# Analysis of Onyike ea (2003)

*Jan van Rongen*

*2018-06-28*

## 1 Summary and conclusions.

We replicate (or try to replicate ) Onyike e.a. 2003 [1]. Quoting from the abstract:

*Obesity was associated with past-month depression in women (odds ratio (OR) = 1.82, 95% confidence interval (CI): 1.01, 3.3) but was not significantly associated in men (OR = 1.73, 95% CI: 0.56, 5.37). When obesity was stratified by severity, heterogeneity in the association with depression was observed. Class 3 (severe) obesity (body mass index > = 40) was associated with past-month depression in unadjusted analyses (OR = 4.98, 95% CI: 2.07, 11.99); the association remained strong after results were controlled for age, education, marital status, physician's health rating, dieting for medical reasons, use of psychiatric medicines, cigarette smoking, and use of alcohol, marijuana, and cocaine.*

I have been able to replicate most of these results using the `survey` package in R [2] by T. Lumley. See also the book by the same author[3].

Even though the numbers can be reconstructed, I do not like the way those conclusions are drawn. The only BMI-category that is really associated with (past month) depression is the over 40 class. It is a large effect but the precision is low (due to the small subsamples). Other attributes that are definetely associated with depression are gebder:Female and maritial status: Separated. And with the use of psychiatric medicin (not shown).

Table 1: Depressed odds ratio's (starred when 1.0 outside CI)

| | OR | 2.5 % | 97.5 % | |
|---|---|---|---|---|
| underweight | 1.02 | 0.43 | 2.44 | |
| overweight | 0.98 | 0.6 | 1.6 | |
| obese 30-35 | 1.28 | 0.65 | 2.52 | |
| obese 35-40 | 1.6 | 0.72 | 3.54 | |
| obese 40+ | 4.42 | 1.73 | 11.29 | * |
| Female | 2.33 | 1.6 | 3.4 | * |
| Never Married | 1.19 | 0.75 | 1.9 | |
| Separated etc | 2.16 | 1.23 | 3.79 | * |

The above table summarize i.m.o. the essence of Onyike et al. article ([1]). These are the estimated odds ratio's for the whole population of the USA (around 1990).

When categories are "expanded", f.i. all obese classes taken together, we get "watered down" results that might still be significant, but do not contain any new information or insight.

## 2 Introduction and methods

We replicate most of the results in ([1]). These results are based on data in the NHANES III survey (1988-1994) ([4]).

Surveys use complex designs in multiple stages for the sampling process. The observations in the sample are not (all) independent. Each observation is weighted in such a way that the sum of the weights is the size of the total population. The wighting process uses known population data on sex, race, age and education.

Sample weights are not the same as precision weights that can be used in most statistical software for linear models. Parameter estimates will be the same, but estimates for errors will be very different. Thus special software is required.

## 2.1 Data, structure and characteristics

The survey design parameters are a pseudo PSU with 2 values and a pseudo stratum with 49 values. For this design 52 replication weights are also present in the data. See the first chapters of [3].

These "design" attributes are:

- "SEQN" is the unique Id,
- "SDPPSU6" is the pseudoPSU,
- "SDPSTRA6" is the pseudo stratum,
- "WTPXRP" (1-52) are the (Fay) replication weights,
- "WTPFEX6" (renamed to weigth) is the sample weight

## 2.2 Standard Error estimates.

In this type of study there are number of possible techniques to estimate variance. On the one hand there is the direct estimate from the original weights, the so called linear method.

On the other hand using various methods so called replication weights can be constructed and the SE is estimated from them.

The latter method has (at least) four useful variants: jacknife, bootstrap, BRR and a modified BRR (Fay).

I will compare all 5. The original article seems to use the Jacknife, even though the data contains replication weights for the modified BRR. It is not explained why this was chosen.

Below we compare the different methods for the statement that *Obesity was associated with past-month depression in women (odds ratio (OR) = 1.82, 95% confidence interval (CI): 1.01, 3.3)*

Table 2: Various CI estimating methods: Obese vs depressed women

| method | OR | 2.5% | 97.5% |
|---|---|---|---|
| linear | 1.82 | 1.021 | 3.25 |
| bootstrap | 1.82 | 0.944 | 3.51 |
| JKn | 1.82 | 1.019 | 3.25 |
| BRR | 1.82 | 0.986 | 3.36 |
| Fay | 1.82 | 1.004 | 3.30 |
| Fay from data | 1.82 | 1.113 | 2.98 |

Clearly the significance is not so clear at all.

# 3 Replications

There are a number of points to consider when using inference on survey data. The first point is whether there are enough observations. Using sample weights lowers the precision, so a simple power calculation is not enough in itself.

The second point is which method to use to calculate the SE's and hence the Confidence Intervals. There are many methods and all have their advantages and disadvantages. In a previous paragraph we already compared those methods.

### 3.1 Data and conversion to a survey design

The NHANES III data was downloaded from the site [4]. The data has been converted using SAS but also as a check there is also a direct conversion in R. Some of the translation of atributes from NHANES III to those in the article is not always obvious.

