Health Economics Review

RESEARCH

Open Access

Diabetes diagnosis based on glucose control levels and time until diagnosis: a regression discontinuity approach to assess the effect on direct healthcare costs

Toni Mora[1] and Beatriz Rodríguez-Sánchez[2*]

Abstract

We estimate the difference in direct healthcare costs of individuals diagnosed with diabetes depending on their glucose level, considering different timespans and subgroups. Using data from administrative registers of 285,450 individuals in Catalonia from 2013 to 2017, we used a fuzzy regression discontinuity design to estimate the causal effect of being diagnosed with diabetes at a given timespan (based on an average glucose value equal to or above 6.5%, the treated group) vs. not (having an average glucose level below the threshold, the control group) on healthcare costs across different timespans (6, 9, 12, 15, 18, 21, and 24 months after the first laboratory test) and distances, in days, between the laboratory test and the doctor's diagnosis. When average glucose level was the only independent parameter and the time until diagnosis was 30 days or less, at the cut-off value (6.5%) healthcare costs were between €3,887 and €5,789 lower for the treated group compared to the control group. Smaller differences were reported as the delay in diagnosis increased, even when additionally controlling for sociodemographic characteristics and health status. Our results highlight the importance of prompt diagnosis and might open the debate about the usefulness of the 6.5% reference value in the blood glucose level as the main diagnostic tool in diabetes.

Highlights

- Values slightly above the blood glucose cut-off are related to lower costs.
- Although always significant, differences in healthcare costs reduce as the delay in diagnosis increases.
- The results hold regardless of the timespan and the considered covariates.
- Our results point towards a diminished effect as the delay in diagnosis increases.
- The findings highlight the role played by physicians in terms of on-time diagnoses.

Keywords Healthcare costs, Diabetes, Glucose level, Administrative data, Fuzzy regression discontinuity

JEL classification H0, H51, I0, I1, I11

*Correspondence:
Beatriz Rodríguez-Sánchez
brodri13@ucm.es
[1]Research Institute for Evaluation and Public Policies (IRAPP), Universitat Internacional de Catalunya (UIC), Carrer de la Immaculada, 22, Barcelona 08017, Spain
[2]Applied Economics, Public Economics and Political Economy, Faculty of Law, Universidad Complutense de Madrid, Plaza Menéndez Pelayo, 4, Madrid 28040, Spain

Introduction

The number of adults with diabetes has substantially increased [1], with 425 million people living with the condition worldwide and this figure projected to reach 629 million individuals by the year 2045 [1, 2]. This substantial increase in the number of people suffering from diabetes can largely be attributed to the effects of adverse lifestyles, population growth and ageing, and the joint effects of these factors [3, 4]. This increase in diabetes prevalence will be accompanied by an increase in diabetes-related care costs [1], which is projected to reach 2.2% of global Gross Domestic Product (GDP) by the year 2050 [5], varying with, amongst other factors, geographical region and age [6]. In Spain, according to the 2019 International Diabetes Federation (IDF) data, the prevalence of diabetes among adults aged from 20 to 79 years old was 13.8% [7], with 9.1% of people having been diagnosed and 4.7% undiagnosed [8]. Although a similar prevalence had traditionally been observed for the case of Catalonia [9], more recent regional data suggests that the prevalence in Catalonia may be slightly higher than the national average due to the aging population [10]. With respect to the incidence of the disease, in both Spain and Catalonia, it is estimated to be around 11 new cases per 1,000 person-years for adults aged 20–79 [8, 9]. The economic burden of diabetes in Spain is significant, with direct medical costs estimated to be around €5.1 billion annually and indirect costs [11–13], amounting to an additional €6.1 billion, with 16% of the national costs being borne by the Catalonian population.

The largest component of diabetes-related medical expenditure is hospital inpatient care which accounts for 43% of the total medical cost [14–16] for three different countries: the United States [14], Italy [15], and Ireland [16]. Additionally, more than one third of diabetes-related care were due to the management of clinical complications in the Italian setting, mainly cardiovascular diseases [15]. The risk of developing cardiovascular complications is, amongst other factors, associated with the degree of long-term glycaemic control in the United States [17, 18], leading to the hypothesis that worse glycaemic control might be associated with increased care costs [19–23] within the United States [19, 20], Italy [21], The Netherlands [22] and Spain [23], more specifically, Catalonia.

Moreover, some authors have pointed towards the cost-effectiveness of the prevention of chronic diseases such as diabetes, but without implying cost savings [24, 25]. However, it seems unclear whether the lifetime medical costs for people with diabetes exceed those of similar individuals without diabetes, with the existing evidence inconclusive. Some authors suggest care costs are higher as life expectancy increases [26, 27] and for those diagnosed with diabetes at younger ages [28], while others have concluded that, despite higher life expectancies, diabetes prevention could lead to sizeable long-term cost savings [29, 30]. To assess the cost implications of a diabetes diagnosis and measures intended to prevent diabetes it is crucial to understand how accumulated healthcare costs increase after diagnosis and whether delays in doctor diagnosis may play a role. This information will improve the understanding of how earlier diagnosis might impact potential medical costs, if at all, and would be important evidence to help design guidelines around the diagnosis of diabetes.

This paper aims to estimate the difference in healthcare costs among the considered individuals ever diagnosed with diabetes across the period of our study (2013–2017) and whose blood glucose level after a laboratory test at the time of diagnosis was equal or above the reference value of 6.5% [31], considering different timespans. General Practitioners (GP) usually diagnose patients whose blood glucose level is equal or above this reference value of 6.5% as being diabetic, but there is discretion in this decision and patients who do not meet this criteria can also be given a diabetes diagnosis. We implement a fuzzy regression discontinuity design (RDD) to estimate the exogenous impact of a diabetes diagnosis on accumulated healthcare costs, obtained from a population administrative database, by comparing individuals just below and just above the pre-determined threshold used for diagnosis. The purpose is to provide evidence on the appropriateness of the commonly used threshold in terms of one specific outcome measure–namely accumulated healthcare costs–and show how, even when a person has already been diagnosed with diabetes, healthcare costs change around this threshold. Although some evidence points towards increasing costs among people with diabetes whose HbA1c values are equal or higher than 6.5% [17–21, 23], limited evidence is available regarding the effect just around the threshold value. We expect to find lower costs right above the cut-off value after a diabetes diagnosis given the onset of the disease and the short-term negative effects that the disease might have had.

We contribute to the existing literature in several significant ways: (i) as far as we are aware, no previous study has assessed the difference in healthcare costs for newly-diabetes diagnosed individuals who are slightly above the reference blood glucose level threshold when diagnosed, compared to those who are slightly below and are not diagnosed with diabetes, who should be comparable in every characteristic apart from the glucose level value; (ii) we apply a regression discontinuity model to evidence a causal impact; (iii) we include different timespans and the additional effect that average glucose control might have on the probability of having a confirmed diagnosis by a doctor, and; (iii) this is a population-based study. The paper is structured as follows: Sect. "Background

literature" provides the reader with existing evidence relating to the scope of the present analysis; Sect. "Data" describes the dataset used, its linkage and the construction of the key variables of interest, and the outcome measures. In Sect. "Methods", the econometric approach is set within the context of our analysis. Then, the descriptive statistics of the sample and the empirical findings are presented in Sect. "Results", which concludes with the discussion of our findings, followed by several appendices that provide supporting data.

Background literature

Rosella et al. [32] found that, during a follow-up of Canadian patients from 2004 until 2012, the average per-person healthcare expenditure was substantially higher for people with diabetes than for those without. There were significant differences in costs during the first year after diagnosis but they stabilised in the following years and significantly declined during the last year of the observed period. These findings are consistent with similar studies in different geographical settings [33, 34]. There is little evidence relating to an "anticipation effect" in terms of increasing healthcare costs before a diabetes diagnosis, although one study suggested that healthcare costs grew in the 5 years before diagnosis and accelerated immediately after diagnosis [35, 36], with the newly-diagnosed subjects spending nearly $9,000 more than their matched counterfactuals, being subjects not diagnosed with diabetes, over the 5 years following diagnosis. However, we were unable to find any studies that analysed the difference in healthcare costs in newly diagnosed diabetes cases that considered multiple timespans since diagnosis.

Diabetes mellitus, in general, can be diagnosed based on any of the following World Health Organisation (WHO) criteria: (i) fasting plasma glucose (FPG) ≥ 7.0 mmol/l (126 mg/dl) or 75 g oral glucose tolerance test (OGTT) with FPG ≥ 7.0 mmol/l (126 mg/dl) and/or 2-hour plasma glucose ≥ 11.1 mmol/l (200 mg/dl); (ii) random plasma glucose ≥ 11.1 mmol/l (200 mg/dl) in the presence of classical diabetes symptoms; or (iii) glycated haemoglobin level (HbA1c) above 6.5% (48 mmol/mol) [31], with the latter the most commonly used threshold for a diabetes diagnosis. There is agreement on the quality and accuracy of glycaemic control as a diagnostic tool among people with diabetes mellitus and an extensive body of evidence in the United States around this topic [14, 30, 37] but few studies have assessed the impact of glucose level on healthcare use and costs in Europe among people with diabetes [21, 23, 38, 39].

The existing evidence points towards increasing healthcare resource utilisation and costs as glycated haemoglobin control worsens among already diagnosed diabetes cases. Indeed, a recent study using population data from Catalonia showed that, compared with individuals with good glycaemic control (which is usually set at an average glucose level equal to or below 6.5%), healthcare costs increased by €428.30 or €395.10 if glucose control was very poor or poor, respectively [23]. Still, there has been recent debate regarding whether the sole use of glycemia as a diagnostic criterion should imply hyperglycaemia as the only therapeutic goal for people with diabetes [40, 41]. Some authors are now proposing a shift in the paradigm to include all non-normal glucose levels and glucose tolerance outcomes; that is, to lower the HbA1c level to 5.7% as the diagnostic criteria for a diagnosis of diabetes, and removing the term 'pre-diabetes' from the current lexicon [41].

