

# Paediatric obesity appears to lower the risk of diabetes if selection bias is ignored

Steven D Stovitz,<sup>1</sup> Hailey R Banack,<sup>2</sup> Jay S Kaufman<sup>3</sup>

<sup>1</sup>Department of Family Medicine and Community Health, University of Minnesota, Minneapolis, Minnesota, USA

<sup>2</sup>Department of Epidemiology and Environmental Health, School of Public Health and Health Professions, University at Buffalo, The State University of New York, Buffalo, New York, USA

<sup>3</sup>Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec, Canada

## Correspondence to

Dr Steven D Stovitz, Department of Family Medicine and Community Health, University of Minnesota, Minneapolis, MN 55455, USA; [stovitz@umn.edu](mailto:stovitz@umn.edu)

Received 14 September 2017

Revised 4 December 2017

Accepted 18 December 2017

## ABSTRACT

**Background** Frustrated with the onslaught of articles reporting fascination with results that appear paradoxical but are merely due to selection bias, we studied the apparent effect of obesity on diabetes risk in youth who had a test for diabetes. We hypothesised that obese subjects would have lower rates of diabetes than non-obese subjects due to selection bias, and consequently, obesity would appear to lower the risk of diabetes.

**Methods** Retrospective cohort study of children (4–9 years), pre-teens (10–12 years) and teenagers (13–19 years). Participation was restricted to those who had a test of haemoglobin A1C along with measured height and weight. Body mass index percentile via the Centers for Disease Control and Prevention age and sex standards was calculated and categorised. The main outcome was A1C%, subsequently categorised at the level for diagnosis of diabetes mellitus ( $\geq 6.5\%$ ).

**Results** The sample consisted of 134 (2%) underweight, 1718 (30%) healthy weight, 660 (12%) overweight and 3190 (56%) obese individuals. 16% ( $n=936$ ) had an A1C  $\geq 6.5\%$ . Overall, healthy weight children had 8.2 times the risk of A1C  $\geq 6.5\%$  (95% CI 5.3 to 12.7) compared with those in the obese category. The relative risk was 13 in pre-teens (95% CI 8.5 to 20.0) and 3.9 in teenagers (95% CI 3.3 to 4.7).

**Conclusions** Healthy weight was associated with a 4–13 times higher relative risk of diabetes mellitus compared with being obese. While apparently shocking, the study's fatal flaw (selection bias) explains the 'paradoxical' finding. Ignoring selection bias can delay advances in medical science.

## INTRODUCTION

Peer-reviewed medical journals are regularly publishing studies where the recruitment of subjects is based on an effect of the exposure of interest, a methodological error that may induce a form of selection bias.<sup>1–3</sup> Selection bias can occur at the study design or analysis stage as a result of conditioning on a common effect of two variables, one of which is the exposure, or a variable associated with the exposure, and the other of which is the outcome, or a variable associated with the outcome.<sup>4</sup> Selection bias may reverse the expected direction of an effect, producing results that appear dramatically contrary to expectations. These surprising results are not due to errors in previous research, random variation or statistical mistakes. Rather, the results are entirely predictable as they are due to a systematic error that operates occultly behind the scenes of the conduct of the rest of the study. If participants are chosen based on an effect of the study's exposure variable,

then that exposure variable can appear protective. The bias in the structure of these studies has been described in detail<sup>4–6</sup> and can create associations that are in the opposite direction from that which occurs in the general population.

Since the results of studies with selection bias appear contrary to expectations, some call the results paradoxical. The most common of these is the so-called 'obesity paradox' which occurs when studies evaluate the effect of obesity among individuals who have a disease caused by obesity, such as heart failure, stroke or myocardial infarction. These studies report that among individuals with the disease the subjects who were obese fare better than those who were not obese.<sup>7–10</sup> While a health risk such as obesity may confer some protective benefits, for example, stronger bones, the apparently protective benefits of obesity are due to selection bias when subjects are chosen based on an effect of obesity. Despite the known methodological explanation for this paradox,<sup>3 11 12</sup> articles reporting evidence of an obesity paradox are on the rise. Our search of EMBASE for the phrase 'obesity paradox' in the title, abstract or keywords found that the first article appeared in the year 2002<sup>13</sup> and there are now >800 full-text peer-reviewed articles.<sup>14</sup> While some of these studies describe the selection bias that causes the apparently paradoxical results, the vast majority express a fascination with the results and call for more research into the possible ways that obesity protects people from the hazards of diseases caused by obesity, such as heart failure.<sup>7–9</sup>

Frustrated with the onslaught of articles expressing fascination with these paradoxical results while not explaining to the reader that the results were, in fact, predictable and due to selection bias, we decided to conduct our own study. We chose to study diabetes mellitus in youth as this is an illness with two well-described mechanisms; type 1, due to lack of insulin production, and type 2, due to lack of insulin sensitivity which is a consequence of obesity.<sup>15</sup> While diabetes in young people was historically thought to be only the result of type 1 diabetes, the paediatric obesity epidemic has led to an increasing number of cases of type 2 diabetes.<sup>16 17</sup>

