Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-025-04962-1 OPEN

Check for updates

More information, better diet?——regression-discontinuity evidence from hypertension diagnoses in China

Qihui Chen [1✉], Juerong Huang[2], Yue Hu[3] & Jiale Bao[3]

Hypertension is diagnosed when one's systolic blood pressure (SBP) reading exceeds 140 mmHg or his/her diastolic blood pressure (DBP) reading exceeds 90 mmHg. Exploiting these diagnostic rules, this paper identifies the causal impact of hypertension diagnoses on one's diet structure by performing a two-dimensional regression discontinuity analysis. Our analysis of a longitudinal dataset on 9355 Chinese adults from a large-scale household survey revealed mixed effects of first-ever hypertension diagnoses based on SBP readings and little impact of DBP-based diagnoses. SBP-based diagnoses help reduce individuals' (already overconsumed) fat and livestock product intake about 3 years later. However, SBP-based diagnoses also undermine individuals' dietary diversity and diet balance by reducing their consumption of foods that should be increased for hypertensive individuals (e.g., fruits). These previously overlooked undesirable effects suggest that even though disease diagnoses can provide patients with accurate and professional health information, patients may misunderstand and misinterpret the information provided, thus leading to unhealthy dietary responses.

[1] Beijing Food Safety Policy & Strategy Research Base, China Agricultural University, Beijing, PR China. [2] College of Economics and Management, Northwest A&F University, Yangling, PR China. [3] College of Economics and Management, China Agricultural University, Beijing, PR China. ✉email: chenqihui2002@126.com; qhchen2013@cau.edu.cn

Introduction

How health information affects food choices has been an active topic in public health and health economics research. Yet, isolating the causal effect of health information from other confounding factors may be challenging. The recent decades have witnessed the emergence of a new strand of research that explores personal medical records, especially information on chronic disease diagnoses, to assess the effect of health information provided during the diagnoses on individuals' food choices (Zhao et al. 2013; Xiang 2016; Bünnings 2017; Oster 2018; Hut and Oster 2020).[1] To the extent that chronic conditions exist before their diagnoses for at least some time, the diagnoses of these conditions provide a valuable opportunity to isolate the impact of health information from that of health conditions.

The diagnoses of cancer, chronic inflammatory diseases, diabetes, and hypertension have been exploited as "teachable moments" for inducing health behavior change in recent studies (Xiang 2016). Most studies on cancer diagnoses followed a cohort of individuals and compared their pre- and post-diagnosis dietary changes. For example, Velentzis et al. (2011) evaluated U.K. patients' dietary intake changes before and after a breast cancer diagnosis. They found statistically significant increases in post-diagnostic consumption of fruits, lean protein sources, vegetables, and whole grains and statistically significant decreases in the intakes of alcoholic drinks, coffee, red meat, refined grains, high-fat and high-sugar products; accompanying changes in supplement and Vitamin intakes were also observed. Lei et al. (2018) tracked a group of early-stage breast cancer patients in China and found that breast cancer diagnoses significantly increased their intake of eggs, fruits, refined grains, nuts, vegetables, and whole grains 18 and 36 months post-diagnosis; conversely, their consumption of coffee, dairy products, processed meat, poultry, red meat, soy foods, and sugary drinks decreased significantly. Fassier et al. (2017) followed 696 French persons with incidentally diagnosed cancer between 2009 and 2016 and observed a statistically significant decrease in intakes of alcoholic beverages, dairy products, meat, soy products, sweetened soft drinks, and vegetables, and an increase in broths and fats/sauces after the diagnosis. As pointed out by Aldossari et al. (2023) in their systematic review, while nearly all studies on cancer diagnosis conducted in the past two decades discovered some positive post-diagnosis changes in dietary and supplement intake, the changes found tend to be quantitatively small and may not have clinical significance.

Compared with cancer-diagnosis studies, studies on diabetes diagnosis adopted a more diverse battery of methods. For example, Kim et al. (2018) applied a propensity score matching analysis to data on 1769 matched pairs of U.S. individuals and found that individuals diagnosed with (pre)diabetes consumed fewer carbohydrates than matched controls without (pre)diabetes. Hinkle et al. (2021) tracked 1731 pregnant U.S. women who received usual prenatal care and compared their nutrition intake before and after a diagnosis of gestational diabetes mellitus (GDM), using women without GDM as a control group. This difference-in-differences (DID) type of analysis revealed significant post-diagnosis reductions in total energy and carbohydrate intakes among women with GDM; their juice and added sugar intake also decreased, whereas their cheese and artificially sweetened beverage consumption increased. These changes were not observed among women without GDM. Mancini et al. (2017) further examined the role of socio-economic factors in inducing dietary changes post-diabetes diagnosis. They tracked 57,304 French women who were free from type-2 diabetes in 1993, following them through 2005, and found that diagnoses of type-2 diabetes were associated with reduced intakes of alcohol, carbohydrate, lipid, and energy; education and whether having a family (with a partner and/or children) also contributed to their dietary changes. Oster (2018) used household scanner data to estimate U.S. households' food purchase response to diabetes diagnoses inferred from purchases of diabetes-related products based on a machine learning approach. The results show statistically significant but relatively small calorie reductions that are concentrated in unhealthy foods.

Studies on the diagnoses of chronic inflammable diseases are far more sparse, and the methods adopted are often ad hoc. Opstelten et al. (2019) compared dietary patterns between a group of Dutch patients with longstanding inflammable bowel diseases (IDBs) and a control group without IDBs. The patient group had higher carbohydrate and animal protein intakes and lower alcohol, dietary fiber, and (unsaturated) fat intakes than the control group, which can be explained by the former's higher consumption of carbonated beverages, meat, and poultry and fewer intakes of dairy products, fruits, and vegetables. Weiss et al. (2023) conducted semi-structured online interviews with Australian patients diagnosed with multiple sclerosis to understand how they made food choices following a multiple sclerosis diagnosis. They identified four post-diagnostic themes from the interviews: moving toward dietary guidelines, modifying dietary fat intake, requiring mental effort, and needing help from dietitians—how patients interpret healthy eating advice played a role in forming different themes.

Among the chronic diseases examined, hypertension received relatively more attention because of several of its unique features. First of all, despite its high global prevalence and relatively easy detection (e.g., through regular checkups),[2] it is estimated that nearly 60% of hypertensive individuals were not diagnosed (Zhou et al. 2021). If not appropriately controlled, hypertension may also cause other conditions, such as kidney disease and stroke, to develop (Mendis et al. 2011). Secondly, hypertension has close connections with one's dietary pattern. In particular, hypertension may result from over-consuming "tasty" high-fat and high-sodium foods (World Health Organization 2022). Given these features, the causal effects of hypertension diagnosis on food choices will inform both hypertension management and the prevention of related illnesses. Finally, hypertension diagnosis has clear thresholds based on continuous biomarker readings, which provides a valuable opportunity to apply a quasi-experimental approach that mimics locally randomized trials—regression discontinuity (RD) design (Lee and Lemieux 2010)—to identify its effect. Under the assumption that persons with blood pressure readings just above and just below the diagnosis cutoffs are comparable in all relevant aspects except their diagnosis results, RD designs focusing on individuals around these cutoffs can yield more convincing estimates of the impact of hypertension diagnoses than conventional before-and-after-diagnosis comparisons (e.g., Neutel and Campbell 2018). The fact that hypertension is usually asymptomatic and many hypertensive individuals are not diagnosed renders its diagnosis an *unexpected* information shock to many individuals, which also helps interpret the identified impact.

However, existing RD studies on the impact of hypertension diagnoses have yielded results that vary substantially across contexts. Dai et al. (2022), for instance, discovered that diagnoses based on systolic blood pressure (SBP) readings significantly lowered Chinese individuals' fat intake; in contrast, diagnoses based on diastolic blood pressure (DBP) readings had essentially no effect. Slade and Kim (2014) found that individuals' responses to hypertension diagnosis differed depending on the timing of the diagnosis. Those with a recent diagnosis were more likely to reduce their intake of nutrients critical to blood pressure management (e.g., sodium). These context-specific findings suggest that more analyses should be conducted to uncover more granular patterns.

One possibility is to expand the "basket" of foods under examination. Nutrients in foods are not consumed in isolation; rather, they are consumed in a complex and interactive manner (de Carvalho et al. 2014). Thus, the whole food basket of foods being consumed relates more closely to one's health than the separate intake of individual food items or nutrients. Yet, in recent studies on dietary and lifestyle changes related to hypertension diagnoses, the specific foods examined are relatively limited and differ greatly across studies. For example, Zhao et al. (2013) examined the intakes of three macronutrients (carbohydrates, protein, and fat) and total energy, and Dai et al. (2022) only used fat intake as a proxy for an unhealthy diet. Little attention has been paid to the overall structure (e.g., diversity and balance) of the foods consumed. This is unfortunate as disease diagnoses might induce substitution between foods consumed, causing changes in one's diet structure even if the amount of food consumed remains unchanged.

In fact, many authoritative organizations have promoted dietary diversity as a fundamental principle of balanced, healthy diets that can reduce the risks of major chronic diseases (United States Agency for International Development 2016; German Nutrition Society 2017; Chinese Nutrition Society 2022). Dietary diversity has also been incorporated into popular evidence-based healthy dietary patterns, such as the DASH (Dietary Approaches to Stop Hypertension) diet and the Mediterranean diet, and is affirmed in most national food-based dietary guidelines (U.S. Department of Agriculture and U.S. Department of Health and Human Services 2020). Various indicators reflecting one's diet structure, e.g., the Diet Quality Index and the Healthy Eating Index, have been devised in influential dietary-pattern studies since the 1990s (Patterson et al. 1994; Kennedy et al. 1995; Reedy et al. 2018; Shams-White et al. 2023). However, applications of such comprehensive indicators in studies of disease-diagnosis impacts have remained limited.

The current study intends to fill this gap by analyzing data from China. Exploiting two explicitly set biomarker cutoffs for hypertension diagnoses—140 and 90 mmHg, respectively, for SBP and DBP readings (Unger et al. 2020), we perform a two-dimensional regression-discontinuity analysis to estimate the impact of one's *first-ever* hypertension diagnosis on his/her diet structure, regarding both dietary diversity and diet balance. Drawing a panel dataset of 9355 adults from a longitudinal survey covering 12 Chinese provinces, the China Health and Nutrition Survey (CHNS), our analysis reveals both desirable and undesirable hypertension-diagnosis effects. Diagnoses based on SBP readings exerted some beneficial impact on Chinese adults' diet (e.g., by reducing their fat intake) about 3 years later. However, SBP-based diagnoses also undermine their diet balance by reducing their intake of foods already being *under*-consumed. Fruits, for example, which have been inadequately consumed by Chinese adults and ought to be consumed more by hypertensive patients (World Health Organization 2021), were consumed even less post-diagnosis. This finding suggests that even though disease diagnoses can provide patients with accurate professional health information, patients may misunderstand and misinterpret this information, thus leading to unhealthy dietary responses.

Material and method

Ethical approval. The analysis reported this study was carried out based on publicly available, deidentified longitudinal datasets from the China Health and Nutrition Survey (CHNS): https://www.cpc.unc.edu/projects/china/data/datasets. The procedures used adhere to the tenets of the Declaration of Helsinki. The study gained ethical approval from the Ethics Committee of the College of Economics and Management at China Agricultural University (CAU-CEM-IRB; approval number: HC0028401A).