However, very little evidence has been reported for the short-term cost consequences of diabetes relating to glucose levels, and what exists reaches different conclusions [21, 42]. Some find little difference between the five categories of HbA1c level, taking as the target HbA1c values below 7% [21], and other work suggests a U-shaped trajectory in diabetes-related costs, falling during the first years after diagnosis and then rising again [39].

Another relevant field of interest to discuss is whether the innovative approach of RDD that we use to model healthcare costs in newly diagnosed diabetes cases has ever been applied to the analysis of either healthcare costs or diabetes, or both. To the best of our knowledge, there is no published work using an RDD to study diabetes-related healthcare costs. We found two studies applying this statistical technique to the blood glucose level cut-off for people with diabetes, but these had as outcomes a set of healthy behaviours [43, 44]. The results suggested a significant improvement in healthy lifestyles, such as obesity reduction or smoking abatement, after a diabetes diagnosis, with differences depending on age and gender. Still, no study has been found that applies the RDD technique to model healthcare costs around a particular diagnostic cut-off value.

Data

Data sources and linkage

We used a large administrative dataset from the Agency for Health Quality and Assessment of Catalonia (AQuAS), which combines information from several providers, although considering different periods, for the whole Catalan adult population during the period 2013–2017, including those diagnosed with any form of diabetes (631,212 individuals). We focused on individuals diagnosed exclusively with diabetes mellitus (622,170 individuals). We further restricted our sample to those newly diagnosed cases who survived for the whole of the considered period. This is important because those who passed away often experienced extremely large values for healthcare costs due to end-of-life medical spending. Likewise, we excluded from the analysis those individuals

who had already been medicated/diagnosed before the date of the first available laboratory test. These exclusions left us with a sample of 300,040 individuals. Finally, we removed any patients with no healthcare use registered (i.e., those with zero healthcare costs) after their laboratory test result. We referred to those individuals who, even being diagnosed and visited across the period, showed no further visits within the public healthcare system. Despite universal coverage, some individuals might prefer to visit through their duplicate coverage condition once diagnosed. Hence, our final working sample consisted of 285,450 individuals.

The AQuAS database contains information on: primary care, hospitalisations and emergency care. The files contain the individual identifier, the visit date (and length of visit in the case of hospitalisations), and all diagnoses and procedures that were administered. In Catalonia, the International Classification of Diseases (ICD-9) was used for diagnostic purposes up to 2017. Diagnoses are shown in an ordinal sense, indicating which was the main diagnosis for each visit and a list of secondary diagnoses. Based on dates and diagnoses amongst the different healthcare providers we were able to identify spells (visits at the same provider related to the same diagnosis within a month).

Additional files provide information about drugs that were dispensed to treat diseases as defined by ATC-7 codes, as well as information related to specific laboratory tests (glucose tolerance tests) that are commonly requested by physicians when diagnosing diabetes. Using ATC-7 codes we were able to identify the following medications used by individuals for the treatment of diabetes: insulins, biguanides, which consisted of metformin only, sulfonylureas, combinations of oral blood glucose lowering drugs (i.e., a combination of biguanides and sulfonylureas such as metformin and vildagliptin, or pioglitazone and alogliptin), alpha glucosidase inhibitors, thiazolidinediones, dipeptidyl peptidase 4 (DPP-4) inhibitors, glucagon-like peptide-1 (GLP-1) analogues, sodium-glucose co-transporter 2 (SGLT2) inhibitors, and other blood glucose lowering drugs, excluding insulins. The specific drugs we identified from their ATC-7 code are listed in Table A1, Appendix.

Costs dataset

There are 2,800 healthcare procedures (HCP) in the dataset, defined and classified according to the International Classification of Diseases, 9th revision, Clinical Modification (ICD-9-CM). The unit cost of each HCP has been imputed using a complete list of public prices approved in 2013 by the Department of Health in Catalonia (Spain) for the Catalan Healthcare Service for primary care services, hospital and specialised services, and psychiatric and mental health services.

These unit prices for hospital and specialised visits have been actualised or incorporated for each year from 2013 until 2020 but very few new healthcare procedures have been added. Therefore, the approved public prices for 2013 were used as the main resource to impute the cost of the HCP. Public prices approved in 2013 for primary care services include unit prices or tariffs for standard primary care services such as GP visits, ambulatory care, and domiciliary care. In the case of hospital and specialised services, most of the HCP prices are set according to the Diagnostic Related Group (DRG), including prices for a wide range of surgical procedures, surgical implants, infection treatment and plastic surgery, etc. Other hospital and specialised services tariffs cover laboratory tests, rehabilitation and physiotherapy, and further tests, procedures, and therapies that support the primary diagnosis.

However, the Catalan Health Service does not provide a price for all HCPs. In such cases where a price is missing, the price of the DRG to which the HCP belongs is assigned. Since groups of hospitals set prices for DRGs, an HCP will have a different price depending on the hospital where it was performed. If an HCP procedure occurs in more than one DRG, the average across all the possible DRGs by groups of hospitals is taken. Therefore, the price inputted for a particular HCP can be either a unique tariff, a tariff differing by a group of hospitals, or an average of different DRG prices across groups of hospitals. We use an online and up-to-date database of reported Spanish healthcare costs called Esalud to calculate those values not provided by the Catalan Health Service. Overall, 93% of the prices were found, with 66% coming from a single tariff (either from Catalonia or from another Spanish region) and 27.28% coming from DRGs and by groups of hospitals. Since the dataset encompasses HCPs from 2013 to 2017, imputed prices have been deflated to 2017 prices, using the corresponding Health Consumer Price Index for each year[1]. Drugs costs were considered from the funder perspective, without discounting co-payments[2]. Individuals with costs above the 99th percentile have been removed, as they may be cases of serious illness and might be regarded as outliers.

Construction of comorbidities and additional covariates

We also considered the presence of comorbidities that can be attributed to a diabetes diagnosis and can be

[1] https://www.ine.es/prensa/ipc_tabla.htm. Instituto Nacional de Estadística.

[2] We then built total costs as those costs related to the total cost generated (visits, drugs and procedures) but also those strictly referred to a visit in which diabetes was included as a diagnosis. Finally, we took the decision to explore differences in the total costs given the diverse comorbidities that arise from this condition, which have been included as covariates in the second regression model specification, which will be explained into more detail in Sect. "Econometric specifications".

considered a proxy of the severity of diabetes worsening, since it has been found that healthcare costs among people with diabetes increase with the number of comorbidities and severity of the complications [23, 45]. We include the following categories of diseases that the literature identifies as being related to diabetes: (i) micro-vascular complications, which refer to retinopathy, nephropathy (renal failure), and neuropathy; (ii) macro-vascular complications, such as coronary heart disease (congestive heart failure and other types of heart failure), myocardial infarction, cerebrovascular disease (stroke and other cerebrovascular accidents) and peripheral artery disease, and other cardiovascular and vascular diseases; (iii) cardiovascular risk factors, which entail hypertension, obesity, and weight loss; and (iv) other diseases, such as dementia.

The data files are linked between all providers via unique personal identifiers to some demographic information: gender, age, drug co-payment level (which is related to socioeconomic status of the individuals), individual nationality, date of death, and the health area the individual belongs to. We also include as an explanatory variable the Adjusted Morbidity Groups (AMG), a new morbidity measurement tool adapted to the Spanish healthcare system which creates 31 mutually exclusive groups, with higher scores indicating worse health status [46].

Methods

Regression discontinuity design

The regression discontinuity design (RDD) has been widely used in recent applied literature as one of the most reliable quasi-experimental designs for identifying, estimating, and interpreting results among treated individuals in a study, i.e., those who receive treatment or intervention. This is particularly true for individuals close to the "local" cut-off applied [47]. In our study, the "local" cut-off is the blood glucose level threshold of 6.5%, used to define a diabetes diagnosis. Compared to other statistical methods, RDD is regarded as the closest to a true experimental design due to its ability to approximate random assignment at the threshold. This design also offers straightforward interpretability, contributing to its increasing popularity among policymakers and policy analysts.

RDD hinges on the assumption that the only systematic difference between individuals just above and just below the threshold is the treatment assignment, which, in our case, is a diabetes diagnosis based on blood glucose levels. The forcing variable, which is the observed blood glucose level, determines the assignment to the treatment group (diagnosis of diabetes) or control group (no diagnosis). This cut-off at 6.5% is critical because it minimizes the potential for manipulation, ensuring that the assignment is as good as random close to the threshold [48–50]. However, in practice, the assignment is not always perfectly adhered to due to factors such as delays in diagnosis or variations in clinical practice. Therefore, in our study, the probability of receiving a diabetes diagnosis at the threshold is less than 100%, leading to what is known as a fuzzy RDD. In a fuzzy RDD, the treatment assignment is not strictly binary but probabilistic, meaning that crossing the threshold influences but does not strictly determine the treatment status. This scenario arises due to "imperfect compliance," where doctor-related factors and timing affect the actual recording of a diabetes diagnosis [51, 52]. Whether the individual is diagnosed with diabetes (allocated to the treatment group) not only depends on whether his/her average glucose control lies above or below the cut-off, but also on the time until the doctor registers this glucose value as a positive diabetes case and other patient characteristics, such as family history of the disease or the presence of other metabolic conditions like hypertension or dyslipidaemia [53]. Our fuzzy RDD approach accommodates these scenarios.

To address this imperfect compliance, we use an instrumental variable approach within the RDD framework to estimate the diagnosis's local average treatment effect (LATE) on healthcare costs. This method leverages the discontinuity at the threshold to identify the causal impact of a diabetes diagnosis among those individuals whose treatment status is influenced by their blood glucose levels. The strength of the fuzzy RDD lies in its ability to account for such real-world complexities while still providing robust causal estimates.

The assumptions underlying our RDD approach include the continuity of potential outcomes at the threshold and the absence of precise manipulation of the forcing variable. We validate these assumptions through diagnostic tests to ensure the reliability of our findings. This comprehensive approach ensures that our results accurately reflect the causal impact of a diabetes diagnosis on healthcare costs.