Diabetes risk is often assessed by a person's level of haemoglobin A1C (A1C), a measure of blood sugar control. As such, having an A1C test is an outcome of either being obese or being ill with symptoms and signs of diabetes. An A1C level  $\geq 6.5\%$  is generally considered diagnostic of diabetes.<sup>15</sup> The purpose of our study was to evaluate the association of A1C level with weight status in children and adolescents ages 4–19 years. We hypothesised that among those



**To cite:** Stovitz SD, Banack HR, Kaufman JS. *J Epidemiol Community Health* Epub ahead of print: [please include Day Month Year]. doi:10.1136/jech-2017-209985

## Research report

with at least one A1C level obese subjects would have lower rates of diabetes than non-obese subjects due to selection bias, and consequently, obesity would appear to lower the risk of diabetes.

## METHODS

## Study sample

This project was a retrospective cohort study involving a chart-review analysis restricted to youth who had an A1C blood test drawn. Clinical data from medical records in the University of Minnesota Clinical Data Repository were compiled by data analysts through the Best Practices Integrative Informatics Consulting and Collaborative Science Core. Participants in our sample were patients who had (1) at least one A1C measure when they were between 4 and 19 years of age between 6 January 2011 and 4 August 2017, and (2) a height and weight measurement associated with this A1C. If there were multiple A1C measurements for one participant, only the first A1C measurement was included in the analysis.

## Measures

The primary exposure in this study is age and sex-specific body mass index (BMI) percentile. BMI percentile was calculated for all participants from measured height and weight using the Centers for Disease Control and Prevention's (CDC's) SAS Program for the 2000 CDC Growth Charts.<sup>18</sup> BMI percentile was broken into categories based on the CDC guidelines for defining childhood weight status where underweight is less than the 5th percentile, healthy weight is from the 5th to less than the 85th percentiles, overweight is from the 85th to less than the 95th percentiles, and obese is at the 95th percentile or greater.

The primary outcome in this study is haemoglobin A1C (%) as collected from the participants' medical records. A1C levels were categorised as either <6.5% or ≥6.5%, the level for diagnosis of diabetes mellitus. Among those with diabetes, we also categorised those with a level ≥8% as having poorly controlled diabetes. Other demographics and sample characteristics examined in this sample included age, race, ethnicity, height, weight and blood pressure.

## Statistical methods

SAS Software V.9.4 (SAS Institute Inc) was used for analysis. Data were examined for outliers and implausible values using histograms and quantile plots. The CDC SAS program for calculating BMI percentile flagged n=24 participants with extreme or biologically implausible values for height, weight and BMI (using the following modified z-score cut-offs: BMI for age <-5 or >13, weight for age <-5 or >13, height for age <-5 or >5). These participants were excluded from the main analysis. The A1C values of the excluded participants were compared with those included via a sensitivity analysis using a two-sample t-test (excluded mean±SD: 5.35±0.72; included mean±SD: 6.04±1.76; P=0.06).

Descriptive statistics were calculated for participant demographics and health characteristics. Pearson's correlation coefficients and simple linear regression were used to test for a linear association between continuous BMI percentile and haemoglobin A1C. Generalised linear regression models were used to compare relative and absolute contrasts of risk of a high A1C (≥6.5% as well as ≥8%) between underweight, healthy weight, overweight and obese participants. We assumed that age would modify the association of weight status and A1C and divided the cohort into children (4–9 years), pre-teens (10–12 years) and teenagers (13–19 years). Within these groups, there was no

significant interaction by age. Results are reported as relative and absolute contrasts, and their 95% CIs for all ages together, as well as for ages 4–9 and 10–12, and 13–19 years, separately.

## RESULTS

## Demographics

Our search strategy identified 9135 unique individuals between the ages of 4 and 19 years who had A1C level measures in our medical records. We excluded 3433 individuals due to height or weight values that were either missing (n=3409) or implausible (n=24). The demographics of the analysed sample of 5702 can be seen in [table 1](#). On average, participants were (mean±SD) 13.0±4.1 years of age, 53% female, 57% white and 85% non-Hispanic. The sample consisted of 134 (2%) underweight, 1718 (30%) healthy weight, 660 (12%) overweight and 3190 (56%) obese individuals. The average A1C level of all participants was 6.04%±1.8%, and 16% of participants (n=936) had an A1C ≥6.5%.

