect to smoking or drinking cessation. We next test the statistical significance of these findings in a regression framework while controlling for a number of potential confounding factors, as described above.

Table 2 reports the first stage RDD estimates of a T2DM diagnosis using different specifications. Specifically, column (1) includes estimates produced by a model with no covariates; column (2) adds a set of observed covariates; column (3) includes time and area FE and, finally, column (4) adds GP FE (preferred specification). Estimates in all specifications show that the parameter of interest is positive and statistically significant. This implies that individuals with a HbA1c value above the cut-off are significantly more likely to be diagnosed with T2DM.[19]

The main estimates of this paper are reported in Table 3.[20] Column (1) displays estimated coefficients for BMI, while columns (2) and (3) report estimated coefficients for smoking and drinking behaviour, respectively. Panel A shows ITT estimates of $H_{i,t}$ on the outcomes of interest. For instance, the estimated coefficient in column (1) implies that one year after presenting a value of the HbA1C equal to or greater than 6.5, patients report a loss of around 0.088 points on the BMI scale. While this may seem a modest reduction, this might be partly explained by the presence of a large number of non-compliers in our sample (around 24%). These are individuals who were diagnosed (potentially because pre-diabetics) without having actually crossed the HbA1c cut-off value. As a consequence, these were assigned to the control group but they might have been recommended to lose weight. To account for compliance, in Panel B we report the corresponding LATE estimates. In this case, the estimated coefficient in Column (1) implies that individuals diagnosed with T2DM exhibit a statistically significant decrease of 1.020 points in their BMI level.[21]

Interestingly, in our data a T2DM diagnosis does not appear to affect either smoking or drinking behaviour.[22] However, we should keep in mind that individuals in our sample report low baseline percentages of smoking and drinking. This is because the administrative data we exploit only include patients that at some point during the period considered for the analysis were seeking medical attention or were diagnosed with a medical condition. This may imply that our estimates on quitting smoking and drinking should be interpreted with caution as they could

[19] To ease the interpretation of the estimates, results in Table 3 are presented in a compact form but we report the full set of ITT and LATE estimates in Tables A1 and A2 in the Appendix.

[20] As a falsification test, we run the same models on body weight and height separately. As expected, we find a significant and negative impact of a diabetes diagnosis corresponding to a weight loss of around - 2.6 kg, (or 3.3% considering the mean body weight of 77.68 kg of individuals in our sample) but a no statistically significant effect on height (see Table A3 in the Appendix).

[21] We also allowed for the possibility that individuals with BMI measures may not be randomly drawn from our sample of patients and explored the presence of a potential sample selection issue following Wooldridge (2010). Accordingly, we estimated probit models for each time period (wave), where the dependent variable was a dummy taking value 1 for not having a BMI measure and including the full set of exogenous regressors and produced corresponding inverse Mills ratios. These were then included in our RDD models and results produced using such models showed no evidence of selection bias.

[22] Due to the reduced size of the resulting estimating sample, we chose not to rely on heterogenous RDD estimates for quitting alcohol consumption. Results are available upon request.

Table 2
RDD Estimates of a T2DM Diagnosis.

	(1) T2DM Diagnosis	(2) T2DM Diagnosis	(3) T2DM Diagnosis	(4) T2DM Diagnosis
$H_{i,t}$ **[0,1]**	0.196***	0.113***	0.095***	0.090***
	(0.032)	(0.025)	(0.026)	(0.025)
$^{H}i,t * {}^{h}i,t$	−0.236***	−0.157***	−0.185***	−0.183***
	(0.025)	(0.017)	(0.019)	(0.018)
Running Variable:				
$^{h}i,t$	0.247***	0.157***	0.183***	0.181***
	(0.025)	(0.017)	(0.019)	(0.018)
Attributes:				
Age		0.010***	0.008***	0.007***
		(0.001)	(0.001)	(0.001)
Age^2		−0.000***	−0.000***	−0.000***
		(0.000)	(0.000)	(0.000)
Female [0,1]		−0.035***	−0.038***	−0.038***
		(0.005)	(0.004)	(0.004)
Active [0,1]		−0.072***	−0.010*	−0.011**
		(0.006)	(0.005)	(0.005)
Onset of T2DM		0.066***	0,075***	0.075***
		(0.066)	(0.005)	(0.005)
Pre-existing conditions:				
Hypertension [0,1]		0.082***	0.055***	0.052***
		(0.005)	(0.004)	(0.004)
Dyslipedimia [0,1]		0.049***	0.033***	0.034***
		(0.004)	(0.004)	(0.004)
Depression [0,1]		−0.020***	−0.0214**	−0.011***
		(0.005)	(0.005)	(0.005)
Time FE			✓	✓
Area FE			✓	✓
GP FE				✓
Observations	39,688	34,359	34,356	34,319

Note: The Table reports RDD estimates of a T2DM diagnosis. Although not shown in the Table, the estimates are conditional on time, area, and GP fixed effects. Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 3
Fuzzy RDD Estimates of the Impact of a T2DM Diagnosis on Lifestyle Behaviours.

	(1) Δ Body Mass Index	(2) Quitting Smoking	(3) Quitting Drinking
	Panel A: ITT Estimates		
$H_{i,t}$ **[0,1]**	−0.087***	−0.003	−0.001
	(0.032)	(0.004)	(0.017)
Observations	13,852	6245	961
	Panel B: LATE Estimates		
T2DM Diagnosis [0,1]	−1.020***	0.013	−0.011
	(0.163)	(0.014)	(0.073)
Running Variable	✓	✓	✓
Covariates	✓	✓	✓
Time FE	✓	✓	✓
Area FE	✓	✓	✓
GP FE	✓	✓	✓
Observations	13,852	6245	961

Note: The Table reports RDD estimates of the impact of a T2DM diagnosis on lifestyle behaviours. Specifically, in Panel A we report the intention-to-treat (ITT) estimates, and in Panel B we report the local average treatment effect (LATE) estimates. Although not shown in the Table, estimates are conditional on a set of covariates, as described in Section 4, as well as time, areas and GP fixed effects (FE). Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

represent the lower bound of the causal impact of health information. Overall, estimates confirm findings displayed in Fig. 3 implying a statistically significant RDD estimate of a T2DM diagnosis on the one-year change in BMI.