The forcing variable: average glucose control

Bearing in mind the information presented in Sect. "Regression discontinuity design", the aim of the RDD is to estimate the parameter τ which denotes the differential value in the outcome between those who lie above or below the threshold, which can be given as:

$$\tau = \tau\left(\bar{X}\right) = E\left\{Y_i(1) - Y_i(0) \middle| X_i = \bar{x}\right\} \quad (1)$$

where $\bar{X}$ is the specified cut-off value (in our study, a glucose level equal to or above 6.5%); $Y_i(1) - Y_i(0)$ represents the potential results for each unit allocated

to the treatment and the control group, respectively; X_i denotes the score for the main independent variable (in our case, blood glucose level).

We define the reference average glucose control as equal to or above 6.5%, as this is the reference value used by the World Health Organization [31] for a diabetes diagnosis. The probability of having a diabetes diagnosis, if solely based on the average glucose level, changes from zero to one when the observed average glucose level reaches (from below) the threshold (6.5%). If the average glucose level is below 6.5%, it would be assumed that the person has a proper level of blood glucose and has no diabetes, whereas equal or above 6.5% diabetes would be presumed[3]. Very close to the threshold, however, we can treat the diagnosis of diabetes as occurring randomly. Thus, by looking at both sides of that sharp cut-off, we are able to isolate exogenous variation and estimate the causal effect of having a newly-diabetes diagnosis on the outcome of interest, in this case healthcare costs.

Econometric specifications

Although all the individuals in our sample were diagnosed with diabetes at some point in the period 2013–2017, we aim to assess the effect on healthcare costs of having a diabetes diagnosis based on average glucose levels above or below the 6.5% threshold. Hence, we focused on individuals whose first laboratory test reported a value equal to or above the cut-off, compared to those below it. As is required in a fuzzy RDD, we applied Two-Stage Least Squares (2SLS) to estimate parametric equations of the following form, taking into account imperfect compliers since having a glucose level equal or above 6.5% does not immediately imply being diagnosed of diabetes:

$$\text{First stage}: \quad P_{it} = \alpha + \delta X_{it} + h(Z_{it}) + \epsilon_{it} \quad (2)$$

$$\text{Second stage}: \quad Y_{it} = \mu + \delta \widehat{P}_{it} + h(Z_{it}) + u_{it} \quad (3)$$

where P_i is a dummy variable that identifies actual participation of individual *i* in the treatment group, i.e. having a diabetes diagnosis. Notice that within both Eqs. (2) and (3), Z_i refers to the average glucose level obtained in the first laboratory test performed [54]. X_i would then refer to the delay until the doctor diagnoses an individual with diabetes. This could either be: (i) number of days between the date of a laboratory test which returned a blood glucose level of at least 6.5%, and a doctor diagnosis or diabetes medication prescription stratified to 30, 45, 60 or 75 days; or (ii) a smaller number of days between a laboratory result below the threshold and then a positive diagnosis in a second laboratory test performed within the next 30, 45, 60, or 75 days.

More precisely, in the second part of the 2SLS performed, we estimate the following equation as the baseline regression model:

$$Y_{it} = \beta_0 + \beta_1 I[HbA1c_{it} \geq 6.5] + \epsilon_{it} \quad (4)$$

where Y_{it} denotes the corresponding accumulated healthcare costs of individual i in the timespan t; and $I[.]$ is an identity function that takes a value of 1 if the individual's glucose level in the first laboratory test performed is above the corresponding threshold or 0 otherwise. ϵ_{it} denotes the error term, which is assumed to be normally distributed.

As the forcing variable is the only independent variable in Eq. (2), we introduced different covariates in a posterior regression model and we accounted for sociodemographic characteristics such as age, gender, and socioeconomic status, as well as some indicators of the individual's health status; we also controlled for the Inverse Mills Ratio (IMR), which accounts for the probability of having a laboratory test performed conditional on individual characteristics, such as age, gender, comorbidities, and the most frequented provider unit for each individual. The final variable accounts for the differences in the use of public-private providers in the performance of laboratory tests given the high percentage of duplicate coverage in Catalonia (around 20% during the considered period). We also adjusted for a list of diabetes-related comorbidities: retinopathy, nephropathy, neuropathy, coronary heart disease (congestive heart failure and other types of heart failure), myocardial infarction, cerebrovascular disease (stroke and other cerebrovascular accidents), peripheral artery disease and other cardiovascular and vascular diseases, hypertension, obesity, weight loss, and dementia.

Hence, the full regression model in the second stage of the 2SLS regression can be specified as follows:

$$\begin{aligned} Y_{it} = &\beta_0 + \beta_1 I[HbA1c_{it} \geq 6.5] + \beta_2' X_{it} \\ &+ \beta_3 IMR_i + \beta_4' comorb_{it} + \varepsilon_{it} \end{aligned} \quad (5)$$

where X_{it} is a vector representing a set of sociodemographic characteristics (age, gender, nationality, drug copayment level, the AMG, and the minimum distance, in days, between the laboratory test measurement and the medical consultation for diagnosis); IMR_i accounts for the Inverse Mills Ratio of having a laboratory test result;

[3] So, if a patient: 55-year-old male attends a routine check-up and an HbA1c of 6.4% is shown, he would be part of the control group. Since HbA1c is below 6.5%, no formal diabetes diagnosis or intensive treatment is initiated. The patient might be advised on general healthy lifestyle habits but does not receive medication or additional monitoring. On the other hand, a patient a 60-year-old female whose routine check-up shows HbA1c of 6.6%, would be part of the treatment group. Since HbA1c is above 6.5%, the patient is diagnosed with diabetes and receives additional interventions, such as medication (e.g., metformin), lifestyle counselling, and possibly more frequent follow-ups.

and $comorb_{it}$ represents the set of comorbidities listed above. We estimated Eqs. (4) and (5) using a local linear estimation within the mean squared error optimal bandwidth and polynomial order following the RDD features specification method proposed by Calonico et al. [55] and weighting the included observations by proximity to the cut-off using triangular kernel and robust inference methods. Equations (4) and (5) were run for each range of delay until diagnosis and time span considered. Clustered standard errors at the basic health area were performed to control for any potential effect of the GP 's higher probability of diagnosis. Sensitivity analyses were performed to confirm the robustness of the observed effects by setting alternative HbA1c threshold values (6.3, 6.7 and 6.9%).

Older people represent around half of those with diabetes, and diabetes prevalence reaches one in every four adults aged 65 years and above [56]. The International Diabetes Federation (IDF) guidelines, last updated in 2017 [57], improve on the previous version from 2012 through several additions, such as the consideration and classification of older people into three groups according to their functional and cognitive status: functionally independent; functionally dependent, if coping with some limitations in the activities of daily living (ADL); and end-of-life care, in the case of older adults having a major illness and whose life expectancy is less than one year. According to the IDF guidelines, once the individual's functional and cognitive status has been evaluated, different glycaemic targets will be established for each patient, to be regarded as "poorly controlled" in the case of having a glycated haemoglobin target above 7–8%, which might also be subject to comorbidities. Hence, we used the same fuzzy regression discontinuity models on the subsample aged 65 years old and above, but increasing the threshold of 6.5% for the average glucose control to 7%, 7.5%, and 8%.

Table 1 Descriptive statistics on sociodemographic characteristics and health status for the whole sample

	Whole sample ***N*** **= 285,450**
Gender: female (%)	45.75
Age	66.86 (12.48)
Drug co-payment level indicator, %	
Exempted	5.04
10% co-payment	69.44
40% co-payment	16.88
50% co-payment	7.66
60% co-payment	0.47
Excluded from co-payment	0.51
Country or geographical are of birth, %	
Spain	92.62
Magreb	2.01
South America	1.14
Eastern Europe	0.77
Others	3.46
Individual health status, AMG	335.27 (20.57)
Comorbidities, %	
Hypertension	63.84
Overweight	46.55
Weight loss	1.38
Myocardial infarction	8.37
Congestive heart failure	4.42
Other heart failure	0.52
Peripheral vascular disease	5.28
Stroke	4.51
Other cerebrovascular diseases	5.51
Other cardiovascular diseases	8.72
Neuropathy	3.31
Retinopathy	9.51
Other vascular diseases	0.19
Dementia	1.74
Renal failure	19.11
Severity according to comorbidities prevalence, %	
Mild	72.46
Moderate	5.50
Severe	22.04

Note Means are presented as its mean value, with the standard deviation within brackets unless indicated otherwise. AMG stands for Adjusted Morbidity Group

Results

Descriptive

Different timespans were selected based on the date of the laboratory test; that is, taking into account the number of months after the performance of the first laboratory test towards a positive diabetes diagnosis and the cumulative healthcare costs of individuals related to diabetes, performing the analyses for a 6-month span, which comprised 159,478 individuals[4]; a 9-month span, with 128,924 individuals; a 12-month span, with 98,528 individuals; a 15-month span, with 79,751 individuals; an 18-month span, with 66,209 individuals; a 21-month span, with 57,300 individuals; and 2 years (24 months) after the test, for a sample of 47,519 individuals.