CHNS data and study sample. Our analysis draws a longitudinal dataset from the CHNS, a project conducted through collaborations between the University of North Carolina and the Chinese Center for Disease Control and Prevention (CDC). The initial survey, conducted in 1989, involved nine Chinese provinces with vastly different natural conditions, cultural characteristics, and levels of socio-economic development: from south to north, Guangxi, Guizhou, Hunan, Hubei, Jiangsu, Henan, Shandong, Liaoning, and Heilongjiang (https://www.cpc.unc.edu/projects/china/about/proj_desc/chinamap). A multistage (county/city→village/urban communities→households) random sampling strategy was employed to select target households for studying nutrition and health issues in China. Follow-up surveys were conducted about every three years (1991, 1993, 1997, 2000, 2004, 2006, 2009, 2011, and 2015). Three municipalities, Beijing, Shanghai, and Chongqing, participated in the project in 2011; three more provinces, Yunnan, Shaanxi, and Zhejiang, joined in 2015. Approximately 25,000 individuals from 4400 households have participated in the CHNS. While we are unable to claim the CHNS to be representative of the Chinese population, it does provide the most widely used data for studying health and nutrition issues in China.[3]

Given the study's purpose to identify the effect of *first-ever* hypertension diagnoses documented in a survey wave on Chinese adults' diet structure observed in the next wave, we focus on data from six waves conducted between 1997 and 2011. We set the baseline at 1997 to circumvent potential sample-selection bias because physical examination records in the previous wave (1993) were missing for about 50% of the observations. The end line, 2011, was chosen because information on respondents' food intakes in 2015 had not been released by the time of writing.

To construct an analytic sample suitable for our analysis, we applied several additional sample restrictions. First, since hypertension is rarely found among nonadults, we dropped observations aged under 18 from the sample. Second, given our interest in estimating the impact of *first-time-ever* hypertension diagnoses observed in a given wave, say, 2004, we excluded 9381 observations with reported previous diagnoses by the previous wave, say, 2000. Thus, in the resultant sample, all individuals with a hypertension diagnosis for the first time in wave t also appeared in wave t-1 with below-threshold SBP and DBP readings. To avoid double counting "treated" individuals in the analysis, we kept only the earliest post-diagnosis observation for each individual diagnosed with hypertension. Finally, to avoid including individuals whose blood pressure readings are much higher or lower than the diagnostic cutoffs in the RD analysis, we kept only those with blood pressure readings within the ±40 mmHg intervals of the respective thresholds in the sample. The final sample consists of 9355 individuals with 20,277 individual-wave observations.

Treatment and assignment variables. The two treatment assignment variables in our RD analysis, individuals' systolic blood pressure (SBP) and diastolic blood pressure (DBP) readings, were collected during physical examinations administered by the CHNS medical team as part of the interview. During the physical examination in each round, medical professionals measured each respondent's SBP and DBP readings on three consecutive days using medical equipment. A respondent would be diagnosed as hypertensive if this person's 3-day average of SBP or DBP reading exceeded the corresponding threshold. Upon

Table 1 Food groups involved in the Diet Balance Index.

(1) Chinese Dietary Guidelines (Chinese Nutrition Society 2008)	(2) Components in the Diet Balance Index (DBI)	(3) Food groups and corresponding intake thresholds involved in the Diet Diversity Score (DDS)
DG-1: Eat a variety of foods, with cereals as the staple and a certain amount of coarse grains.	DBI-1: Diet variety (all groups in column 3)	
	DBI-2: Grains	DDS-1: Rice and rice products (25 g) DDS-2: Wheat and wheat products (25 g) DDS-3: Corn, coarse grains, starchy roots, and related products (25 g)
DG-2: Consume plenty of vegetables, fruits, and tubers.	DBI-3: Vegetables and fruits	DDS-4: Dark-colored vegetables (25 g) DDS-5: Light-colored vegetables (25 g) DDS-6: Fruits (25 g)
DG-3: Consume milk, soybean, and their products every day.	DBI-4: Soybean and dairy products	DDS-7: Soybean and soybean products (5 g) DDS-8: Milk and dairy products (25 g)
DG-4: Consume proper amounts of fish, poultry, eggs, and lean meat.	DBI-5: Animal protein	DDS-9: Red meat (livestock products) (25 g) DDS-10: Poultry and games (25 g) DDS-11: Egg (25 g) DDS-12: Aquatic products (25 g)
DG-5: Reduce cooking oil; choose a light diet that is also low in salt. DG-6: If you drink alcoholic beverages, do so in limited amounts.	DBI-6: Cooking oil, salt, and alcoholic beverages	
DG-7: Avoid overeating and exercise every day to maintain healthy body weight.	n/a	
DG-8: Rationally distribute the daily food intake among the three meals. If you take snacks, do so properly.	n/a	
DG-9: Drink sufficient water every day; rationally choose beverages.	DBI-7: Drinking water	
DG-10: Avoid unsanitary and spoiled foods.	n/a	

Source: He et al. (2009). "Drinking water" is not available in the CHNS data.

completion of the physical examination, the CHNS medical team notified each respondent about the results, along with his/her SBP or DBP readings, but the team did not explicitly inform those hypertensive respondents about the severity of the condition, nor was there any mention of specific prescriptions for hypertension treatment during the survey (Zhao et al. 2013).

Outcome variables. A unique feature of the CHNS that is important for our study is that the surveys collected a wealth of information on sampled respondents' food consumption regarding over 1500 food items and/or local dishes during three consecutive interview days. This feature enables us to construct dietary indicators that reflect multiple dimensions regarding one's diet structure.

Macronutrient intakes. Matching information on each respondent's intake of specific food items/local dishes with the nutrients they contain provided in *Chinese Food and Nutrition Tables* (Yang 2002, 2004), the CHNS team calculated each respondent's daily intake of carbohydrates, protein, fat, and total calories. While nutrient-intake indicators capture important aspects of one's diet structure, they are too crude for detecting more granular patterns. For example, red meat and fish are both excellent protein sources, but fish consumption is generally considered healthier (English et al. 2004; Kolahdooz et al. 2010). To capture changes in one's diet structure induced by hypertension diagnoses more granularly, indicators reflecting structural features beyond nutrient intakes are needed.

Dietary diversity score (DDS). One such indicator is the dietary diversity score. The DDS counts how many food groups a person consumed on the previous day out of a set of essential groups recommended by official dietary guidelines. Originally developed based on guidelines provided by the Food and Agriculture Organization of the United Nations (2011), the DDS was later adapted to fit the context of China by the Chinese CDC and the Chinese Nutrition Society (2008, 2022). The Chinese version of DDS involves 12 essential food groups (Table 1, column 3).[4] If a person i's intake of foods in group j, $food_{ij}$(say, "animal protein"), in the past 24 h attains its guideline-recommended level, g_j, this group will receive a score of one; a score of zero will be assigned otherwise. The total DDS score sums across all 12 groups for each respondent i:

$$DDS_i = \sum_{j=1}^{12} 1\left(food_{ij} \geq g_j\right), \tag{1}$$

with a higher score indicating a more diverse diet structure.[5]

Although the DDS serves as a "quick and easy" diet-structure measure, information on whose components can be easily collected in many types of surveys, including household-based (Kassie et al. 2020) and school-based surveys (Chen et al. 2020), this indicator has two major limitations. First, it does not consider the quantity of food consumed above or below the guideline-recommended level, lacking the power to measure the degree of over-eating and under-eating behaviors. Second, it does not penalize the sorts of food consumption that can potentially undermine one's health, such as the consumption of alcoholic beverages, added salt, and cooking oil. Fortunately, the wealth of information collected by the CHNS enables us to develop a more comprehensive indicator to address both limitations.

Diet balance index (DBI). The diet-balance index was originally designed based on the guiding principles for constructing indicators widely adopted in dietary patterns studies in Western countries, such as the Diet Quality Index and Healthy Eating

Table 2 Summary statistics by hypertension status.

Sample	(1) All		(2) Hypertensive (SBP ≥ 140 or DBP ≥ 90)		(3) Non-hypertensive (SBP < 140 & DBP < 90)		(4) Difference = (3)-(2)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SE
A. Outcomes (t + 1)								
Calorie (kcal)	2193.22	[733.54]	2113.01	[782.89]	2209.24	[722.24]	−96.23***	(14.73)
Carbohydrate (g)	312.81	[113.36]	295.93	[109.27]	316.18	[113.86]	−20.25***	(2.28)
Fat (g)	72.60	[49.09]	71.94	[59.08]	72.73	[46.84]	−0.80	(0.99)
Protein (g)	66.36	[23.97]	64.13	[24.11]	66.81	[23.92]	−2.68***	(0.48)
Diet Diversity Score (DDS)	5.00	[1.30]	4.98	[1.31]	5.00	[1.29]	−0.02	(0.03)
Diet Balance Index (DBI)	−17.15	[8.91]	−17.13	[9.12]	−17.16	[8.87]	0.03	(0.18)
DBI-O: over-consumption	14.41	[5.10]	14.37	[5.18]	14.42	[5.08]	−0.05	(0.10)
DBI-U: under-consumption	−31.54	[5.78]	−31.49	[5.81]	−31.55	[5.78]	0.07	(0.12)
B. Pre-determined factors (t)								
Age (year)	46.79	[14.32]	54.56	[13.83]	45.29	[13.93]	9.27***	(0.27)
Male (1 = yes)	0.47	[0.50]	0.51	[0.50]	0.47	[0.50]	0.05***	(0.01)
Education (years)	6.88	[4.15]	5.97	[4.36]	7.06	[4.08]	−1.09***	(0.08)
Rural residents (1 = yes)	0.71	[0.46]	0.71	[0.45]	0.71	[0.46]	0.001	(0.01)
Household income per capita (CPI-adjusted)	8056.03	[10,838.57]	8490.77	[10,483.47]	7972.58	[10,903.73]	518.19*	(208.12)
Household size	3.84	[1.53]	3.69	[1.61]	3.96	[1.52]	−0.23***	(0.03)
Working (1 = yes)	0.69	[0.46]	0.58	[0.49]	0.71	[0.45]	−0.14***	(0.01)
Married (1 = yes)	0.84	[0.37]	0.84	[0.36]	0.84	[0.37]	−0.002	(0.01)
Han (1 = yes)	0.86	[0.35]	0.88	[0.32]	0.85	[0.35]	0.03***	(0.01)
N	20,277		3270		17,007			

Data source: China Health and Nutrition Survey (1997, 2000, 2004, 2006, 2009, and 2011).
Hypertensive individuals refer to those with SBP readings ≥ 140 mmHg or with DBP readings ≥ 90 mmHg. Non-hypertensive individuals refer to those with SBP readings < 140 mmHg and DBP readings < 90 mmHg.
Standard deviations (errors) are reported in brackets (parentheses).
***$p < 0.01$; *$p < 0.1$.

Index (e.g., Reedy et al. 2018; Shams-White et al. 2023). Tailored to the context of China by experts from the Chinese CDC (He et al. 2009), the DBI consists of seven sub-indicators involved in the *Chinese Dietary Guidelines*.

As shown in Table 1, the first sub-indicator (DBI-1) mirrors the DDS, capturing one's diet diversity as a dimension of diet balance (column 2). The second through seventh sub-indicators (DBI-2 ~ 7) separately capture a person's over- and under-consumption of animal products, condiments/alcoholic beverages, dairy/soybean products, drinking water, grains, and vegetables/fruits (column 2),[6] each having 1 ~ 3 even more specific sub-indicators (column 3).

In constructing the DBI, we first assign a group-specific score, s_j, to a person *i*'s consumption of a given food group *j*, according to the criteria provided by the Chinese CDC; the criteria differ by the energy intake levels recommended for individuals in different categories of gender, age, and daily activity intensity. Supplementary Appendix Table A1 shows the criteria for an adult with a recommended energy intake level of 2000 kcal/day.[7] The assigned score $s_j > 0$ if person *i*'s intake of food group j is higher than the recommended level and $s_j \leq 0$ if it falls below that level. All food group-specific scores are then added together to produce a total DBI score:

$$DBI_i = \sum_{j=1}^{6} s_j, \qquad (2)$$

with $DBI_i = 0$ suggesting a *balanced* diet, $DBI_i > 0$ indicating over-consumption of at least some foods, and $DBI_i < 0$ suggesting underconsumption of some foods.

