ARTICLE

Education and incident type 2 diabetes: quantifying the impact of differential exposure and susceptibility to being overweight or obese

Jimmi Mathisen[1] · Aksel K. G. Jensen[2,3] · Ingelise Andersen[4] · Gregers S. Andersen[5] · Ulla A. Hvidtfeldt[6] · Naja H. Rod[1]

Abstract

Aims/hypothesis Educational inequality in type 2 diabetes incidence is evident in many high-income countries. Previous studies have shown that differential exposure to being overweight/obese across educational groups may partly explain this inequality. Whether differential susceptibility to being overweight/obese across educational groups contributes to this inequality has been investigated less frequently, even though it is a plausible mechanism. The two mechanisms may even be highly intertwined. In this longitudinal cohort study, we investigated the simultaneous contribution of differential exposure and differential susceptibility to being overweight/obese to educational inequality in type 2 diabetes incidence.

Methods The study population comprised 53,159 Danish men and women aged 50–64 years at baseline who were followed for a mean of 14.7 years. We estimated rate differences of type 2 diabetes by education level per 100,000 person-years. Using counterfactual mediation analysis, these rate differences were decomposed into proportions attributable to differential exposure, differential susceptibility and all other pathways, respectively. We compared this approach with conventional approaches to mediation and interaction analysis.

Results Compared with a high level of education, a low education level was associated with 454 (95% CI 398, 510) additional cases of type 2 diabetes, and a medium education level with 316 (CI 268, 363) additional cases. Differential exposure to being overweight/obese accounted for 37% (CI 31%, 45%) of the additional cases among those with a low education level and 29% (CI 24%, 36%) of the additional cases among those with a medium education level. Differential susceptibility accounted for 9% (CI 4%, 14%) and 6% (CI 3%, 10%) of the additional cases among those with a low and medium education level, respectively. Compared with the counterfactual approach, the conventional approaches suggested stronger effects of both mechanisms.

Conclusions/interpretation Differential exposure and susceptibility to being overweight/obese are both important mechanisms in the association between education and type 2 diabetes incidence.

Keywords BMI · Diabetes mellitus, type 2 · Educational status · Health status disparities · Mediation analysis · Obesity · Overweight · Socioeconomic factors · Susceptibility

✉ Jimmi Mathisen
   jima@sund.ku.dk

[1] Section of Epidemiology, Department of Public Health, University of Copenhagen, Oester Farimagsgade 5, 1353 Copenhagen, Denmark

[2] Section of Biostatistics, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[3] Division of Epidemiology and Biostatistics, School of Public Health, University of California Berkeley, Berkeley, CA, USA

[4] Section of Social Medicine, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[5] Steno Diabetes Center Copenhagen, Capital Region, Gentofte, Denmark

[6] Danish Cancer Society Research Center, Copenhagen, Denmark

Research in context

What is already known about this subject?

- Incidence of type 2 diabetes in high-income countries is usually higher in groups with less education. Being overweight/obese is often more common among the less educated

- This differential exposure to being overweight/obese across educational groups partly explains educational inequality in type 2 diabetes

- Whether being overweight/obese has a stronger association with type 2 diabetes among the less educated (differential susceptibility) has received less attention

What is the key question?

- To what extent do differential exposure and differential susceptibility to being overweight/obese contribute to educational inequality in type 2 diabetes incidence?

What are the new findings?

- We found evidence of differential susceptibility to being overweight and obese, as being overweight/obese caused more cases of type 2 diabetes among people with a lower education level

- Differential exposure and susceptibility to being overweight/obese are both important mechanisms contributing to educational inequality in type 2 diabetes incidence

How might this impact on clinical practice in the foreseeable future?

- Being overweight/obese might be a stronger risk factor for type 2 diabetes in individuals with a low level of education, as compared with those with higher levels of education

Abbreviations

DCH	Diet, Cancer and Health
MIE	Mediated interactive effect
NDR	National Diabetes Register
OGLD	Oral glucose-lowering drug
PNDE	Pure natural direct effect
PNIE	Pure natural indirect effect
SIC	Social Inequality in Cancer

Introduction

Educational inequality in type 2 diabetes is evident in many high-income countries, in which less education is associated with markedly higher incidence of type 2 diabetes [1, 2]. The aetiology of type 2 diabetes is multifactorial and complex, but being overweight and obese are key modifiable risk factors [3, 4]. Since being overweight and obese are often more prevalent among people with a lower education level in high-income countries [5], it is plausible that an interrelationship between education level and being overweight/obese could explain the educational inequality in type 2 diabetes incidence.

This interrelationship can be studied within a theoretical framework of the social basis of disparities in health [6, 7], previously used by the WHO in their work on social determinants of health [8]. The framework points to two key mechanisms (Fig. 1): (1) differential exposure, in which the distribution of being overweight and obese is different across educational levels. More cases of type 2 diabetes develop in groups with lower education levels because they are more exposed to being overweight and obese; (2) differential susceptibility, in which the effect of being overweight and obese on the risk of type 2 diabetes is different across educational levels. More cases of type 2 diabetes develop in groups with lower education levels because being overweight and obese is more harmful in these groups. The increased susceptibility is caused by simultaneous exposure to other complementary risk factors for type 2 diabetes, such as lack of physical activity, poor diet, smoking and chronic stress, which all
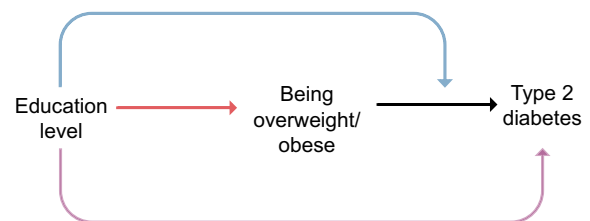
Fig. 1 The proposed mechanisms of differential exposure (red arrow) and differential susceptibility (blue arrow) to being overweight and obese across educational levels, leading to educational inequality in type 2 diabetes incidence. All other pathways from education level to type 2 diabetes, not working through being overweight and obese, are depicted by (purple arrow)