When looking at sociodemographic characteristics and individual health status (Table 1), people who have ever been diagnosed with diabetes mellitus throughout the period considered are, on average, 66.9 years old (age range: 18 to 103 years old); 45.8% of the sample are females and 92.6% are Spanish. 69.4% are eligible for a 10% co-payment in terms of pharmaceutical provision,

[4] The number of observations drops from the initial sample of 275,450 individuals since the timespans refer to subjects who were observed during the same period of time before and after the first laboratory test. In order for a patient to be included in the 6-month span they must have had their test at least 6 months before the end of your time frame but if they are included in the 24-month span they must have had their test at least 24 months before the end and therefore there are less patients to consider as your time span increases.

according to their income, whereas 5% are exempted from co-payment. With respect to their health status, nearly 64% of the individuals have hypertension, 46.6% are overweight, 8.4% have ever had a myocardial infarction, and 19.1% suffer from nephropathy. Based on the comorbidities suffered, almost three quarters of the sample are classified as mild severity, whereas nearly 25% of the sample are classified as severe.[5]

Table 2 below shows the mean and median 6-month healthcare costs after the first positive diabetes diagnosis laboratory test performed, as well as the number of visits, depending on the blood glucose level. Although minor and non-significant differences were observed in average terms, higher total costs were incurred by positive diabetes mellitus cases, with HbA1c ≥ 6.5%, with respect to those whose HbA1c was below the threshold (€8,594 vs. €8,433). The main differences were found in the costs of procedures during specialist visits, which reached nearly €4,521 for the former group and nearly €4,357 for the latter; and costs of medical visits, being €3,426 for those with HbA1c ≥ 6.5% and €3,386 for those with values below the cut-off point. Drug costs, including both diabetes-related and non-diabetes-related drugs, were approximately €50 higher for the control group compared to those with blood glucose levels above the threshold, although this difference was not statistically significant.

If we based a conclusion on data from the full period, the statistically significant differences in costs of procedures might explain the higher total accumulated costs among the individuals with HbA1c ≥ 6.5% (Table 3). In fact, total costs since the first laboratory test were almost €300 higher for people with HbA1c ≥ 6.5% (€7,634) than for those whose glucose level is below 6.5% (€7,336), and the largest differences is costs due to procedures (€4,001 vs. €3,745). Moreover, as can be seen from Fig. 1, when comparing kernel density functions for the treated and the untreated individuals, the dotted and continuous lines follow the same pattern, indicating no differences between those two groups. The red lines, which correspond to the 6-month span, are generally above the blue and the brown lines, implying that higher healthcare costs are concentrated over the shortest timespans.

Table 2 Descriptive statistics on healthcare use and costs for those treated (HbA1c ≥ 6.5%) vs. controls (HbA1c < 6.5%). Span = 6 months

	Average glucose control (HbA1c) < 6.5% *N* = 86,704		**Average glucose control (HbA1c) ≥ 6.5% *N* = 72,774**	
	Mean	**Median**	**Mean**	**Median**
Number of medical visits	53.05	44.00	53.51**	44.00
Cost of medical visits	3,386.10	2,369.39	3,426.28**	2,365.76
Cost of procedures	4,356.75	0.00	4,521.05**	0.00
Drugs cost	690.52	0.00	646.33	0.00
Total costs during the first 6 months after the laboratory test	8,433.36	2,651.25	8,593.66	2,602.78
Accumulated costs during the 6 months before the laboratory test[a]	467.57 (5,316.12)		305.13*** (3,573.81)	

[a] Mean and standard deviation within brackets

The stars represent the tests for statistical differences between means, which denote: *: p-value < 0.10; **: p-value < 0.05; ***: p-value < 0.01

Table 3 Descriptive statistics on healthcare use and costs for those treated (HbA1c ≥ 6.5%) vs controls (HbA1c < 6.5%). Span = 24 months

	Average glucose control (HbA1c) < 6.5% N = 27,079		**Average glucose control (HbA1c) ≥ 6.5% N = 20,440**	
	Mean	**Median**	**Mean**	**Median**
Number of medical visits	43.87	35.00	43.99	35.00
Cost of medical visits	2,815.28	1,816.86	2,820.69	1,822.56
Cost of procedures	3,745.23	0.00	4,001.21**	0.00
Drugs cost	775.22	0.00	812.31	0.00
Total costs during the first 6 months after the laboratory test	7,335.73	1,918.02	7,634.21	1,919.12
Accumulated costs during the 6 months before the laboratory test[a]	1,016.23 (7,340.44)		805.76*** (5,898.62)	

[a] Mean and standard deviation within brackets

The stars represent the tests for statistical differences between means, which denote: *: p-value < 0.10; **: p-value < 0.05; ***: p-value < 0.01

Econometric results

Figure 2 shows our benchmark results, with each panel representing the regression discontinuity plot for each timespan considered in the analysis. The y-axis represents accumulated healthcare costs over the different timespans, and include costs relating to visits and procedures in primary care, hospitalisations and emergency care, as well as prescription drug use. The x-axis denotes the glucose level value of the laboratory test result, measured by HbA1c, at the time of diabetes diagnosis. We plotted two local first-order polynomial regressions estimated for individuals either side of the 6.5% blood glucose level threshold. We illustrate the 6.5% threshold with a vertical line in the figure. The dots are the average surplus of the individuals in each of the bins within that glucose level interval. Our main result is illustrated by the discontinuous jump in accumulated healthcare costs right at the threshold, showing a higher discontinuity in total healthcare costs at the cut-off point as the timespan increases.

Table A2, Appendix, shows the coefficients obtained for the first stage of the fuzzy RD regression models

[5] The mean differences test across the set of considered spans found no differences. Individuals were very balanced based on treatment.

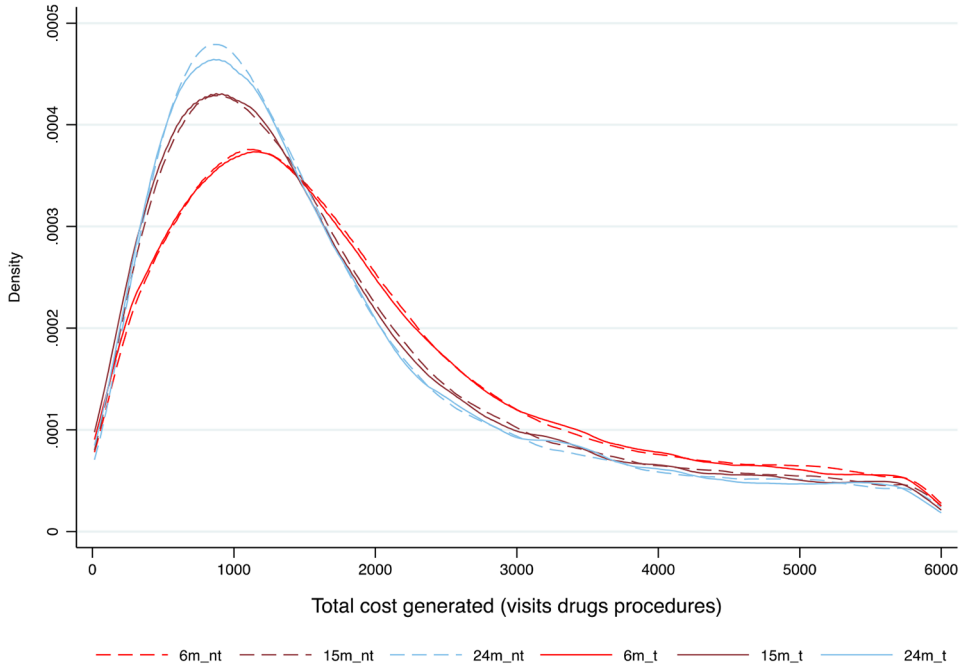

Fig. 1 Average healthcare costs across timespans. *Note*: the dotted lines refer to the non-treated individuals (those whose average glucose control level does not lie above the cut-off of 6.5%), whereas the continuous lines denote the treated individuals. The timespans considered are 6 months, 15 months and 24 months since the first laboratory test

performed. The results show that a glucose level found to be equal to or above 6.5% in a laboratory test is significantly associated with a higher probability of being diagnosed with diabetes.

Table 4 shows the results from the second stage of the regression discontinuity models performed, in which the cut-off value of the forcing variable, average glucose level, was set at 6.5%. Accumulated healthcare costs after the first laboratory test differ significantly between those immediately above the threshold and those whose glucose level was below the cut-off, regardless of the delay in the doctor's diagnosis, the timespan observed, and the covariates included. In fact, if we account for the forcing variable as the only independent parameter (Model 1), we observe that the difference in healthcare costs between the treated (individuals with diabetes whose first laboratory test result was equal to or above 6.5%) and the controls (individuals with diabetes but with average glucose levels below 6.5%) was -€3,886.8 at 6 months after the laboratory test if the doctor's diagnosis was within 30 days, -€5,789.4 at 9 months, -€5,064.8 at 12 months, -€4,185 at 15 months, -€4,139.4 at 18 months, -€4,840.5 at 21 months, and -€5,343.7 at 24 months. These numbers show that healthcare costs were lower among the treated, i.e., newly-diagnosed diabetes cases, than among the control group. The difference in healthcare costs between the treated and the controls reduces as the delay in diagnosis increases, reaching its minimum after 75 days. For this group, the differences are: -€2,668.5 after 6 months, -€3,931 after 9 months, -€3,656.1 after 12 months, -€2,809.3 after 15 months, -€3,041.7 after 18 months, -€3,646.7 21 months after the first laboratory test, and -€4,024.7 at 24 months.

The coefficients show changes of magnitude in Model 2, but not with respect to statistical significance nor regarding the lower costs incurred by the treated, when additional covariates–including sociodemographic and health status variables as well as the IMR–are introduced. For example, the average difference in total healthcare costs between the treated and the control groups slightly decreases to -€3,437.6 during the first 6 months after the initial laboratory test result, again reaching the peak at 9 months after the first laboratory test, where the accumulated healthcare costs of those individuals with diabetes whose first laboratory test result was equal to or above 6.5% were lower by -€4,536.5 than their comparators if the time until a confirmed diagnosis was equal to 30 days or less and falling afterwards. The difference in direct healthcare costs between the treated and the controls at the cut-off value was smallest when there was the longest delay in diagnosis (75 days) at -€1,997.2[6].

[6] As the Table A3, Appendix, shows, the results found for the accumulated total healthcare costs particularly hold for medical visits, regardless of the time span considered, the delay until diagnosis, and the covariates included. Significant results pointing to lower costs among those in the treatment group (those slightly above the cut-off) compared to the control group (those slightly below the threshold) were found for medical procedure costs only at the longest timespan. Barely any significant effect was found for drug costs.