Note that some food groups, such as "grains," have both positive and negative scores, which means both over- and under-consumption of these foods could be captured in the DBI. In contrast, some other foods, such as "salt" and "cooking oil," have only non-negative scores, meaning that according to the *Chinese Dietary Guidelines*, they could be over-consumed but would never be considered under-consumed. Foods such as "fish" and "vegetables" have only non-positive scores—the *Guidelines* always encourage the consumption of these foods.

Note also that over-consumption and under-consumption of foods may cancel each other out so that the resultant overall DBI is close to zero (suggesting a balanced diet). To address this problem, we compute two DBI sub-indicators that assess these two dimensions separately: the over-consumption score (DBI-O) adds up the first-step positive group-specific scores, and the under-consumption score (DBI-U) sums the negative scores. The ranges of these sub-indicator scores reflect the "fair" (1 ~ 6), "low" (7 ~ 13), "modest" (14 ~ 19), and "high" (20 ~ 32) levels of food over-consumption, and "fair" (−1 ~ −12), "low" (−13 ~ −24), "modest" (−25 ~ −36), and "high" (−37 ~ −60) levels of under-consumption (He et al. 2009).

Intakes of essential food groups. Finally, to help understand and interpret the results obtained based on the dietary diversity score and diet balance index, the amount and incidence of one's intake of individual food groups involved in the construction of these two indicators, such as animal protein, dairy products, and grains, are also examined in some analyses below.

Statistical models. A diagnosis of hypertension is "triggered" when one's blood pressure (systolic or diastolic blood pressure) reading, an intrinsically continuous treatment assignment variable (or "running/forcing" variable), passes a medically specified cutoff smoothly. This diagnostic rule suggests a regression discontinuity approach to estimating the impact of hypertension diagnosis. Under the identifying assumption that persons with blood pressure readings just above and just below the diagnostic cutoffs are similar in all aspects except the diagnostic results, a well-implemented RD analysis can effectively eliminate the influence of potential confounding factors. Yet, unlike standard RD designs that involve a single cutoff, hypertension diagnoses involve two cut-offs, 140 mmHg for SBP readings and 90 mmHg for DBP readings, which call for a two-dimensional RD approach.

A two-dimensional RD design can be implemented in multiple ways. The first views one set of blood pressure readings as the treatment assignment variable, ignoring the other set. For example, Zhao et al. (2013) focus on SBP readings while ignoring DBP readings:

$$\delta^{S|Dall} = \lim_{s\nearrow 140} E\left[y_{it}\middle|S_{i,t-1}=140\right] - \lim_{s\searrow 140} E\left[y_{it}\middle|S_{i,t-1}=140\right], \tag{3}$$

where y_{it} denotes an outcome variable (reflecting an essential aspect of diet structure) of individual i recorded in survey wave t; $S_{i,t-1}$ denotes i's SBP readings measured in wave t-1; $\lim_{x\nearrow c} E\left[y|x=c\right]$ and $\lim_{x\searrow c} E\left[y|x=c\right]$ are, the left- and right-hand limits of y, respectively, as the assignment variable x approaches the diagnostic cut-off c continuously.

Figure 1, plotting SBP readings against DBP readings, illustrates this design: observations in quadrants II and IV form the "treated" group; those in quadrants I and III serve as the "control" group—estimation is based on observations close to the 140-mmHg cutoff say, within the (130,150) interval. In effect, the estimator $\delta^{S|D\_all}$ (Eq. (2)) identifies the impact of hypertension diagnosis based on the SBP rule, *averaging over DBP readings* at the 140-mmHg SBP cutoff.

To avoid possible contamination of DBP-based hypertension diagnoses, Dai et al. (2022) adopted a refined approach, excluding observations with DBP readings above the 90-mmHg DBP cutoff (Fig. 1, quadrants I and II) when estimating Eq. (3):

$$\begin{aligned}\delta^{S|D<90} &= \lim_{s\nearrow 140} E[y_{it}|S_{i,t-1}=140; D_{i,t-1}<90] \\ &= \lim_{s\searrow 140} E[y_{it}|S_{i,t-1}=140; D_{i,t-1}<90],\end{aligned} \tag{4}$$

where $D_{i,t-1}$ is person i's DBP readings in wave t-1. Similar to the estimator $\delta^{S|D\_all}$ defined in Eq. (3), $\delta^{S|D<90}$ defined in Eq. (4) is unable to identify the impact of DBP-based diagnoses.

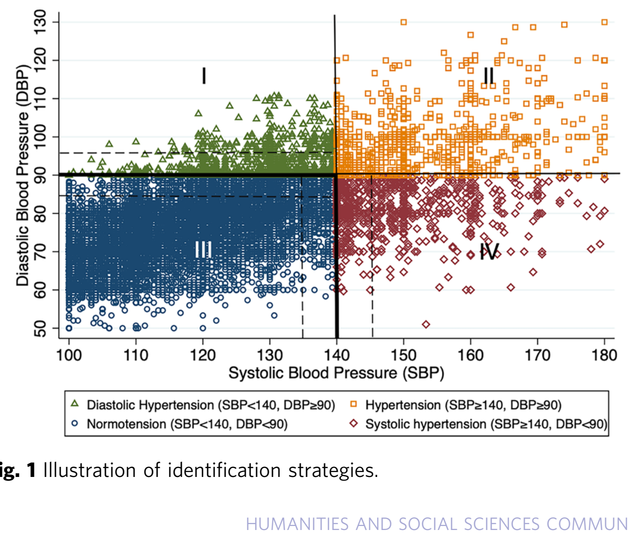


Fig. 1 Illustration of identification strategies.

Another approach, proposed by Wong et al. (2013), identifies the impacts of both DBP- and SBP-based diagnoses. In essence, this method performs two *separate* RD analyses, each based on one diagnostic cutoff, excluding observations diagnosed with hypertension based on the other cutoff. For example, one first performs a one-dimensional RD analysis based on the $\delta^{S|D<90}$ estimator (Eq. (4)), using SBP readings as the running variable while excluding observations with above-90-mmHg DBP readings. One then performs another RD analysis using DBP readings as the running variable while dropping observations with above-140-mmHg SBP readings:

$$\begin{aligned}\delta^{D|S<140} &= \lim_{d\nearrow 90} E[y_{it}|D_{i,t-1}=90; S_{i,t-1}<140] \\ &= \lim_{d\searrow 90} E[y_{it}|D_{i,t-1}=90; S_{i,t-1}<140].\end{aligned} \tag{5}$$

Estimations based on the estimators $\delta^{S|D<90}$ (Eq. (4)) and $\delta^{D|S<140}$ (Eq. (5)) naturally exclude observations who are hypertensive by *both* rules (Fig. 1, quadrant II).

The fourth method, devised by Imbens and Zajonc (2011), exploits both diagnostic rules simultaneously by defining the running variable based on the "distance" to the diagnostic "boundary" determined by both diagnostic rules (Fig. 1: thickened polyline). Formally, for individual i, whose blood pressure was measured in wave t, the distance from his/her blood pressure readings ($Dist_{it}$) to the diagnostic boundary (b) is:

$$Dist_{it} = \begin{cases} \sqrt{(S_{it}-140)^2 + (D_{it}-90)^2}, & if\ S_{it} \geq 140,\ D_{it} \geq 90; \\ D_{it} - 90, & if\ S_{it} < 140,\ D_{it} \geq 90; \\ -\min(140 - S_{it}, 90 - D_{it}), & if\ S_{it} < 140,\ D_{it} < 90; \\ S_{it} - 140, & if\ S_{it} \geq 140,\ D_{it} < 90; \end{cases} \tag{6}$$

Correspondingly, the associated RD estimator is a one-dimensional design:

$$\delta^{SD} = \lim_{s\nearrow 0} E\left[y_{it}\middle|Dist_{i,t-1}=b\right] - \lim_{s\searrow 0} E\left[y_{it}\middle|Dist_{i,t-1}=b\right]. \tag{7}$$

Estimates obtained using all four estimators discussed above will be reported below for comparison purposes. However, our preferred estimators are $\delta^{S|D<90}$ defined in (4) and $\delta^{D|S<140}$ defined in (5), because they enable us to learn how Chinese adults may respond to the SBP and DBP diagnostic rules differently.

All RD estimates reported in the following section are non-parametric ones obtained with STATA 17. The optimal bandwidth for each estimation is chosen to minimize the mean squared error (MSE) at the corresponding diagnostic cutoff (Imbens and Kalyanaraman 2012; Calonico et al. 2014). To account for the potential influence of unobserved local medical conditions, we cluster standard errors at the community (urban community or rural village) level.[8]

Empirical results

Descriptive findings. Table 2 portrays the profile of sampled adults with all individual-wave observations (Panel B). Slightly more than half of the sampled adults are female. The average sampled adult was 46.8 years old, with slightly less than seven years of education, and lived in a household with a per capita income of about 8000 Chinese yuan (roughly 1300 U.S. dollars at 2015 prices). Nearly 70% of all individuals held a rural residential permit (*Hukou*); about the same proportion worked in the survey year.

Panel A of the table reports summary statistics of diet-related outcomes for hypertensive (column 2) and non-hypertensive adults (column 3) and their differences (column 4). While

Table 3 Non-parametric RD estimates of hypertension-diagnosis effects on daily nutrient intake (individuals ≥ 18 years old).

Estimator	Mean [SD]	RD estimates based on SBP readings (cutoff: SBP = 140)		RD estimates based on DBP readings (cutoff: DBP = 90)		RD estimates based on the distance to the joint cutoff (SBP = 140, DBP = 90)
	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Non-hypertensive (SBP < 140, DBP < 90)	Diastolic normal (DBP < 90)	Full	Systolic normal (SBP < 140)	Full	Full
A. Energy (kcal)	2209.24	−57.335	−46.668	−97.036	23.963	45.152
	[722.24]	(78.649)	(65.697)	(144.083)	(77.995)	(79.879)
Optimal bandwidth		16.004	13.158	5.991	12.054	23.033
B. Carbohydrates (g)	316.18	23.005	14.902	0.926	4.937	4.763
	[113.86]	(15.776)	(11.930)	(17.195)	(12.526)	(18.656)
Optimal bandwidth		13.414	13.524	7.995	11.772	16.332
C. Fat (g)	72.73	−12.855***	−9.441***	−7.419	−0.753	−2.389
	[46.84]	(4.394)	(3.312)	(11.842)	(6.463)	(4.664)
Optimal bandwidth		12.455	15.938	6.665	11.155	22.270
D. Protein (g)	66.81	−4.326	−3.838	2.313	1.320	1.358
	[23.92]	(3.136)	(2.426)	(3.506)	(1.885)	(2.940)
Optimal bandwidth		14.681	13.721	5.600	15.153	20.721
N	14,850	15,680	17,817	16,054	17,817	17,817

Data were drawn from the China Health and Nutrition Survey (1997, 2000, 2004, 2006, 2009, and 2011). Analytical samples are restricted to observations within ±40 mmHg intervals of the respective cutoffs. The sample size (*N*) is the same for all panels in each column of the table; the analytical sample size for each estimate is determined by the optimal bandwidth chosen by minimizing the mean squared error (Calonico et al. 2014). Non-parametric RD estimates using the triangular kernel and the associated optimal bandwidths are reported. Standard deviations are reported in brackets. Robust standard errors are reported in parentheses, clustered at the community level.
***$p < 0.01$.

hypertensive adults consume significantly less food and nutrition (especially carbohydrates and protein) than non-hypertensive adults, no systematic differences in diet structure are observed between the two groups (column 4). Yet, it should be noted that these contrasts were observed without addressing the potential founding effect of other factors. As shown in panel B, hypertensive and non-hypertensive adults differ significantly in many demographic and socio-economic dimensions, such as gender, age, and years of education, which may contribute to the differences in their food choices. To address unobserved confounding and thus infer causation from hypertension diagnosis to one's food choices and diet structure, we apply the two-dimensional RD design developed in section 2.5 in the analysis reported below.