Table 4
Fuzzy RDD – LATE Estimates: Heterogeneity by Health Status.

	(1) **Δ Body Mass Index**	(2) **Quitting Smoking**
	No Pre-existing Conditions:	
T2DM Diagnosis [0,1]	−0.749*	0.032
	(0.447)	(0.026)
Observations	2019	1432
	Hypertension:	
T2DM Diagnosis [0,1]	−0.957***	−0.013
	(0.191)	(0.020)
Observations	9371	3110
	Dyslipidemia:	
T2DM Diagnosis [0,1]	−1.085***	−0.006
	(0.214)	(0.026)
Observations	8062	3523
	Depression:	
T2DM Diagnosis [0,1]	−1.796***	0.007
	(0.587)	(0.039)
Observations	2305	936

Note: The Table reports RDD estimates of the impact of a T2DM diagnosis on lifestyle behaviours. Although not shown in the Table, estimates are conditional on a set of covariates as described in Section 4, as well as time, areas and GP fixed effects (FE). Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

5.3. Heterogeneity analysis

In Table 4, we investigate heterogeneous effects and explore whether and to what extent the short-term impact of a T2DM diagnosis differ by health status by looking at patients with no pre-existing conditions versus those with hypertension, dyslipidaemia and depression.[23] With respect to BMI, our LATE estimates do not show a statistically significant effect of the diagnosis among individuals without pre-existing conditions, however they do display strong and statistically significant effects for those with previous conditions. Interestingly, the effects on BMI are larger for patients previously diagnosed with depression, leading to a reduction of 1.796 points on the BMI scale.

Overall, our findings suggest that, following a T2DM diagnosis, the more substantial BMI reductions occur among individuals with comorbidities. Furthermore, we test for other sources of heterogeneous effects, such as gender (men vs women); broad age groups (being younger or older than 60 years old); employment status (active vs not active); and living arrangements (living vs not living alone) (see Table 5). Results show that overall the estimated effects are similar across the different sub-groups. Yet, the effects of a T2DM diagnosis on BMI appear stronger among those active in the labour market and those living alone. On the other hand, the effect of a T2DM diagnosis on quitting smoking seems to be stronger among those younger than 60 years old.

Finally, Table 6 shows results by initial body weight status. While the estimated effects are generally consistent across different initial weight categories, they seem to be slightly larger among individuals who were initially classified as obese.

5.4. Robustness checks

Our results appear to be robust to a battery of further checks. Table 7 reports RDD LATE estimates obtained using alternative cut-off values of

[23] Corresponding estimates for ITT effects also show non-statistically significant effects for different values of the biomarkers and are available upon request.

Table 5
Fuzzy RDD – LATE Estimates: Heterogeneity by Socio-Demographic Characteristics.

	(1) Δ **Body Mass Index**	(2) **Quitting Smoking**
	Men:	
T2DM Diagnosis [0,1]	−0.907***	0.003
	(0.230)	(0.017)
Observations	6353	4764
	Women:	
T2DM Diagnosis [0,1]	−1.105***	0.011
	(0.230)	(0.022)
Observations	7499	1481
	Age < 60:	
T2DM Diagnosis [0,1]	−1.086***	0.036**
	(0.308)	(0.015)
Observations	3570	3369
	Age ≥60:	
T2DM Diagnosis [0,1]	−0.960***	−0.027
	(0.182)	(0.025)
Observations	10,282	2876
	Active:	
T2DM Diagnosis [0,1]	−1.263***	0.040*
	(0.350)	(0.021)
Observations	2730	2953
	Inactive:	
T2DM Diagnosis [0,1]	−0.924***	−0.023
	(0.194)	(0.019)
Observations	11,122	3292
	Living Alone:	
T2DM Diagnosis [0,1]	−2.667***	0.107
	(0.760)	(0.091)
Observations	1636	656
	Not Living Alone:	
T2DM Diagnosis [0,1]	−0.909***	0.005
	(0.155)	(0.013)
Observations	12,216	5589

Note: The Table reports RDD estimates of the impact of a T2DM diagnosis on lifestyle behaviours. Although not shown in the Table, estimates are conditional on a set of covariates, as described in Section 4, as well as time, areas and GP fixed effects (FE). Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 6
Fuzzy RDD – LATE estimates: Heterogeneity by Weight Category & Lifestyle Behaviour.

	(1) Δ **Body Mass Index**	(2) **Quitting Smoking**
	Healthy Weight:	
T2DM Diagnosis [0,1]	−0.981**	0.021
	(0.390)	(0.037)
Observations	1298	604
	Overweight:	
T2DM Diagnosis [0,1]	−0.882***	0.027
	(0.195)	(0.027)
Observations	5805	1882
	Obese:	
T2DM Diagnosis [0,1]	−1.189***	−0.002
	(0.268)	(0.015)
Observations	6738	3752

Note: The Table reports RDD estimates of the impact of a T2DM diagnosis on lifestyle behaviours. Weight categories are calculated following the standard WHO definition. Although not shown in the Table, estimates are conditional on a set of covariates, as described in Section 4, as well as on time, area and GP fixed effects (FE). Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 7
RDD Estimates of the Impact of a T2DM Diagnosis on Life style Behaviours. Alternative Cut-offs.

