Homework 2

A comparison of different survey methods analysis studying the association between subjective well-being and body mass index: an experience from the 1997 Belgian health interview survey

Taught by

Pr. Dr. Geert Molenberghs

By
Stijn Van Herck
Thomas Bullen
Tom Louwies
Soutrik Banerjee

Master of Science in Biostatistics: 1^{st} year
2010 – 2011

Table of Contents

List of Tables

1. Introduction

The Health Interview Survey (HIS) was conducted for the first time in Belgium in 1997. The primary goal of this survey was to know the health status of the residents in Belgium, including subjective perception and objective measures of health. This information was essential for the governing bodies to bring about a change in the existing health policies as well as to prioritise the domains in the field of health, which needed a thrust.

The detailed description of the HIS survey methodology is mentioned elsewhere *[van Oyen et al., 1997]*. In brief, the survey was conducted by the Department of Epidemiology of the Institute of Hygiene and the Institute of Statistics by splitting Belgium into three regions (*i.e.*, Flanders, Wallonia and Brussels), which in turn was divided into 12 provinces (or strata); the Liège province was again divided into two parts: one for the French speaking and the other for the German speaking communities. An overall sample size of 10,000 interviewees was targeted, which was split into 3,500 (Flanders), 3,500 (Wallonia) and 3,000 (Brussels) among the regions. In a multistage design, a systematic sampling by ordering the communities (Primary Sampling Units) from large to small enabled to select a sample of communities within each province (nested within a region). At the second stage, from the households (Secondary Sampling Units) list obtained from the national register, a total of 50 interviews (and not households) were chosen per community by clustered age-stratified systematic sampling; the number of interviews was based on the results of an earlier pilot survey to assess the feasibility of this survey. In this case, a cluster of four similar households (based on the criteria: statistical sector, size of the household and age of the principal resident, respectively) was selected (in three separate age strata based on the age of the principal resident) as reserves in order to replace a non-participating household and hence to minimise the non-participation bias of a household. The households cluster list was 'scrambled' and then randomised before the fieldwork. In the third stage, within a household, a cluster of maximum four individuals (Tertiary Sampling Units) were interviewed (*i.e.*, typically the principal resident, his or her spouse, and two of the children). Children less than 15 years old were interviewed by proxy.

In this present paper, we analysed the impact of the three factors, namely, having a stable general practitioner (SGP), VOEG: a questionnaire measuring subjective perception of health, and General Health Questionnaire (GHQ) on the body mass index (BMI), keeping into account the multistage design of the survey during the analysis, namely, strata (provinces), cluster (households) and final weights of the subjects, performed by domain (regions) and overall. We have also compared the results obtained from the survey-specific SAS procedure with a non-survey procedure SAS procedure.

2. Methods

Exposure variables

The participants were asked if they had an **SGP** (reference category) or not.

VOEG: originally, a 21- and 13-item Dutch questionnaire that assessed the subjective health perceived by the respondent *[Dirken, 1967]*. A lower score indicated better outcome. The '0' and '1' states were coalesced as the reference category and the rest (higher scores) were classified as the other category, in order that this derived variable is binary *[Gecková et al., 2003]*. This questionnaire was translated into French and German for the respective native speakers.

GHQ: originally, 60 and 30-item questionnaire developed by Goldberg (1972). In this survey, a shortened 12-item questionnaire was used. A lower score indicated better outcome. This version was dichotomised as having '0' or '1' complaint as the reference category and the rest (higher scores) were classified as the other category.

The subjects with missing exposure variables were not included in the study.

Outcome variable

The **BMI** of an individual is equal to weight (kg) divided by height (m^2), originally defined by the Belgian Polymath Adolphe Quetlet between 1830-1850 (also called Quetelet's index). It was trichotomised into 'normal' (reference), 'overweight' and 'obese' categories *[WHO]*:

- Normal: 18.5 to 25
- Overweight: 25+ to 30
- Obese: 30+

We eliminated the subjects having less than normal BMI (*i.e.*, < 18.5), because these potentially pathological subjects might have a relationship with the exposure variable(s) in this study in a different way than the normal BMI subjects.