Main findings

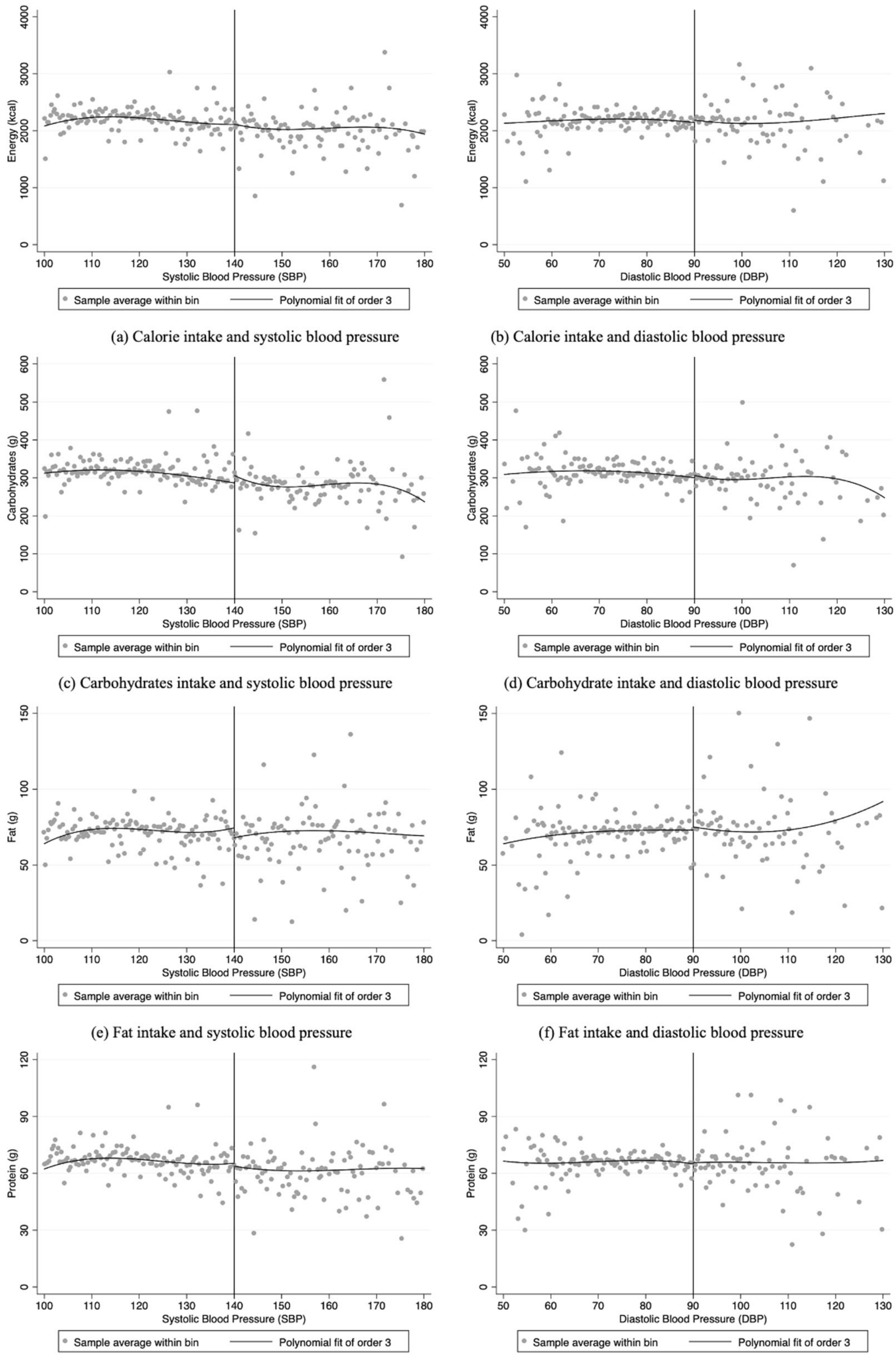
Macronutrient intakes: replication analysis. Table 3 reports the estimated impacts of hypertension diagnoses on Chinese adults' macronutrient intakes. While these variables have been examined by Zhao et al. (2013), our reanalysis is still informative. For one, Zhao et al. adopted only the estimator $\delta^{S|D\_all}$ (Eq. (2)), which identifies the effects of SBP-based hypertension diagnoses, ignoring potential contamination from DBP-based diagnoses. In contrast, we applied all four estimators discussed in section 2.5, which helps assess the robustness of the findings of Zhao et al. (2013). For another, our analysis incorporates two more recent waves (2009 and 2011) of the CHNS data so that we can see whether their findings, if valid, persist over time.

Table 3, column 2, presents RD estimates based on the $\delta^{S|D<90}$ estimator (Eq. (3)), which exploits the "jump" in the number of diagnosed hypertension cases at the SBP cutoff (140 mmHg) to identify the effects of SBP-based diagnoses among Chinese adults with below 90-mmHg BDP readings. Notably, first-ever SBP-based hypertension diagnoses *reduced* BDP-normal adults' fat intake significantly by 12.9 g/day about 3 years later (row C). By contrast, no significant impact of SBP-based hypertension diagnosis was detected for other indicators (protein, carbohydrate, or calorie intake). Figure 2, left panels (a, c, e, and g), visualizes the above findings.

This pattern aligns with the findings of Zhao et al. (2013). The estimates reported in Table 3, column (2), are comparable to those reported in column (3), which are obtained based on the estimator $\delta^{S|D\_all}$ (2) adopted by Zhao et al. (2013). The similarity between estimates reported in these two columns suggests that potential contamination of DBP-based diagnoses is unlikely to be a serious concern when estimating the effects of SBP-based diagnoses. Corroborating this finding, DBP-based estimates (Table 3, column 4) suggest that Chinese adults' macronutrient intake hardly responds to the DBP rule.[9] Also, the Imbens-Zajonc estimator $\delta^{S\&D}$ (Eq. (6)), which incroporates both diagnostic rules, does no find any notable hypertension-diagnosis effect (Table 3, column 6). Intuitively, since sampled individuals are not responsive to DBP-based diagnoses, the analytic sample contains too much "noise" along the DBP dimension, rendering the effects of SBP-based diagnoses much less pronounced.

Diet structure. While the above analyses suggest that SBP-based hypertension diagnoses improve one's diet somewhat (through a reduction in fat intake), the nutrition-intake outcomes examined may be too crude for detecting more granular structural patterns. In particular, fat may come from different foods, but the above analyses lack the power to pinpoint whether the reduction in fat intake was indeed driven by dietary changes that benefit one's health. Thus, we resort to the two diet-structure indicators constructed above, DDS and BDI, below.

Table 4, column (2), presents RD estimates based on the $\delta^{S|D<90}$ estimator (Eq. (2)), i.e., the impact of SBP-based diagnoses on dietary diversity scores (DDS) for those with normal BDP. The estimate suggests that an SDP-based diagnosis *reduced* an average diastolic-normal adult's diet diversity by 0.6 out of the 12 recommended food groups (column 2, row A). Given non-hypertensive adults' average DDS of 0.5, this drop is substantial. Echoing the drop in diet



(a) Calorie intake and systolic blood pressure
(b) Calorie intake and diastolic blood pressure
(c) Carbohydrates intake and systolic blood pressure
(d) Carbohydrate intake and diastolic blood pressure
(e) Fat intake and systolic blood pressure
(f) Fat intake and diastolic blood pressure
(g) Protein intake and systolic blood pressure
(h) Protein intake and diastolic blood pressure

Fig. 2 Relationships between blood pressure and daily nutrient intake. a Calorie intake and systolic blood pressure; **b** calorie intake and diastolic blood pressure; **c** carbohydrates intake and systolic blood pressure; **d** carbohydrates intake and diastolic blood pressure; **e** fat intake and systolic blood pressure; **f** fat intake and diastolic blood pressure; **g** protein intake and systolic blood pressure; **h** protein intake and diastolic blood pressure.

Table 4 Non-parametric RD estimates of the effects of hypertension diagnoses on diet structure (individuals ≥ 18 years old).

	(1)	(2)	(3)	(4)	(5)	(6)
Estimator	**Mean [SD]**	**RD estimates based on SBP readings (cutoff: SBP = 140)**		**RD estimates based on DBP readings (cutoff: DBP = 90)**		**RD estimates based on the distance to the cutoff boundary (SBP = 140, DBP = 90)**
Sample	**Non-hypertensive (SBP < 140, DBP < 90)**	**Diastolic normal (DBP < 90)**	**Full**	**Systolic normal (SBP < 140)**	**Full**	**Full**
A. Dietary Diversity Score (DDS)	5.00	−0.610***	−0.478***	0.159	−0.036	−0.329*
	[1.29]	(0.179)	(0.149)	(0.241)	(0.162)	(0.199)
Optimal bandwidth		11.650	12.455	4.856	10.41	14.466
N	14,873	15,704	17,847	16,081	17,847	17,847
B. Diet Balance Index (DBI)	−17.16	−2.542**	−2.378**	−0.801	−0.196	−1.639
	[8.87]	(1.150)	(0.991)	(1.730)	(0.977)	(1.304)
Optimal bandwidth		12.106	12.212	5.169	11.115	14.907
N	14,277	15,074	17,147	15,444	17,147	17,147
C. DBI-O: over-consumption	14.42	−0.029	−0.328	−1.125	−0.269	−0.404
	[5.08]	(0.543)	(0.501)	(0.918)	(0.504)	(0.666)
Optimal bandwidth		14.468	13.789	5.302	11.152	17.495
N	14,277	15,074	17,147	15,444	17,147	17,147
D. DBI-U: under-consumption	−31.55	−2.278***	−1.870***	0.521	0.039	−0.816
	[5.78]	(0.805)	(0.655)	(1.188)	(0.710)	(0.853)
Optimal bandwidth		12.31	12.786	4.685	10.486	15.447
N	14,616	15,432	17,541	15,804	17,541	17,541

Data were drawn from the China Health and Nutrition Survey (1997, 2000, 2004, 2006, 2009, and 2011). Analytical samples are restricted to observations within ±40 mmHg intervals of the respective cutoffs. The actual sample size used for each estimate is determined by the optimal bandwidth chosen by minimizing the mean squared error to produce a lower bias (Calonico et al., 2014). Non-parametric RD estimates using the triangular kernel and the associated optimal bandwidth are reported. Standard deviations are reported in brackets. Robust standard errors are reported in parentheses, clustered at the community level.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

diversity, column (2), row B, shows that an SDP-based hypertension diagnosis reduced an average Chinese adult's overall diet balance (DBI) score by 2.5 points. Given the significant reduction in one's fat intake resulting from SDP-based diagnoses, these drops in diet-structure indicators are unsurprising.

However, separately examining the over-consumption (column 2, row C) and under-consumption (column 2, row D) dimensions of diet balance reveals a more worrisome pattern. The reduction in the overall DBI (column 2, row B) is almost entirely driven by that in DBI-U (under-consumption) scores (column 2, row C). In other words, SBP-based hypertension diagnoses reduced one's consumption of foods that *were already inadequately consumed*, a finding that has largely been overlooked in previous studies. The left panels of Fig. 3 visualize these effects. Estimates based on the Zhao et al. estimator, $\delta^{S|D\_all}$ (Eq. (2)), reveal a similar pattern (column 3).

Reporting estimates of $\delta^{S|D<140}$ based on the DBP rule (Eq. (3)), Table 4, column (4) reveals no significant effects of hypertension diagnosis at the 90-mmHg cutoff. This finding largely aligns with those of Dai et al. (2022), that while Chinese adults reduced their fat intake and smoking incidence significantly upon SBP-based diagnoses, they were hardly responsive to SBP-based diagnoses. Again, the Imbens-Zajonc estimator, $\delta^{S\&D}$ (Eq. (6)), which uses the distance to the joint boundary to define "diagnosis," fails to detect any significant effect.