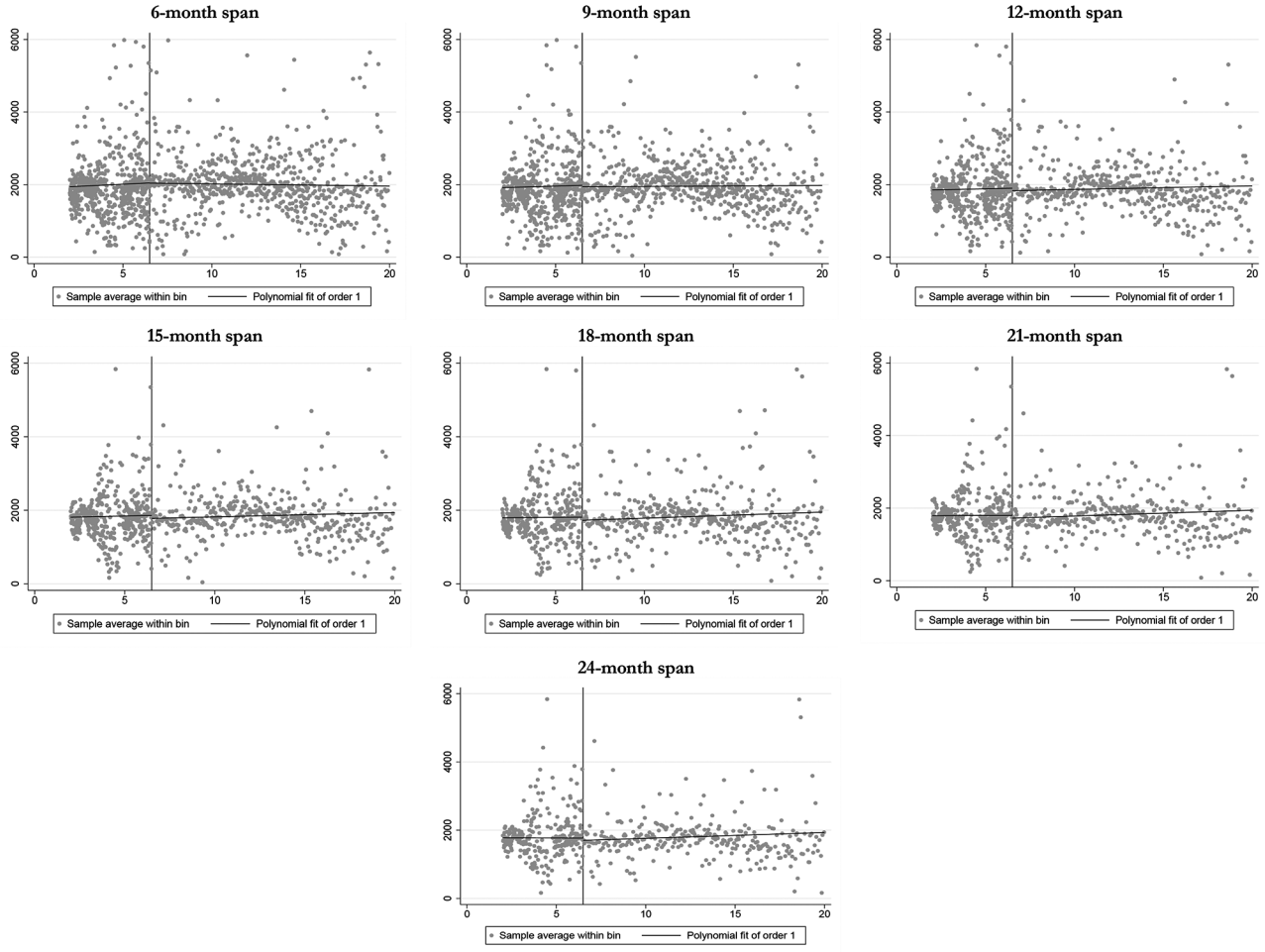

Fig. 2 Effect of average glucose level on accumulated healthcare costs at the discontinuity point (6.5%) across different timespans. Note: (1) The y-axis plots accumulated healthcare costs over the different timespans, which include visits and procedures in primary care, hospitalisations and emergency care, as well as drugs use, and where the number of days between the laboratory and the doctor's diagnosis was 60 days. The x-axis refers to the glucose level value, measured by HbA1c, at the time of diabetes diagnosis- The vertical line denotes the threshold value of 6.5%. (2) The dots are bin averages. (3) The solid line represents a first-order polynomial regression

Furthermore, Figure A1, Appendix, plots the results for the second stage of the fuzzy RD across different timespans and diagnosis delays, depending on the sample size considered.[7] The figure shows that the results are partially consistent across timespans and dropped outliers, especially qualitatively. In addition, Figure A2, Appendix, shows the results when varying the HbA1c threshold to 6.3, 6.7 and 6.9, including the reference results of 6.5% in the Figure A2 as well. There was no significant effect in terms of direct healthcare costs when the threshold was 6.9%, regardless of the delay in diabetes diagnosis and the length of follow-up, as well as for follow-up periods above 12 months when the threshold was set at 6.7%. Still, the results obtained for 6, 9 and 12-month follow-ups confirm, regardless of the delay in diagnosis, the results obtained in the main analysis (HbA1c threshold of 6.5%) become larger in magnitude when the threshold is higher. When setting the HbA1c threshold below the reference value of 6.5%, exactly at 6.3%, the results show that those individuals whose HbA1c values are slightly above the threshold value of 6.3% show higher accumulated healthcare costs than those slightly below. The results obtained when the threshold is set at 6.3% are indeed consistent with our main results concerning the diminishing effect around the threshold value as the delay in diagnosis increases.

Table A4, Appendix, reports the results from the second stage regressions performed on the subsample of older adults (those aged 65 years and above) applying different average glucose level thresholds (7%, 7.5% and 8%). The results show statistical differences when using 7% as the threshold value but only across the 6-month

[7] 97% after removing outliers in healthcare costs above the 97th percentile, 98% removing outliers above the 98th percentile, and 99%, being the actual sample assessed in the current paper, when outliers above the 99th percentile of direct healthcare costs were removed.

Table 4 Effects of average glucose level on healthcare costs from the fuzzy regression discontinuity models excluding outliers >99%. Forcing variable: distance from cut-off HbA1c = 6.5%

	Model 1[a]				Model 2[b]			
	30-day delay	45-day delay	60-day delay	75-day delay	30-day delay	45-day delay	60-day delay	75-day delay
				6-month span				
Differences in total healthcare costs for treated (HbA1c ≥ 6.5%) vs. controls (HbA1c < 6.5%)	-3,886.8** (1,910.6)	-3,234.2** (1,573.5)	-2,950.9** (1,453.2)	-2,668.5** (1,314.5)	-3,437.6** (1,661.1)	-2,737.2** (1,357.2)	-2,491.5** (1,238.8)	-2,262.6** (1,121.1)
Number of observations, left of the cut-off	85,889	85,889	85,889	85,889	85,794	85,794	85,794	85,794
Number of observations, right of the cut-off	72,009	72,009	72,009	72,009	71,924	71,924	71,924	71,924
				9-month span				
Differences in total healthcare costs for treated (HbA1c ≥ 6.5%) vs. controls (HbA1c < 6.5%)	-5,789.4*** (1,809.4)	-4,733.1*** (1,467.8)	-4,266.2*** (1,314.9)	-3,931*** (1,211.3)	-4,536.5*** (1,663.5)	-3,586.1*** (1,311.6)	-3,236.3*** (1,176.3)	-3,030*** (1,108.4)
Number of observations, left of the cut-off	70,690	70,690	70,690	70,690	70,610	70,610	70,610	70,610
Number of observations, right of the cut-off	56,945	56,945	56,945	56,945	56,876	56,876	56,876	56,876
				12-month span				
Differences in total healthcare costs for treated (HbA1c ≥ 6.5%) vs. controls (HbA1c < 6.5%)	-5,064.8*** (1,907.3)	-4,265.3*** (1,559.8)	-3,893.9*** (1,417.9)	-3,656.1*** (1,328.8)	-3,243.5** (1,573	-2,480.2** (1,225.3)	-2,277.7** (1,110.2)	-1,997.2** (1,008.1)
Number of observations, left of the cut-off	54,677	54,677	54,677	54,677	54,618	54,618	54,618	54,618
Number of observations, right of the cut-off	42,850	42,850	42,850	42,850	42,802	42,802	42,802	42,802
				15-month span				
Differences in total healthcare costs for treated (HbA1c ≥ 6.5%) vs. controls (HbA1c < 6.5%)	-4,185** (1,705.3)	-3,078.1** (1,353.1)	-2,834.8** (1,218.2)	-2,809.3** (1,161.3)	-3,503.5** (1,538.1)	-2,962.1** (1,291.8)	-2,711.3** (1,179.7)	-2,544.6** (1,113.7)
Number of observations, left of the cut-off	44,576	44,576	44,576	44,576	44,526	44,526	44,526	44,526
Number of observations, right of the cut-off	34,361	34,361	34,361	34,361	34,320	34,320	34,320	34,320
				18-month span				
Differences in total healthcare costs for treated (HbA1c ≥ 6.5%) vs. controls (HbA1c < 6.5%)	-4,319.4*** (1,479.3)	-3,242.5*** (1,166.7)	-3,272.6*** (1,122)	-3,041.7*** (1,046.7)	-3,728.8*** (1,373.7)	-3,145.1*** (1,182.1)	-2,842.1** (1,099.8)	-2,640** (1,031)
Number of observations, left of the cut-off	37,035	37,035	37,035	37,035	36,993	36,993	36,993	36,993
Number of observations, right of the cut-off	28,506	28,506	28,506	28,506	28,468	28,468	28,468	28,468
				21-month span				
Differences in total healthcare costs for treated (HbA1c ≥ 6.5%) vs. controls (HbA1c < 6.5%)	-4,840.5*** (1,500.6)	-4,227.2*** (1,307.9)	-3,864.9*** (1,190)	-3,646.7*** (1,130.1)	-3,734.7*** (1,347.4)	-3,211.9*** (1,168.7)	-2,899.3*** (1,067.6)	-2,723.5*** (1,013.1)
Number of observations, left of the cut-off	32,179	32,179	32,179	32,179	32,144	32,144	32,144	32,144
Number of observations, right of the cut-off	24,544	24,544	24,544	24,544	24,510	24,510	24,510	24,510
				24-month span				
Differences in total healthcare costs for treated (HbA1c ≥ 6.5%) vs. controls (HbA1c < 6.5%)	-5,343.7*** (1,433)	-4,574.5*** (1,217.8)	-4,269*** (1,135.2)	-4,024.7*** (1,074.8)	-3,471.1*** (1,262.3)	-3,005.9*** (1,094)	-2,793.3*** (1,015.4)	-2,656.7*** (968.43)
Number of observations, left of the cut-off	26,834	26,834	26,834	26,834	26,803	26,803	26,803	26,803
Number of observations, right of the cut-off	20,211	20,211	20,211	20,211	20,187	20,187	20,187	20,187

a Model 1 only includes the forcing variable, which refers to the average glucose (HbA1c) level through a laboratory test, with the cut-off value being 6.5

b Model 2 adjusts for age and its square, gender, being Spaniard, drug co-payment level according to the TSI, the adjusted morbidity group, the comorbidities listed in Table 1 and the inverse Mills ratio (IMR), which denotes the probability of having a laboratory test conditional on individual characteristics and the most frequent provider unit