Other variables

The domain were the 3 regions (Flanders, Wallonia, Brussels) and overall (Belgium).

The strata were the 12 provinces including the Liège province split into 2 communities.

The clusters were the households, the sampling frame for this study.

The variable 'wfin' was the final weight carried by each individual, which is the inverse of that individual's selection probability from the total population of Belgian residents eligible for this study.

Statistical methods

Given the objective of the survey to compare different (within) survey and (between) non-survey methods, a consistent and comparable model including all the three exposure variables, was fitted against the trichotomised outcome variable, instead of fitting one exposure variable at a time.

At a first glance, a more parsimonious model – proportional odds model was chosen. But owing to the poor fit (discussed later) for certain regions, this approach was abandoned and a more general baseline category logit model was fitted *[Agresti, 2002]*.

The model,

$$log\frac{\pi_j(\boldsymbol{x})}{\pi_J(\boldsymbol{x})} = \alpha_j + \boldsymbol{\beta}_j'\boldsymbol{x},\ \ j = 1, 2, \cdots, J-1.$$

The above GLM multinomial model computes the logit of the estimated probability (π_j) for each of the J – 1 categories (of the outcome variable) with respect to the J-th (reference) category (of the outcome variable), assuming that the outcome variable follows a multinomial distribution with 2 degrees of freedom (here), and a log canonical link to relate the mean of the outcome to the linear predictor vector (***x***). We exponentiated the log odds ratios to obtain the adjusted odds ratios for each exposure variable.

The survey-specific procedure SAS PROC SURVEYLOGISTIC was used for this purpose *[SAS, v9.1.3]*. In addition SAS PROC LOGISTIC was also used (discussed later).

3. Results

Table 1. Parameter estimates (95% CIs) of the predictors in the fitted multivariable baseline category logit models.