Given the relative performance of all four estimators in detecting the effects of hypertension diagnoses, we base our analyses below on $\delta^{D|S<90}$, the DBP-based estimator.

Consumption of specific foods. To understand what changes in the consumption of specific foods led to the drops in DDS and DBI scores, Table 5 explores how SBP-based hypertension diagnoses impact the incidence (column 5) and amount (column 2) of one's consumption of specific food groups. Two important findings emerge. First, as column (5) shows, SBP-based diagnoses (at least marginally) significantly *reduced* the incidence of one consuming the following five food groups: fruits, light-colored vegetables, livestock products, rice, other grains, and soybean products,[10] which helps explain the drops in DDS and DBI scores. Note that these reductions in food consumption (except perhaps that in livestock consumption) were not detected in previous studies, nor were they captured in our findings above based on nutrient-intake indicators (Table 3). Note also that while SBP-based diagnoses *raised* Chinese adults' wheat consumption (column 2), this increase did not change their diet diversity notably (column 5), suggesting that such increases occurred at the "intensive margin" rather than around the recommended intake threshold.

Second, and more importantly, except for the reduction in livestock consumption, other hypertension diagnosis-induced changes in food intake are *undesirable* in general. As Table 5, column (7) indicates, livestock products were originally *over-consumed* among non-hypertensive individuals; in contrast, most other foods, e.g., light-colored vegetables, fruits, and soybean products, were initially *under-consumed*. As such, while the reduction in livestock consumption may be deemed desirable, drops in vegetable, fruit, and soybean consumption could worsen hypertensive individuals' health—in fact, eating more fruits and vegetables is recommended by the World Health Organization (2021) for hypertension prevention and management. The increase in wheat intake may also be undesired, as grains had already been *over-consumed* among non-hypertensive individuals (column 7). These undesirable effects imply that unlike what has been almost uniformly concluded by previous studies, (SBP-based) hypertension diagnoses may *not* necessarily lead to healthier (eating) behavior.

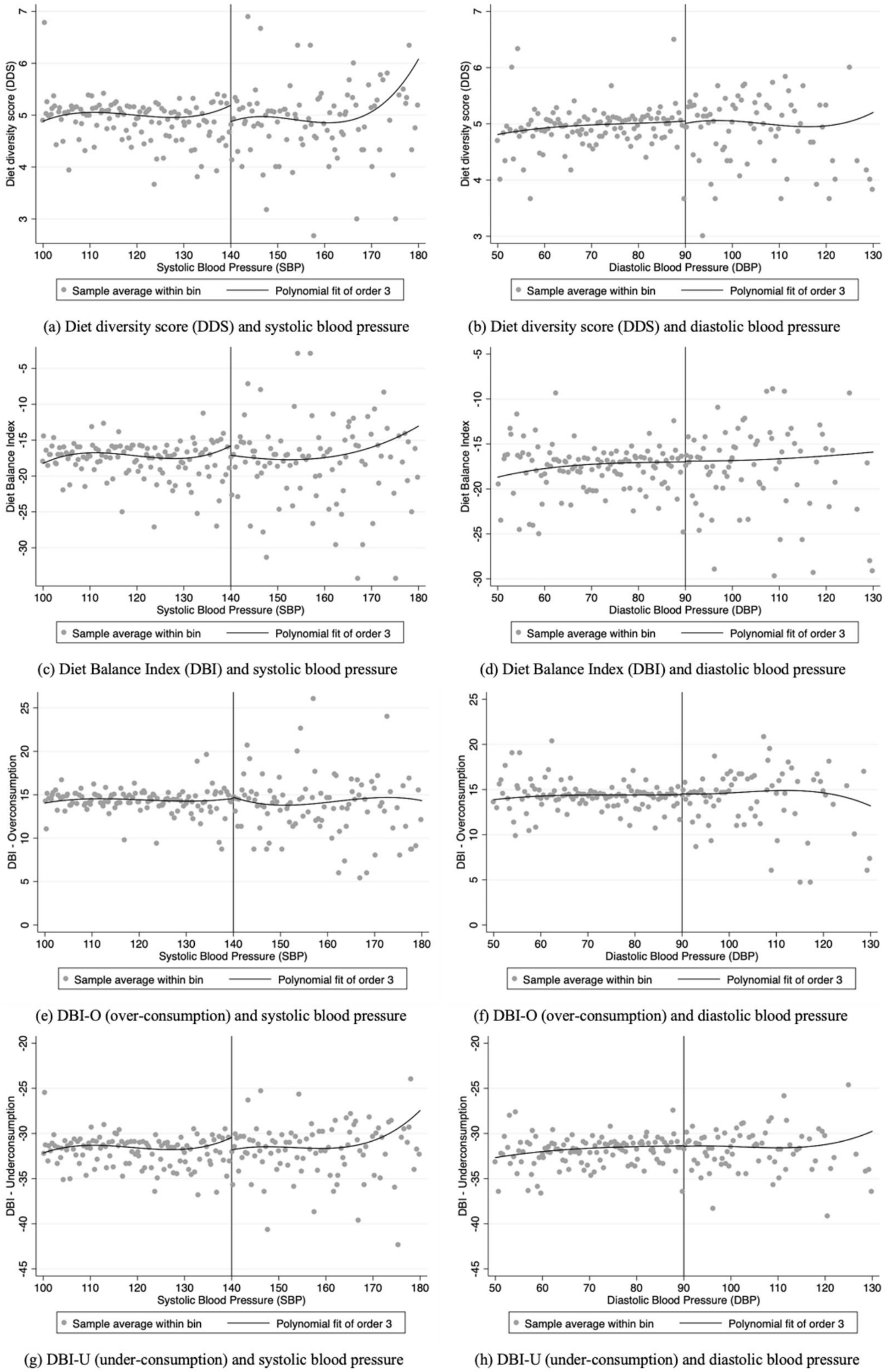


Fig. 3 Relationships between blood pressure and diet structure. **a** Diet diversity score (DDS) and systolic blood pressure; **b** diet diversity score (DDS) and diastolic blood pressure; **c** diet balance index (DBI) and systolic blood pressure; **d** diet balance index (DBI) and diastolic blood pressure; **e** DBI-O (over-consumption) and systolic blood pressure; **f** DBI-O (over-consumption) and diastolic blood pressure; **g** DBI-U (under-consumption) and systolic blood pressure; **h** DBI-U (under-consumption) and diastolic blood pressure.

Table 5 Non-parametric RD estimates of the effects of hypertension diagnosis on daily consumption of food groups.

	Amount consumed			Food groups in the DDS			DBI components		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimator	**Mean [SD]**	**RD estimates (cutoff: SBP = 140)**	**Optimal bandwidth**	**Mean [SD]**	**RD Estimates (cutoff: SBP = 140)**	**Optimal bandwidth**	**Mean [SD]**	**RD Estimates (cutoff: SBP = 140)**	**Optimal bandwidth**
Sample	**Non-hypertensive (SBP < 140, DBP < 90)**	**Diastolic normal (DBP < 90)**		**Non-hypertensive (SBP < 140, DBP < 90)**	**Diastolic normal (DBP < 90)**		**Non-hypertensive (SBP < 140, DBP < 90)**	**Diastolic normal (DBP < 90)**	
DBI-1: Dietary variety								−0.610*** (0.179)	
DBI-2: Grains (g)									
Rice	259.93 [176.46]	−19.964 (23.446)	16.36	0.82 [0.34]	−0.078* (0.046)	19.20	5.16 [4.97]	0.782 (0.477)	16.50
Wheat	143.01 [157.00]	52.079** (22.833)	14.16	0.62 [0.41]	0.050 (0.053)	17.68			
Other grains	68.37 [85.93]	−4.009 (9.366)	16.98	0.43 [0.39]	−0.092* (0.051)	12.09			
DBI-3: Vegetables & fruits (g)									
Dark vegetables	60.36 [86.00]	13.717 (9.423)	13.33	0.33 [0.37]	0.042 (0.042)	14.23	−1.74 [1.37]	0.057 (0.150)	17.30
Light-colored veg.	269.82 [160.89]	−1.325 (19.445)	12.90	0.88 [0.23]	−0.037* (0.021)	16.48			
Fruits	35.90 [83.55]	−17.316* (9.640)	15.03	0.17 [0.33]	−0.075** (0.037)	16.32	−5.27 [1.52]	−0.288* (0.172)	16.64
DBI-4: Dairy & bean products (g)									
Dairy products	8.09 [36.09]	−9.622 (5.881)	12.53	0.04 [0.18]	−0.027 (0.026)	15.14	−5.83 [0.75]	−0.093 (0.099)	17.57
Soybean products	45.13 [58.70]	−7.455 (6.487)	16.18	0.30 [0.32]	−0.068* (0.038)	13.83	−4.21 [1.89]	−0.461** (0.227)	13.44
DBI-5: Animal protein (g)									
Livestock	214.36 [185.24]	−47.968** (22.414)	14.25	0.75 [0.36]	−0.141*** (0.048)	13.34	1.73 [2.77]	−0.945** (0.394)	12.43
Poultry	20.84 [40.87]	−5.284 (5.212)	12.39	0.17 [0.29]	−0.040 (0.039)	11.74			
Aquatic products	27.18 [47.11]	−8.689 (5.950)	12.95	0.18 [0.27]	−0.039 (0.035)	15.66	−3.33 [1.04]	−0.157 (0.140)	14.35
Egg	24.53 [30.05]	−4.150 (3.550)	16.32	0.31 [0.33]	−0.033 (0.040)	16.10	−2.07 [2.16]	−0.240 (0.251)	16.91
DBI-6: Oil, condiments & alcoholic beverages (g)									
Oil	77.25 [77.51]	−11.651 (7.863)	15.12				2.54 [1.60]	−0.215 (0.201)	12.72
Salt	16.95 [15.64]	0.267 (1.560)	17.29				2.67 [1.51]	0.169 (0.191)	14.71
Alcohol	10.05 [47.32]	1.631 (5.066)	14.04				0.20 [0.76]	−0.020 (0.079)	16.80

Table 5 (continued)

	Amount consumed			Food groups in the DDS			DBI components		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimator	Mean [SD]	RD estimates (cutoff: SBP = 140)	Optimal bandwidth	Mean [SD]	RD Estimates (cutoff: SBP = 140)	Optimal bandwidth	Mean [SD]	RD Estimates (cutoff: SBP = 140)	Optimal bandwidth
Sample	Non-hypertensive (SBP < 140, DBP < 90)	Diastolic normal (DBP < 90)		Non-hypertensive (SBP < 140, DBP < 90)	Diastolic normal (DBP < 90)		Non-hypertensive (SBP < 140, DBP < 90)	Diastolic normal (DBP < 90)	
Observations in the full analytical sample	14,819		15,564			15,432			15,432

Data were drawn from the China Health and Nutrition Survey (1997, 2000, 2004, 2006, 2009, 2011). Analytical samples are restricted to observations within ±40 mmHg intervals of the respective cutoffs. The actual sample size is determined by the optimal bandwidth chosen by minimizing the mean squared error to produce lower biases (Calonico et al. 2014). Non-parametric RD estimates using the triangular kernel and the associated optimal bandwidth are reported. Standard deviations are reported in brackets. Robust standard errors are reported in parentheses, clustered at the community level.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Robustness checks. Before exploring further, this section rules out several potential threats to our identification strategy to strengthen the credibility of our findings discussed so far.