## Continuous relationship between BMI percentile and A1C

Among all participants, there was a weak negative correlation between BMI percentile and A1C (r=-0.20, P<0.0001). Linear regression revealed that for each 10% increment in BMI percentile A1C was 0.13% lower (95% CI -0.14 to -0.11). The negative correlation occurred for participants in each of the three groups; ages 4–9 years (r=-0.22; β = -0.12 (95% CI -0.16 to -0.09)), ages 10–12 years (r=-0.27; β = -0.15 (95% CI -0.18 to -0.12)) and ages 13–19 years (r=-0.17; β = -0.12 (95% CI -0.14 to -0.10)).

## Weight category risk of diabetes mellitus (A1C≥6.5%)

As seen in [table 2](#) and [figure 1](#), in all age groups, those in the obese weight category had the lowest risk for an A1C ≥6.5%, and those in the healthy weight category had the highest risk for an A1C ≥6.5%. Children in the healthy weight category had 8.2 times (95% CI 5.3 to 12.7) the risk of A1C ≥6.5% than those in the obese category. In pre-teens, the relative risk (RR) was 13 (95% CI 8.5 to 20.0), and in teenagers the RR was 3.9 (95% CI=3.3 to 4.7).

These RR differences are especially notable given the prevalence of A1C ≥6.5% in our cohort. The absolute risk of an A1C ≥6.5% in healthy weight pre-teens was 38.9%, whereas in obese pre-teens the absolute risk of an A1C ≥6.5% was 3.0%. Consequently, the risk difference (RD) was 35.9% (number needed to treat=1/RD=2.8), suggesting that only three healthy weight pre-teens would need to become obese to prevent one case of diabetes. The number needed to treat (from healthy to obese) in children was 3.8, and in the teenagers was 4.2.

## Risk of elevated A1C levels (≥8%) among those with diabetes mellitus

As seen in [table 3](#), among those with diabetes, overall 77% had an A1C level ≥8%, signifying poor control. This percentage was generally similar between the weight categories with the exception of the underweight subjects who had a higher percentage with poor control in the child and teenage groups (RR=1.3 and 1.4, respectively).

## DISCUSSION

In this study of 5702 children and adolescents, we found that obesity was *negatively* associated with A1C levels and the risk of diabetes, that is, the higher the BMI percentile, the lower the risk of A1C ≥6.5%. In our cohort, subjects who were of a

**Table 1** Sample demographics and characteristics

Values are mean±SD (range) or n (%)	
N	5702
Age (years)	13.04±4.10 (4–19)
Children: 4–9	1205 (21.13%)
Pre-teens: 10–12	1194 (20.94%)
Teenagers: 13–19	3303 (57.93%)
Sex	
Female	3018 (52.93%)
Male	2684 (47.07%)
Race	
White	3254 (57.07%)
Asian	317 (5.56%)
American Indian or Alaska Native	103 (1.81%)
Black or African-American	942 (16.52%)
Native Hawaiian or other Pacific Islander	14 (0.25%)
More than one race or other race	93 (1.63%)
Unknown	979 (17.17%)
Ethnicity	
Hispanic or Latino	524 (9.19%)
Not Hispanic or Latino	4869 (85.39%)
Unknown	309 (5.42%)
Encounters in chart review time period (n)	
1	4061 (71.22%)
2–3	909 (15.94%)
4+	732 (12.84%)
Haemoglobin A1C (%)	6.04±1.76 (3.90–15.90)
<5.7	3928 (68.89%)
5.7 to <6.5	838 (14.70%)
6.5 to <8.0	240 (4.21%)
≥8.0	696 (12.21%)
Height (inches)	61.14±7.68 (36.34–79.09)
Weight (lbs)	154.87±70.87 (27.20–615.30)
BMI percentile	81.27±27.48 (0.00–100.00) Median: 96.62
Underweight (<5th percentile)	134 (2.35%)
Healthy weight (5th to <85th percentile)	1718 (30.13%)
Overweight (85th to <95th percentile)	660 (11.57%)
Obese (≥ 95th percentile)	3190 (55.95%)
BMI z-score	1.39±1.27 (–10.63–5.36)
Maximum blood glucose (±20 days from A1c measure, mg/dL)	121.64±89.87 (50–811)
N (%) with max. glucose value	3504 (61.45%)
SBP (mm Hg)	113.98±12.67 (78–204)
DBP (mm Hg)	69.08±9.27 (36–113)
N (%) with SBP/DBP values	5590 (98.04%)

BMI, body mass index; DBP, diastolic blood pressure; SBP, systolic blood pressure.

‘healthy weight’ had the *highest* risk of diabetes. Among children, pre-teens and teenagers, a healthy weight was associated with a 4–13 times higher risk of diabetes compared with those who were obese.