Analysis	SAS Procedure	Parameter	Belgium (95% CI)	Brussels (95% CI)	Flanders (95% CI)	Wallonia (95% CI)
SRS	Surveylogistic	SGP_{obese}	**0.390 (0.278-0.547)**	**0.547 (0.352-0.851)**	**0.436 (0.200-0.954)**	**0.202 (0.088-0.463)**
		$SGP_{overwt.}$	**0.752 (0.632-0.895)**	1.028 (0.812-1.301)	**0.360 (0.220-0.587)**	0.719 (0.504-1.025)
		$VOEG_{obese}$	**0.649 (0.541-0.779)**	**0.613 (0.412-0.910)**	**0.628 (0.466-0.847)**	**0.746 (0.558-0.996)**
		$VOEG_{overwt.}$	**0.826 (0.734-0.929)**	**0.719 (0.569-0.910)**	0.907 (0.755-1.090)	0.819 (0.663-1.012)
		GHQ_{obese}	1.073 (0.916-1.257)	1.029 (0.757-1.398)	1.122 (0.831-1.514)	1.114 (0.877-1.413)
		$GHQ_{overwt.}$	**1.227 (1.096-1.373)**	1.186 (0.968-1.453)	1.092 (0.895-1.332)	**1.395 (1.157-1.681)**
SRS	Surveylogistic*	SGP_{obese}	**0.390 (0.278-0.547)**	**0.547 (0.352-0.851)**	**0.436 (0.200-0.954)**	**0.202 (0.088-0.463)**
		$SGP_{overwt.}$	**0.752 (0.632-0.895)**	1.028 (0.812-1.301)	**0.360 (0.220-0.587)**	0.719 (0.504-1.025)
		$VOEG_{obese}$	**0.649 (0.541-0.779)**	**0.613 (0.412-0.910)**	**0.628 (0.466-0.847)**	**0.746 (0.558-0.996)**
		$VOEG_{overwt.}$	**0.826 (0.734-0.929)**	**0.719 (0.569-0.910)**	0.907 (0.755-1.090)	0.819 (0.663-1.012)
		GHQ_{obese}	1.073 (0.916-1.257)	1.029 (0.757-1.398)	1.122 (0.831-1.514)	1.114 (0.877-1.413)
		$GHQ_{overwt.}$	**1.227 (1.096-1.373)**	1.186 (0.968-1.453)	1.092 (0.895-1.332)	**1.395 (1.157-1.681)**
STRAT	Surveylogistic	SGP_{obese}	**0.390 (0.278-0.547)**	**0.547 (0.352-0.851)**	**0.436 (0.200-0.953)**	**0.202 (0.088-0.463)**
		$SGP_{overwt.}$	**0.752 (0.632-0.895)**	1.028 (0.812-1.301)	**0.360 (0.220-0.587)**	0.719 (0.504-1.025)
		$VOEG_{obese}$	**0.649 (0.541-0.779)**	**0.613 (0.412-0.910)**	**0.628 (0.466-0.847)**	**0.746 (0.558-0.996)**
		$VOEG_{overwt.}$	**0.826 (0.734-0.929)**	**0.719 (0.569-0.910)**	0.907 (0.755-1.090)	0.819 (0.663-1.012)
		GHQ_{obese}	1.073 (0.916-1.257)	1.029 (0.757-1.398)	1.122 (0.831-1.514)	1.114 (0.877-1.413)
		$GHQ_{overwt.}$	**1.227 (1.096-1.373)**	1.186 (0.968-1.453)	1.092 (0.895-1.332)	**1.395 (1.157-1.681)**
STRAT	Surveylogistic*	SGP_{obese}	**0.390 (0.278-0.547)**	**0.547 (0.352-0.851)**	**0.436 (0.200-0.953)**	**0.202 (0.088-0.463)**
		$SGP_{overwt.}$	**0.752 (0.632-0.895)**	1.028 (0.812-1.301)	**0.360 (0.220-0.587)**	0.719 (0.504-1.025)
		$VOEG_{obese}$	**0.649 (0.541-0.779)**	**0.613 (0.412-0.910)**	**0.628 (0.466-0.847)**	**0.746 (0.558-0.996)**
		$VOEG_{overwt.}$	**0.826 (0.734-0.929)**	**0.719 (0.569-0.910)**	0.907 (0.755-1.090)	0.819 (0.663-1.012)
		GHQ_{obese}	1.073 (0.916-1.257)	1.029 (0.757-1.398)	1.122 (0.831-1.514)	1.114 (0.877-1.413)
		$GHQ_{overwt.}$	**1.227 (1.096-1.373)**	1.186 (0.968-1.453)	1.092 (0.895-1.332)	**1.395 (1.157-1.681)**