Validity of identification assumptions. The fundamental identifying assumption underlying our RD analysis is that all factors relevant to hypertension except the diagnosis result vary continuously as the two running variables, DBP and SBP readings, smoothly pass their respective cutoffs. If this assumption holds, we will expect no discontinuities in pre-determined factors at the diagnostic cutoffs. To check this, we performed a series of RD regressions treating available pre-determined demographic and socio-economic characteristics, including a male dummy, age (years), education (years), household size, income, and residential area (an "urban" dummy), as outcome variables. Supplementary Appendix Fig. A1 visualizes the results, revealing no notable "jumps" in these characteristics at either diagnostic cutoff. Although we are unable to perform similar tests for the balancedness in unobserved characteristics at the cutoffs, the balancedness in observed characteristics shown in the Figure strongly supports our RD findings' credibility.

Also needed for a valid RD design is the "no perfect manipulation" condition (Lee and Lemieux 2010): i.e., individuals cannot deliberately alter the value of the running variable to be "just treated" or "just untreated." This condition rules out any "clustering" in DBP and SBP readings on either side of their respective cutoffs. Reassuringly, the densities of both sets of readings plotted in Supplementary Appendix Fig. A2 reveal no clustering around either cutoff.

Alternative bandwidths. The RD results reported in Tables 3–5 were estimated based on optimal bandwidths determined by the method provided by Calonico et al. (2014). However, these estimates may still be subject to bias if the chosen bandwidths are too wide or too narrow, such that they affect the estimated limits of the outcome variable on the two sides of the diagnostic cutoff. To see how the choice of bandwidth influences our estimates, we re-performed RD estimations for diet-structure indicators within *half* (Supplementary Appendix Table A2, columns 1 and 3) and *twice* (columns 2 and 4) of the optimal bandwidth.[11] The results remain similar to our original estimates reported in Table 4: our RD estimates are robust to the choice of bandwidth.

Measurement errors. The above analysis assumes away potential measurement errors in blood pressure readings and the resulting diagnosis results.[12] Measurement errors in blood pressure readings mainly take two forms: random errors and non-random errors—the latter could be caused by the "white coat" effect (i.e., one being nervous when seeing a doctor) or relatively minor unreported health conditions.[13] While random and non-random measurement errors will lead to different misclassifications of individuals into the "hypertensive" and "non-hypertensive" groups, they will impact our RD estimates in a similar way.

If the measurement errors involved are random errors, they will be balanced out on the two sides of the diagnostic cutoff in a properly implemented RD design. However, with random measurement errors occurring, some non-hypertensive individuals would be misclassified as "hypertensive," and some hypertensive ones would be misclassified as "non-hypertensive." These misclassifications will create a bias similar to the conventional "attenuation bias" that drives the RD estimates toward zero, as some hypertensive individuals in the "hypertensive" group are now compared with the hypertensive individuals misassigned to the "non-hypertensive" group, cancelling some of the effect of hypertension diagnoses. Nevertheless, this bias would not invalidate our findings because if such measurement errors exist, we should have obtained estimates that are larger in magnitude, which, in fact, strengthens our findings.

The issue of non-random measurement errors seems more complicated, but their influence on our RD estimates is similar to that of random errors. Take the white-coat effect, for example. With such an effect occurring, blood pressure readings are likely to be systematically higher than their values in the absence of this effect. As such, some actually non-hypertensive individuals may be classified as "hypertensive." Again, this will lead to a bias that "attenuates" RD estimates because some correctly classified non-hypertensive individuals will be compared with those non-hypertensive individuals misassigned to the "hypertensive" group, thereby canceling out part of the hypertension-diagnosis effects.

To help assess how such an effect may have impacted our RD estimates, at least partially, we redefined the assignment variables in our RD design and re-estimated our RD models. Note that if there is a white-coat effect, it is most likely to occur on the first of the three interview dates when sampled respondents and the CHNS medical team first met, and the blood pressure readings subject to this effect should be higher than otherwise. Indeed, our data show that the average blood pressure readings were the highest on the first date.[14] Given this reasoning, we redefined the assignment variables (and the associated hypertension diagnosis results) based on the following alternative definitions: the average of blood pressure readings on the second and the third interview dates, that of the two lowest readings, and the median of three readings. Note that these modifications bring some actually non-hypertensive individuals misassigned to the "hypertensive" group back to the "non-hypertensive" group, thereby mitigating the white-coat effect to some extent. As shown in Supplementary Appendix Table A3, based on SBP readings (columns 2–4), the new estimates for the four diet-structure indicators obtained using the new assignment variables are slightly larger than the original estimates (reproduced in column 1), signaling the existence of the white-coat effect; however, the effect is small; meanwhile, as with the original estimates based on DBP readings (column 5), the new estimates are largely insignificant (columns 6–8). More generally, as discussed above, even if blood pressure readings provided in the CHNS data are measured with errors, our benchmark RD estimates can still serve as meaningful lower bounds of the true hypertension-diagnosis effects on Chinese adults' diet structure.

Alternative estimation approach. The CHNS's panel structure provides valuable opportunities to apply other empirical methods, such as the difference-in-differences (DID) approach (Chen et al. 2022), to complement our RD analysis. The DID approach compares dietary changes between individuals diagnosed with hypertension and those without.[15] Unlike RD designs, which identify "local" effects around the diagnosis thresholds, the DID approach may identify more "global" effects and thus has two advantages: it can incorporate more observations (not necessarily those near the diagnosis thresholds) in the analysis, and consequently, its findings have stronger external validity, as it speaks to a broader population. Thus, we performed DID estimations on the diet-structure indicators based on both SBP and DBP readings. The DID estimates, adjusted for a basic set of personal socio-demographic characteristics (Supplementary Appendix Table A4 for SBP-based diagnosis and Supplementary Appendix Table A5 for DBP-based diagnosis), are qualitatively similar to our RD estimates (Table 4) but mostly statistically insignificant, suggesting that the effects of hypertension diagnoses are indeed local ones. This also supports the interpretation of these effects as (health information" (as opposed to "health condition") shocks. Of course, the lack of effects on a broader population could also be due to the fact that our outcome measures were observed about 3 years after the diagnosis, which may have introduced noises in the hypertension diagnosis-diet change relationship. Data with a shorter interval between the diagnosis and the observation of outcome measures may be able to reveal a clearer broader impact.

Working channels. The checks reported above help strengthen our findings of the impacts of (SBP-based) hypertension diagnoses on one's diet structure. Yet, do these effects represent the effects of *information* provided during health checkups? Some alternative channels are worth examining.

Other chronic conditions. One possibility is that our estimates of SBP-based-diagnosis effects capture mostly the influence of health conditions other than hypertension. As noted above, if not properly controlled, hypertension can cause other health conditions, such as heart diseases, to develop (Mendis et al. 2011), which may also affect one's food choices (Krämer et al. 2021; Neutel and Campbell 2018). Since our RD design captures the impacts of (SBP-based) hypertension diagnoses about 3 years post-diagnosis, our estimates may partly reflect the influence of other chronic illnesses developed during these post-diagnosis years. To check this possibility, we performed a set of RD estimations, treating the incidences of chronic diseases other than hypertension recorded in the CHNS data (apoplexy, asthma, diabetes, and myocardial infarction) as outcome variables. The results reveal only negligible correlations between these conditions and SBP-based diagnoses (Table 6). Although it is impossible to check for other chronic illnesses (e.g., gastric and liver diseases) that were not recorded in the CHNS data, the findings reported in Table 6 greatly reduce the concern about other chronic conditions being a potential channel.

Non-medical channels. The impacts of (SBP-based) hypertension diagnoses reported in Tables 3–5 may also work through non-medical channels. During the 3-or-so-year gap between the diagnosis and observing the outcomes examined above, the diagnosis may have caused non-medical factors that may impact food choices (e.g., work status, income, and dining locations) to change. This subsection examines a number of such channels. Regarding one's diet plans, Table 7 shows that, on average, SBP-based hypertension diagnoses reduced 0.2 meals one had in the past three days (panel A) but did not change one's dining location (panel B). Regarding one's work-related behaviors, the results suggest that SBP-based diagnoses exerted virtually no impact on the likelihood of working (panel C), labor intensity (panel D), or income (panel E). Simply put, among all potential channels examined here, only the reduction in daily meals helps explain the changes in one's diet structure found in Tables 3–5.

Together, the findings reported in Tables 6 and 7 support the interpretation of the effects of hypertension diagnoses found above as information shocks, as minimal changes in chronic conditions and socio-economic behavior in response to the diagnoses were found.

Heterogenous responses. Another way to understand the potential channels of the hypertension-diagnosis effects is to explore possible heterogeneous patterns in the impact across subsamples with different characteristics such as gender, age, education, and location. Table 8 reports the results. While the general patterns discussed above (e.g., reductions in dietary diversity and balance) remain, male, younger, better-educated, and rural residents are more responsive to hypertension diagnoses than their respective

Table 6 Continuity in the incidences of chronic conditions at the SBP cutoff.

Estimator	(1) Mean [SD]	(2) Non-parametric RD estimates based on SBP readings (cutoff: SBP = 140)
Sample	Non-hypertensive	Diastolic normal (DBP < 90)
A. Diabetes	0.011	−0.022
	[0.106]	(0.022)
Optimal bandwidth		13.793
N	16,710	17,619
B. Myocardial Infarction	0.003	−0.013
	[0.052]	(0.009)
Optimal bandwidth		14.867
N	16,875	17,788
C. Apoplexy	0.004	0.003
	[0.060]	(0.011)
Optimal bandwidth		17.361
N	16,766	17,672
D. Asthma	0.010	0.001
	[0.102]	(0.019)
Optimal bandwidth		13.510
N	6210	6622
E. Ashma missing (=1 if yes)	0.635	0.059
	[0.481]	(0.053)
Optimal bandwidth		15.797
N	17,007	17,925

Data were drawn from the China Health and Nutrition Survey (1997, 2000, 2004, 2006, 2009, and 2011). Analytical samples are restricted to observations within ±40 mmHg intervals of the respective cutoffs. The actual sample size is determined by the optimal bandwidth chosen by minimizing the mean squared error (Calonico et al. 2014). Non-parametric RD estimates (with the triangular kernel) and the associated optimal bandwidth are reported in columns (2) and (3). Standard deviations are reported in brackets. Robust standard errors are reported in parentheses, clustered at the community level.

Table 7 Potential channels of the effects of SBP-based hypertension diagnosis.

Estimator	(1) Mean [SD]	(2) Non-parametric RD estimates based on SBP readings (cutoff: SBP = 140)
Sample	Non-hypertensive	Diastolic normal (DBP < 90)
A. Number of meals had in the past three days	8.71	−0.196**
	[0.94]	(0.092)
Optimal bandwidth		16.13
N	14,934	15,769
B. Proportion of meals eaten at home	0.92	0.011
	[0.19]	(0.019)
Optimal bandwidth		14.36
N	14,934	15,737
C. Working (=1, if yes)	0.67	−0.043
	[0.47]	(0.060)
Optimal bandwidth		14.82
N	15,129	15,970
D. Labor intensity (=1, if "High" or "Very high")	0.44	−0.001
	[0.50]	(0.070)
Optimal bandwidth		13.40
N	14,622	15410
E. Household income (log)	9.86	−0.161
	[1.46]	(0.141)
Optimal bandwidth		16.622
N	16,676	17,564

Notes: Data were drawn from the China Health and Nutrition Survey (1997, 2000, 2004, 2006, 2009, and 2011). Analytical samples are restricted to observations within ±40 mmHg intervals of the respective cutoffs. The actual sample size is determined by the optimal bandwidth chosen by minimizing the mean squared error to produce a lower bias (Calonico et al. 2014). Non-parametric RD estimates using the triangular kernel and the associated optimal bandwidth are reported in column (2). Other categories of labor intensity include "very light," "light," and "moderate." Standard deviations are reported in brackets. Robust standard errors are reported in parentheses, clustered at the community level.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

counterparts. The stronger reactions of rural residents might be driven by their more risk-averse nature; those of male, younger, and better-educated individuals may be due to their more abundant resources for adjusting dietary plans. Nevertheless, they failed to utilize these resources better to improve their diet structure by seeking more professional information.

