Mortality risk of obesity and underweight is overestimated with self-reported body mass index

Faeh, David; Roh, Lucienne; Paccaud, Fred; Chiolero, Arnaud

DOI: https://doi.org/10.1097/EDE.0000000000000009

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-88496
Published Version

Originally published at:
Faeh, David; Roh, Lucienne; Paccaud, Fred; Chiolero, Arnaud (2014). Mortality risk of obesity and underweight is overestimated with self-reported body mass index. Epidemiology, 25(1):156-158.
DOI: https://doi.org/10.1097/EDE.0000000000000009
revision Clinical Modification. We considered the years before and after 2006 as “before” or “after” vaccine introduction. Data on HPV vaccine coverage, estimated by the National Institutes of Health based on the vaccination in the age group 13–17 years, were retrieved from the Centers for Disease Control Web site.11

We included both sexes because, from the beginning of 2011, the Advisory Committee on Immunization Practices has recommended the routine use of the quadrivalent (HPV) vaccine in young men and young women. We analyzed data on the rates of disease and vaccination considering the variation between calendar years in women of all ages and patients of both sexes 17 years of age and younger because these groups were most likely to be exposed to the HPV vaccine.

As shown in Figure A, there was no increase in hospitalizations after the introduction of the HPV vaccine, considering women of all ages. Similarly, there was no increase when we considered only patients 17 years of age and younger from the National Inpatient Sample and the Kids’ Inpatient Sample (Figure B).

We carried out a further analysis of data from the Nationwide Emergency Department Sample for people admitted to the emergency department with systemic lupus erythematosus as their first diagnosis. There was an increase in the number of emergency department admissions for women (Figure A). However, this appeared to apply only to older patients because it was not reproduced among those younger than 17 years (Figure B).

We could not obtain a reliable estimate of lupus hospitalization from the National Hospital Discharge Survey database because of the small number of patients who possibly had received the HPV vaccine. We found no evidence of an increase in the number of hospitalizations or emergency department admissions because of lupus in patient groups exposed to HPV vaccine. This is consistent with results from studies of patients receiving the HPV vaccine4 and with recently published guidelines from the European League Against Rheumatism.3

ACKNOWLEDGMENT
The financial support by Regione Lombardia (Monitoraggio degli Eventi Avversi in Pediatria project) is gratefully acknowledged.

Paolo Pellegrino*
Carla Carnovale*
Valentia Perrone
Dionigi Salvati
Marta Gentili
Stefania Antoniazzi
Emilio Clementi
Sonia Radice

REFERENCES

O bese persons (those with a body mass index [BMI] ≥30 kg/m2) tend to underestimate their weight, leading to an underestimation of their true (measured) BMI and obesity prevalence.1,2 In contrast, underweight people (BMI <18.5 kg/m2) tend to report themselves heavier, resulting in a higher BMI compared with measured BMI and an underestimation of underweight prevalence.1

Less is known about biases in the estimate of mortality risk associated with these body weight categories using...
self-reported data. It has been shown that the mortality risk of obesity based on self-reported BMI can be overestimated, ie, biased away from the null. Underweight is associated with an increased mortality risk and, because underweight persons tend to overestimate BMI, one might intuitively expect that the mortality risk of underweight based on self-reported BMI would be underestimated, ie, biased toward the null (the opposite of the effect in obese persons). Is that a misleading intuition?

Consider a cohort of 1000 men, of whom 10% were underweight at baseline, 70% had normal weight or were overweight, and 20% were obese, based on measured BMI. After 5 years, 36 men died (Table). The relative risk (RR) of death was higher in underweight (RR = 2.0) and obese men (RR = 1.5) compared with that in normal weight/overweight based on self-reported data. The RR of mortality was higher in underweight men (RR = 2.0) and obese men (RR = 2.0) compared with that in normal weight and overweight men combined. We assume that the mortality risk was higher in the lowest (<17.5 kg/m²) category among underweight persons and in the highest BMI category (≥31 kg/m²) among obese persons.

Participants also reported their weight and height at baseline to compute self-reported BMI. We assume that underweight men (measured BMI <18.5 kg/m²) overestimated their BMI by 1 unit, on average. Consequently, men with measured BMI 17.5–18.4 kg/m² were classified as “normal weight” based on self-reported BMI, while in truth they were underweight (eAppendix, eFigure, http://links.lww.com/EDE/A735). Thus, using self-reported data, the prevalence of underweight was underestimated. Men with self-reported BMI <18.5 kg/m² had a measured BMI <17.5 kg/m². Therefore, these men had a higher mortality compared with men with true BMI <18.5 kg/m² (8% vs. 6%).