Analysis	SAS Procedure	Parameter	Belgium (95% CI)	Brussels (95% CI)	Flanders (95% CI)	Wallonia (95% CI)
Weighted	Surveylogistic	SGP_{obese}	**0.302 (0.190-0.482)**	**0.454 (0.270-0.763)**	**0.291 (0.121-0.697)**	**0.216 (0.069-0.676)**
		$SGP_{overwt.}$	**0.574 (0.443-0.744)**	0.916 (0.678-1.237)	**0.337 (0.188-0.605)**	0.746 (0.465-1.197)
		$VOEG_{obese}$	**0.654 (0.502-0.852)**	0.722 (0.430-1.212)	**0.640 (0.434-0.945)**	0.756 (0.505-1.132)
		$VOEG_{overwt.}$	0.854 (0.721-1.012)	0.821 (0.584-1.154)	0.874 (0.694-1.100)	0.864 (0.642-1.161)
		GHQ_{obese}	0.998 (0.787-1.267)	0.966 (0.628-1.485)	0.923 (0.618-1.381)	1.163 (0.850-1.591)
		$GHQ_{overwt.}$	1.099 (0.933-1.294)	1.275 (0.982-1.656)	0.939 (0.736-1.198)	**1.347 (1.039-1.747)**
Weighted	Surveylogistic*	SGP_{obese}	**0.302 (0.190-0.482)**	**0.454 (0.270-0.763)**	**0.291 (0.121-0.697)**	**0.216 (0.069-0.676)**
		$SGP_{overwt.}$	**0.574 (0.443-0.744)**	0.916 (0.678-1.237)	**0.337 (0.188-0.605)**	0.746 (0.465-1.197)
		$VOEG_{obese}$	**0.654 (0.502-0.852)**	0.722 (0.430-1.212)	**0.640 (0.434-0.945)**	0.756 (0.505-1.132)
		$VOEG_{overwt.}$	0.854 (0.721-1.012)	0.821 (0.584-1.154)	0.874 (0.694-1.100)	0.864 (0.642-1.161)
		GHQ_{obese}	0.998 (0.787-1.267)	0.966 (0.628-1.485)	0.923 (0.618-1.380)	1.163 (0.850-1.591)
		$GHQ_{overwt.}$	1.099 (0.933-1.294)	1.275 (0.982-1.656)	0.939 (0.736-1.198)	**1.347 (1.039-1.747)**
CLUST	Surveylogistic	SGP_{obese}	**0.390 (0.271-0.559)**	**0.547 (0.342-0.876)**	**0.436 (0.180-1.061)**	**0.202 (0.089-0.459)**
		$SGP_{overwt.}$	**0.752 (0.629-0.900)**	1.028 (0.809-1.306)	**0.360 (0.217-0.596)**	0.719 (0.503-1.028)
		$VOEG_{obese}$	**0.649 (0.539-0.781)**	**0.613 (0.408-0.921)**	**0.628 (0.468-0.844)**	**0.746 (0.555-1.001)**
		$VOEG_{overwt.}$	**0.826 (0.733-0.930)**	**0.719 (0.571-0.906)**	0.907 (0.755-1.091)	0.819 (0.658-1.019)
		GHQ_{obese}	1.073 (0.916-1.257)	1.029 (0.763-1.387)	1.122 (0.836-1.506)	1.114 (0.873-1.421)
		$GHQ_{overwt.}$	**1.227 (1.094-1.376)**	1.186 (0.967-1.454)	1.092 (0.888-1.343)	**1.395 (1.156-1.682)**
CLUST	Surveylogistic*	SGP_{obese}	**0.390 (0.271-0.559)**	**0.547 (0.342-0.876)**	**0.436 (0.180-1.061)**	**0.202 (0.089-0.459)**
		$SGP_{overwt.}$	**0.752 (0.629-0.900)**	1.028 (0.809-1.306)	**0.360 (0.217-0.596)**	0.719 (0.503-1.028)
		$VOEG_{obese}$	**0.649 (0.539-0.781)**	**0.613 (0.408-0.921)**	**0.628 (0.468-0.844)**	**0.746 (0.555-1.001)**
		$VOEG_{overwt.}$	**0.826 (0.733-0.930)**	**0.719 (0.571-0.906)**	0.907 (0.755-1.091)	0.819 (0.658-1.019)
		GHQ_{obese}	1.073 (0.916-1.257)	1.029 (0.763-1.387)	1.122 (0.836-1.506)	1.114 (0.873-1.421)
		$GHQ_{overwt.}$	**1.227 (1.094-1.376)**	1.186 (0.967-1.454)	1.092 (0.888-1.343)	**1.395 (1.156-1.682)**
ALL	Surveylogistic	SGP_{obese}	**0.302 (0.187-0.488)**	**0.454 (0.264-0.780)**	**0.291 (0.117-0.721)**	**0.216 (0.069-0.676)**
		$SGP_{overwt.}$	**0.574 (0.438-0.752)**	0.916 (0.679-1.236)	**0.337 (0.185-0.614)**	0.746 (0.455-1.122)
		$VOEG_{obese}$	**0.654 (0.505-0.848)**	0.722 (0.428-1.216)	**0.640 (0.440-0.931)**	0.756 (0.506-1.130)
		$VOEG_{overwt.}$	0.854 (0.720-1.013)	0.821 (0.587-1.148)	0.874 (0.692-1.103)	0.864 (0.639-1.167)
		GHQ_{obese}	0.998 (0.790-1.262)	0.966 (0.650-1.436)	0.923 (0.625-1.364)	1.163 (0.845-1.601)