	(1) Δ **Body Mass Index**	(2) **Quitting Smoking**	(3) **Quitting Drinking**
	Cut-off: HbA1C = 5		
T2DM Diagnosis [0,1]	−0.535*	0.039	−0.013
	(0.275)	(0.049)	(0.163)
Observations	13,852	6245	961
	Cut-off: HbA1C = 4.5		
T2DM Diagnosis [0,1]	−1.620	0.117	0.038
	(1.619)	(0.083)	(0.196)
Observations	13,852	6245	961
	Cut-off: HbA1C = 4		
T2DM Diagnosis [0,1]	−4.489	0.117**	0.300
	(5.317)	(0.059)	(0.183)
Observations	13,852	6245	961

Note: The Table reports RDD estimates on the outcomes of interest, when using a placebo cut-off, for which no effects should be observed. Although not shown in the Table, estimates are conditional on a set of covariates, time, areas and GP fixed effects (FE). Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

the biomarker (4, 4.5, and 5 percent). Our findings confirm that when using HbA1c cut-offs of 4 and 4.5 percent, estimated effects are not statistically significant.[24] In addition, when using a cut-off value of 5 percent results appear to be weakly statistically significant. This is consistent with medical guidelines and suggests that GPs could advise patients to change their lifestyle before the actual diabetes diagnosis, especially those patients whose blood sugar level is increasing and are presenting other risk factors related to diabetes. However, the corresponding effect is only marginally significant (at 10 percent significance level) and half the size if compared to the one identified in our RDD baseline model.[25]

Table 8 reports fuzzy RDD LATE estimates obtained using a non-parametric local randomization approach focusing on a small neighbourhood around the cut-off (Cattaneo et al., 2020). For comparison purposes, Panel A of Table 8 presents the benchmark LATE RDD estimates obtained using the parametric approach. In Panels B to F, we report a series of non-parametric LATE RDD estimates using only observations within the intervals defined at varying bandwidths around the HbA1c threshold of 6.5 percent. The findings appear to confirm a statistically significant decrease in BMI following a T2DM diagnosis (the smaller the bandwidth, the closer the estimated coefficients to the baseline estimates). No causal short-term impact is found for either smoking or drinking cessation.[26]

Furthermore, Table 9 shows results obtained by estimating the fuzzy RDD parametrically using different polynomial orders, ranging from a polynomial of order 1 (Panel A or baseline estimate) to a polynomial of order 4 (Panel D). Once again, estimates allowing for different flexibility

[24] The table also reports a statistically significant coefficient for quitting smoking at the 4% threshold. An exploration of our data as well as of clinical guidelines (especially those related to the health treatments corresponding to different levels of blood sugar) appears to suggest that this might be a spurious correlation due to the low sample size. However, we cannot exclude this might be the result of systematic doctors' advice that might not be reported in our data or suggested by clinical guidelines.

[25] These estimates were obtained using a non-parametric fuzzy RDD estimation method based on the Stata command rdrandinf (Cattaneo et al., 2020), assuming untransformed outcomes, a uniform kernel function to weight the observations and fixed margins or a complete randomization mechanism.

[26] Although issues around non-random heaping in the running variable might not be a concern in our study (because of the accuracy in the measurement of the biomarker), we also run a robustness check around this following Barreca et al. (2011, 2016). As expected, an exploration of the distribution of the running variable reveals no evidence of heaping in the data (corresponding graphs are available upon request).

Table 8
Fuzzy RDD Estimates. Non parametric Approach.

	(1) Δ Body Mass Index	(2) Quitting Smoking	(3) Quitting Drinking
Panel A: Baseline			
T2DM Diagnosis [0,1]	−1.002***	0.010	−0.010
	(0.162)	(0.014)	(0.074)
Observations	13,852	6245	961
Panel B: Bandwidth = 2			
T2DM Diagnosis [0,1]	−0.687***	−0.008	−0.019
	(0.184)	(0.0129)	(0.067)
Observations	12,621	5404	824
Panel C: Bandwidth = 1.75			
T2DM Diagnosis [0,1]	−0.722***	−0.008	−0.007
	(0.194)	(0.016)	(0.006)
Observations	12,115	5075	765
Panel D: Bandwidth = 1.5			
T2DM Diagnosis [0,1]	−0.737***	−0.008	−0.019
	(0.198)	(0.020)	(0.040)
Observations	11,546	4778	709
Panel E: Bandwidth = 1.25			
T2DM Diagnosis [0,1]	−0.870***	−0.004	0.014
	(0.233)	(0.009)	(0.017)
Observations	10,298	4188	618
Panel F: Bandwidth = 1			
T2DM Diagnosis [0,1]	−0.930**	0.004	0.021
T2DM Diagnosis [0,1]	(0.250)	(0.005)	(0.032)
Observations	9258	3682	543

Note: The Table reports non-parametric RDD estimates of the impact of a T2DM diagnosis on lifestyle behaviours. Estimates are obtained using STATA routine recommended by Cattaneo et al. (2020) for the case of discrete running variable, namely rdrandinf. Each coefficient in the table report the effect of be-ing diagnosed with T2DM on lifestyle behaviours. Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 9
Fuzzy RDD Estimates. Different Polynomials.

	(1) Δ Body Mass Index	(2) Quitting Smoking	(3) Quitting Drinking
Panel A: Baseline			
T2DM Diagnosis [0,1]	−1.020***	0.013	−0.011
	(0.163)	(0.014)	(0.073)
Observations	13,852	6245	961
Panel B: Quadratic			
T2DM Diagnosis [0,1]	−1.056***	−0.002	−0.076
	(0.244)	(0.016)	(0.113)
Observations	13,852	6245	961
Panel C: Cubic			
T2DM Diagnosis [0,1]	−0.719***	−0.003	−0.039
	(0.247)	(0.016)	(0.104)
Observations	13,852	6245	961
Panel D: Quartic			
T2DM Diagnosis [0,1]	−0.693***	−0.014	0.011
	(0.239)	(0.020)	(0.112)
Observations	13,852	6245	961

Note: The Table reports RDD estimates of the impact of a T2DM diagnosis on lifestyle behaviours, when considering different polynomials of the running variable. Although not shown in the Table, estimates are conditional on a set of covariates, as described in Section 4, as well as time, areas and GP fixed effects (FE). Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

of the running variable are qualitatively the same.[27]