Since obesity is known to be a risk factor for diabetes, and yet in our study appeared protective, some would call our results ‘paradoxical’. However, our results are simply a textbook example of catastrophic selection bias. Consistent with the common

definition of ‘selection’, we selected our participants based on a criterion (having an A1C drawn) that led to a biased sample. Selection bias can occur at various points in a study’s design or analysis when there is conditioning on a common effect of two variables, one of which is the exposure, or a variable associated with the exposure, and the other of which is the outcome, or a variable associated with the outcome.<sup>4</sup> The causal inference literature refers to the common effect of two variables as a ‘collider’; hence, the synonymous term ‘collider stratification bias’ for this phenomenon.<sup>6</sup> The name derives from the causal structure when two or more variables ‘collide’ at a particular node (variable) in a directed acyclic graph (DAG). In our study, obesity and illness ‘collide’ at the node representing a laboratory test for A1C, thus conditioning (via restriction) on having an A1C test leads them to be inversely associated in the study sample, even if they are unassociated/independent in the general population.<sup>6</sup> Participants in our study were only included in the analytic cohort if their clinicians determined that a blood test for A1C level should be obtained, meaning that they had some type of clinical indication to warrant such a test. Many clinicians draw an A1C level when evaluating obese children and adolescents because recommendations include the consideration of testing for diabetes.<sup>19</sup> Clinicians also draw an A1C level when children have symptoms of diabetes and appear ill, regardless of bodyweight status. This is the key to understanding the selection bias<sup>4–6</sup>: ordering the A1C is an effect of either being obese or appearing ill with symptoms and signs of diabetes.

We have drawn out a DAG representing our study design (see [figure 2A](#)). Since all subjects in our study had an A1C drawn, the analysis is conditioned on having had an A1C test by restricting the study population. A box around the A1C node represents this conditioning. Conditioning on having an A1C test induces an inverse and non-causal association between ‘obesity’ and ‘appearing ill’. We have used ‘appearing ill’ as an example of an unmeasured common cause of having an A1C test drawn and diabetes. Future studies could attempt to measure and then adjust for variables that are common causes of having A1C test drawn and diabetes, such as child thirst or urinary output. It is important to note that the simplified DAG presented in [figure 2A](#) and the structural diagram in [figure 2Bb](#) do not include potential confounders of the exposure–outcome relationship. For additional DAGs depicting various study design structures representing selection bias, we refer readers to other publications.<sup>4 20</sup> Since having symptoms and signs of diabetes is a much stronger risk factor for diabetes than obesity, in our study design, obesity paradoxically appears protective.

Our study had all of the appearances of a high-quality medical publication. It was conducted at a major academic medical setting and abided by all of the tenets of scientific guidelines. We have attached our Strengthening the Reporting of Observational Studies in Epidemiology checklist as an addendum. Consider the apparent strengths of our study; we had a large sample size with 5702 different individuals. Thus, our findings, while wrong, are unlikely to be due to chance. The subjects were from a diverse population of urban and rural settings seen in the clinical setting. Therefore, we could tout that our results are likely generalisable to other clinic populations of children and adolescents. Finally, the P values were really, really small (<0.0001), and so many readers may think that there must be some truth in the findings.

Since paediatric obesity is clearly associated with decreased insulin sensitivity and a higher risk for chronic diseases, including diabetes mellitus,<sup>21–23</sup> we could write a discussion section filled

## Research report

**Table 2** Risk of high A1C level ( $\geq 6.5\%$ ) by weight status categories

BMI category	A1C <6.5% n (row %)	A1C $\geq 6.5\%$ n (row %)	RR (95% CI)	P value	RD (95% CI)	P value
All participants (n=5702)*						
Underweight	105 (78.4%)	29 (21.64%)	3.6 (2.5 to 5.1)	<0.0001	15.7% (8.6% to 22.7%)	<0.0001
Healthy weight	1164 (67.8%)	554 (32.3%)	5.5 (4.6 to 6.3)	<0.0001	26.3% (23.9% to 28.6%)	<0.0001
Overweight	498 (75.5%)	162 (24.6%)	4.1 (3.4 to 5.0)	<0.0001	18.6% (15.2% to 21.9%)	<0.0001
Obese	2999 (94.0%)	191 (6.0%)	Ref		Ref	
Children						
Ages 4–9 (n=1205)						
Underweight	26 (76.5%)	8 (23.5%)	6.6 (3.2 to 13.7)	<0.0001	20.0% (5.6% to 34.3%)	0.0064
Healthy weight	302 (70.6%)	126 (29.4%)	8.2 (5.3 to 12.7)	<0.0001	25.9% (21.3% to 30.4%)	<0.0001
Overweight	92 (72.4%)	35 (27.6%)	7.7 (4.7 to 12.7)	<0.0001	24.0% (16.1% to 31.9%)	<0.0001
Obese	594 (96.4%)	22 (3.6%)	Ref		Ref	
Pre-teens						
Ages 10–12 (n=1194)						
Underweight	17 (73.9%)	6 (26.1%)	8.7 (3.9 to 19.4)	<0.0001	23.1% (5.1% to 41.1%)	0.0118
Healthy weight	165 (61.1%)	105 (38.9%)	13.0 (8.5 to 20.0)	<0.0001	35.9% (30.0% to 41.8%)	<0.0001
Overweight	102 (77.9%)	29 (22.1%)	7.4 (4.4 to 12.4)	<0.0001	19.2% (11.9% to 26.4%)	<0.0001
Obese	747 (97.0%)	23 (3.0%)	Ref		Ref	
Teenagers						
Ages 13–19 (n=3303)						
Underweight	62 (80.5%)	15 (19.5%)	2.4 (1.5 to 3.9)	0.0003	11.4% (2.5% to 20.3%)	0.0125
Healthy weight	697 (68.3%)	323 (31.7%)	3.9 (3.3 to 4.7)	<0.0001	23.6% (20.5% to 26.7%)	<0.0001
Overweight	304 (75.6%)	98 (24.4%)	3.0 (2.4 to 3.8)	<0.0001	16.3% (11.9% to 20.7%)	<0.0001
Obese	1658 (91.9%)	146 (8.1%)	Ref		Ref	