*: p-value < 0.10; **: p-value < 0.05; ***: p-value < 0.01

Clustered standard errors within brackets at the basic health area. A triangular kernel-weighted local polynomial smoothing has been applied, which gives more weight to points close to the threshold

timespan. When the number of days between the laboratory test and the doctor's diagnosis was 30 days or less, the difference in accumulated healthcare costs at the cut-off value of 7% between the treated and the controls was -€3,249.5, decreasing to -€2,126.3 after 45 days and remaining stable at 60 days (-€2,018.5) and 75 days (-€2,135.3) following diagnosis. Lower costs were found for the treated than the control group. No statistically significant results were found for other timespan, days until the doctor's diagnosis after the first laboratory test, or threshold value.

Discussion

In this paper, we explore the effect of glucose levels, conditional on having been diagnosed with diabetes by a doctor, on healthcare costs across different timespans.

Our evidence, stemming from a fuzzy RDD that exploits guidelines for diagnosing diabetes based on a threshold recommended by official health bodies, including the World Health Organization [31], suggests that a diagnosis just above the recommended threshold of 6.5% is statistically significantly related to lower accumulated healthcare costs, regardless of the timespan considered, the delay in doctor diagnosis, and the covariates included, than a positive diagnosis just below the threshold. Overall, the results obtained show that accumulated healthcare costs, compared to those within the first six months of the first laboratory test, are higher for the control group at longer timespans, reaching their peak with the timespan of nine months (€5,789.4 if the forcing variable was the only independent variable and €4,536.5 in case of the full model). This shows that those who were diagnosed and had a blood glucose level below 6.5% (the controls) had higher costs than those newly diagnosed with higher blood glucose levels (the treated). However, our results do point towards a diminished effect around the threshold value as the delay in diagnosis increases.

This shows that an earlier diagnosis is associated with potential savings, since we observe the effect throughout different timespans, with savings being observed up to two years. Using different time spans (from 6 months up to 24 months) ensures that our results do not entirely depend on a specific version of the functional form of the forcing variable or are driven by data points that are far from the threshold [49, 58]. In general, irrespective of whether the raw model (which only includes the forcing variable) or the full specification is used, the results are consistent in that smaller differences occur as the delay between the laboratory test and the doctor diagnosis increases and over longer timespans.

These results could be due to the targeted interventions used by physicians to help patients manage their glucose levels. The literature has already shown the beneficial effects of lowering HbA1c levels among already diagnosed and uncontrolled diabetes cases on healthcare costs, stating that average savings were substantially higher in the first year ($2,503) than in the second-year of follow-up ($1,690), compared with people with diabetes whose glucose level increased [19]. Indeed, intensive glycaemic control programmes have been found in the existing literature to be very cost-effective [59, 60], leading to a £258 cost reduction per patient with diabetes in the United Kingdom [61]. Promoting a proper blood glucose level among people who are not diagnosed with diabetes but show symptoms of diabetes might lead to higher potential savings.

However, the comparability of the results obtained by other authors with our findings is limited, because of the different comparison groups used or the delays in diabetes diagnosis that we consider. Other studies group individuals depending on the blood glucose level attained [20, 21, 23], as we do, for example, in the summary statistics tables (e, g, good and poor glucose control). They do report higher accumulated healthcare costs, though not statistically significant differences, among people with diabetes whose average glucose level is equal to or above 6.5% relative to their comparators. However, when performing the regression analyses, we assess local average treatment effects around the threshold value assumed, without being able to infer whether the effect remains steady as the blood glucose value lies further from the cut-off. Actually, when varying the HbA1c threshold value in sensitivity analyses, we found that results remained consistent when the threshold value was set at 6.7%, but, at lower threshold values (HbA1c being equal to 6.3%), results shifted to the opposite direction: those who are slightly above the threshold value of 6.3% have higher accumulated healthcare costs than those slightly below. However, since regression discontinuity designs estimate local average treatment effects (LATE), changing the threshold implies that different subsets of the populations are effectively being analysed. So, when lowering the threshold to 6.3%, individuals previously in the control group (6.4-6.49%) now become the treatment group in this new threshold specification. This shift means that the composition of both groups has changed, which can lead to different estimates (i.e. different treatment responses or different baseline health expenditures). For instance, healthcare providers might escalate interventions more aggressively for those closer to 6.5% (seeing them at higher risk), which could affect cost trajectories. Such cost trajectories may not be monotonic across the HbA1c range and those closer to 6.3% might be receiving different interventions (e.g., lifestyle changes rather than medications), affecting cost patterns differently.

It is worth noting that thresholds different from the 6.5% WHO value are considered depending on individual characteristics (such as age or body mass index) or cognitive and functional impairment in daily practice, such as those the International Diabetes Federation published in its 2017 guideline [57]. For example, in 2019 the Spanish Society of Endocrinology and Nutrition (SEEN) published a report on the integrated approach towards diabetes mellitus [62]. Within this guidance, the SEEN established that the objective in diabetes patients should be to reach an average level of glucose (HbA1c) below 7%, although several individual characteristics might require a more (HbA1c < 6.5%) or less (HbA1c < 8-8.5%) demanding target, such as frailty status, hypoglycaemia odds, comorbidities, body mass index, self-care ability, level of resources, and life expectancy. When we varied the cut-off value of the average glucose control to 7%, 7.5%, and 8% among those aged 65 years and above, (the age group with the highest diabetes prevalence [49]), our findings

showed almost no significant results in terms of accumulated healthcare costs for the treated versus the controls. We cannot say that this removes the need for varying targets, but suggests any different healthcare costs amongst this group are driven by variables that we could not control for in our analysis, such as the functional status of older people with diabetes. Indeed, previous results have shown that functional impairment is a stronger predictor of hospital admission than individual characteristics or diabetes itself among older adults [63], resulting in a three-fold increase in care costs among older dependent individuals compared with independent older adults [64]. Previous studies have also reported that lifetime medical costs among people with diabetes differ when distinguishing by gender, age [28, 32], and age at diagnosis [28], pointing towards lower annual excess medical spending when diagnosed at a later age. Moreover, empirical evidence suggests that physician adherence to evidence-based guidelines varies significantly [65, 66], depending on organizational arrangement [65] as physicians working in group practice seem to be more prone adhere to evidence-based guidelines and request the necessary laboratory tests than physicians in individual practices; and general practitioner (gender, experience and number of patients in the roster) and patient's characteristics (age and comorbidity level) [65, 66]. In addition, financial and institutional incentives (i.e. pay-for-performance or pay-for-compliance programs) can further shape adherence to guidelines [67], with delays in diagnoses in settings without such incentives and treatment assignment occurring below the 6.5% threshold in healthcare settings with such incentives, weakening the discontinuity and leading to imperfect compliance with the treatment assignment, motivating the fuzzy RDD approach that we have applied in this study. Hence, physician heterogeneity in guideline adherence represents a potential source of variation in our results. However, we have controlled for the probability of requesting laboratory results by the most frequent provider unit in our model. By acknowledging these differences and exploring their implications, we enhance the robustness of our conclusions regarding the cost effects of diabetes diagnosis. Future research could further investigate how economic and institutional factors shape guideline adherence and patient outcomes in similar quasi-experimental settings.

A natural question that arises is why our results matter and how useful they can be in informing policy decisions. The general consensus is that the prevalence of diabetes mellitus is a serious, pandemic-level issue in Western countries, due to its rapid increase and projections for further growth in the coming years. It represents a significant economic burden for healthcare systems and is a healthcare policy priority [5, 14, 28]. Therefore, it is necessary to develop new efficient therapy strategies, together with appropriate prevention measures, rapid diagnosis tools, and better management of the disease and its associated risk factors. The findings reported in this paper could be especially relevant in terms of practice, as the time between the laboratory test result and the doctor diagnosis seems to play a key role in the difference in healthcare costs between the two compared groups. Preventive measures promoted by primary healthcare professionals, especially those responsible for the diabetes diagnosis, and good care of hyperglycaemia, among other steps, should be adopted as routine. Our results suggest that diabetes diagnosis itself might not be driving bigger cost differentials among older populations, which could point towards higher effectiveness of diabetes prevention strategies among younger populations. Moreover, our findings call into question whether it is appropriate to use 6.5% as the average glucose control to confirm a new diabetes diagnosis, since those individuals slightly above the threshold have lower accumulated healthcare costs than those individuals slightly below it. It is true that the difference decreases when controlling for sociodemographic characteristics and health status, but the sociodemographic factors play less of a role than when studying delays in diagnosis after a laboratory test. Hence, the physician plays an essential role from an economic point of view. Further analyses could also assess the role of glucose control levels on healthcare costs across different timespans, as other studies performed in a Catalonian setting have suggested that poorer glucose control leads to higher healthcare costs [23]. However, these authors did not evaluate the effect that maintenance of glucose control throughout different time periods.