tend to accumulate in individuals with less education [9, 10]. The concept of differential susceptibility is closely linked with the sufficient-component-cause model [11], in which the effect of one cause is dependent on the presence of other interacting causes of the same disease. These mechanisms are not mutually exclusive [6, 7, 12] and may work simultaneously to create and uphold educational inequality in type 2 diabetes incidence.

Previous studies have found that differential exposure to being overweight/obese could partly explain the association between education level and type 2 diabetes incidence [13–21]. Most often, however, these studies assess differential exposure using methods inspired by Baron and Kenny [22], typically comparing estimates from regression models with and without adjustment for BMI [13–20]. This is often referred to as the difference method [23]. This method adequately estimates the strength of a mediator (e.g. being overweight/obese) for an association (e.g. between education level and type 2 diabetes) when used in linear regression models without interactions between the exposure and the mediator [24, 25]. However, the method is typically not viable for nonlinear models or in settings where the effect of a mediator is different across exposure categories (i.e. differential susceptibility) [25]. This limitation generally also applies to other so-called traditional approaches to mediation analysis, such as the product-of-coefficient method [23].

Counterfactual-based approaches address some of these limitations [23]. Nonetheless, we have only identified one study that used such an approach to address the relationship between education level, being overweight/obese and diabetes incidence. The study found that a sizeable part of the educational inequality in diabetes was mediated through BMI [21], but it did not assess a potential interaction between education level and BMI. Therefore, to our knowledge, no previous study has assessed the impact of differential susceptibility to being overweight/obese on the educational inequality in type 2 diabetes incidence.

Knowledge of the impact of differential susceptibility to being overweight and obese can be used to guide policies and interventions aimed at reducing rates of type 2 diabetes [7]. Specific educational groups could be targeted if they are found to be more susceptible to the adverse effects of being overweight and obese. Furthermore, interventions that would reduce the number of people who are overweight and obese to the same extent in all educational groups would also provide a greater reduction in the rates of type 2 diabetes among the groups that are more susceptible.

In this study, we aimed to assess the educational inequality in type 2 diabetes incidence, and to quantify the contributions of differential exposure and susceptibility to being overweight and obese to this association.

Methods

In this longitudinal cohort study, we applied a counterfactual approach to mediation analysis based on a marginal structural model [26] and a three-way decomposition of effects [27]. This approach enabled simultaneous estimation of mediation (differential exposure) and interaction (differential susceptibility). We compared these results with estimates from a conventional approach to mediation analysis, specifically a difference method comparison of regression model estimates [23]. This method can only assess mediation, so we, therefore, used a separate method to assess interaction, namely an additive joint effects model [28]. We refer to these methods as conventional approaches throughout the paper.

Study population The data used was previously collected for the Diet, Cancer and Health (DCH) study, which was a population-based prospective cohort of 57,053 Danish men and women who entered the study between 1993 and 1997 [29]. All participants were between 50 and 64 years old at baseline and were born in Denmark [29]. Access to DCH study data was obtained through the Social Inequality in Cancer (SIC) cohort, which is a consortium of several Danish population-based cohorts linked to social and health registers [30]. Following the exclusion of 1247 individuals because of a cancer diagnosis before baseline or because of previous participation in another cohort in the consortium, data on 55,806 participants in the DCH study were available through the SIC cohort. We excluded 2647 participants who had missing information on key variables, had a BMI below 18.5 or had diabetes at baseline. The study population available for the analyses comprised 53,159 individuals (see the flowchart in electronic supplementary material [ESM] Fig. 1).

The DCH study was approved by the Regional Ethical Committees in Copenhagen and Aarhus ([KF] 01–345/93) and the Danish Data Protection Agency (1991–1110-635). The participants gave informed consent to participation in the study and informed consent to search for personal data in medical registers.

Education Education was categorised into three levels according to highest educational attainment: low (primary or lower secondary education), medium (upper secondary, vocational or technical education) and high (tertiary education of short- and medium-cycle non-university programmes and long-cycle university programmes). Information was drawn from the Population Education Register 1 year prior to baseline in the DCH study.

Overweight and obesity Participants were categorised as being overweight and obese using BMI (weight in kg/[height in m]^2) calculated from baseline anthropometric measurements conducted by a laboratory technician. Participants were

divided into three weight categories: normal weight (18.5–24.9 kg/m^2), overweight (25–29.9 kg/m^2) and obese (≥30 kg/m^2). People with a BMI below 18.5 kg/m^2 were excluded because it is unclear whether diabetes in these individuals is pathophysiologically similar to type 2 diabetes [31].