A potential concern regarding our results is whether part of the weight loss identified by our models might be driven by the effects of medications. To investigate this issue in the absence of data on the use of medications, we exploit information about the time elapsed since a T2DM diagnosis and estimate Equations (3) and (4) by restricting the sample to recently diagnosed patients, i.e., including only those diagnosed within the previous year. This is because within the Spanish system diagnosed patients are first recommended to follow a specific diet and an exercise regimen and only when these do not seem to be effective (and in any case not before a period between 3 and 6 months), they are eventually recommended the use of medications (specifically, metformin). Hence, the effects identified in the sample of recently diagnosed individuals are less likely to be affected by a potential pharmaceutical treatment. Table 10 reports estimates of this exercise and shows that the estimated LATE RDD coefficients for BMI are still statistically significant and similar to those reported in Table 3. Moreover, evidence shows that metformin (the medication eventually recommended) does not appear to cause weight losses (its effect is mostly neutral and/or limited, ADA, 2022). Thus, although we cannot categorically exclude that part of the impact we identify on BMI may be driven by pharmaceutical treatments, based on our robustness checks, healthcare guidelines and available evidence, it might be reasonable to assume our results should not be greatly affected by medications. Moreover, the estimated effects by weight categories shown in Table 6 appear stronger for those who were initially obese, also suggesting that the weight loss observed after a T2DM diagnosis is more likely to be affected by behavioural changes.

As an additional robustness check, we produce estimates excluding patients with hypertension and dyslipidaemia and allowing for heterogenous treatment effects (see Tables A4-5 in the Appendix).

5.5. Long-term impacts of a T2DM diagnosis

Another advantage brought by the use of our longitudinal administrative dataset is that it allows examining both short- and long-term impacts of a T2DM diagnosis. Hence, we also investigate whether the observed reductions in BMI following a T2DM diagnosis are sustained over time, namely 1, 2, and 3 years after the diagnosis. Accordingly, the estimation of this model implies the use of at least four consecutive BMI measurements.

Results reported in Table 11 suggest that the estimated BMI reductions after a T2DM diagnosis are observed beyond year 1 and sustained over time. More specifically, we find statistically significant decreases of 1.169, 1.519, and 2.207 on the BMI scale for 1, 2, and 3 years after the T2DM diagnosis, respectively. This appears to imply that the health information included in a T2DM diagnosis might not just have a statistically significant casual impact in the short-term but that it can also have a longer-term impact on BMI reduction. Such sustained causal impact also appears to increase over time with the largest reduction in BMI occurring three years after the diagnosis. This may be the result of a cumulated effect driven by consistent weight losses throughout the years for diagnosed patients. [28]

6. Conclusions

We exploit the exogenous cut-off of blood sugar in the diagnosis of T2DM provided by a reliable and accurate biomarker (HbA1c) via a fuzzy RDD and measure its causal impact on objectively measured BMI, quitting smoking and quitting drinking. This builds on and directly extends the emerging economics literature on the causal effects of health information on lifestyle behaviours by focusing on T2DM, a major public

[27] Note that we estimate long-term effects on a sample of recently diagnosed individuals. We do so as the identification of the long-term impacts on lifestyle behaviours using the more general sample of patients (including individuals with longer time elapsed from a T2DM diagnosis) might be confounded by other factors including other reasons/treatments for losing weight, that we might not be able to fully account for. Accordingly, this might have an impact on their weight loss over and above the one of a T2DM diagnosis. Corresponding estimates obtained using the more general sample appears to confirm this as they are quantitatively larger (these are available upon request).

Table 10
Fuzzy RDD Estimates. Recently Diagnosed Individuals.

	(1) Δ Body Mass Index	(2) Quitting Smoking	(3) Quitting Drinking
T2DM Diagnosis	−0.923**	0.089	−0.073
	(0.372)	(0.070)	(0.076)
	Running Variable:		
HbA1c (%)	0.028	−0.005	0.002
	(0.052)	(0.006)	(0.004)
	Attributes:		
Age	0.065***	0.000	−0.004
	(0.024)	(0.002)	(0.007)
Age^2	−0.001***	−0.000	0.000
	(0.000)	(0.000)	(0.000)
Female [0,1]	−0.131**	0.003	−0.011
	(0.052)	(0.006)	(0.028)
Not Living Alone [0,1]	−0.009	−0.000	0.029
	(0.095)	(0.008)	(0.029)
Active [0,1]	−0.026	−0.007	−0.019
	(0.092)	(0.005)	(0.026)
Immigrant [0,1l	0.377	0.017	−0.500
	(0.336)	(0.031)	(0.365)
	Pre-existing Conditions:		
Hypertension [0,1]	−0.004	−0.003	0.036**
	(0.061)	(0.006)	(0.017)
Dyslipedimia [0,1]	−0.010	0.002	0.011
	(0.051)	(0.005)	(0.015)
Depression [0,1]	−0.025	−0.006	0.056
	(0.068)	(0.005)	(0.042)
Observations	3266	2374	364

Note: The Table reports RDD estimates of the impact of a T2DM diagnosis on lifestyle behaviours, for the sample of individuals not diagnosed at time *t*. Although not shown in the Table, estimates are conditional on a set of covariates, time, areas and GP fixed effects (FE). Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 11
Fuzzy RDD Estimates of Body Mass Index. Long term Effects.

	(1) 1 Year After Diagnosis	(2) 2 Years After Diagnosis	(3) 3 Years After Diagnosis
T2DM Diagnosis	−1.169**	−1.519**	−2.207***
	(0.521)	(0.622)	(0.805)
Observations	1535	1210	899

Note: The Table reports RDD estimates of the long-term effect of a T2DM diagnosis on body mass index, for the sample of individuals not diagnosed at time *t*. Although not shown in the Table, estimates are conditional on a set of covariates, time, areas and GP fixed effects (FE). Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

health concern whose prevalence is projected to increase worldwide. Importantly, we make use of comprehensive longitudinal medical records from Spain, including several diagnosed health conditions and relevant sociodemographic variables and allowing the identification of the long-term impacts of a diagnosis.