Analysis	SAS Procedure	Parameter	Belgium (95% CI)	Brussels (95% CI)	Flanders (95% CI)	Wallonia (95% CI)
		$GHQ_{overwt.}$	1.099 (0.929-1.299)	1.275 (0.977-1.665)	0.939 (0.728-1.211)	**1.347 (1.048-1.733)**
ALL	Surveylogistic*	SGP_{obese}	**0.302 (0.187-0.488)**	**0.454 (0.264-0.780)**	**0.291 (0.117-0.721)**	**0.216 (0.069-0.676)**
		$SGP_{overwt.}$	**0.574 (0.438-0.752)**	0.916 (0.679-1.236)	**0.337 (0.185-0.614)**	0.746 (0.455-1.122)
		$VOEG_{obese}$	**0.654 (0.505-0.848)**	0.722 (0.428-1.216)	**0.640 (0.440-0.931)**	0.756 (0.506-1.130)
		$VOEG_{overwt.}$	0.854 (0.720-1.013)	0.821 (0.587-1.148)	0.874 (0.692-1.103)	0.864 (0.639-1.167)
		GHQ_{obese}	0.998 (0.790-1.262)	0.966 (0.650-1.436)	0.923 (0.625-1.364)	1.163 (0.845-1.601)
		$GHQ_{overwt.}$	1.099 (0.929-1.299)	1.275 (0.977-1.665)	0.939 (0.728-1.211)	**1.347 (1.048-1.733)**
SRS	Logistic	SGP_{obese}	**0.390 (0.278-0.547)**	**0.547 (0.353-0.849)**	**0.436 (0.200-0.952)**	**0.202 (0.088-0.462)**
		$SGP_{overwt.}$	**0.752 (0.632-0.895)**	1.028 (0.812-1.301)	**0.360 (0.221-0.586)**	0.719 (0.504-1.026)
		$VOEG_{obese}$	**0.649 (0.540-0.780)**	**0.613 (0.412-0.910)**	**0.628 (0.464-0.850)**	**0.746 (0.556-0.999)**
		$VOEG_{overwt.}$	**0.826 (0.734-0.929)**	**0.719 (0.567-0.912)**	0.907 (0.756-1.089)	0.819 (0.664-1.010)
		GHQ_{obese}	1.073 (0.914-1.259)	1.029 (0.756-1.399)	1.122 (0.828-1.520)	1.114 (0.875-1.417)
		$GHQ_{overwt.}$	**1.227 (1.096-1.373)**	1.186 (0.966-1.455)	1.092 (0.896-1.330)	**1.395 (1.158-1.679)**
Weighted	Logistic	SGP_{obese}	**0.302 (0.298-0.307)**	**0.454 (0.442-0.466)**	**0.291 (0.283-0.298)**	**0.216 (0.209-0.223)**
		$SGP_{overwt.}$	**0.574 (0.570-0.579)**	**0.916 (0.903-0.929)**	**0.337 (0.332-0.342)**	**0.746 (0.736-0.756)**
		$VOEG_{obese}$	**0.654 (0.650-0.658)**	**0.722 (0.706-0.738)**	**0.640 (0.635-0.646)**	**0.756 (0.748-0.764)**
		$VOEG_{overwt.}$	**0.854 (0.851-0.857)**	**0.821 (0.809-0.832)**	**0.874 (0.870-0.878)**	**0.864 (0.857-0.870)**
		GHQ_{obese}	0.998 (0.993-1.004)	**0.966 (0.949-0.984)**	**0.923 (0.916-0.931)**	**1.163 (1.152-1.173)**
		$GHQ_{overwt.}$	**1.099 (1.094-1.103)**	**1.275 (1.260-1.291)**	**0.939 (0.934-0.944)**	**1.347 (1.338-1.357)**

n = 7,703; * Population level correction (Belgian population is considered to be equal to 10,000,000).

