Transnational University Limburg is a unique collaboration of two universities in two countries: the University of Hasselt and Maastricht University.

UHASSELT
KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

UHASSELT
KNOWLEDGE IN ACTION

Maastricht University

Faculty of Sciences
School for Information Technology

Master of Statistics

Masterthesis

The impact on quality of life assessment of differences in meaning of words between Dutch-speaking patients in Belgium and the Netherlands

Michael Obimpeh

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics, specialization Biostatistics

SUPERVISOR :
Mevrouw Robin BRUYNDONCKX

SUPERVISOR :
Prof.dr. Samuel COENEN

2017
2018

Maastricht University

Faculty of Sciences
School for Information Technology

Master of Statistics

Masterthesis

The impact on quality of life assessment of differences in meaning of words between Dutch-speaking patients in Belgium and the Netherlands

Michael Obimpeh

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics, specialization Biostatistics

SUPERVISOR :

Mevrouw Robin BRUYNDONCKX

SUPERVISOR :

Prof.dr. Samuel COENEN

Abstract

Introduction and Objective

The main objective of this study is to determine the impact on the quality of life assessment of the difference in meaning of words between patients in Flanders, Belgium and patients in the Netherlands .This is due to the fact that these two groups of patients share a common language known as Dutch.The Dutch EuroQol EQ-5D is used to measure the quality of life.Both patients in the Netherlands and patients in Flanders, Belgium completed the Dutch EuroQol EQ-5D version where ' walking about' is translated with "lopen" which would be interpreted as running by Flemish patients and as walking about by Dutch patients.However, due to the slight difference in meaning of words between the two groups of patients, we are to determine if this difference affects their quality of life after accounting for other explanatory variables corresponding to disease severity.

Statistical Methodology

The endpoint of this study is Mobility, which is an ordered categorical variable hence Proportional odds and Partial proportional odds models were fitted.Due to the large number of explanatory variables,univariate models were fitted in bid to obtain important covariates.They were subjected to stepwise regression to select variables which will improve the fit of the model.The proportional odds assumption of the models were violated hence Partial proportional odds model was finally settled upon and used for the final analysis.

Results

Two different models were fitted to the data after selecting of important explanatory variables.One with an interaction term between country and gender and a second without an interaction term .Both models were fitted with partial proportionals odds since the assumption of proportionality was violated.We therefore realized that,in model 1 after accounting for important variables ,the probability to be in a lowest category of Mobility is 50% higher for patients from the Netherlands than for patients from Belgium however, this effect is statistically not significant.Also in model 2,females from the Netherlands are approximately 3 times as likely to respond in the lowest category of Mobility as compared to females from Belgium and this is statistically significant.

Discussion

Based on model 1, we don't need to know what country a patient comes from in order to predict his or her mobility score.However, from model 2,conditional on this optimal prediction of the mobility score, scores from patients in Belgium and the Netherlands differ among female patients.We thereby conclude that there is a logic in the fact that Flemish responders have more problems with 'running' than Dutch responders have with 'walking', and this is true only for female responders.

Contents

List of Figures

List of Tables

1 Introduction

It is an acknowledgeable fact that, illness as a burden on a patient cannot be fully described by measures of disease status ,(Muldon et al., 2014).Quality of life measurements therefore tend to describe fully how a patient's way of life such as ones ability to carry out certain activities as a result of the effect of a disease or severity of a sickness. Quality of life is therefore a useful and unique self reported assessment of a patient's status and has been substantially researched and has also been involved in important clinical decision making, (Lin et al.,2017).

Measurement of quality of life is on the rise to subjectively measure the health state of a patient due to the impact and effects of a disease. Uses of quality of life measurements for over a period have shown that there is an increasing acceptance of its measurement and uses with regards to how patients assess themselves as to how satisfied they are, such as effects of a disease condition and being able to perform certain activities, (Alison and Higginson, 2001). Furthermore these quality of life assessments are known to be patient specific and also has its limitations as they are being subjectively measured and often reflects patients differences in the categories of patients who assess their quality of life. A main drive and objective of this thesis stems from this fact.Two groups of patients from different countries filling out a quality of life questionnaire, there is bound to be a likely difference in these two groups patients such as the difference in meaning of certain words which are familiar to both groups of patients.

The endpoint or the outcome of this study is a quality of life measured by the Dutch EuroQol EQ-5D.The EuroQol EQ-5D is a standardized measure of health status developed by the EuroQol Group in order to provide a simple, generic measure of health for clinical and economic appraisal(Reenen et al.,2015). The EuroQol EQ-5D was designed for self completion by respondents. This is a 5 dimension with three levels of health status to be completed by the correspondents.The five dimensions contained in the data are,mobility, self-care, usual activities, pain/discomfort and anxiety/depression.Each dimension has 3 levels such as no problems, some problems and extreme problems.These are coded as 1,2 and 3 respectively . The respondent is asked to indicate his or her health state by ticking in the box against the most appropriate statement in each of the 5 dimensions.