Weight status definitions via BMI percentiles with underweight = <5th, healthy weight = 5th to <85th, overweight = 85th to <95th and obese  $\geq 95$ th.

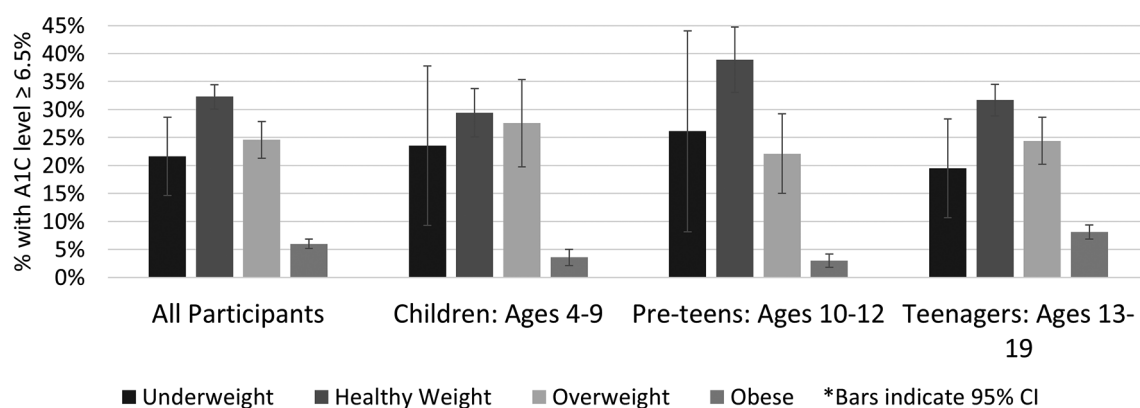
\*Crude results for all participants are reported here, even though there is a statistically significant interaction between age and weight status category (P value<0.0001).

BMI, body mass index; RD, risk difference; Ref, reference category; RR, relative risk.

with headline-grabbing quotes describing our shocking results. Imagine news outlets getting hold of ‘Healthy weight is a risk factor for diabetes’. Since the risk (probability) differences might be attributed to causal effects, then, given the absolute RDs of 25.9%, 35.9% and 23.6% in the children, pre-teens and teenagers, respectively, industries selling high caloric foods and beverages might put together slick presentations demonstrating that clinicians need to help only 3–4 healthy weight youth become obese to prevent a case of diabetes; a huge return on investment.

We could write a section on our study’s limitations, with a few perfunctory comments such as the fact that heights and weights were performed and recorded by clinic personnel and

subject to routine mismeasurements. Also, we might include the standard comments that BMI is an imperfect measure of obesity. Yet all of these strengths and limitations are trumped by the selection bias. Given the selection bias, there are no strengths that should make us more confident, nor limitations that should make us more uncertain about the results. As explained by others,<sup>24</sup> selection bias does not lead to a mere limitation of generalisability, but rather affects validity because the associations in the sample do not have a causal interpretation. In the present example, selection bias was unavoidable as only a selected group of participants had an A1C test. Had we evaluated every child seen at our clinics between the ages of 4 and 19 years, there would have been no selection bias.