Anyway, after data conversion we build the designs as given by Lumley and others [3].

```
##
## Data for table 2: 8773 rows, 66 cols
```

```
##
## Data for others : 8410 rows, 69 cols
```

### 3.2 Table 1: only a power test is not enough.

Table 1 contains sizes of classes and from there a beta which is probably meant to be the power of the tests. However in a sample weighted survey the power on the unweighted data is not a very useful measure. In surveys we need more observations for the same precision. See [3] page 6.

### 3.3 Table 2: population distribution does not provide much extra information

This table, for some reason, uses a slightly larger dataset than the rest of the analysis. The given SE are percentages, and can be replicated using the jAcknife method.

Two small examples. In the first we also give the design effect - the decrease in precision caused by the non random sampling:

| | mean | SE | deff |
|---|---|---|---|
| sexF | 0.506 | 0.006 | 1.39 |
| sexM | 0.494 | 0.006 | 1.39 |

| | F | M |
|---|---|---|
| age_class15-19 | 17.26 | 17.567 |
| age_class20-24 | 20.07 | 19.117 |
| age_class25-29 | 19.55 | 20.665 |
| age_class30-34 | 21.50 | 22.780 |
| age_class35-39 | 21.62 | 19.871 |
| se1 | 1.23 | 0.927 |
| se2 | 1.24 | 0.989 |
| se3 | 1.14 | 1.028 |
| se4 | 1.40 | 1.225 |
| se5 | 1.15 | 1.143 |

### 3.4 Table 3: distribution of BMI_classes versus depressed is missing SE and dEFF

This table firstr gives the number of observations but then uses weights for the percentages. It is an omission not to look at the SE's and the DEFF, because they indicate whether the following inferences will make sense. A strange omission, as table 2 does have the SE's.

The `910` in the table is a typo, should be *981*.

An example:

| BMI | Sex | % depressed | SE | dEFF |
|---|---|---|---|---|
| normal | M | 1.670 | 0.369 | 1.64 |
| underweight | M | 1.824 | 1.385 | 1.05 |
| overweight | M | 1.374 | 0.472 | 1.97 |
| obese 30-35 | M | 1.884 | 0.844 | 1.48 |
| obese 35-40 | M | 0.825 | 0.831 | 1.05 |
| obese 40+ | M | 11.540 | 10.150 | 6.46 |
| normal | F | 3.816 | 0.724 | 3.11 |
| underweight | F | 3.815 | 1.628 | 1.45 |
| overweight | F | 4.011 | 0.595 | 1.01 |
| obese 30-35 | F | 4.967 | 1.614 | 3.29 |
| obese 35-40 | F | 6.794 | 2.625 | 3.09 |
| obese 40+ | F | 13.032 | 4.237 | 3.18 |

| BMI | Sex | % depressed | SE | dEFF |
|---|---|---|---|---|
| normal | M | 1.67 | 0.369 | 1.64 |
| underweight | M | 1.82 | 1.385 | 1.05 |
| overweight | M | 1.37 | 0.472 | 1.97 |
| obese | M | 2.85 | 1.431 | 4.23 |
| normal | F | 3.82 | 0.724 | 3.11 |
| underweight | F | 3.81 | 1.628 | 1.45 |
| overweight | F | 4.01 | 0.595 | 1.01 |
| obese | F | 6.74 | 1.407 | 3.41 |

| BMI | % depressed | SE | dEFF |
|---|---|---|---|
| normal | 2.79 | 0.443 | 3.00 |
| underweight | 3.24 | 1.227 | 1.44 |
| overweight | 2.42 | 0.426 | 1.77 |
| obese | 5.13 | 1.234 | 5.19 |

The standard errors from the above tables (not given in O_2003) indicate that the precision of these means are sometimes low and that most p-values given for this table in [1] must be too small.

```
##
## P-value (Chi_sq) depressed vs obese: 0.0163
```

## 3.5 Table 4: we find very different SE's for smaller cases

The survey package has at least seven ways to caluclate the SE's of odds ratio's, but I have not found *one* method that replicates *all* results in a consistent way. As mentioned above, the following results are similar to the JKn method. However there are considerable difference for subsets.

Table 8: All: odds ratios with JKn

| | NR | OR | 2.5 % | 97.5 % |
|---|---|---|---|---|
| underweight | 301 | 1.167 | 0.492 | 2.77 |
| overweight | 2297 | 0.863 | 0.534 | 1.40 |
| obese | 1658 | 1.883 | 1.035 | 3.43 |
| obese 30-35 | 981 | 1.283 | 0.650 | 2.53 |
| obese 35-40 | 410 | 1.756 | 0.777 | 3.97 |
| obese 40+ | 267 | 4.980 | 2.070 | 11.98 |

Note: only very small differences.