Covariates Potential confounders were identified based on prior knowledge and the method of directed acyclic graphs [32]. ESM Fig. 2 shows a simplified depiction of the assumed causal relationships between education level, being overweight/obese and type 2 diabetes. We identified age, sex, ethnicity, family socioeconomic position, family history of type 2 diabetes and fetal/neonatal development as potential baseline confounders. Data were not available on familial and early-life factors. Accordingly, we adjusted our analyses for sex, ethnicity and age (underlying timescale in the additive hazards model). Ethnicity was assessed by country of origin and dichotomised into Danish if participants descended from Danish parents and non-Danish if participants descended from foreign-born parents not holding Danish citizenship. We acknowledge that country of origin is an imperfect proxy for ethnicity. Information was drawn from the Danish Civil Registration System.

We also identified several potential confounders of the association between being overweight/obese and type 2 diabetes, namely cohabitation status, chronic stress exposure, physical activity levels, dietary factors, alcohol intake and smoking. However, since these factors were also associated with education level, they were simultaneously mediators and confounders (so-called intermediate confounders). Controlling for these factors was necessary to reduce confounding in analysis of the relationship between being overweight/obese and type 2 diabetes. Doing so, however, would block some of the direct effect of education level on type 2 diabetes. Therefore, both adjusting and not adjusting for these may have led to biased estimates, possibly even in the same direction [24]. Hence, these factors were omitted in the main analysis but were included in a sensitivity analysis (with the exception of chronic stress exposure, for which no data was available).

Type 2 diabetes Incident type 2 diabetes was ascertained as the date of inclusion in the National Diabetes Register (NDR), which requires meeting one of several diabetes-related criteria in central Danish health registries [33]. Briefly, the criteria were registration of: (1) diabetes diagnosis defined by ICD-8 (www.wolfbane.com/icd/icd8h.htm) and ICD-10 (http://apps.who.int/classifications/icd10/browse/2016/en) codes (ICD-9 was never used in Denmark); (2) use of chiropody (foot therapy) as a diabetic patient; (3) second purchase of oral glucose-lowering drug (OGLD) or insulin; and (4) five instances of blood glucose measurement within 1 year or two instances of blood glucose measurements in 5 consecutive years. Previous

studies have indicated that up to 20% of cases in the NDR might be false positives owing to the inclusion criteria of five blood glucose measurements during 1 year [34]. We conducted a sensitivity analysis of this issue by excluding people registered by the blood glucose measurement criteria. Also, the registration dates are systematically delayed for more than 30% of the registrants in NDR before 1997 [34]. However, incidence data are complete from 1995 [33]. We conducted two sensitivity analyses of this issue by excluding people registered before 1995 and 1997, respectively.

The participants were followed from the date of study entry into the DCH study until the date of incident diabetes, death, emigration or end of follow-up (31 December 2012), whichever came first. Follow-up ended because the NDR was discontinued at the end of 2012. Information about migration and death, used for censoring purposes, was drawn from the Danish Civil Registration System. Due to emigration, 1.2% were lost to follow-up. The mean follow-up time was 14.7 years (SD: 4.1).

Statistical analyses In both the counterfactual and the conventional approaches, we applied Aalen's semiparametric additive hazards regression model [35]. We estimated the absolute rate differences in incident cases of type 2 diabetes per 100,000 person-years between groups with different educational levels (reference: high education) and in different BMI categories (reference: normal-weight BMI category).

In the conventional approach, we applied the difference method to estimate the proportion of the rate differences of type 2 diabetes between educational levels that could be attributed to being overweight/obese. Here, mediation through being overweight/obese is quantified as the difference between rate differences of type 2 diabetes estimated with and without adjustment for BMI. Also, we applied an additive joint effects model [28] to estimate the proportion of the educational rate differences of type 2 diabetes that could be attributed to an interaction between education level and BMI. Here, interaction is quantified as deviation from additivity of effects [36], i.e. the extent to which the rate difference in a doubly exposed group (to e.g. low education level and being obese) deviates from the sum of rate differences in the separately exposed groups ([e.g. low education level, being normal weight] + [high education level, being obese]).

In the counterfactual approach, we defined the effects of education level acting through being overweight/obese on type 2 diabetes using a nested counterfactual framework [23]. Briefly, we estimated the rate differences of counterfactually changing a given educational level from the reference level (high education) to an exposure level (medium or low education), which corresponds to the total effect of education level on type 2 diabetes. We applied the marginal structural approach by Lange et al. [26] and the three-way

decomposition of effects by VanderWeele [27] to decompose the total effect of education level into three effects: (1) the pure natural direct effect (PNDE) of education level on type 2 diabetes; (2) the pure natural indirect effect (PNIE) of education level on type 2 diabetes mediated through BMI; and (3) the mediated interactive effect (MIE) of an interaction between education level and BMI on type 2 diabetes. The PNIE corresponds to differential exposure and the MIE corresponds to differential susceptibility [12]. The PNDE corresponds to all other (unaccounted) pathways than through being overweight/obese. The analytical framework in the counterfactual approach is further outlined in the ESM Methods. In an auxiliary analysis, we also stratified this analysis by sex.

Age was the underlying timescale in all additive hazards models. The time-constant hazard difference assumption was tested for education level and BMI and was not violated. We used robust standard errors to compute 95% CIs for the decomposed effects. To obtain 95% CI for the total effect and the proportions of the decomposed effects, we repeated the analysis on 10,000 bootstrapped samples. The statistical technique and its implementation in R are described elsewhere [37, 38]. All analyses were conducted using SAS 9.4 [39] and R 3.5.0. [40].