**Figure 1** Risk of A1C level  $\geq 6.5\%$  by body mass index category and age stratification.



**Table 3** Proportion of diabetics\* with A1C  $\geq 8\%$  by weight status† and age groups

	Children: ages 4–9 years	Pre-teens: ages 10–12 years	Teenagers: ages 13–19 years	Overall
Underweight	8/8 (100%)	4/6 (67%)	14/15 (93%)	26/29 (90%)
Healthy weight	87/126 (69%)	75/105 (71%)	256/323 (79%)	418/554 (76%)
Overweight	24/35 (69%)	21/29 (72%)	77/98 (79%)	122/162 (75%)
Obese	16/22 (73%)	14/23 (61%)	100/146 (69%)	130/191 (68%)
Overall	135/191 (71%)	114/163 (70%)	447/582 (77%)	723/936 (77%)

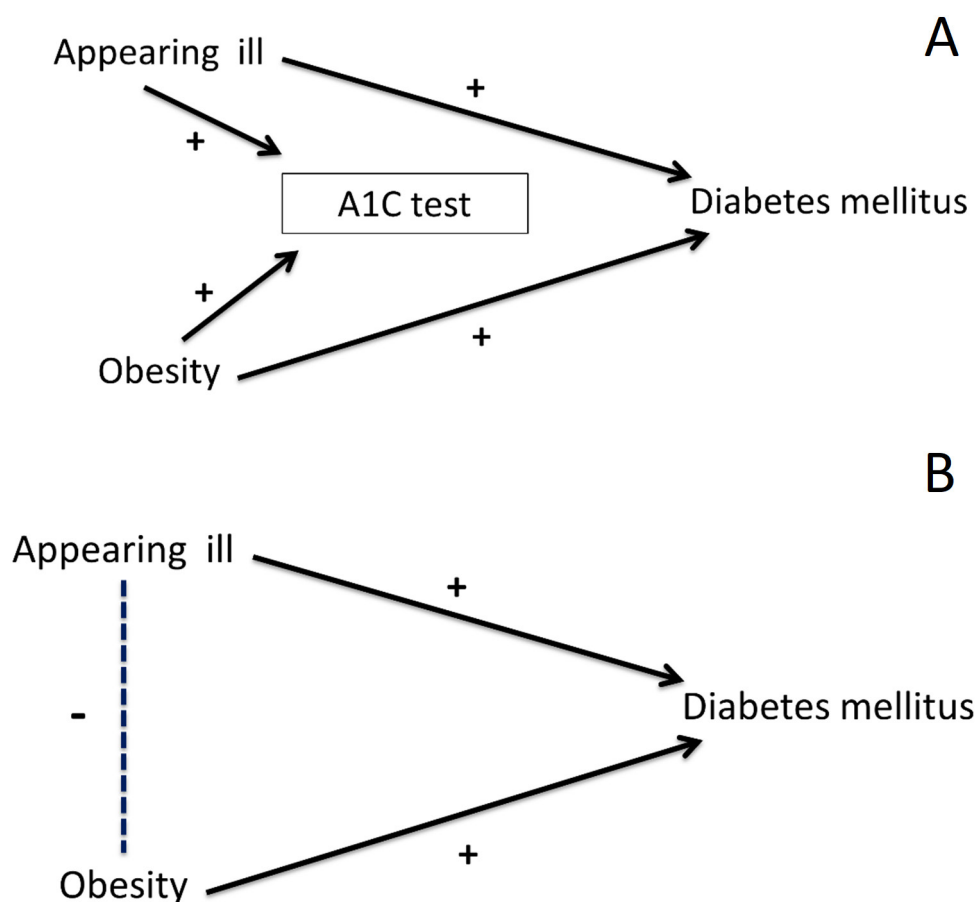
\*Diabetics defined as those with an A1C  $\geq 6.5\%$ .

†Weight status definitions via body mass index percentiles with underweight= $<5$ th, healthy weight= $5$ th to  $<85$ th, overweight= $85$ th to  $<95$ th and obese  $\geq 95$ th.

Furthermore, there are analytic techniques to account for differential selection, such as inverse probability weighting, that could have been used to adjust for this form of bias.<sup>25</sup>

Unfortunately, selection bias has rarely proved a barrier to publication. There is an ever-increasing body of literature on studies that are structurally equivalent to the design of our study, whereby participants are restricted to an effect of the exposure and the results are labelled as a ‘paradox’. In our study, the

non-obese subjects had another cause of their diabetes, that is, lack of insulin production. In the studies of patients with other ramifications of obesity,<sup>7–10</sup> the patients who were not obese, by definition, had other causes for their illness. For example, heart failure can be due to genetic issues, infectious cardiomyopathy or cancer. If these other causes are more dangerous than obesity, then those who contracted heart failure due to obesity appear protected. Examples of selection bias have also



**Figure 2** (A) Directed acyclic graph of assumed causes and effects in our study design. The box around ‘A1C test’ signifies that this variable was fixed by requiring an A1C test as a necessary condition for being included in the study. The variable ‘appearing ill’ was unmeasured but assumed as a reason why clinicians would order an A1C test in youth who were not obese. (B) The associational structure that results from conditioning on an A1C test, generating an inverse and non-causal association between ‘obesity’ and ‘appearing ill’. Although the true causal relation between obesity and diabetes mellitus remains positive, the observed relationship becomes distorted by the inverse association generated between obesity and ‘appearing ill’. If the magnitude of this inverse association is sufficiently strong, this bias can push the observed association below the null so that obesity appears to have a protective effect on diabetes mellitus.