GHQ score: general health questionnaire (binary); VOEG score: Dutch subjective health score (binary); SGP: stable general practitioner (binary); CI: confidence interval; SRS: simple random sampling; STRAT: stratified sampling (provinces are the strata); CLUST: cluster sampling (households are the clusters).

Body Mass Index: 'normal' category is the reference category; SGP: 'yes' category is the reference category; VOEG score: '0 or 1' coalesced categories from the full VOEG scale is the reference category; GHQ score: '0 or 1' coalesced categories from the GHQ-12 scale is the reference category.

<u>Summary</u>:
A total of 7,703 subjects were selected in this analysis.

In general in the baseline category logit models with all the 3 variables included in the models, the likelihood of having an SGP increased with increased BMI. Again, the likelihood of a very good subjective perception of health (VOEG score) decreased with increased BMI. The impact of the GHQ score was equivocal without a clear pattern.

<u>Comparison between the different survey procedures</u>:
In Table 1, we have compared the different methods, namely, stratified analysis, weighted analysis, cluster analysis and all of these combined together, with the **method of reference** (*i.e.*, **SRS**). Survey-specific SAS procedure (*i.e.*, PROC SURVEYLOGISTIC) was used for this purpose. We also compared the SRS outputs with the 'SRS' outputs of the SAS PROC LOGISTIC procedure; the means were same and 95% CIs were very close (sometimes a slight increase, on other times a slight decrease was noted) (*cf.* also subsection below for the weighted analysis comparison).
First, we observed that in this data, the ***stratified analysis*** results did <u>not differ</u> (except in one case for the 3rd decimal place of the 95% CI) from those of the reference method.
Second, the ***weighted analysis*** differed considerably from the SRS method, for both means (= increase or decrease) and 95% CIs. The general trend was a <u>decrease in precision</u> (except on three occasions, a reverse trend was observed).
Third, the ***clustered analysis*** did not alter the mean estimates, but often <u>decreased the precision</u> compared to the SRS analysis (although on some occasions, a reverse trend was observed). However, this decrease in precision was usually <u>less pronounced</u> than that observed in the weighted analysis when compared to the reference method.
Fourth, in the ***combined analysis*** (*i.e.*, stratified, weighted, cluster), the means were the same as those of the weighted analysis. The precision sometimes increased a little and on other occasions decreased a little compared to that of the weighted analysis. But again (like for the weighted analysis), compared this time to the SRS analysis (the reference method), the general trend was that the <u>precision decreased</u> in the combined analysis (except on two occasions, a reverse trend was observed).
Fifth, the ***correction for the Belgian population*** did <u>not have any impact</u> (except in one case for the 3rd decimal place of the 95% CI, a very minute increase in precision) from its corresponding uncorrected method.

<u>Comparison with the non-survey procedures</u>:
The weighted logistic regression gave very narrow standard errors. However, in a survey context, the standard errors computed by a non-survey procedure like PROC LOGISTIC is not appropriate in a weighted analysis *[Cassell]*. For this reason, we do not pay too

much importance to the very precise 95% CIs obtained with the weighted logistic regression, although the mean estimates were not different from a survey method procedure.

4. Discussion

The 1997 Belgian health interview survey data analysis provides evidence that with increasing body mass index (BMI), the probability of having a stable general practitioner increases in the Belgian population, given that for normally weighted individuals, the proportion of having a stable general practitioner is already quite high (> 85%). The second important conclusion was that the subjective perception of well-being (yes / no) as assessed by the Dutch VOEG questionnaire decreases with increasing BMI. This decrease in feeling healthy is roughly two and half fold from the normally weighted to obese individuals. Finally, the association between general health-related complaints (yes / no) as assessed by the General Health Questionnaire (GHQ) and BMI does not have a clear pattern in this study. One possible explanation for the latter finding might be that conditional on the VOEG score and having a stable general practitioner, the GHQ questionnaire does not have a clear trend with the BMI. Another explanation could be that the GHQ itself does not lend to any further information on subjective health related to BMI that the VOEG score is able to extract.