To the best of our knowledge, this is the first paper that applies a fuzzy RDD to study the local average treatment effects around the glucose control value threshold on accumulated healthcare costs in newly-diagnosed diabetes individuals. Our analysis is enriched by using a large administrative dataset, which contains information on 285,450 individuals from 2013 to 2017, allowing us to perform the analyses on timespans up to two years after the first diagnostic laboratory test was performed. Moreover, we have information on primary care, hospitalisations and emergency care, as well as medications, which enables us to provide reliable estimates on healthcare costs. However, there are also some limitations that should be mentioned. The main limitations of our study are related to the availability of data, mainly due to the lack of data on the number of specialist visits, and laboratory tests (such as a regular blood analysis) that take place in addition to those checking blood glucose levels. Future analyses could assess longer follow-up periods in order to check whether the pattern observed within our timeframe remains. Additionally, we could use, instead

of the crude average glucose control values, the changes in the main variable of interest within a particular range (for example, ± 2 points) and the effect on accumulated healthcare costs. Although there are some other clinical criteria for diagnosing diabetes, the use of HbA1c has several advantages: firstly, it reflects long-term glucose control [68, 69], offering a more comprehensive assessment of an individual's blood sugar management; secondly, the HbA1c test does not require fasting [70, 71], making it more convenient for patients and reducing the likelihood of non-compliance due to fasting requirements and, lastly, HbA1c levels are not influenced by short-term factors such as stress or illness [72], which can temporarily alter blood glucose levels, thus providing a more stable and reliable measure of chronic glycemia. In contrast, and in addition to its higher usage cost [73], the estimated prevalence might be underreported if some hemoglobinopathies and conditions affecting red blood cell turnover, such as anaemia, are present [74, 75]. As it is not influenced by short-term factors, it may not detect recent changes in glucose levels [76, 77]. Additionally, it should be noted that the unique structure of the Catalonian healthcare system, with universal public coverage, may limit the generalizability of our findings to regions that rely more heavily on private insurance. The healthcare system in Catalonia, which is predominantly public and universal, provides a unique context for examining health outcomes. In contrast, regions with mixed or private healthcare models—where access to care is often influenced by insurance coverage or out-of-pocket costs—could see significant differences in both healthcare access and outcomes. For instance, individuals in regions with a higher proportion of private healthcare might face barriers to accessing preventive services or diabetes management, leading to a higher prevalence of uncontrolled risk factors such as hypertension, obesity, and poor glycaemic control. Additionally, the population's socioeconomic status could play a larger role in healthcare access in systems where private insurance is more prevalent, potentially exacerbating health inequalities. Exploring these differences in further research could enhance the relevance of our study, shedding light on how public healthcare systems, in contrast to those with stronger private sectors or varying levels of public coverage, might affect the management of chronic diseases, healthcare accessibility, and overall population health.

Still, this paper raises important points and provides new avenues for researchers interested in analysing the relationships and trade-offs between clinical targets and the role of the physicians themselves.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13561-025-00613-y.

Supplementary Material 1

Acknowledgements
We thank Antonio Valero from Hospital Clínic of Barcelona and the Agency for Health Quality and Assessment of Catalonia for the access to the dataset.

Author contributions
The authors equally contributed to the manuscript.

Funding
Toni Mora gratefully acknowledges the financial support from the PID2021-124067OB-C21.

Data availability
The data that support the findings of this study are available from the Agency for Health Quality and Assessment of Catalonia but restrictions apply to the availability of these data, which were used under licence for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the Agency for Health Quality and Assessment of Catalonia.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no conflicts of interest.

Received: 27 March 2023 / Accepted: 11 March 2025
Published online: 24 March 2025

References

1. OECD. Health at a glance 2019: OECD indicators. Paris: OECD Publishing; 2019.
2. IDF. IDF diabetes atlas eighth edition 2017. International Diabetes Federation; 2017. http://www.diabetesatlas.org.
3. NCD Risk Factor Collaboration. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. Lancet. 2016;387:1513–30.
4. OECD. Cardiovascular disease and diabetes: policies for better health and quality of care. Paris: OECD Health Policy Studies, OECD Publishing; 2015.
5. Bommer C, Sagalova V, Heesemann E, Manne-Goehler J, Atun R, Bärnighausen T, Vollmer S. Global economic burden of diabetes in adults: projections from 2015 to 2030. Diabetes Care. 2018;41(5):963–70.
6. Williams R, Karuranga S, Malanda B, Saeedi P, Basit A, Besançon S, Colagiuri S. Global and regional estimates and projections of diabetes-related health expenditure: results from the international diabetes federation diabetes atlas. Diabetes Res Clin Pract. 2020;162:108072.
7. Atlas D. International diabetes federation. *IDF diabetes atlas*. 9th ed. Brussels, Belgium: International Diabetes Federation; 2019.
8. Soriguer F, Goday A, Bosch-Comas A, Bordiú E, Calle-Pascual A, Carmena R, Vendrell J. Prevalence of diabetes mellitus and impaired glucose regulation in Spain: the Di@ Bet. Es study. Diabetologia. 2012;55:88–93.
9. Regidor E, Franch J, Seguí M, Serrano R, Rodríguez-Artalejo F, Artola S. Traditional risk factors alone could not explain the excess mortality in patients with diabetes: a National cohort study of older Spanish adults. Diabetes Care. 2012;35(12):2503–9.
10. Enquesta de salut de Catalunya. 2022. Barcelona: Direcció General de Planificació en Salut; 2023.
11. Crespo C, Brosa M, Soria-Juan A, Lopez-Alba A, Lopez-Martinez N, Soria B. Direct cost of diabetes mellitus and its complications in Spain. Seccaid study: Spain estimated cost Ciberdem-Cabimer in diabetes. Value Health. 2013;16(7):A436.