## Research report

appeared in other medical specialties. In the obstetrical literature, causes of small or pre-term babies (maternal smoking and pre-eclampsia) paradoxically appear protective when observing only babies below a certain weight or gestational age.<sup>2</sup> Lack of understanding selection bias sets back advancements in medical science. For example, just a few decades ago, authors struggled in vain to provide a physiological explanation for the positive effects of maternal smoking on newborn health.<sup>20 26</sup> Now, due to the articles claiming an obesity paradox, researchers are starting to question the standard the clinical recommendation for obese subjects with heart failure to lose weight.<sup>8 27</sup>

Of course, with diabetes there are considered to be two types, with type 1 traditionally thought to occur in the young and thinly built, and type 2 understood to be more common in the older and obese. In fact, there is much more overlap than those classic divisions suggest. A database of >5500 adolescents with type 1 diabetes found that 23% were overweight and 13% were obese.<sup>28</sup> Furthermore, in youth with type 2 diabetes, the mean age of onset is 13.5 years.<sup>29</sup> The one stereotypical feature that seems to prevail is that few, if any, type 2 diabetics are lean. The fact that obesity is, on a population basis, more common in type 2 than in type 1 diabetes creates the observed inverse correlation with illness in our study when one selects only those considered at risk by their clinician.

Were our example not so absurd and contrary to existing knowledge, it might pass for a credible finding, and children might be recommended for fast-food gavage as a means to stave off diabetes. While paediatric obesity, similar to smoking, as a cause of disease prevention appears absurd, adult obesity as a health benefit seems more plausible. Obesity is associated with increased muscle mass and bone strength. Yet, the structure of the selection bias is the same in our study as in studies reporting that adult obesity is protective against outcomes that it causes. Risk factors can appear paradoxically protective if researchers restrict their studies to selected populations and disregard the bias that affects all of the results. The punchline here is that selected samples can contain non-causal associations that reverse the true

causal relations present in the general population. Researchers who ignore this simple and repeatable fact will generate harmful misinformation that is no joke.

**Acknowledgements** The authors thank Samantha Carlson, MPH for assistance with the statistical analysis. Support was provided by Research Services, Department of Family Medicine and Community Health, Medical School, University of Minnesota.

**Contributors** All authors contributed to the concept and design of the study, as well as the analysis and manuscript writing.

**Funding** HRB received funding from the Canadian Institute of Health Research Banting Postdoctoral Fellowship Award.

**Competing interests** None declared.

**Ethics approval** Our study involved deidentified analysis of pre-existing data and was determined by the University of Minnesota's Institutional Review Board to be exempt for the need for further review.

**Provenance and peer review** Not commissioned; externally peer reviewed.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