Results

Descriptive characteristics stratified by educational level for the 53,159 participants are shown in Table 1. Education was associated with being overweight and obese at baseline as more people in the high level of education group were normal weight and more people in the low level of education group were obese. There were more women in the low-level education group compared with the medium- and high-level education groups. A total of 7564 participants developed type 2 diabetes during follow-up (Table 2) and the proportion who developed type 2 diabetes was higher in groups with lower education level and markedly higher in groups with higher BMI.

The results of the assessment of differential exposure using the conventional approach (Table 2) show that low education level was associated with 455 (CI 398, 511) additional cases of type 2 diabetes and medium education level was associated with 316 (CI 269, 360) additional cases per 100,000 person-years, as compared with high education level. According to the difference method, 34% of the additional cases of type 2 diabetes in the medium-level education group and 45% in the low-level education group could be attributed to differential exposure to being overweight and obese across educational groups (without accounting for differential susceptibility).

The result of the assessment of differential susceptibility using the conventional approach (Table 3) show an overall additive interaction between education level and BMI ($p = 0.003$). Compared with the associations in the high-level education group, being overweight was associated with 14% of additional cases of type 2 diabetes and obesity was associated with 19% of additional cases in the low-level education group. In the group with a medium level of education, being overweight was associated with 13% of additional cases of type 2 diabetes and obesity was associated with 19% of additional cases. Thus, without accounting for differential exposure, the joint effects method suggests that between 13% and 19% of the educational inequality in type 2 diabetes incidence can be attributed to differential susceptibility to being overweight/obese across educational groups.

Table 4 shows the results from the counterfactual approach, which simultaneously quantifies differential exposure and susceptibility. Low-level education was associated with 454 (CI 398, 510) additional cases of type 2 diabetes and medium-level education was associated with 316 (CI 268, 363) additional cases per 100,000 person-years, as compared with a high level of education. These rate differences were essentially similar to the rate differences shown in Table 2, with only modelling variation. Decomposing the rate differences showed that differential exposure to being overweight and obese accounted for 37% (CI 31%, 45%) of the additional cases of type 2 diabetes in the low-level education group and 29% (CI 24%, 36%) in the medium-level education group. Differential susceptibility to being overweight/obese accounted for 9% (CI 4%, 14%) of the additional cases of type 2 diabetes in the low-level education group and 6% (CI 3%, 10%) of the addition cases in the medium-level education group.

In an auxiliary analysis, we stratified the counterfactual analysis by sex (ESM Table 1). Here, the overall educational inequality in type 2 diabetes incidence was larger for men than for women, but the proportion attributable to differential exposure to being overweight/obese was larger for women. The proportions attributable to differential susceptibility were similar. We also conducted a number of sensitivity analyses. First, we adjusted the counterfactual analyses for the intermediate confounders that were omitted from the main analysis, namely physical activity level, dietary factors, alcohol intake, smoking and cohabitation status. Descriptive characteristics are shown in ESM Table 2. Adjusting for these factors showed somewhat similar results to those without adjustment (ESM Table 3). However, as expected, the effects of differential exposure and susceptibility to being overweight/obese were stronger and the effect of other pathways was weaker compared with the analysis without adjustment. Further, excluding people registered in the NDR by the blood glucose measurement criteria did not substantially influence the results (data not shown). Finally, excluding people registered in the NDR with type 2 diabetes before 1995 or 1997 did not influence the results (data not shown).

Table 1 Descriptive characteristics for the 53,159 participants by level of education

Characteristic	Total population	Education level		
		Low	Medium	High
n (%)	53,159 (100)	14,875 (28)	24,027 (45)	14,257 (27)
Age (years), median (IQR)	56 (53–60)	57 (53–61)	56 (53–60)	55 (52–59)
Sex, n (%)				
Female	27,852 (52)	9443 (63)	11,593 (48)	6816 (48)
Male	25,307 (48)	5432 (37)	12,434 (52)	7441 (52)
Ethnicity, n (%)				
Danish	53,109 (99.9)	14,863 (99.9)	24,005 (99.9)	14,241 (99.9)
Non-Danish	50 (0.09)	12 (0.08)	22 (0.09)	16 (0.11)
BMI category, n (%)[a]				
Normal weight	23,234 (44)	5803 (39)	10,109 (42)	7322 (51)
Overweight	22,352 (42)	6255 (42)	10,502 (44)	5595 (39)
Obese	7573 (14)	2817 (19)	3416 (14)	1340 (9)

People with missing data for education, BMI, age, ethnicity or sex were excluded. Please see ESM Fig. 1

[a] Normal weight: 18.5–24.9 kg/m^2 ; overweight: 25–29.9 kg/m^2 ; obese, ≥30 kg/m^2

Discussion

In this longitudinal cohort study of 53,159 men and women, we found educational inequality in type 2 diabetes incidence. In accordance with the applied theoretical framework, we hypothesised that at least some of this inequality could be attributed to the mechanisms of differential exposure and susceptibility to being overweight and obese across educational levels. We found that roughly one-third of the inequality could be attributed to differential exposure to being

Table 2 Numbers and proportions of cases of type 2 diabetes and rate differences in incident cases of type 2 diabetes per 100,000 person-years (95% CI) by education level and BMI