We assume that obese men underestimated their BMI by 1 unit, on average. Therefore, men with true BMI 30.0–31.0 kg/m² were classified as normal weight/overweight based on self-reported BMI although they were in truth obese. Using self-reported data, the prevalence of obesity was underestimated. Men with self-reported BMI ≥30.0 kg/m² had a measured BMI ≥31.0 kg/m². Accordingly, they had a higher mortality compared with men with true BMI ≥30.0 kg/m² (6% vs. 4.5%).

Using self-reported BMI to define body weight categories, the risk of underweight relative to normal/overweight was (4/50)/[(2 + 21 + 3)/(50 + 700 + 100)] = 2.6, which was higher than the RR (2.0) obtained with measured BMI. The RR of obesity was (6/100)/[(2 + 21 + 3)/(50 + 700 + 100)] = 2.0, which was also higher than the RR (1.5) obtained with measured BMI.

Although self-reports lead to an overestimation of BMI by underweight persons and an underestimation by obese persons, the mortality (or disease) risk in both obese and underweight persons is likely overestimated when self-reported data are used to categorize people.

Our assumptions are simplistic. The errors in weight estimations depend on sex, age, and other characteristics, and the direction of the bias will depend on how people in the various categories estimate their weight. The possible effects of exposure misclassification on the estimation of RR are complex and barely generalizable. Still, as shown, the effects of such misclassification can be counterintuitive. Direction and magnitude of bias should be evaluated carefully for each situation. The overestimation of obesity risk based on self-reports has been demonstrated and recently corroborated in a meta-analysis. Research is needed to confirm the bias entailed by defining underweight with self-reported data.

David Faeh
Lucienne Roh
Institute of Social and Preventive Medicine (ISPM)
University of Zurich
Zurich, Switzerland
david.faeh@uzh.ch

Fred Paccaud
Arnaud Chiolero
Institute of Social and Preventive Medicine (ISPM)
University Hospital Center of Lausanne
Lausanne, Switzerland

REFERENCES
3. Chiolero A, Peytremann-Bridevaux I, Paccaud F. Associations between obesity and health conditions may be overestimated if self-reported body mass index is used (commentary). Obes Rev. 2007;8:307–326.
5. Flegal KM, Graubard BI, Williamson DF, Gail MH. Excess deaths associated with un-

Number Allowed to Diagnose

To the Editor:

In diagnostic research, diagnostic procedures must be able to discriminate between diseased and disease-free patients. Such discrimination is usually expressed as a combination of sensitivity (se) and specificity (sp). Having two criteria makes it more cumbersome to compare diagnostic modalities than therapeutic regimes, which can be summarized by a single endpoint such as overall survival. Analogous to the number needed to treat (NNT) in treatment trials,1 the number needed to diagnose2 has been proposed as a single summary statistic for diagnostic tests. However, as pointed out by Habibzadeh,3 the number needed to diagnose lacks clinical utility. Moreover, it seeks to construct an analogy to treatment trials—an analogy that, in our view, just does not exist.

Turning to a variant that is both clinically interpretable and useful, Habibzadeh4 has introduced the number needed to misdiagnose (NNM), which is a measure of sensitivity and the number needed to diagnose lacks clinical utility. Moreover, it seeks to construct an analogy to treatment trials—an analogy that, in our view, just does not exist.

To the Editor:

In diagnostic research, diagnostic procedures must be able to discriminate between diseased and disease-free patients. Such discrimination is usually expressed as a combination of sensitivity (se) and specificity (sp). Having two criteria makes it more cumbersome to compare diagnostic modalities than therapeutic regimes, which can be summarized by a single endpoint such as overall survival. Analogous to the number needed to treat (NNT) in treatment trials,1 the number needed to diagnose2 has been proposed as a single summary statistic for diagnostic tests. However, as pointed out by Habibzadeh,3 the number needed to diagnose lacks clinical utility. Moreover, it seeks to construct an analogy to treatment trials—an analogy that, in our view, just does not exist.

Turning to a variant that is both clinically interpretable and useful, Habibzadeh4 has introduced the number needed to misdiagnose (NNM), which is a measure of sensitivity and the number needed to diagnose lacks clinical utility. Moreover, it seeks to construct an analogy to treatment trials—an analogy that, in our view, just does not exist.

To the Editor:

In diagnostic research, diagnostic procedures must be able to discriminate between diseased and disease-free patients. Such discrimination is usually expressed as a combination of sensitivity (se) and specificity (sp). Having two criteria makes it more cumbersome to compare diagnostic modalities than therapeutic regimes, which can be summarized by a single endpoint such as overall survival. Analogous to the number needed to treat (NNT) in treatment trials,1 the number needed to diagnose2 has been proposed as a single summary statistic for diagnostic tests. However, as pointed out by Habibzadeh,3 the number needed to diagnose lacks clinical utility. Moreover, it seeks to construct an analogy to treatment trials—an analogy that, in our view, just does not exist.