The survey-specific procedure PROC SURVEYLOGISTIC was the principal SAS procedure used in this study. The stratified analysis did not differ from the simple random sampling estimates in mean and precision. However, the clustered analysis somewhat decreased the precision without any change in the mean estimates. The weighted analysis resulted in a change of mean estimates, and also resulted in a further decrease in precision. The impact of weighted analysis again carried on most into the combined analysis, where all these different survey aspects were accounted for simultaneously.

The interpretation of the above results is as follows: first, the unchanged effect of the stratified analysis in the overall group (*i.e.*, Belgium) (with respect to simple random sampling) points towards the fact that the within strata variability might not be too different from the between strata variability of BMI. Second, in the clustered analysis, a somewhat decrease in precision could indicate towards a positive intraclass correlation within-cluster (households). Finally, weighted or combined analysis is important in this study when population-level (*i.e.*, representativity issue) inferences are sought accounting for the survey design, albeit it comes at a loss in precision.

Limitations of the study

We started fitting first with a parsimonious modelling approach for an ordinal outcome – proportional odds model. In the SRS analysis for the regions – Flanders and Brussels, the model fit was not rejected (p = 0.21 and p = 0.07, respectively). But the model fit was poor for Belgium (overall) and the Brussels region ($p \leq 0.01$ in both). So we used a more general approach – baseline category logit model that treated the outcome as a nominal variable, comparing the adjusted odds of obese and overweight categories each time with respect to the 'normal' (reference) category of the outcome – Body Mass Index (BMI), for the three exposure variables. This also impacted our selection of SAS procedures in this analysis (*cf.* paragraph below). For consistency reasons and comparability issues, we restricted all the models to baseline category logits applicable to multinomial distributions.

The SAS PROC GENMOD procedure is used for fitting Generalised Linear Models (GLM) *[McCullagh & Nelder, 1989]* with correlated / clustered data through the so-called Generalised Estimating Equations (GEE). However in our study with a trichotomised outcome, this was not possible, because although PROC GENMOD can fit a GEE proportional odds model (link = cumlogit), but it does not accommodate a baseline category logit model (link = glogit). As discussed above that due to the poor fit of the proportional odds model in case of certain regions, we had to abandon this approach in the current study as a general model for all analyses. This restricted our analysis to using the SAS PROC LOGISTIC (or PROC CATMOD) procedure (instead of the PROC GENMOD), which cannot accommodate clustered data; whereas PROC GENMOD carries out a moment-based (not maximum likelihood) estimation through an exchangeable working correlation matrix (currently four options available) in case of clustered data to yield the parameter estimates of a marginal model *[Liang & Zeger, 1986]*. This signified that only an 'SRS'-like analysis and 'weighted'-like analysis could be carried out to compare with the survey-specific procedure results.