We find a short-term reduction in body weight as a response to a T2DM diagnosis of around 1 point on the BMI scale with stronger effects among patients with pre-existing health conditions such as hypertension, depression or obesity. Larger estimates for these individuals may be expected as they could be more reactive to avoid potentially severe consequences. Slightly stronger effects are also found for individuals still active in the labour market and those living alone. Differently from most previous studies, we also provide evidence of statistically significant long-term impacts of a T2DM diagnosis on BMI up to three years since the diagnosis. Our results are confirmed by a series of robustness checks, including parametric estimations with different polynomials, as well as non-parametric estimations with several bandwidths. Our findings differ from the ones of recent studies (e.g. Alalouf et al., 2019; Kim et al., 2019; Iizuka et al., 2021) as they might be partly driven by older individuals with diabetes-related health conditions. Moreover, they could also reflect institutional incentives about adherence to treatments and monitoring diabetes within a typical European health care system.

Overall, our study provides causal evidence suggesting that health information, in the form of a T2DM diagnosis, may have a causal impact on weight reduction both in the short- and long-run. This appears to be consistent with the ability of individuals to correctly process health information when making health-related decisions. However, the same information provided by a T2DM diagnosis does not seem to causally affect smoking or drinking. Yet, this might not be entirely surprising as weight management, through a low carbohydrate diet and an exercise regimen, is the main recommended strategy for an adequate glycaemic control among patients with initial stage T2DM (Mata et al., 2013; NICE, 2020).

Furthermore, we should keep in mind that our data are not capable of capturing the intensive margin of smoking and drinking, and so we cannot observe whether individuals diagnosed with T2DM would eventually *reduce* the daily number of cigarettes smoked or units of alcohol. More generally, it is also worth noting that we employ administrative data including individuals seeking medical advice at some point during the period examined in our analysis. As such, some of these patients present already low baseline prevalence of smoking and (especially) drinking as they might have been already advised to quitting smoking and drinking before a potential T2DM diagnosis. Thus, our estimates on lifestyle behaviours might represent the lower bound of the true impact of a T2DM diagnosis on these behaviours and should be considered with caution.

From a policy perspective, the institutional setting provided by the Spanish National Health Service recommending constant follow-up visits to diagnosed patients, suggests that implementing strategies facilitating frequent contacts between patients and doctors, could promote adherence to health treatments and secure sustained health improvements. Finally, since a significant proportion of individuals affected by T2DM are *undiagnosed* (e.g. up to 40 percent of individuals with diabetes in Europe; Rojo-Martínez et al., 2020), promoting frequent check-ups among individuals at risk of diabetes, could help providing timely health information to patients and reduce avoidable health care costs (WHO, 2016; Liang et al., 2019).

Our study has some potential limitations. Firstly, although our data includes a measure of employment status, as in most administrative datasets, specific information on education and income is not directly available, therefore limiting the scope of our heterogeneity analysis. Secondly, the lack of information on the specific medications prescribed to patients, prevents us from controlling for their potential effects, although these are expected to be limited in this case. Despite such limitations, our paper offers novel causal evidence on the impact of the health information provided by a T2DM diagnosis on body weight and key lifestyle behaviours both in the short-term and over time.

Data availability

The authors do not have permission to share data

Acknowledgments

we gratefully acknowledge financial support from the Ministry of Science, Innovation and Universities (grant number PID2019-105688RB-I00). Joan Gil is supported by the Ministry of Economics and Competitiveness (grant number PID2020-114040RB-100). Eugenio Zucchelli also acknowledges support from the Tomás y Valiente Fellowship funded by the Madrid Institute for Advanced Study (MIAS), Universidad Autónoma de Madrid, and grants SI1/PJI/2019-00326 and H2019/HUM-5793 funded by the Community of Madrid.

Appendix

To explore whether some patients, especially those with known risk factors leading to T2DM, may be advised to reduce their weight and change their health-behaviours before a potential T2DM diagnosis, we explore the consistency of our estimates when excluding, respectively, patients with hypertension and dyslipidaemia. Our findings (Table A4) are in line with baseline estimates and corroborate the substantial causal effect of health information, in the form of a T2DM diagnosis on weight reduction. Finally, to allow for heterogeneous treatment effects, we implemented a multiple cut-off approach (van der Klaauw, 2002), and included in our regression a series of indicators corresponding to the cut-offs for pre-diabetes ($HbA1c_{i,t} \geq 5.7$) and uncontrolled diabetes ($HbA1c_{i,t} \geq 7$). As shown in Table A5, results confirm the relevance of the information brought about by an actual diagnosis in affecting weight loss.

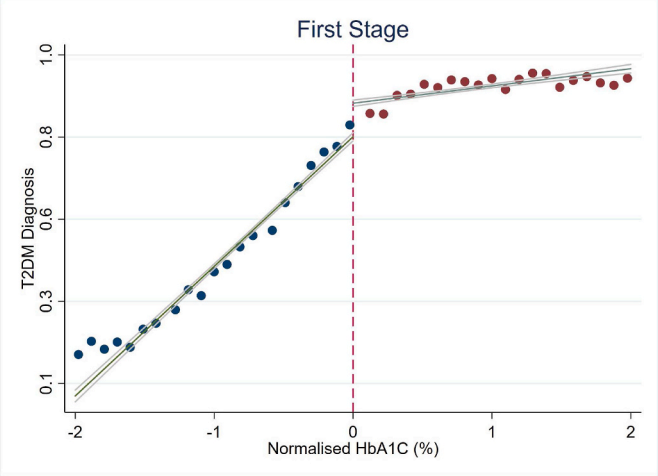

Fig. A.1. First Stage
Note: The figure shows local polynomial estimates of the probability of being diagnosed with T2DM at time $t + 1$, as a function of the running variable at time t.