Turning to a variant that is both clinically interpretable and useful, Habibzadeh4 has introduced the number needed to misdiagnose (NNM), which is a measure of sensitivity and the number needed to diagnose lacks clinical utility. Moreover, it seeks to construct an analogy to treatment trials—an analogy that, in our view, just does not exist.

To the Editor:

In diagnostic research, diagnostic procedures must be able to discriminate between diseased and disease-free patients. Such discrimination is usually expressed as a combination of sensitivity (se) and specificity (sp). Having two criteria makes it more cumbersome to compare diagnostic modalities than therapeutic regimes, which can be summarized by a single endpoint such as overall survival. Analogous to the number needed to treat (NNT) in treatment trials,1 the number needed to diagnose2 has been proposed as a single summary statistic for diagnostic tests. However, as pointed out by Habibzadeh,3 the number needed to diagnose lacks clinical utility. Moreover, it seeks to construct an analogy to treatment trials—an analogy that, in our view, just does not exist.

Turning to a variant that is both clinically interpretable and useful, Habibzadeh4 has introduced the number needed to misdiagnose (NNM), which is a measure of sensitivity and the number needed to diagnose lacks clinical utility. Moreover, it seeks to construct an analogy to treatment trials—an analogy that, in our view, just does not exist.

In summary, we recommend using the term “number allowed to diagnose” instead of “number needed to misdiagnose.” Furthermore, we suggest including weights for sensitivity and specificity in the computations if there

between this and the NNT: for the NNT “good” values are small values because only a small number of patients needs to be treated for one life to be saved or one failure to be avoided. In contrast, with the NAD, large values are better—we can expect a large number of patients to benefit from a correct diagnosis before a misdiagnosis occurs. With a prevalence of p, the NAD (earlier NNM) is given by

\[ NAD = \frac{1}{Pr(\text{misclassification})} = \frac{1}{(1-se)p + (1-sp)(1-p)}. \]

In the case of a population-based accuracy study that enables the estimation of both prevalence and accuracy, we can rewrite the number allowed to diagnose as

\[ NAD = \frac{1}{TP+FN+FP+TN} = \frac{1}{1 - \frac{TP+TN}{TP+FN+FP+TN}} = \frac{1}{1 - \text{Accuracy}}. \]

with TP, FN, FP, and TN indicating the true positive, false negative, false positive, and true negative test results, respectively. Habibzadeh has pointed out a limitation of the number needed to misdiagnose, in that it treats the false-positive and false-negative test results equally despite their quite different consequences for the patient. In many applications, however, this assumption of equal importance is not reasonable. An alternative weighting can be introduced by assigning the costs \( c_1 \) and \( c_0 \) to the false-positive and false-negative test results, respectively. Then the number allowed to diagnose can be generalized to the number of subjects needed to be diagnosed before we can expect the overall misclassification cost to equal the cost of misclassifying one randomly selected patient. The latter is equal to

\[ c_1 \times p + c_0 \times (1-p), \]

and the expected misclassification cost in \( N \) subjects is equal to

\[ N \left[ (c_1 \times (1-se)p) + (c_0 \times (1-sp)(1-p)) \right]. \]

When requiring these numbers to be equal or, alternatively, requiring their ratio to equal one, multiplication by \( N \) gives a cost-weighted version of the NAD:

\[ NAD_{\text{cost}} = \frac{c_1 \times p + c_0 \times (1-p)}{c_1 \times (1-se)p + c_0 \times (1-sp)(1-p)}. \]

It should be noted that the NAD-cost depends only on the cost through the ratio \( c_1/c_0 \) (which can be easily seen by dividing both the numerator and the denominator by \( c_0 \)). Even when not thinking in terms of cost, but requiring sensitivity to be \( x \) times more important than specificity, we can choose \( c_1 = x \) and \( c_0 = 1 \). It is also possible to choose \( c_1 \) and \( c_0 \) so that they sum to \( 1 \).5 Values for the NAD-cost vary considerably when cost ratios \( c_1/c_0 \) other than 1 are investigated (Table). Finally, assuming equal cost \( (c = c_1 = c_0) \), the NAD-cost simplifies to the NAD.

In summary, we recommend using the term “number allowed to diagnose” instead of “number needed to misdiagnose.” Furthermore, we suggest including weights for sensitivity and specificity in the computations if there

### Table

<table>
<thead>
<tr>
<th>Cost Ratio ( c_1/c_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
</tr>
<tr>
<td>0.05</td>
</tr>
<tr>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
</tr>
<tr>
<td>0.5</td>
</tr>
<tr>
<td>0.8</td>
</tr>
<tr>
<td>0.9</td>
</tr>
<tr>
<td>0.95</td>
</tr>
<tr>
<td>0.99</td>
</tr>
</tbody>
</table>

*Value corresponding to the value in Habibzadeh’s example.*

The authors have no conflicts of interest to declare.