12. Lopez-Bastida J, Boronat M, Moreno JO, Schurer W. Costs, outcomes and challenges for diabetes care in Spain. Globalization Health. 2013;9:1–9.
13. Mata-Cases, M., Casajuana, M., Franch-Nadal, J., Casellas, A., Castell, C., Vinagre,I.,… Bolíbar, B. (2016). Direct medical costs attributable to type 2 diabetes mellitus: a population-based study in Catalonia, Spain. The European Journal of Health Economics, 17, 1001–1010.
14. American Diabetes Association. Economic costs of diabetes in the US in 2017. Diabetes Care. 2018;41(5):917–28.
15. Bruno, G., Karaghiosoff, L., Merletti, F., Costa, G., De Maria, M., Panero, F.,… Gnavi, R. (2008). The impact of diabetes on prescription drug costs: the population-based Turin study. Diabetologia, 51(5), 795–801.
16. O'Neill KN, McHugh SM, Tracey ML, Fitzgerald AP, Kearney PM. Health service utilization and related costs attributable to diabetes. Diabet Med. 2018;35(12):1727–34.
17. Huang ES, Liu JY, Moffet HH, John PM, Karter AJ. Glycemic control, complications, and death in older diabetic patients: the diabetes and aging study. Diabetes Care. 2011;34(6):1329–36.
18. Stolar M. Glycemic control and complications in type 2 diabetes mellitus. Am J Med. 2010;123(3):S3–11.
19. Bansal M, Shah M, Reilly B, Willman S, Gill M, Kaufman FR. Impact of reducing glycated hemoglobin on healthcare costs among a population with uncontrolled diabetes. Appl Health Econ Health Policy. 2018;16(5):675–84.
20. Boye KS, Lage MJ, Thieu VT. The association between HbA1c and 1-Year Diabetes-Related medical costs: A retrospective claims database analysis. Diabetes Therapy. 2022;13(2):367–77.
21. Degli Esposti, L., Saragoni, S., Buda, S., Sturani, A., & Degli Esposti, E. (2013). Glycemic control and diabetes-related health care costs in type 2 diabetes; retrospective analysis based on clinical and administrative databases. Clinico-Economics and outcomes research: CEOR, 5, 193.
22. Janssen, L. M. M., Hiligsmann, M., Elissen, A. M. J., Joore, M. A., Schaper, N. C.,Bosma, J. H. A.,… Evers, S. M. A. A. (2020). Burden of disease of type 2 diabetes mellitus: cost of illness and quality of life estimated using the Maastricht Study. Diabetic Medicine, 37(10):1759–1765.
23. Mata-Cases M, Rodríguez-Sánchez B, Mauricio D, Real J, Vlacho B, Franch-Nadal J, Oliva J. The association between poor glycemic control and health care costs in people with diabetes: a population-based study. Diabetes Care. 2020;43(4):751–8.
24. Cohen JT, Neumann PJ, Weinstein MC. Does preventive care save money? Health economics and the presidential candidates. N Engl J Med. 2008;358(7):661–3.
25. Maciosek MV, Coffield AB, Flottemesch TJ, Edwards NM, Solberg LI. Greater use of preventive services in US health care could save lives at little or no cost. Health Aff. 2010;29(9):1656–60.
26. Eddy DM, Schlessinger L, Kahn R. Clinical outcomes and cost-effectiveness of strategies for managing people at high risk for diabetes. Ann Intern Med. 2005;143(4):251–64.
27. Russell LB. Preventing chronic disease: an important investment, but don't count on cost savings. Health Aff. 2009;28(1):42–5.
28. Zhuo X, Zhang P, Barker L, Albright A, Thompson TJ, Gregg E. The lifetime cost of diabetes and its implications for diabetes prevention. Diabetes Care. 2014;37(9):2557–64.
29. Colagiuri S, Walker AE. Using an economic model of diabetes to evaluate prevention and care strategies in Australia. Health Aff. 2008;27(1):256–68.
30. Zhuo X, Zhang P, Gregg EW, Barker L, Hoerger TJ, Pearson-Clarke T, Albright A. A nationwide community-based lifestyle program could delay or prevent type 2 diabetes cases and save $5.7 billion in 25 years. Health Aff. 2012;31(1):50–60.
31. World Health Organisation. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia. report of a WHO/IDF consultation; 2006.
32. Rosella, L. C., Lebenbaum, M., Fitzpatrick, T., O'Reilly, D., Wang, J., Booth, G.L.,… Wodchis, W. P. (2016). Impact of diabetes on healthcare costs in a population-based cohort: a cost analysis. Diabetic Medicine, 33(3):395–403.
33. Sabale, U., Bodegård, J., Sundström, J., Östgren, C. J., Nilsson, P., Johansson, G.,… Henriksson, M. (2015). Healthcare utilisation and costs following newly diagnosed type-2 diabetes in Sweden: a follow-up of 38,956 patients in a clinical practice setting. Primary Care Diabetes, 9(5):330–337.
34. Visaria, J., Iyer, N. N., Raval, A. D., Kong, S. X., Hobbs, T., Bouchard, J.,… Willey,V. J. (2020). Healthcare costs of diabetes and microvascular and macrovascular disease in individuals with incident type 2 diabetes mellitus: a ten-year longitudinal study.Clinicoeconomics and Outcomes Research: CEOR, 12, 423.
35. Khan T, Yang J, Wozniak G. Trends in medical expenditures prior to diabetes diagnosis: the early burden of diabetes. Popul Health Manage. 2021;24(1):46–51.
36. Zhou X, Shrestha SS, Shao H, Zhang P. Factors contributing to the rising National cost of glucose-lowering medicines for diabetes during 2005–2007 and 2015–2017. Diabetes Care. 2020;43(10):2396–402.
37. Wang Y, Park J, Li R, Luman E, Zhang P. National trends in Out-of-Pocket costs among US adults with diabetes aged 18–64 years: 2001–2017. Diabetes Care. 2021;44(11):2510–7.
38. Gil J, Li Donni P, Zucchelli E. Uncontrolled diabetes and health care utilisation: a bivariate latent Markov model approach. Health Econ. 2019;28(11):1262–76.
39. Rodríguez-Sánchez B, Feenstra TL, Bilo HJ, Alessie RJ. Costs of people with diabetes in relation to average glucose control: an empirical approach controlling for year of onset cohorts. Eur J Health Econ. 2019;20(7):989–1000.
40. Schwartz SS, Epstein S, Corkey BE, Grant SF, Gavin III, Aguilar JR, R. B., Herman ME. A unified pathophysiological construct of diabetes and its complications. Trends Endocrinol Metabolism. 2017;28(9):645–55.
41. Schwartz SS, Rachfal AW, Corkey BE. The time is now for new, lower diabetes diagnostic thresholds. Trends Endocrinol Metabolism. 2022;33(1):4–7.
42. McBrien, K. A., Manns, B. J., Chui, B., Klarenbach, S. W., Rabi, D., Ravani, P.,… Clement, F. (2013). Health care costs in people with diabetes and their association with glycemic control and kidney function. Diabetes Care, 36(5):1172–1180.
43. Gaggero A. The effect of type 2 diabetes diagnosis in the elderly. Econ Hum Biology. 2020;37:100830.
44. Gaggero A, Gil J, Jiménez-Rubio D, Zucchelli E. (2021). Health information and lifestyle behaviours: the impact of a diabetes diagnosis. IZA Discussion Paper No. 14106.
45. Kjellberg J, Tikkanen CK, Bagger M, Gæde P. Short-term societal economic burden of first-incident type 2 diabetes-related complications–a nationwide cohort study. Expert Rev PharmacoEcon Outcomes Res. 2020;20(6):577–86.
46. Monterde D, Vela E, Clèries M. Adjusted morbidity groups: A new multiple morbidity measurement of use in primary care. Atencon Primaria. 2016;48(10):674–82.
47. Calonico S, Cattaneo MD, Farrell MH, y, Titiunik R. (2017). rdrobust: Software for regression-discontinuity designs. The Stata Journal. 17(2):372–404.
48. Cook TD. Waiting for life to arrive: a history of the regression-discontinuity design in psychology, statistics and economics. J Econ. 2008;142(2):636–54.
49. Lee DS, Lemieux T. Regression discontinuity designs in economics. J Econ Lit. 2010;48(2):281–355.
50. Abadie A, Cattaneo MD. Econometric methods for program evaluation. Annual Rev Econ. 2018;10:465–503.
51. Imbens GW, Wooldridge JM. Recent developments in the econometrics of program evaluation. J Econ Lit. 2009;47(1):5–86.
52. Hausman C, Rapson DS. Regression discontinuity in time: considerations for empirical applications. Annual Rev Resource Econ. 2018;10:533–52.
53. Gaggero A, Gil J, Jiménez-Rubio D, Zucchelli E. Sick and depressed? The causal impact of a diabetes diagnosis on depression. Health Econ Rev. 2023;13(1):38.
54. Imbens GW, Lemieux T. Regression discontinuity designs: A guide to practice. J Econ. 2008;142(2):615–35.
55. Calonico S, Cattaneo MD, Titiunik R. Robust nonparametric confidence intervals for regression-discontinuity designs. Econometrica. 2014;82(6):2295–326.
56. Soriguer, F., Goday, A., Bosch-Comas, A., Bordiú, E., Calle-Pascual, A., Carmena,R.,… Vendrell, J. (2012). Prevalence of diabetes mellitus and impaired glucose regulation in Spain: the Di@ bet. es Study. Diabetologia, 55(1), 88–93.
57. Cho NH, Colagiuri S, Distiller L et al. for the IDF Working group. (2017). Managing Older People with Type 2 Diabetes: Global Guideline. http://www.idf.org/sites/default/files/IDFGuideline-for-older-people-T2D.pdf. (Accessed 10th February 2022).
58. Angrist JD, Pischke JS. Mostly harmless econometrics. Princeton University Press; 2008.
59. Li R, Zhang P, Barker LE, Chowdhury FM, Zhang X. Cost-effectiveness of interventions to prevent and control diabetes mellitus: a systematic review. Diabetes Care. 2010;33(8):1872–94.
60. Liebl A, Khunti K, Orozco-Beltran D, Yale JF. Health economic evaluation of type 2 diabetes mellitus: a clinical practice focused review. Clin Med Insights: Endocrinol Diabetes. 2015;8:CMED–S20906.
61. Clarke, P., Gray, A., Adler, A., Stevens, R., Raikou, M., Cull, C.,… Holman, R. (2001).Cost-effectiveness analysis of intensive blood-glucose control with

metformin in overweight patients with type II diabetes (UKPDS No. 51). Diabetologia, 44(3), 298–304.
62. Reyes-García, R., Moreno-Pérez, Ó., Tejera-Pérez, C., Fernández-García, D., Bellido-Castañeda,V., de la Torre Casares, M. L.,… Mezquita-Raya, P. (2019). Documento de abordaje integral de la diabetes tipo 2. Endocrinología, Diabetes y Nutrición, 66(7):443–458.
63. Rodríguez-Sánchez B, Cantarero-Prieto D. Socioeconomic differences in the associations between diabetes and hospital admission and mortality among older adults in Europe. Econ Hum Biology. 2019;33:89–100.
64. Sinclair A, Dunning T, Rodriguez-Mañas L. Diabetes in older people: new insights and remaining challenges. Lancet Diabetes Endocrinol. 2015;3(4):275–85.
65. Fantini MP, Compagni A, Rucci P, Mimmi S, Longo F. General practitioners' adherence to evidence-based guidelines: a multilevel analysis. Health Care Manage Rev. 2012;37(1):67–76.
66. McKinlay JB, Link CL, Freund KM, Marceau LD, O'Donnell AB, Lutfey KL. Sources of variation in physician adherence with clinical guidelines: results from a factorial experiment. J Gen Intern Med. 2007;22:289–96.
67. Bruni ML, Nobilio L, Ugolini C. Economic incentives in general practice: the impact of pay-for-participation and pay-for-compliance programs on diabetes care. Health Policy. 2009;90(2–3):140–8.
68. Committee TIE. International expert committee report on the role of the A1C assay in the diagnosis of diabetes. Diabetes Care. 2009;32(7):1327.
69. American Diabetes Association. Standards of medical care in diabetes—2022 abridged for primary care providers. Clin Diabetes. 2022;40(1):10–38.
70. Nathan DM, Kuenen J, Borg R, Zheng H, Schoenfeld D, Heine RJ, A1c-Derived Average Glucose (ADAG) Study Group. Translating the A1C assay into estimated average glucose values. Diabetes Care. 2008;31(8):1473–8.
71. Colagiuri S. Glycated haemoglobin (HbA1c) for the diagnosis of diabetes mellitus—practical implications. Diabetes Res Clin Pract. 2011;93(3):312–3.
72. Little RR, Rohlfing C, Sacks DB. The National glycohemoglobin standardization program: over 20 years of improving hemoglobin A1c measurement. Clin Chem. 2019;65(7):839–48.
73. Stark Casagrande S, Fradkin JE, Saydah SH, Rust KF, Cowie CC. The prevalence of meeting A1C, blood pressure, and LDL goals among people with diabetes, 1988–2010. Diabetes Care. 2013;36(8):2271–9.
74. Sacks DB. A1C versus glucose testing: a comparison. Diabetes Care. 2011;34(2):518.
75. Ang SH, Thevarajah M, Alias Y, Khor SM. Current aspects in hemoglobin A1c detection: a review. Clin Chim Acta. 2015;439:202–11.
76. Saudek CD, Herman WH, Sacks DB, Bergenstal RM, Edelman D, Davidson MB. A new look at screening and diagnosing diabetes mellitus. J Clin Endocrinol Metabolism. 2008;93(7):2447–53.
77. Bennett CM, Guo M, Dharmage SC. HbA1c as a screening tool for detection of type 2 diabetes: a systematic review. Diabet Med. 2007;24(4):333–43.

Publisher's note