The SAS PROC GLIMMIX procedure has the advantage over the SAS PROC NLMIXED procedure in that the weights can be included easily in its code. However, for a binary or in our case, a three-category outcome, the results can be biased as it uses a linearisation of the outcome by a Taylor-series expansion (of first degree), thus producing a 'pseudo-data' to fit the model. The procedure works relatively well with normally (or *t*-) distributed data or even for a count data (~ Poisson or negative binomial distribution), but not so well for a binary or 'ternary' outcome. Further, we estimated by Restricted Penalised Quasi-Likelihood method (default in PROC GLIMMIX), which performs better in the presence of significant random-effects than the other available options *[Molenberghs & Verbeke, 2005]*. Furthermore, the 'empirical' and '_residual_'options (or 'lsmeans' statement) could not be used in case of an ordinal outcome (~ multinomial distribution) with PROC GLIMMIX, meaning that the precision might be overestimated and R side random-effects could not be estimated, respectively. Finally, the estimates from a PROC GLIMMIX have a household-level interpretation (conditional) as opposed to a marginal interpretation with other procedures that do not accommodate random-effects, given that that for non-Gaussian data, the 'marginalised' estimates by integrating out the household-specific random-effects will most likely not be the same as those obtained from a marginal model without random-effects. In conclusion, we cannot trust the estimates by PROC GLIMMIX in case of ordinal data with three categories. A final caveat in this regard was that in spite of these inconveniences just stated, on running PROC GLIMMIX for more than 24 hours with a generalised logit link function, the model did not converge; using Cholesky option, increasing the maximum number of interations, changing the type of variance-covariance matrix or the iteration algorithm did not help in achieving model convergence (*e.g.*, this is also cited in SAS forums). The convergence was however achieved using the cumulative logit link function for a proportional odds model (not shown), which was not the right kind of model for this data as mentioned earlier. This means we are also unable to present the PROC GLIMMIX result outputs for comparison with the survey-specific result outputs.

5. References

1. Agresti A. *Categorical Data Analysis*, 2nd Ed., New York: John Wiley & Sons, Inc., 2002.
2. Cassell DL. Wait, wait, don't tell me… you're using the wrong proc. *Proc. of SUGI 31 (paper 193).*
3. Dirken JM. *Arbeid en Stress; het vaststellen van aanpassingsproblemen in werksituaties [Work and stress: assessing problems in the work situation]*, Tjeenk Willink, Groningen, the Netherlands, 1967.
4. Goldberg DP. *The detection of psychiatric illness by questionnaire*. London, Oxford University Press, 1972.
5. Gecková A, van Dijk JP, Stewart R, Groothoff JW, Post D. Influence of social support on health among gender and socio-economic groups of adolescents. *Eur J Pub Health* 2003;**13**:44-50.
6. Liang KY, Zeger SL. Models for Longitudinal Data: A Generalized Estimating Equation Approach. *Biometrics* 1986;**44**:1049-60.
7. McCullagh P, Nelder JA. *Generalised Linear Models*, London: Chapman Hall, 1989.
8. Molenberghs G, Verbeke G. *Models for Repeated Discrete Data*. New York: Springer-Verlag.
9. *SAS / STAT User's Guide*, version 9.1.3, Cary, NC, USA.
10. van Oyen H, Tafforeau J, Hermans H, Quataert P, Schiettecatte E, Lebrun L, Bellamammer L. The Belgian Health Interview Survey. *Arch Public Health* 1997;**55**:1-13.
11. World Health Organisation. BMI classification.

6. Appendix to selected SAS codes

A. Clustered, stratified, weighted analysis by domain corrected for the total (approximate) Belgian population

```
PROC SURVEYLOGISTIC DATA = bmi2 TOTAL = 10000000;
  TITLE 'B-C logit model: ALL (by region)';
  BY regionch;
  STRATA province;
  CLUSTER hh;
  CLASS sgp (PARAM = REF REF = '1') ghqbin voegcatn;
  MODEL bmicatn (REF = '1') = sgp voegcatn ghqbin / LINK = GLOGIT;
  WEIGHT wfin;
RUN;
```

B. Clustered, weighted analysis by domain using Generalised Linear Mixed Model (Restricted Penalised Quasi-Likelihood)

```
PROC GLIMMIX DATA = bmi2 METHOD = RSPL;
  NLOPTIONS MAXITER = 250 TECHNIQUE = NEWRAP;
  TITLE 'B-C logit model: GLMM (by region) WT CS';
  BY regionch;
  CLASS sgp ghqbin voegcatn hh;
  MODEL bmicatn (ORDER = FREQ REF = FIRST) = sgp voegcatn ghqbin / S
                DDFM = KR LINK = GLOGIT DIST = MULTINOMIAL ODDSRATIO;
  RANDOM int / SUBJECT = hh TYPE = CS GROUP = bmicatn;
  WEIGHT wfin;
RUN;
```