Variable	Type 2 diabetes, n (%)[a]	Unadjusted RD (95% CI)[b]	Adjusted RD (95% CI)[c]	Mutually adjusted RD (95% CI)[d]
n(%)	7564 (14)	–	–	–
Education level				
High	1466 (10)	0 (ref.)	0 (ref.)	0 (ref.)
Medium	3635 (15)	313 (266, 360)	316 (269, 363)	207 (159, 254)
Low	2463 (17)	392 (336–448)	455 (398, 511)	250 (195, 307)
BMI category				
Normal weight	1589 (7)	0 (ref.)	0 (ref.)	0 (ref.)
Overweight	3476 (16)	568 (528, 607)	518 (478, 558)	499 (459, 540)
Obese	2499 (33)	1958 (1863, 2054)	1934 (1836, 2032)	1897 (1799, 1995)
Part of RD attributable to differential exposure to being overweight/obese[e]				
Medium level of education	–	–	–	109
Low level of education	–	–	–	205

N = 53,159

[a] Percentages calculated as number of cases of type 2 diabetes in each group divided by number of individuals in each group: high-level education, n = 14,257; medium-level education, n = 24,027; low-level education, n = 14,875; normal-weight, n = 23,234; overweight, n = 22,352; obese, n = 7573

[b] Separate modelling of education and BMI, with no adjustment for confounders

[c] Separate modelling of education and BMI, with adjustment for potential confounders (age, sex, ethnicity)

[d] Combined modelling of education and BMI with mutual adjustment and adjustment for potential confounders (age, sex, ethnicity)

[e] Conventional approach

RD, rate difference; ref., reference

Table 3 Rate differences in incident cases of type 2 diabetes per 100,000 person-years (95% CI) in groups jointly categorised by education and BMI

Variable	BMI category		
	Normal weight	Overweight	Obese
Education level			
High	0 (ref.)	437 (368, 506)	1580 (1377, 1783)
Medium	134 (89, 180)	658 (599, 717)	2117 (1974, 2260)
Low	170 (115, 225)	706 (631, 780)	2152 (1991, 2313)
Part of RD attributable to differential susceptibility to being overweight/obese[a]			
Medium level of education×being overweight	–	–	87 (−5, 179)
Medium level of education×being obese	–	–	402 (154, 651)
Low level of education×being overweight	–	–	99 (−5, 202)
Low level of education×being obese	–	–	401 (148, 655)

N = 53,159

Data are adjusted for age, sex and ethnicity

[a] Conventional approach

p value for interaction term between education level and being overweight/obese = 0.003

RD, rate difference; ref., reference

overweight/obese and that a smaller, yet considerable, proportion could be attributed to differential susceptibility to being overweight/obese across educational levels. These findings are consistent with the employed theoretical framework as both mechanisms play important roles for the educational inequality in type 2 diabetes. Comparing counterfactual and conventional approaches, we found that the mechanisms of differential exposure and differential susceptibility to being overweight and obese were not mutually exclusive and that they should be considered simultaneously to fully understand the impact of being overweight/obese on the educational inequality in type 2 diabetes incidence.

The magnitude of the educational inequality in type 2 diabetes incidence in this study is in accordance with other studies from high-income countries [1, 2]. Previous studies have found that between 8% [17] and 64% [18] of the association between education level and type 2 diabetes could be attributed to differential exposure to being overweight/obese [13–21]. Our findings corroborate this evidence and underscore that differential exposure to being overweight/obese is an important mechanism in the association between education level and type 2 diabetes incidence.

The impact of the potential interaction between education level and being overweight/obese on type 2 diabetes incidence have previously been assessed on a multiplicative scale [15, 16, 41, 42]. From a public health perspective, it has been argued that interactions should be assessed as deviation from additivity (i.e. interaction on an additive scale), since this

Table 4 Natural effects model estimating the total and decomposed rate differences in incident cases of type 2 diabetes per 100,000 person-years due to counterfactual changes in educational level (95% CI)

Variable	RD (95% CI)	Proportion of total effect (95% CI)
High→medium level of education		
Total effect of education level	316 (268, 363)	100%
Differential exposure to being overweight/obese	92 (83, 102)	29% (24, 36)
Differential susceptibility to being overweight/obese	20 (8, 32)	6% (3, 10)
Other pathways (direct effect)	204 (158, 249)	65% (59, 69)
High→low level of education		
Total effect of education level	454 (398, 510)	100%
Differential exposure to being overweight/obese	168 (149, 186)	37% (31, 45)
Differential susceptibility to being overweight/obese	41 (17, 64)	9% (4, 14)
Other pathways (direct effect)	246 (193, 298)	54% (48, 59)

N = 53,159

Data are adjusted for age, sex and ethnicity

RD, rate difference

approach directly identifies the subgroups in which most cases could potentially be prevented [28, 36]. To our knowledge, the finding of an additive interaction between education level and being overweight/obese on the risk of type 2 diabetes (i.e. differential susceptibility) is novel, and it illustrates the complex interplay between social factors and being overweight/obese that underlies the educational inequalities in type 2 diabetes.

We also compared conventional and counterfactual approaches to mediation and interaction analysis and found that the estimates of both differential exposure and susceptibility were weaker when evaluated simultaneously in the counterfactual approach than when evaluated separately using the conventional approaches. The finding that a mediator could be attributable to a larger part of an association when conventional methods are applied, compared with when counterfactual methods are applied, has previously been shown in other settings within social epidemiology [43]. It is, therefore, important to highlight that the choice of methodological approach may affect the evaluation of the importance of differential exposure and differential susceptibility to a particular risk factor. It is essential for the understanding of social inequalities in health to acknowledge that these two mechanisms are not mutually exclusive and that analyses of these mechanisms should consequently accommodate them simultaneously.