## REFERENCES

- 1 Dahabreh IJ, Kent DM. Index event bias as an explanation for the paradoxes of recurrence risk research. *JAMA* 2011;305:822–3.
- 2 Ananth CV, Schisterman EF, Confounding SEF. Confounding, causality, and confusion: the role of intermediate variables in interpreting observational studies in obstetrics. *Am J Obstet Gynecol* 2017;217:167–75.
- 3 Banack HR, Kaufman JS. Does selection bias explain the obesity paradox among individuals with cardiovascular disease? *Ann Epidemiol* 2015;25:342–9.
- 4 Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15:615–25.
- 5 Westreich D. Berkson's bias, selection bias, and missing data. *Epidemiology* 2012;23:159–64.
- 6 Cole SR, Platt RW, Schisterman EF, et al. Illustrating bias due to conditioning on a collider. *Int J Epidemiol* 2010;39:417–20.
- 7 Lavie CJ, De Schutter A, Milani RV. Body composition and the obesity paradox in coronary heart disease: can heavier really be healthier? *Heart* 2015;101:1610–1.
- 8 Lavie CJ, Sharma A, Alpert MA, et al. Update on Obesity and Obesity Paradox in Heart Failure. *Prog Cardiovasc Dis* 2016;58:393–400.
- 9 Sharma A, Lavie CJ, Borer JS, et al. Meta-analysis of the relation of body mass index to all-cause and cardiovascular mortality and hospitalization in patients with chronic heart failure. *Am J Cardiol* 2015;115:1428–34.
- 10 Goel K, Lopez-Jimenez F, De Schutter A, et al. Obesity paradox in different populations: evidence and controversies. *Future Cardiol* 2014;10:81–91.
- 11 Lajous M, Bijon A, Fagherazzi G, et al. Body mass index, diabetes, and mortality in French Women. *Epidemiology* 2014;25:10–14.
- 12 Preston SH, Stokes A. Obesity Paradox. *Epidemiology* 2014;25:454–61.
- 13 Gruber L, Weissman NJ, Waksman R, et al. The impact of obesity on the short-term and long-term outcomes after percutaneous coronary intervention: the obesity paradox? *J Am Coll Cardiol* 2002;39:578–84.
- 14 Stovitz SD, Banack HR, Kaufman JS. Structural Bias in Studies of Cardiovascular Disease: Let's Not Be Fooled by the "Obesity Paradox". *Can J Cardiol* 2017.
- 15 American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2014;37(Suppl 1):S81–S90.
- 16 Pulgaron ER, Delamater AM. Obesity and type 2 diabetes in children: epidemiology and treatment. *Curr Diab Rep* 2014;14:508.
- 17 Mayer-Davis EJ, Lawrence JM, Dabelea D, et al. Incidence trends of type 1 and type 2 diabetes among youths, 2002–2012. *N Engl J Med* 2017;376:1419–29.
- 18 Centers for disease control and prevention. A SAS program for the 2000 CDC growth charts (ages 0 to <20 years). <https://www.cdc.gov/nccddp/dnpao/growthcharts/resources/sas.htm> (accessed 10 Aug 2017).
- 19 Krebs NF, Himes JH, Jacobson D, et al. Assessment of child and adolescent overweight and obesity. *Pediatrics* 2007;120 Suppl 4(Suppl4):S193–S228.
- 20 Hernández-Díaz S, Schisterman EF, Hernán MA. The birth weight "paradox" uncovered? *Am J Epidemiol* 2006;164:1115–20.
- 21 Friedemann C, Heneghan C, Mahtani K, et al. Cardiovascular disease risk in healthy children and its association with body mass index: systematic review and meta-analysis. *BMJ* 2012;345:e4759.
- 22 Steinberger J, Moran A, Hong CP, et al. Adiposity in childhood predicts obesity and insulin resistance in young adulthood. *J Pediatr* 2001;138:469–73.
- 23 Sinha R, Fisch G, Teague B, et al. Prevalence of impaired glucose tolerance among children and adolescents with marked obesity. *N Engl J Med* 2002;346:802–10.
- 24 Flanders WD, Klein M. Properties of 2 counterfactual effect definitions of a point exposure. *Epidemiology* 2007;18:453–60.

## What is already known on this subject

- Selection bias can occur when inclusion criteria are based on a consequence of exposure and may produce results that appear in the opposite direction of the true causal effects.
- The medical literature is full of examples where selection bias produces counterintuitive results that are labelled a 'paradox' with a call for more research.
- Diabetes in children and adolescents is a result of either insulin resistance (caused by obesity) or lack of insulin production.

## What this study adds

- A clear example where selection bias produces results that are counterintuitive and nonsensical.
- Not a whole lot to our understanding of the link between paediatric obesity and diabetes.
- A good strategy for generating headline-grabbing results: study only those with a defined outcome and then look for a cause of the outcome that appears 'paradoxically' to be protective.

- 25 Banack HR, Kaufman JS. The obesity paradox: understanding the effect of obesity on mortality among individuals with cardiovascular disease. *Prev Med* 2014;62:96–102.
- 26 Yerushalmy J. The relationship of parents' cigarette smoking to outcome of pregnancy—implications as to the problem of inferring causation from observed associations. *Am J Epidemiol* 1971;93:443–56.
- 27 Charnigo R, Guglin M. Obesity paradox in heart failure: statistical artifact, or impetus to rethink clinical practice? *Heart Fail Rev* 2017;22:13–23.
- 28 Minges KE, Whittemore R, Weinzimer SA, *et al.* Correlates of overweight and obesity in 5529 adolescents with type 1 diabetes: The T1D Exchange Clinic Registry. *Diabetes Res Clin Pract* 2017;126:68–78.
- 29 Rosenbloom AL, Silverstein JH, Amemiya S, *et al.* International society for pediatric and adolescent diabetes. Ispad clinical practice consensus guidelines 2006–2007. Type 2 diabetes mellitus in the child and adolescent. *Pediatr Diabetes* 2008;9:512–26.



# Paediatric obesity appears to lower the risk of diabetes if selection bias is ignored

Steven D Stovitz, Hailey R Banack and Jay S Kaufman

*J Epidemiol Community Health* published online January 26, 2018

---

Updated information and services can be found at:

<http://jech.bmj.com/content/early/2018/01/25/jech-2017-209985>

---

*These include:*

## References

This article cites 27 articles, 4 of which you can access for free at:  
<http://jech.bmj.com/content/early/2018/01/25/jech-2017-209985#ref-list-1>

## Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

---

## Notes

---

To request permissions go to:

<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:

<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:

<http://group.bmj.com/subscribe/>