Discussion

Contributions to the existing literature. Public health researchers and health economists are increasingly interested in how health information affects food choices. The recent decades have witnessed the emergence of a new strand of research that explores medically-provided health information, particularly that on chronic disease diagnoses, to identify the causal impact of health information shocks on one's food choice (Zhao et al. 2013; Bünnings 2017; Oster 2018; Hut and Oster 2020).

The present study intends to contribute to this emerging literature from the perspective of hypertension diagnosis. Exploiting the rule that a person is diagnosed with hypertension if his/her blood pressure attains an explicitly set diagnostic cutoff, we implemented a two-dimensional RD design to estimate the impact of hypertension diagnosis on Chinese adults' food choices, focusing on the first-ever diagnosis. Our analysis, analyzing a dataset of 9355 Chinese adults, reveals both desirable and undesirable effects of hypertension diagnoses. While SBP-based diagnoses exert some desirable effects by reducing Chinese adults' fat intake and livestock consumption about 3 years later, it also reduces their consumption of foods that ought to increase for hypertensive individuals (e.g., fruits)—an undesirable effect often overlooked in existing studies. These effects (particularly reductions in food intake) mainly work through the decreased number of one's daily meals but not changes in chronic conditions or work-related behaviors.

Although two other RD studies, those of Zhao et al. (2013) and Dai et al. (2022), have investigated the effect of hypertension diagnosis on food-related outcomes in the Chinese context, our study extends their analyses in two major ways. First, both existing studies focused on a narrow set of nutrient intakes. Zhao et al. (2013) applied a one-dimensional RD design (similar to Eq. (2)) to estimate the impacts of SBP-based hypertension diagnosis on one's carbohydrate, fat, protein, and total calorie intakes. Dai et al. (2022) adopted a two-dimensional RD framework similar to ours to identify hypertension-diagnosis impacts on fat intake as a measure of health behavior, along with hypertension awareness, medication, and smoking cessation. By comparison, we examined the entire diet structure involving 12 essential food groups based on more comprehensive measures. In particular, the Diet Balance Indices we constructed allow us to examine both over- and under-consumption of foods that the two existing studies were unable to discover. Unlike these existing studies, which found only *desirable* effects of hypertension diagnosis, we also discovered some *undesirable* effects (e.g., reductions in fruit

Table 8 Heterogeneous effects of SBP-based hypertension diagnoses on diet structure (diastolic normal individuals ≥ 18 years old).

	Gender		Age		Education		Location	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Subsamples	Female	Male	< 50 years	≥ 50 years	< 9 years	≥ 9 years	Rural	Urban
A. Dietary Diversity Score (DDS)	−0.472**	−0.831***	−0.613	−0.570***	−0.447**	−0.854**	−0.601***	−0.379
	(0.207)	(0.257)	(0.383)	(0.196)	(0.182)	(0.338)	(0.199)	(0.332)
Optimal bandwidth	11.890	11.116	8.176	12.281	12.270	9.861	12.676	12.294
N	8582	7122	9113	6591	8307	7276	11,860	3844
B. Diet Balance Index (DBI)	−1.729	−4.417**	−3.983*	−1.941	−1.764	−3.736*	−2.395*	−0.766
	(1.301)	(1.867)	(2.349)	(1.235)	(1.265)	(2.074)	(1.233)	(2.335)
Optimal bandwidth	14.524	10.423	9.852	12.968	12.457	10.985	14.992	12.136
N	8239	6835	8703	6371	8025	6937	11,469	3605
C. DBI-O: over-consumption	0.130	−0.607	−0.920	0.307	−0.272	0.502	−0.127	0.962
	(0.663)	(1.015)	(1.323)	(0.627)	(0.710)	(1.083)	(0.586)	(1.298)
Optimal bandwidth	15.237	9.936	9.034	13.994	11.514	12.505	16.962	12.746
N	8239	6835	8703	6371	8025	6937	11,469	3605
D. DBI-U: under-consumption	−1.588	−1.841*	−2.751	−2.050**	−1.249	−4.067***	−2.154**	−1.304
	(0.967)	(1.027)	(1.815)	(0.896)	(0.769)	(1.538)	(0.905)	(1.457)
Optimal bandwidth	11.314	16.894	8.166	12.819	15.720	10.301	13.146	13.443
N	8430	7002	8942	6490	8164	7155	11,697	3735

Data were drawn from the China Health and Nutrition Survey (1997, 2000, 2004, 2006, 2009, and 2011). Analytical samples are restricted to observations within ±40 mmHg intervals of the SBP cutoffs and are diastolic normal (DBP < 90 mmHg). The actual sample size is determined by the optimal bandwidth chosen by minimizing the mean squared error (Calonico et al. 2014). Non-parametric RD estimates (with the triangular kernel) and the associated optimal bandwidth are reported in columns (2) and (3). Standard deviations are reported in brackets. Robust standard errors are reported in parentheses, clustered at the community level.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

consumption), stressing the need to rethink carefully how health information may be provided to avoid undesirable dietary outcomes.

Secondly, compared to the two existing studies, our study performed more channel analyses to strengthen causal inference. More specifically, we ruled out other major chronic conditions (e.g., asthma, diabetes, and heart attack), reduced workload, and income change as cofounders of the effects of health information shocks and identified the reduction in the number of daily meals as a potential working channel. In addition to deepening our understanding of how hypertension diagnosis impacts individuals' food choices, these extensions also enhance the validity of our own and existing studies' findings.

Practical implications. Two policy implications are immediate. First, the beneficial effects of (SBP-based) hypertension diagnoses necessitate regular and timely health checkups, which allow for early detection and in-time management of hypertension and other related health conditions. Since health checkups incur costs, subsidies may be provided to encourage Chinese residents, especially aged and retired persons, to seek hypertension screening. Indeed, the beneficial effects of health screening on lifestyle choices have long been documented (Larsen et al. 2007; Kim et al. 2017). Reassuringly, China has recently begun to provide free health checkups for adults above 65, which includes hypertension screening, in local community clinics.[16]

Second, the overall negative hypertension-diagnosis effects on one's dietary diversity and diet balance suggest that urging the general public to adopt healthier dietary plans requires more than regular checkups—nutrition education matters (Kattelmann 2014; Weiss et al. 2023). The undesirable hypertension-diagnosis effects may be due to individuals' misunderstanding of blood pressure control practices. Health and food authorities shall disseminate more easily accessible and understandable information on the food choice-nutrition-health nexus to the public. Particularly, social media may be utilized to educate the public about which foods hypertensive individuals should consume more (e.g., whole grains, vegetables, and fruits) and what should be consumed less (e.g., red meat, fat, and salt) (World Health Organization 2021).

Limitations. Finally, this study has a number of limitations. Firstly, to circumvent reverse-causality issues, we estimated the impacts of hypertension diagnoses in a given survey wave on one's food choices observed in the next wave. But, given the design of CHNS, there is a 2–4-year gap between two consecutive waves. Since many things can happen within 3 or so years, they could distort the observed hypertension diagnosis-food choice relationship. Even though we have ruled out a large number of possible factors, future studies employing more frequently observed outcomes are likely to be fruitful. Secondly, unlike diagnoses performed in hospitals, those performed during household surveys (even by medical professionals) might not be deemed authoritative by the respondents. The respondents might not fully believe the health information provided during the diagnoses, which might blur the "hypertension diagnosis-food intake" relationship. Lastly, the empirical method we adopted, albeit with strong internal validity, limits the external validity of our findings. In particular, RD designs only identify the effects of hypertension diagnoses around the diagnostic cutoffs but have little power to generalize these effects to individuals with other blood pressure readings, which in turn limits the scope of the policy implications derived above. Future studies shall consider combining RD designs with other methods to increase the external validity of empirical findings.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 23 July 2024; Accepted: 28 April 2025;
Published online: 10 May 2025

Notes

1 A more traditional approach estimates the impact of general health information, particularly, knowledge about public health issues, such as nutrition labels, on one's diet through choice experiments or contingent valuation studies (Chang and Just 2007; Chern et al. 1995; Crutchfield et al. 2001; Kim and Chern 1999; Kim et al. 2000; Lin and Yen 2011; Moon et al. 2005, 2011; Øvrum et al. 2012; Variyam et al. 1998). Most studies examining the effect of general health information found that it increased one's intake of healthy foods while reducing one's consumption of unhealthy foods. For example, Kim et al. (2000) found that provided with more health information, U.S. food-label users tend to consume lower proprtion of calories from cholesterol, fat, and sodium and a higher proportion of fiber than non-users. Kim and Chern (1999) found that the provision of general health information reduced Japanese individuals' consumption of hog grease, palm oil, and tallow and increased their fish oil consumption. Similarly, Øvrum et al. (2012) found in Norway that the provision of diet-related health information raised consumers' willing-to-pay for low-fat and organic cheese. However, other scholars found much less significant effects of general health information. For example, Roosen et al. (2009) found that "high-toxic-level" warnings altered French households' fish consumption only slightly in the short term; the effect disappeared three months after warning exposure. Krämer et al. (2021) studied the effect of a nutrition information intervention on household dietary behavior in India but found no significant effects.

2 Globally, the number of 30–79-year-olds diagnosed with hypertension has doubled in the last two decades, reaching 1728 million in 2019 (Zhou et al. 2021).

3 By the time of writing, there had been more than 2000 research papers and reports written based on the CHNS data. Source: (https://www.cpc.unc.edu/projects/china/publications?b_start:int¼0 (accessed: Jan 25, 2025).

4 We use the 2008 *Guidelines* rather thant the 2022 ones to construct the DDS because the former reflected more closely the food consumption patterns in our study period.

5 While the official definition of DDS requires data on one's food intake on only one day (24 h), CHNS provides daily food-intake data for three consecutive days for each sample individual. Thus, we first calculated DDS for each day and then averaged the scores over three days. The three-day averaging helps eliminate daily fluctuations in one's food intakes, providing a more accurate measure of his/her usual dietary structure.

6 Information on water intake is not avaible in the data. Thus, six out of seven sub-indicators were used to construct the DBI. Yet, since water drinking is unlikely to be affected by hypertension diagnoses, it seems safe to leave it out.

7 He et al. (2009) provided detailed criteria for individuals with other recommended energy intake levels.

8 The results remain very similar when the standard errors are clustered at the blood pressure reading or individual level. Detailed results are not reported but available upon request.

9 The lack of effect found around the DBP cutoff aligns with the findings of Dai et al. (2022), who examined fat intake and other non-food outcomes. They offered several explanations for the lack of effects. For example, individuals may consider the top number (SBP reading) more important than the bottom number (DBP) in their physical examination report. Another reason is that those around the DBP threshold generally have tighter work schedules and more competitive jobs, preventing them from responding actively to DBP-based diagnosis results.

10 Alcohol consumption also dropped as a result of SBP-based diagnoses, which also helps explain the drop in DBI scores, but not that in the DDS, as it is not involved in the DDS calculation.

11 Detailed results on other variables are not reported but are made available upon request.

12 We thank an anonymous reviewer for pointing out this issue.

13 To the extent that these slight conditions cause some individuals to be misassigned to either the "hypertensive" group or the "non-hypertensive" group, they play the role of non-random measurement errors.

14 The average SBP readings were 120.177 mmHg, 120.014 mmHg, and 119.982 mmHg on the three interview dates. The average DBP readings were 77.981 mmHg, 77.943 mmHg, and 77.919 mmHg on the three interview dates.