Table A.1
ITT Estimates

	(1) Δ Body Mass Index	(2) Quitting Smoking	(3) Quitting Drinking
Above [0,1]	−0.087***	−0.003	−0.001
	(0.032)	(0.004)	(0.017)
Above * HbA1c	0.195***	−0.005*	0.002
	(0.035)	(0.003)	(0.014)
	Running Variable:		
HbA1c (%)	−0.164***	0.005*	−0.001
	(0.029)	(0.003)	(0.013)
	Attributes:		
Age	0.046**	−0.003***	0.001
	(0.020)	(0.001)	(0.008)
Age^2	−0.000***	0.000***	−0.000
	(0.000)	(0.000)	(0.000)
Female [0,1]	−0.084***	−0.009**	−0.021
	(0.022)	(0.004)	(0.025)
Not Living Alone [0,1]	−0.035	0.003	0.015
	(0.034)	(0.005)	(0.028)
Active [0,1]	0.007	0.001	0.002
	(0.048)	(0.003)	(0.026)
Immigrant [0,1]	0.410**	−0.013	−0.302
	(0.161)	(0.017)	(0.213)
Onset of T2DM	0.016***	−0.001	0.002
	(0.004)	(0.000)	(0.002)
	Pre-existing Conditions:		
Hypertension [0,1]	−0.034	−0.003	0.033*
	(0.024)	(0.003)	(0.019)
Dyslipedimia [0,1]			

(*continued on next page*)

Table A.1 (*continued*)

	(1) Δ Body Mass Index	(2) Quitting Smoking	(3) Quitting Drinking
	0.038*	−0.001	−0.003
	(0.023)	(0.004)	(0.015)
Depression [0,1]	0.033	0.002	0.054*
	(0.036)	(0.004)	(0.028)
Observations	13,852	6245	961

Note: The Table reports intention-to-treat (ITT) estimates of a T2DM diagnosis. Although not shown in the Table, the estimates are conditional on time, area, and GP fixed effects. Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table A.2
LATE Estimates

	(1) Δ Body Mass Index	(2) Quitting Smoking	(3) Quitting Drinking
T2DM Diagnosis [0,1]	−1.020***	0.013	−0.011
	(0.163)	(0.014)	(0.073)
	Running Variable:		
HbA1c (%)	0.012	−0.000	0.001
	(0.017)	(0.001)	(0.006)
	Attributes:		
Age	0.054***	−0.003***	Am
	(0.020)	(0.001)	(0.008)
Age^2	−0.000***	0.000***	−0.000
	(0.000)	(0.000)	(0.000)
Female [0,1]	−0.114***	−0.008*	−0.023
	(0.020)	(0.004)	(0.027)
Not Living Alone [0,1]	−0.069*	0.003	0.014
	(0.037)	(0.005)	(0.028)
Active [0,1]	0.008	0.001	0.002
	(0.047)	(0.003)	(0.026)
Immigrant [0,1]	0.411**	−0.013	−0.301
	(0.179)	(0.017)	(0.211)
Onset of T2DM	0.075***	−0.001	0.003
	(0.012)	(0.001)	(0.008)
	Pre-existing Conditions:		
Hypertension [0,1]	−0.005	−0.003	0.033*
	(0.027)	(0.003)	(0.018)
Dyslipedimia [0,1]	0.066***	−0.001	−0.003
	(0.022)	(0.004)	(0.016)
Depression [0,1]	0.016	0.002	0.055*
	(0.036)	(0.004)	(0.030)
Observations	13,852	6245	961

Note: The Table reports local average treatment effect (LATE) estimates of a T2DM diagnosis. Although not shown in the Table, the estimates are conditional on time, area, and GP fixed effects. Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table A.3
Fuzzy RDD. LATE Estimates for Body Weight and Height

	(1) Δ Body Weight (Kg)	(2) Δ Body Height (cm)
T2DM Diagnosis [0,1]	−2.612***	−0.049
	(0.410)	(0.086)
	Running Variable:	
HbA1c (%)	0.036	0.009
	(0.042)	(0.007)
	Attributes:	
Age	0.146**	−0.010
	(0.062)	(0.008)
Age^2	−0.001***	0.000
	(0.000)	(0.000)
Female [0,1]	−0.211***	0.027**
	(0.052)	(0.012)
Not Living Alone [0,1]	−0.198**	−0.018
	(0.091)	(0.019)
Active [0,1]	0.069	−0.005
	(0.124)	(0.017)
Immigrant [0,1]	1.090**	0.037
	(0.447)	(0.098)
Onset of T2DM	0.194***	0.007
	(0.032)	(0.005)
	Pre-existing Conditions:	
Hypertension [0,1]	−0.024	−0.002
	(0.072)	(0.016)

(*continued on next page*)

Table A.3 (*continued*)

	(1) Δ Body Weight (Kg)	(2) Δ Body Height (cm)
Dyslipedimia [0,1]	0.172***	0.008
	(0.055)	(0.013)
Depression [0,1]	−0.002	−0.030*
	(0.087)	(0.018)
Observations	13,886	13,855

Table A.4
Fuzzy RDD Estimates of the Impact of a T2DM Diagnosis on Life Style Behaviours

	(1) Δ Body Mass Index	(2) Quitting Smoking
	Panel A: Benchmark	
T2DM Diagnosis [0,1]	−1.020***	0.013
	(0.163)	(0.014)
Observations	13,852	6245
	Panel B: Excluding Hypertension = 1	
T2DM Diagnosis [0,1]	−1.027***	0.041*
	(0.291)	(0.023)
Observations	4481	3135
	Panel C: Excluding Dyslipidemia = 1	
T2DM Diagnosis [0,1]	−1.017***	0.041**
	(0.249)	(0.017)
Observations	5790	2722

Note: The Table reports RDD estimates of the impact of a T2DM diagnosis on lifestyle behaviours. Although not shown in the Table, estimates are conditional on a set of covariates, as described in Section 4, as well as time, areas and GP fixed effects (FE). Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table A.5
Fuzzy RDD Estimates of the Impact of a T2DM Diagnosis on Lifestyle Behaviours