Table 9: Females odds ratio's

| | NR | OR | 2.5 % | 97.5 % |
|---|---|---|---|---|
| underweight | 202 | 1.00 | 0.382 | 2.62 |
| overweight | 1095 | 1.05 | 0.645 | 1.72 |
| obese | 1084 | 1.82 | 1.019 | 3.25 |
| obese 30-35 | 597 | 1.32 | 0.606 | 2.86 |
| obese 35-40 | 285 | 1.84 | 0.711 | 4.75 |
| obese 40+ | 202 | 3.78 | 1.668 | 8.55 |

Notes: Two classes have 1.28 and 1.75 for odds rate that should be as above 1.32 and 1.84. CI's are the same, so this is a copying error.

Table 10: Males odds ratio's

| | NR | OR | 2.5 % | 97.5 % |
|---|---|---|---|---|
| underweight | 99 | 1.09 | 0.129 | 9.28e+00 |
| overweight | 1202 | 0.82 | 0.347 | 1.94e+00 |
| obese | 574 | 1.73 | 0.524 | 5.71e+00 |
| obese 30-35 | 384 | 1.13 | 0.403 | 3.17e+00 |
| obese 35-40 | 125 | 0.49 | 0.000 | 3.14e+07 |
| obese 40+ | 65 | 7.68 | 0.121 | 4.86e+02 |

Notes: big differences in CI for last two classes.

## 3.6 Table 5

I have difficulty finding all the extra attributes. Cannot find the medical diet f.i. So decided to use a part of the information.

Lumley [3] discusses in chapter 5.3 the pro's and con's of using weights in this regression. So we compare both and note that the conclusions are the same.

Also note that the "reference person" - a 15-19 year old male, married and poorly educated, with normal BMI, does not occur in the data.

Table 11: glm using weights and design

| | OR | 2.5 % | 97.5 % | |
|---|---|---|---|---|
| bmi_class_2underweight | 1.01 | 0.41 | 2.46 | |
| bmi_class_2overweight | 0.97 | 0.59 | 1.6 | |
| bmi_class_2obese 30-35 | 1.29 | 0.66 | 2.52 | |
| bmi_class_2obese 35-40 | 1.69 | 0.75 | 3.77 | |
| bmi_class_2obese 40+ | 4.51 | 1.74 | 11.71 | * |
| sexF | 2.37 | 1.63 | 3.45 | * |
| age_class20-24 | 1.34 | 0.68 | 2.65 | |
| age_class25-29 | 1.11 | 0.46 | 2.67 | |
| age_class30-34 | 1.14 | 0.52 | 2.47 | |
| age_class35-39 | 1.42 | 0.67 | 2.99 | |
| raceBlack | 0.71 | 0.45 | 1.1 | |
| raceHispanic | 0.82 | 0.44 | 1.5 | |
| raceOther | 0.2 | 0 | 13081757.64 | |
| edu_class9-11 | 0.81 | 0.43 | 1.5 | |
| edu_class12 | 0.51 | 0.26 | 0.98 | |
| edu_class12+ | 0.53 | 0.25 | 1.1 | |
| marit_classNever Married | 1.24 | 0.71 | 2.15 | |
| marit_classSeparated etc | 2.13 | 1.23 | 3.71 | * |

Table 12: normal glm without weights

| | OR | 2.5 % | 97.5 % | |
|---|---|---|---|---|
| bmi_class_2underweight | 1.24 | 0.64 | 2.21 | |
| bmi_class_2overweight | 1.23 | 0.91 | 1.67 | |
| bmi_class_2obese 30-35 | 1.02 | 0.66 | 1.53 | |
| bmi_class_2obese 35-40 | 1.22 | 0.67 | 2.07 | |
| bmi_class_2obese 40+ | 2.28 | 1.32 | 3.75 | * |
| sexF | 2.45 | 1.85 | 3.29 | * |
| age_class20-24 | 1.2 | 0.79 | 1.81 | |
| age_class25-29 | 0.91 | 0.56 | 1.45 | |
| age_class30-34 | 1.01 | 0.62 | 1.65 | |
| age_class35-39 | 1.09 | 0.67 | 1.79 | |
| raceBlack | 0.64 | 0.46 | 0.88 | |
| raceHispanic | 0.86 | 0.63 | 1.18 | |
| raceOther | 0.22 | 0.01 | 1.03 | |
| edu_class9-11 | 1.33 | 0.89 | 2.04 | |
| edu_class12 | 0.97 | 0.64 | 1.5 | |
| edu_class12+ | 0.87 | 0.56 | 1.37 | |
| marit_classNever Married | 1.18 | 0.85 | 1.65 | |
| marit_classSeparated etc | 2.24 | 1.54 | 3.22 | * |

# 4 Literature

[1] Onyike e.a. *Is obesity associated with major depression?* (Am J Epidemiol. 2003)
[2] http://r-survey.r-forge.r-project.org/survey/
[3] Complex Surveys; A Guide to Analysis Using R; Thomas Lumley. ISBN 978-0-470-28430-8, Wiley, 2009.
[4] https://wwwn.cdc.gov/nchs/nhanes/nhanes3/Default.aspx