Strengths We used longitudinal data with a clear temporal ordering and separation between education level, being overweight/obese and type 2 diabetes incidence. The large observational cohort, with a long follow-up period and a large number of cases, gave us sufficient power to investigate mediation and interaction simultaneously. With this setup, we employed counterfactual mediation methods to disentangle the contribution of differential exposure and differential susceptibility to being overweight/obese in the association between education level and type 2 diabetes incidence, which is a novel approach in this setting. In addition, we assessed differential exposure and differential susceptibility on an absolute scale and, thus, directly estimated the number of type 2 diabetes cases that could potentially be prevented.

Limitations Identification of natural direct and indirect effects requires that four non-confounding assumptions are fulfilled, and at least two of these are likely violated in our study. First, the effect estimates of educational level on type 2 diabetes may be inflated due to unmeasured confounding from early-life factors, such as family socioeconomic position [44], family history of type 2 diabetes [45] and fetal/neonatal factors [46].

Second, intermediate confounders of the relationship between being overweight/obese and type 2 diabetes may bias the estimated contributions of differential exposure and

susceptibility (but not the overall association between education level and type 2 diabetes). Factors including cohabitation status, chronic stress exposure, physical activity level, dietary factors, alcohol intake and smoking are associated with both being overweight/obese and type 2 diabetes incidence. In our data, those with a lower education level were less physically active, smoked more, drank less and had a lower intake of fruit and vegetables than those with a high education level, while cohabitation was most common among those with a medium level of education (ESM Table 2). A measure of chronic stress exposure was not available. Adjusting for these variables showed somewhat similar results to those without adjustment (ESM Table 3). However, as expected, the effect of education level through other pathways (i.e. the direct effect) was weaker, and the effects of differential exposure and susceptibility to being overweight/obese were stronger. While we may not conclude that the true estimate lies between the unadjusted and adjusted estimates, it is encouraging that we still find evidence of differential exposure and susceptibility when adjusting for these variables.

Selection bias may have attenuated both the overall association between education level and type 2 diabetes and the proportion of the effect attributable to differential exposure and susceptibility. Participants in the DCH study had higher education [29] and a lower mortality rate [47] than non-participants. Our study population is, thus, disproportionately comprised of higher-educated and healthier individuals than the source population, which may affect the translatability of our findings to other settings.

Bias in the measurement of being overweight/obese is also a potential limitation. The BMI of the participants was likely to fluctuate throughout the follow-up period, yet we only had baseline measures available. One study has shown that models using either single or longitudinal measures of BMI did not differ substantially in explaining occupational inequality in type 2 diabetes incidence [48]. This suggests that measurement bias caused by having only one measurement of BMI might not be severe in our study.

The measurement of type 2 diabetes incidence might also be subject to bias. First, it has been suggested that 26% of individuals who would clinically qualify for a type 2 diabetes diagnosis in Denmark are not diagnosed [49]. The proportion is possibly higher in groups with a lower level of education [49], which could lead to differential misclassification, resulting in an underestimation of the association between education level and type 2 diabetes. Next, several validity issues have been reported with the NDR, including that up to 20% of cases included might be false-positive cases and that registration dates before 1997 are systematically delayed [34]. We assessed these issues in separate analyses and did not find a substantial impact on the results (data not included). Additionally, the NDR identifies patients with type 2 diabetes by the purchase of OGLDs, among other criteria. These drugs

can be prescribed to patients with impaired glucose tolerance and/or impaired fasting glucose who do not meet the diagnostic criteria for type 2 diabetes, resulting in a false-positive case. However, the OGLD criteria has been shown to have a high positive predictive value of 95% when validated against medical records [33]. A new diabetes register based on Danish national registries has been developed [50] but was not available for this study.

Conclusions and implications The finding of differential susceptibility to being overweight and obese (i.e. that being overweight and obese cause more cases of type 2 diabetes in groups with lower levels of education) might inform practitioners and policymakers. Practitioners may include information on education level or risk factors associated with education level to more accurately assess type 2 diabetes risk in patients. Policymakers may take into account that obesity-reducing interventions could have a higher impact on reducing type 2 diabetes rates if groups with lower education levels are specifically targeted. Such knowledge is especially useful in settings with sparse resources. Additionally, a population-wide reduction in the number of people who are overweight/obese may benefit groups with lower education levels more than groups with higher education levels in terms of lowering type 2 diabetes rates.

Future studies will need to replicate the finding of differential susceptibility to being overweight and obese. It would also be important to disentangle the mechanism of differential susceptibility to investigate whether it is the lower education level in itself or the associated contributing risk factors that are responsible for the higher rate of cases among those with low education levels. In addition, studies that include data on early-life determinants of education level, being overweight or obese, and type 2 diabetes might shed further light on the causal mechanisms behind these associations. This study has also shown the potential of the recent developments in counterfactual mediation analysis and we suggest that other researchers implement these methods in their inquiry into the social inequalities in type 2 diabetes.