	(1) Δ **Body Mass Index**	(2) **Quitting Smoking**	(3) **Quitting Drinking**
T2DM Diagnosis [0,1]	−1.276***	0.018	−0.023
	(0.328)	(0.021)	(0.109)
Pre-Diabetes [HbA1c ≥ 5.7]	0.082	−0.003	0.004
	(0.076)	(0.007)	(0.028)
Uncontrolled Diabetes [HbA1c ≥ 7]	0.005	0.001	−0.001
	(0.053)	(0.005)	(0.024)
	Running Variable:		
HbA1c (%)	0.011	−0.000	0.001
	(0.028)	(0.001)	(0.008)
	Attributes:		
Age	0.057***	−0.003***	0.002
	(0.020)	(0.001)	(0.008)
Age^2	−0.000***	0.000***	−0.000
	(0.000)	(0.000)	(0.000)
Female [0,1]	−0.122***	−0.008*	−0.024
	(0.023)	(0.004)	(0.027)
Not Living Alone [0,1]	−0.076**	0.003	0.013
	(0.038)	(0.005)	(0.027)
Active [0,1]	0.009	0.001	0.002
	(0.047)	(0.003)	(0.026)
Immigrant [0,1l	0.408**	−0.012	−0.299
	(0.183)	(0.017)	(0.210)
Onset of T2DM	0.089***	−0.002	0.004
	(0.021)	(0.001)	(0.011)
	Pre-existing Conditions:		
Hypertension [0,1]	0.002	−0.003	0.033*
	(0.030)	(0.003)	(0.019)
Dyslipedimia [0,1]	0.072***	−0.001	−0.002
	(0.025)	(0.004)	(0.017)
Depression [0,1]	0.011	0.002	0.056*
	(0.037)	(0.004)	(0.029)
Observations	13,852	6245	961

Note: The Table reports RDD estimates of the impact of a T2DM diagnosis on lifestyle behaviours. Although not shown in the Table, estimates are conditional on a set of covariates, as described in Section 4, as well as time, areas and GP fixed effects (FE). Robust standard errors are clustered on the running variable. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

References

Alalouf, M., Miller, S., Wherry, L.R., 2019. What Difference Does a Diagnosis Make? Evidence from Marginal Patients (No. W26363). National Bureau of Economic Research.

Allais, O., Etilé, F., Lecocq, S., 2015. Mandatory labels, taxes and market forces: an empirical evaluation of fat policies. J. Health Econ. 43, 27–44.

American Diabetes Association (ADA), 2020. Standards of medical care in diabetes. Diabetes Care 3 (1).

American Diabetes Association Professional Practice Committee (ADA 2022). Pharmacologic Approaches to Glycemic Treatment: Standards of Medical Care in Diabetes— Diabetes Care Volume 45, Suppl. 1, January 2022.

Arni, P., Dragone, D., Goette, L., Ziebarth, N.R., 2021. Biased health perceptions and risky health behaviors—theory and evidence. J. Health Econ. 76, 102425.

Barreca, A.I., Guldi, M., Lindo, J.M., Waddell, G.R., 2011. Saving babies? Revisiting the effects of very low birth weight classification. Q. J. Econ. 126 (4), 2217, 2123.

Barreca, A.I., Lindo, J.M., Waddell, G.R., 2016. Heaping-induced bias in regression-discontinuity designs. Econ. Inq. 54 (1), 268–293.

Becker, G.S., Murphy, K.M., 1988. A theory of rational addiction. J. Polit. Econ. 96 (4), 675–700.

Bernal-Delgado, E., Garcia-Armesto, S., Oliva, J., Sanchez Martinez, F.I., Repullo, J.R., Pena-Longobardo, L.M., Ridao-Lopez, M., Hernandez-Quevedo, C., 2018. Spain: health system review. Health Syst Transit 20 (2), 1–179. PMID: 30277216 2018 May.

Bhargava, S., Loewenstein, G., Sydnor, J., 2017. Choose to lose: health plan choices from a menu with dominated option. Q. J. Econ. 132 (3), 1319–1372.

Brown, D.J., Schrader, L.F., 1990. Cholesterol information and shell egg consumption. Am. J. Agric. Econ. 72 (3), 548–555.

Burris, J.L., Studts, J.L., DeRosa, A.P., Ostroff, J.S., 2015. Systematic review of tobacco use after lung or head/neck cancer diagnosis: results and recommendations for future research. Cancer Epidemiology and Prevention Biomarkers 24 (10), 1450–1461.

Carrasco-Garrido, P., Gil de Miguel, A., Hernández-Barrera, V., Jiménez-García, R., 2007. Health profiles, lifestyles and use of health resources by the immigrant population resident in Spain. Eur. J. Publ. Health 17 (5), 503–507.

Cattaneo, M.D., Jansson, M., Ma, X., 2018. Manipulation testing based on density discontinuity. STATA J. 18 (1), 234–261.

Cattaneo, M.D., Idrobo, N., Titiunik, R., 2020. A Practical Introduction to Regression Discontinuity Designs (Elements in Quantitative and Computational Methods for the Social Sciences). Cambridge University Press.

Cawley, J., Ruhm, C.J., 2011. The economics of risky health behaviors. In: Handbook of Health Economics, vol. 2. Elsevier, pp. 95–199.

Chern, W.S., Loehman, E.T., Yen, S.T., 1995. Information, health risk beliefs, and the demand for fats and oils. Rev. Econ. Stat. 555–564.

Chong, S., Ding, D., Byun, R., Comino, E., Bauman, A., Jalaludin, B., 2017. Lifestyle changes after a diagnosis of type 2 diabetes. Diabetes Spectr. 30 (1), 43–50.

Cook, W., 2019. The effect of personalised weight feedback on weight loss and health behaviours: evidence from a regression discontinuity design. Health Econ. 28 (1), 161–172.

Dai, T, Jiang, S, Liu, X, Sun, A, 2022. The Effects of a Hypertension Diagnosis on Health Behaviors: A Two-Dimensional Regression Discontinuity Analysis. Health Economics.

Fichera, E., von Hinke, S., 2020. The response to nutritional labels: evidence from a quasi-experiment. J. Health Econ. 72, 102326.

Gaggero, A., 2020. The effect of type 2 diabetes diagnosis in the elderly. Econ. Hum. Biol. 37, 1–23.

Gelman, A., Imbens, G., 2018. Why high-order polynomials should not Be used in regression discontinuity designs. J. Bus. Econ. Stat. 37 (3), 447–456.