15 We thank an anonymous reviewer for suggesting this approach.

16 Available at: https://www.gov.cn/zhengce/zhengceku/2022-01/18/content_5669095.htm (accessed December 15, 2024).

References

Aldossari A, Sremanakova J, Sowerbutts AM et al. (2023) Do people change their eating habits after a diagnosis of cancer? A systematic review. J Hum Nutr Dietetics J Br Dietetic Assoc 36(2):566–579

Bünnings C (2017) Does new health information affect health behaviour? The effect of health events on smoking cessation. Appl Econ 49(10):987–1000

Calonico S, Cattaneo MD, Titiunik R (2014) Robust non-parametric confidence intervals for regression-discontinuity designs. Econometrica 82(6):2295–2326

Chang HH, Just DR (2007) Health information availability and the consumption of eggs: are consumers Bayesians? J Agric Resour Econ 32(1):77–92

Chen Q, Pei C, Huang J, Tian G (2022) Public health insurance and enrollees' diet structure in rural China. Heliyon 8(5):e09382

Chen Q, Pei C, Zhao Q (2020) Intrahousehold flypaper effects—Quasi-experimental evidence from a school-based randomized trial in rural northwestern China. Econ Lett 191:109134

Chern WS, Loehman ET, Yen ST (1995) Information, health risk beliefs, and the demand for fats and oils. Rev Econ Stat 77(3):555–564

Chinese Nutrition Society (2008) Chinese dietary guidelines. Tibet People Press, Lhasa. (in Chinese)

Chinese Nutrition Society (2022) Chinese dietary guidelines 2022. People's Medical Publishing House, Beijing. (in Chinese)

Crutchfield S, Kuchler F, Variyam JN (2001) The economic benefits of nutrition labeling: a case study for fresh meat and poultry products. J Consum Policy 24(2):185–207

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 Econ 31(4):574–596

de Carvalho KMB, Dutra ES, Pizato N, Gruezo ND, Ito MK (2014) Diet quality assessment indexes. Rev De Nutricao Braz J Nutr 27(5):605–617

English DR, MacInnis RJ, Hodge AM, Hopper JL, Haydon AM, Giles GG (2004) Red meat, chicken, and fish consumption and risk of colorectal cancer. Cancer Epidemiol Biomark Prev 13(9):1509–1514

Food and Agriculture Organization of the United Nations (2011) Guidelines for measuring household and individual dietary diversity, Version 4. Available online: [http://agrobiodiversityplatform.org/files/2011/05/guidelines_MeasuringHousehold.pdf] (accessed: May 20, 2022)

Fassier P, Zelek L, Lecuyer L, Bachmann P, Touillaud M, Druesne-Pecollo N, Galan P, Cohen P, Hoarau H, Latino-Martel P, Kesse-Guyot E, Baudry J, Hercberg S, Deschasaux M, Touvier M (2017) Modifications in dietary and alcohol intakes between before and after cancer diagnosis: Results from the prospective population-based NutriNet-Sante cohort. Int J Cancer 141:457–470

German Nutrition Society (2017) 10 guidelines of the German Nutrition Society (DGE) for a wholesome diet. Available online: https://www.dge.de/ernaehrungspraxis/vollwertige-ernaehrung/10-regeln-der-dge/en/

Hinkle SN, Li M, Grewal J, Yisahak SF, Grobman WA, Newman RB, Wing DA, Grantz KL, Zhang C (2021) Changes in diet and exercise in pregnant women after diagnosis with gestational diabetes: findings from a longitudinal prospective cohort study. J Acad Nutr Dietetics 121(12):2419–2428

He Y, Zhai F, Yang X, Ge K (2009) To revise the Chinese diet balance index. Acta Nutrimenta Sin 31(6):532–536. In Chinese

Hut S, Oster E (2020) Changes in household diet: determinants and predictability. NBER Working Papers 24892. National Bureau of Economic Research

Imbens G, Kalyanaraman K (2012) Optimal bandwidth choice for the regression discontinuity estimator. Rev Econ Stud 79:933–959

Imbens G, Zajonc T (2011) Regression discontinuity design with multiple forcing variables. J Econ 142:615–635

Kassie M, Fisher M, Muricho G, Diiro G (2020) Women's empowerment boosts the gains in dietary diversity from agricultural technology adoption in rural Kenya. Food Policy 95:101957

Kattelmann K (2014) What is effective nutrition education? J Nutrition Educ Behav 46(6), 457

Kennedy ET, Ohls J, Carlson S, Fleming K (1995) The healthy eating index: design and applications. J Am Dietetic Assoc 95(10):1103–1108

Kim SR, Chern WS (1999) Alternative measures of health information and demand for fats and oils in Japan. J Consum Aff 33(1):92–109

Kim SY, Nayga RM, Capps OC (2000) The effect of food label use on nutrient intakes: an endogenous switching regression analysis. J Agric Resour Econ 25(1):215–231

Kim D, Koh K, Swaminathan S, Trivedi AN (2018) Association of diabetes diagnosis with dietary changes and weight reduction. Expert Rev Pharmacoeconomics Outcomes Res 18(5):543–550

Kim H, Lee S, Lim W (2017) Knowing is not half the battle: impacts of the national health screening program in Korea. J Health Econ 65:1–14

Kolahdooz F, van der Pols JC, Bain CJ et al. (2010) Meat, fish, and ovarian cancer risk: results from 2 Australian case-control studies, a systematic review, and meta-analysis. Am J Clin Nutr 91(6):1752–1763

Krämer M, Kumar S, Vollmer M (2021) Anemia, diet, and cognitive development: impact of health information on diet quality and child nutrition in rural India. J Econ Behav Organ 190:495–523

Larsen I, Grotmol T, Alemndingen K, Hoff G (2007) Impact of colorectal cancer screening on future lifestyle choices: a three-year randomized controlled trial. Clin Gastroenterol Hepatol 5:477–483

Lee D, Lemieux T (2010) Regression discontinuity designs in economics. J Economic Lit 48:281–355
Lei Y-Y, Ho SC, Cheng A, Kwok C, Cheung KL, He Y-Q, Lee C-K, Lee C, Yeo W (2018) Dietary changes in the first 3 years after breast cancer diagnosis: a prospective Chinese breast cancer cohort study. Cancer Manag Res 10:4073–4084
Lin BH, Yen ST (2011) Consumer knowledge, food label use and grain consumption in the US. Appl Econ 40:437–448
Mancini FR, Affreta A, Dowa C, Balkaua B, Bihane H, Clavel-Chapelona F, Boutron-Ruaulta M-C, Bonneta F, Fagherazzi G (2017) Educational level and family structure influence the dietary changes after the diagnosis of type 2 diabetes: evidence from the E3N study. Nutr Res 44:9–17
Mendis S, Puska P, Norrving B (2011) Global atlas on cardiovascular disease prevention and control. World Health Organization. Available online: https://www.who.int/publications/i/item/9789241564373
Moon W, Balasubramanian SK, Rimal A (2005) Perceived health benefits and soy consumption behavior: two-stage decision model approach. J Agric Resour Econ 30(2):315–332
Moon W, Balasubramanian SK, Rimal A (2011) Health claims and consumers' behavioral intentions: the case of soy-based food. Food Policy 36:480–489
Neutel CI, Campbell NR (2018) Changes in lifestyle after hypertension diagnosis in Canada. Can J Cardiol 24(3):199–204
Opstelten JL, de Vries JHM, Wools A, Siersema PD, Oldenburg B, Witteman BJM (2019) Dietary intake of patients with inflammatory bowel disease: a comparison with individuals from a general population and associations with relapse. Clin Nutr 38(4):1892–1898
Oster E (2018) Diabetes and diet: purchasing behavior change in response to health information. Am Econ J Appl Econ 10(4):308–348
Øvrum A, Alfnes F, Almli VL, Rickertsen K (2012) Health information and diet choices: results from a cheese experiment. Food Policy 37:520–529
Patterson RE, Haines PS, Popkin BM (1994) Diet quality index: capturing a multidimensional behavior. J Am Dietetic Assoc 94:57–64
Reedy J, Lerman JL, Krebs-Smith SM, Kirkpatrick SI et al. (2018) Evaluation of the healthy eating index-2015. J Acad Nutr Dietetics 118(9):1622–1633
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
Shams-White MM, Pannucci TE, Lerman JL, Herrick KA, Zimmer M, Mathieu KM, Stoody EE, Reedy J (2023) Healthy Eating Index-2020: review and update process to reflect the dietary guidelines for Americans, 2020–2025. J Acad Nutr Dietetics 123(9):1280–1288
Slade AN, Kim H (2014) Dietary responses to a hypertension diagnosis: evidence from the National Health and Nutrition Examination Survey (NHANES) 2007–2010. Behav Med 40(1):1–13
Unger T, Borghi C, Charchar F et al. (2020) 2020 International Society of Hypertension Global Hypertension Practice Guidelines. Hypertension 75(6):1334–1357
United States Agency for International Development (2016) Why is dietary diversity important? https://www.fantaproject.org/node/1199
U.S. Department of Agriculture and U.S. Department of Health and Human Services (2020) Dietary Guidelines for Americans 2020–2025. https://www.dietaryguidelines.gov/resources/2020-2025-dietary-guidelines-online-materials
Variyam JN, Blaylock J, Smallwood D (1998) Informational effects of nutrient intake determinants on cholesterol consumption. J Agric Resour Econ 23(1):110–125
Velentzis LS, Keshtgar MR, Woodside JV et al. (2011) Significant changes in dietary intake and supplement use after breast cancer diagnosis in a UK multicentre study. Breast Cancer Res Treat 128(2):473–482
Weiss, H, Russell, RD, Black, L, Begley, A (2023) Interpretations of healthy eating after a diagnosis of multiple sclerosis: a secondary qualitative analysis. Brit Food J 125(8):2918–2930
World Health Organization (2021) Guideline for the pharmacological treatment of hypertension in adults. World Health Organization, Geneva
World Health Organization (2022) Toolkit for developing a multisectoral action plan for noncommunicable diseases. World Health Organization, Geneva
Wong VC, Steiner PM, Cook TD (2013) Analyzing regression-discontinuity designs with multiple assignment variables: A comparative study of four estimation methods. J Educ Behav Stat 38(2):107–141
Xiang XL (2016) Chronic disease diagnosis as a teachable moment for health behavior changes among middle-aged and older adults. J Aging Health 28(6):995–1015
Yang Y (2002) China food composition 2002. Book 1. Peking University Medical Press, Beijing. (in Chinese)
Yang Y (2004) China food composition 2004. Book 2. Peking University Medical Press, Beijing. (in Chinese)
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
Zhou B, Carrillo-Larco RM, Danaei G, Riley LM, Paciorek CJ, Stevens GA et al. (2021) Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. Lancet 398(10304):957–980

Acknowledgements
The authors thank Professors Chengfang Liu, Yanjun Ren, and Zhihao Zheng, as well as participants of the 2022 annual meeting of the Beijing Agricultural Economists Association, for their critical yet constructive comments on earlier drafts of this paper. Financial support from the National Social Science Foundation of China (grant number: 22&ZD113) and the 2115 Talent Development Program of China Agricultural University is gratefully acknowledged. This analysis of this study was performed based on datasets from the China Health and Nutrition Survey. The research team is grateful to the National Institute of Nutrition and Food Safety, the China Center for Disease Control and Prevention, the Carolina Population Center, the University of North Carolina at Chapel Hill, the National Institutes of Health (NIH; R01-HD30880, DK056350, and R01-HD38700), and the Fogarty International Center, NIH, for financial support for the CHNS data collection and analysis files since 1989.

Author contributions
QC and JH—designed the study; YH, JH, and JB—acquired the data, performed the research and performed data analysis. QC—provided supervision and advice on data analysis and interpretation of findings. JH, QC, JB, and YH—wrote the manuscript and revised the manuscript. All authors have read and approved the final manuscript.

Competing interests
The authors declare no competing interests.

Ethical approval
Approval was obtained from the Institutional Review Board of the College of Economics and Management at China Agricultural University (CAU-CEM-IRB; approval number: HC0028401A). The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Informed consent
Informed consent was obtained from all participants and/or their legal guardians.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-025-04962-1.

Correspondence and requests for materials should be addressed to Qihui Chen.