References

1. Agardh E, Allebeck P, Hallqvist J, Moradi T, Sidorchuk A (2011) Type 2 diabetes incidence and socio-economic position: a systematic review and meta-analysis. Int J Epidemiol 40(3):804–818. https://doi.org/10.1093/ije/dyr029

2. Espelt A, Borrell C, Roskam AJ et al (2008) Socioeconomic inequalities in diabetes mellitus across Europe at the beginning of the 21st century. Diabetologia 51(11):1971–1979. https://doi.org/10.1007/s00125-008-1146-1

3. Chatterjee S, Khunti K, Davies MJ (2017) Type 2 diabetes. Lancet 389(10085):2239–2251. https://doi.org/10.1016/S0140-6736(17)30058-2

4. The InterAct Consortium (2012) Long-term risk of incident type 2 diabetes and measures of overall and regional obesity: the EPIC-InterAct case-cohort study. PLoS Med 9(6):e1001230. https://doi.org/10.1371/journal.pmed.1001230

5. Cohen AK, Rai M, Rehkopf DH, Abrams B (2013) Educational attainment and obesity: a systematic review. Obes Rev 14(12):989–1005. https://doi.org/10.1111/obr.12062

6. Diderichsen F, Evans T, Whitehead M (2001) The social basis of disparities in health. In: Evans T, Whitehead M, Diderichsen F, Bhuiya A, Meg M (eds) Challenging inequities in health: from ethics to action. Oxford University Press, New York, NY, pp 12–23

7. Diderichsen F, Hallqvist J, Whitehead M (2019) Differential vulnerability and susceptibility: how to make use of recent development in our understanding of mediation and interaction to tackle health inequalities. Int J Epidemiol 48(1):268–274. https://doi.org/10.1093/ije/dyy167

8. Blas E, Sivasankara Kurup A, World Health Organization (2010) Equity, social determinants and public health programmes. Available from: https://apps.who.int/iris/handle/10665/44289. Accessed: 12 August 2019

9. Kelly SJ, Ismail M (2015) Stress and type 2 diabetes: a review of how stress contributes to the development of type 2 diabetes. Annu Rev Public Health 36:441–462. https://doi.org/10.1146/annurev-publhealth-031914-122921

10. Noble N, Paul C, Turon H, Oldmeadow C (2015) Which modifiable health risk behaviours are related? A systematic review of the clustering of smoking, nutrition, alcohol and physical activity ('SNAP') health risk factors. Prev Med 81:16–41. https://doi.org/10.1016/j.ypmed.2015.07.003

11. Rothman KJ (1976) Causes. Am J Epidemiol 104(6):587–592. https://doi.org/10.1093/aje/kwx099

12. Nordahl H, Lange T, Osler M et al (2014) Education and cause-specific mortality: the mediating role of differential exposure and vulnerability to behavioral risk factors. Epidemiology 25(3):389–396. https://doi.org/10.1097/EDE.0000000000000080

13. Maty SC, Everson-Rose SA, Haan MN, Raghunathan TE, Kaplan GA (2005) Education, income, occupation, and the 34-year incidence (1965–99) of type 2 diabetes in the Alameda County study.

Int J Epidemiol 34(6):1274–1281. https://doi.org/10.1093/ije/dyi167

14. Lidfeldt J, Li TY, Hu FB, Manson JE, Kawachi I (2007) A prospective study of childhood and adult socioeconomic status and incidence of type 2 diabetes in women. Am J Epidemiol 165(8):882–889. https://doi.org/10.1093/aje/kwk078

15. Krishnan S, Cozier YC, Rosenberg L, Palmer JR (2010) Socioeconomic status and incidence of type 2 diabetes: results from the black Women's health study. Am J Epidemiol 171(5):564–570. https://doi.org/10.1093/aje/kwp443

16. Lee TC, Glynn RJ, Peña JM et al (2011) Socioeconomic status and incident type 2 diabetes mellitus: data from the Women's health study. PLoS One 6(12):e27670. https://doi.org/10.1371/journal.pone.0027670

17. Demakakos P, Marmot M, Steptoe A (2012) Socioeconomic position and the incidence of type 2 diabetes: the ELSA study. Eur J Epidemiol 27(5):367–378. https://doi.org/10.1007/s10654-012-9688-4

18. Sacerdote C, Ricceri F, Rolandsson O et al (2012) Lower educational level is a predictor of incident type 2 diabetes in European countries: the EPIC-InterAct study. Int J Epidemiol 41(4):1162–1173. https://doi.org/10.1093/ije/dys091

19. Espelt A, Borrell C, Palència L et al (2013) Socioeconomic inequalities in the incidence and prevalence of type 2 diabetes mellitus in Europe. Gac Sanit 27(6):494–501. https://doi.org/10.1016/j.gaceta.2013.03.002

20. Steele CJ, Schöttker B, Marshall AH et al (2017) Education achievement and type 2 diabetes—what mediates the relationship in older adults? Data from the ESTHER study: a population-based cohort study. BMJ Open 7(4):e013569. https://doi.org/10.1136/bmjopen-2016-013569

21. Smith PM, Smith BT, Mustard CA, Lu H, Glazier RH (2013) Estimating the direct and indirect pathways between education and diabetes incidence among Canadian men and women: a mediation analysis. Ann Epidemiol 23(3):143–149. https://doi.org/10.1016/j.annepidem.2012.12.012

22. Baron RM, Kenny DA (1986) The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 51(6):1173–1182

23. VanderWeele TJ (2015) Explanation in causal inference: methods for mediation and interaction. Oxford University Press, New York, NY

24. Richiardi L, Bellocco R, Zugna D (2013) Mediation analysis in epidemiology: methods, interpretation and bias. Int J Epidemiol 42(5):1511–1519. https://doi.org/10.1093/ije/dyt127

25. Kaufman JS, MacLehose RF, Kaufman S (2004) A further critique of the analytic strategy of adjusting for covariates to identify biologic mediation. Epidemiol Perspect Innov 1(1):4. https://doi.org/10.1186/1742-5573-1-4