Goldstein, D.E., Little, R.R., Lorenz, R.A., Malone, J.I., Nathan, D., Peterson, C.M., Sacks, D.B., 2004. Tests of glycaemia in diabetes. Diabetes Care 27 (7), 1761–1773.

Grossman, M., 1972. On the concept of health capital and the demand for health. J. Polit. Econ. 80 (2), 223–255.

Hahn, J., Todd, P., van der Klaauw, W., 2001. Identification and estimation of treatment effects with regression-discontinuity design. Econometrica 69 (1), 201–209.

Ho-Pham, L.T., Nguyen, U.D., Tran, T.X., Nguyen, T.V., 2017. Discordance in the diagnosis of diabetes: comparison between HbA1c and fasting plasma glucose. PLoS One 12 (8), e0182192.

Iizuka, T., Nishiyama, T., Chen, B., Eggleston, K., 2021. False alarm? Estimating the marginal value of health signals. J. Publ. Econ. 195, 104368.

Imbens, G.W., Lemieux, T., 2008. Regression discontinuity designs: a guide to practice. J. Econom. 142 (2), 615–635.

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.

Jazieh, A.R., Foraida, M., Ghouse, M., Khalil, M.M., Kopp, M., Savidge, M., 2006. The impact of cancer diagnosis on the lifestyle and habits of patients served at a Veterans Administration Hospital. J. Cancer Educ. 21 (3), 147–150.

Jo, A., Coronel, B.D., Coakes, C.E., Mainous III, A.G., 2019. Is there a benefit to patients using wearable devices such as Fitbit or Health Apps on Mobiles? A systematic review. Am. J. Med. 132 (12), 1394–1400.

Kettlewell, N., 2020. Policy choice and product bundling in a complicated health insurance market: do people get it right? J. Hum. Resour. 55 (2), 566–610.

Khan, M.A.B., Hashim, M.J., King, J.K., Govender, R.D., Mustafa, H., Al Kaabi, J., 2020. Epidemiology of type 2 diabetes–global burden of disease and forecasted trends. Journal of Epidemiology and Global Health 10 (1), 107.

Kim, S.R., Chern, W.S., 1999. Alternative measures of health information and demand for fats and oils in Japan. J. Consum. Aff. 33, 92–109.

Kim, H.B., Lee, S.A., Lim, W., 2019. Knowing is not half the battle: impacts of information from the National health screening program in Korea. J. Health Econ. 65, 1–14.

Kolesár, M., Rothe, C., 2018. Inference in regression discontinuity designs with a discrete running variable. Am. Econ. Rev. 108 (8), 2277–2304.

Lee, D.S., Card, D., 2008. Regression discontinuity inference with specification error. J. Econom. 142 (2), 655–674.

Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. J. Econ. Lit. 48 (2), 281–355.

Liang, Y.W., Chang, H.F., Lin, Y.H., 2019. Effects of health-information-based diabetes shared care program participation on preventable hospitalizations in Taiwan. BMC Health Serv. Res. 19 (1), 890.

Lyons, T.-J., Basu, A., 2012. Biomarkers in diabetes: hemoglobin A1c, vascular and tissue markers. Transl. Res. 159 (4), 303–312.

Mata, M., Cos, F.X., Morros, R., Diego, L., Barrot, J., Berengué, M., et al., 2013. Abordatge de la Diabetes Mellitus Tipus 2. 2a. Edició. Institut Català de la Salut 2013 (accessible at. https://scientiasalut.gencat.cat/handle/11351/4514.

McCrary, J., 2008. Manipulation of the running variable in the regression discontinuity design: a density test. J. Econom. 142 (2), 698–714.

National Institute for Health and Care Excellence (NIHR), 2020. Type 2 Diabetes in Adults. Management.

Oster, Emily, 2018. Diabetes and diet: purchasing behavior change in response to health information. Am. Econ. J. Appl. Econ. 10 (4), 308–348.

Patel, M.S., Asch, D.A., Volpp, K.G., 2015. Wearable devices as facilitators, not drivers, of health behavior change. JAMA 313 (5), 459–460.

Rojo-Martínez, G., Valdés, S., Soriguer, F., Vendrell, J., Urrutia, I., Pérez, V., Ortega, E., Ocón, P., Montanya, E., Menéndez, E., Lago-Sampedro, A., 2020. Incidence of diabetes mellitus in Spain as results of the nation-wide cohort di@bet.es study. Sci. Rep. 10 (1), 1–9.

Roosen, J., Marette, S., Blanchemanche, S., Verger, P., 2009. Does health information matter for modifying consumption? A field experiment measuring the impact of risk information on fish consumption. Rev. Agric. Econ. 31 (1), 2–20.

Seuring, T., Serneels, P., Suhrcke, M., Bachmann, M., 2020. Diabetes, employment and behavioural risk factors in China: marginal structural models versus fixed effects models. Econ. Hum. Biol. 39, 100925.

Slade, A.N., 2012. Health investment decisions in response to diabetes information in older Americans. J. Health Econ. 31 (3), 502–520.

Spanish Ministry of Health, Social Services and Equality, 2012. Diabetes Strategy of the National Health System.

Thistlethwaite, D.L., Campbell, D.T., 1960. Regression-discontinuity analysis: an alternative to the ex-post facto experiment. Journal of Education Phycology 51 (6), 309–317.

Trochim, W.M.K., 1984. Research Design for Program Evaluation: the Regression Discontinuity Approach. Sage Publications, Beverly Hills.

Van der Klaauw, W., 2002. Estimating the effect of financial aid offers on college enrolment: a regression-discontinuity approach. Int. Econ. Rev. 43 (4), 1249–1287.

Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data, second ed. The MIT Press.

World Health Organization, 2016. Assessing Health Services Delivery Performance with Hospitalizations for Ambulatory Care Sensitive Conditions. WHO Regional Office for Europe.

Zhao, M., Konishi, Y., Glewwe, P., 2013. Does information on health status lead to a healthier lifestyle? Evidence from China on the effect of hypertension diagnosis on food consumption. J. Health Econ. 32 (2), 367–385.