26. Lange T, Vansteelandt S, Bekaert M (2012) A simple unified approach for estimating natural direct and indirect effects. Am J Epidemiol 176(3):190–195. https://doi.org/10.1093/aje/kwr525

27. VanderWeele TJ (2013) A three-way decomposition of a total effect into direct, indirect, and interactive effects. Epidemiology 24(2):224–232. https://doi.org/10.1097/EDE.0b013e318281a64e

28. Rod NH, Lange T, Andersen I, Marott JL, Diderichsen F (2012) Additive interaction in survival analysis: use of the additive hazards model. Epidemiology 23(5):733–737. https://doi.org/10.1097/EDE.0b013e31825fa218

29. Tjønneland A, Olsen A, Boll K et al (2007) Study design, exposure variables, and socioeconomic determinants of participation in diet, Cancer and health: a population-based prospective cohort study of 57,053 men and women in Denmark. Scand J Public Health 35(4):432–441. https://doi.org/10.1080/14034940601047986

30. Nordahl H, Hvidtfeldt UA, Diderichsen F et al (2014) Cohort profile: the social inequality in Cancer (SIC) cohort study. Int J Epidemiol 43(6):1750–1758. https://doi.org/10.1093/ije/dyu003

31. George AM, Jacob AG, Fogelfeld L (2015) Lean diabetes mellitus: an emerging entity in the era of obesity. World J Diabetes 6(4):613–620. https://doi.org/10.4239/wjd.v6.i4.613

32. Greenland S, Pearl J, Robins JM (1999) Causal diagrams for epidemiologic research. Epidemiology 10(1):37–48

33. Carstensen B, Kristensen JK, Marcussen MM, Borch-Johnsen K (2011) The National Diabetes Register. Scand J Public Health 39(Suppl 7):58–61. https://doi.org/10.1177/1403494811404278

34. Green A, Sortsø C, Jensen PB, Emneus M (2014) Validation of the Danish National Diabetes Register. Clin Epidemiol 7:5–15. https://doi.org/10.2147/CLEP.S72768

35. McKeague I, Sasieni P (1994) A partly parametric additive risk model. Biometrika 1(3):501–514

36. Greenland S, Lash TL, Rothman KJ (2008) Concepts of interaction. In: Rothman KJ, Greenland S, Lash TL (eds) Modern epidemiology, 3rd edn. Lippincott Williams & Wilkins, Philadelphia, PA, pp 71–84

37. Lange T, Hansen JV (2011) Direct and indirect effects in a survival context. Epidemiology 22(4):575–581. https://doi.org/10.1097/EDE.0b013e31821c680c

38. Lange T, Hansen KW, Sørensen R, Galatius S (2017) Applied mediation analyses: a review and tutorial. Epidemiol Health 39:e2017035. https://doi.org/10.4178/epih.e2017035

39. SAS Institute Inc (2002) SAS/STAT. SAS Institute Inc., Cary, NC, USA

40. R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

41. Shang X, Li J, Tao Q et al (2013) Educational level, obesity and incidence of diabetes among Chinese adult men and women aged 18–59 years old: an 11-year follow-up study. PLoS One 8(6):e66479. https://doi.org/10.1371/journal.pone.0066479

42. Wang A, Stronks K, Arah OA (2014) Global educational disparities in the associations between body mass index and diabetes mellitus in 49 low-income and middle-income countries. J Epidemiol Community Health 68(8):705–711. https://doi.org/10.1136/jech-2013-203200

43. Naimi AI, Schnitzer ME, Moodie EEM, Bodnar LM (2016) Mediation analysis for health disparities research. Am J Epidemiol 184(4):315–324. https://doi.org/10.1093/aje/kwv329

44. Tamayo T, Herder C, Rathmann W (2010) Impact of early psychosocial factors (childhood socioeconomic factors and adversities) on future risk of type 2 diabetes, metabolic disturbances and obesity: a systematic review. BMC Public Health 10(1):525. https://doi.org/10.1186/1471-2458-10-525

45. The InterAct Consortium (2013) The link between family history and risk of type 2 diabetes is not explained by anthropometric, lifestyle or genetic risk factors: the EPIC-InterAct study. Diabetologia 56(1):60–69. https://doi.org/10.1007/s00125-012-2715-x

46. Stein AD, Obrutu OE, Behere RV, Yajnik CS (2019) Developmental undernutrition, offspring obesity and type 2 diabetes. Diabetologia 62(10):1773–1778. https://doi.org/10.1007/s00125-019-4930-1

47. Larsen SB, Dalton SO, Schüz J et al 2012) Mortality among participants and non-participants in a prospective cohort study. Eur J Epidemiol 27(11):837–845. https://doi.org/10.1007/s10654-012-9739-x

48. Stringhini S, Tabak AG, Akbaraly TN et al (2012) Contribution of modifiable risk factors to social inequalities in type 2 diabetes: prospective Whitehall II cohort study. BMJ 345:e5452. https://doi.org/10.1136/bmj.e5452

49. Holm AL, Andersen GS, Jørgensen ME, Diderichsen F (2018) Is the rule of halves framework relevant for diabetes care in Copenhagen today? A register-based cross-sectional study. BMJ Open 8(11):e023211. https://doi.org/10.1136/bmjopen-2018-023211

50. Carstensen B, Jørgensen ME (2019) A Danish diabetes register. Version 37. Steno Diabetes Center Copenhagen, Gentofte, Denmark. Available from: http://bendixcarstensen.com/DMreg/Reg2016.pdf. Accessed: 6 January 2020

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.