Often, we ask questions with regards to whether quality of life measurements truly reflect the real health status of the patients who fill them. Differences could arise if different patients fill them with respect to how well they understand the meaning of the questions being asked. These differences could have an effect on how a patient answers his questionnaire, and in the long run could have an impact on the quality of life a assessed by the patient.
The Dutch from the Netherlands and the Belgians from Flanders,Belgium share a common language which is Dutch language.However, there are nuisances and slight differences in the meaning of certain words between these two groups of people.For instance words such as "lopen', "zich somber voelen" ,"Angst / depressie" could be interpreted differently by patients from Belgium and patients from the Netherlands. With regards to this project, patients in the Netherlands and patients in Flanders, Belgium completed the Dutch EuroQol EQ-5D version where ' walking about' is translated with "lopen" which would be interpreted

as running by Flemish patients and as walking about by Dutch patients. Hence when patients in Flanders, Belgium and patients in the Netherlands are asked about their ability to walk , there is likely to be differences in their scoring of their ability to "lopen" on the quality of life questionnaire. The main objective of this project is to determine the impact on the quality of life assessment of the difference in meaning of the word "lopen" between patients in Flanders, Belgium and patients in the Netherlands after accounting for other covariates such as other elements of the quality of life assessment or the severity of symptoms.

1.1 Log-Linear Models for agreement

The general objective of this study is to assess the agreement between scores given by patients in Flanders, Belgium and scores given by patients in the Netherlands through the linear model framework. According to (Hunt et al., 2015) agreement between two different raters with regards to their scoring of a particular question is normally measured or determined using the Cohen's Kappa statistic. Cohen's Kappa Statistics is usually used to measure inter rater agreement when the measurement is qualitative. The Cohen's Kappa Statistics however have limitations as it is highly dependent on the underlying distributions of the condition or situation been scored and also merely measures agreement between the two rater without accounting for important explanatory variables such as disease severity and patients or raters characteristics. The degree of agreement measured by Kappa values with regards to different groups of people which assumes different distributions are burdensome to compare with statistical precision (Hunt et al.,2015).

Log linear models for agreement, as used in a study by (Hunt et al., 2015), was originally developed by (Agresti,1988) and these models have been applied to health and behavioral research to assess the agreement between two categories of raters with respect to their scoring. Unlike the Cohen's Kappa statistic , the Log linear modeling of agreement between two raters depends less on the underlying distribution of the the condition which is being scored and makes room for assessing agreement between two groups of patients.Below is a formulation of a log linear model of agreement,in order to run this model the outcome which is a three level is broken down into a 3 by 3 contingency table and then modeled as a count data.

$$log(Y_{ij}) = \lambda_0 x_0 + \lambda_1 x_1 + \lambda_2 x_2 + \delta x_3 + \beta x_4$$

With regards to the model above, Y_{ij} represents the expected counts in a 3 by 3 contigency table, the intercept covariate, x_0, which is a constant always takes a value of one. The covariate x_1 takes a value of zero when neither Patient assigns a score of one, a value of one when either Patients, but not both, assigns a score of one, and a value of two when both raters assign a score of one, to EuroQol ED Mobility measurement. Similarly, x_2 takes a value of zero when neither rater assigns a score of 2, a value of one when only one rater assigns a score of 2, and a value of two when both assign a score of two to EuroQol ED Mobility measurement.

The covariate corresponding to δ, x_3, equals one when the two patients assign equal scores

to the EuroQol ED Mobility , that is, when they agree exactly, and zero otherwise.
The covariate corresponding to β, x_4, equals the product of the row number and column number and ranges from 1 to 9 for the three by three table considered here.

Before this model can be adjusted to account for other explanatory variables,the contingency table has to be stratified to adjust for each covariate making it extremely complex in the model building since there are quite a number of explanatory variables. Moreover , for such analysis to be adjusted for continues variables ,such variables have to be categorized which could also lead to loss of information.In light of these challenges, Multinomial models were proposed to assess the agreement between the two groups of patients after accounting for important explanatory variables.Since the response is ordinal in nature ,cumulative logistic regression models such as proportional and partial proportional models were considered in the next section.

2 Methodology

The data for this project was obtained from Genomics to Combact Resistance against Antibiotics in Community-acquired LRTI in Europe (GRACE study). The data was collected by an European project on the management of community-acquired lower respiratory tract infections in primary care that assessed patients' quality of life using the EuroQol EQ-5D.The EuroQol EQ-5D is a quality of life measurement which contains the main endpoint of the study known as "Mobilty".The endpoint has three categories namely "no problems","some problems" and "confined to bed".These levels are coded 1,2 and 3 respectively.
The data also comes along with several explanatory variables particularly corresponding to disease severity as well as patient measured characteristics such as age, country of origin and gender to mention but a few.These explanatory variables are a combination of categorical and continues variables.A total of 724 questionnaires administered to 395 patients from Belgium and 327 patients from Netherlands.There a total of 212 variables in the data with quite some missing values in the explanatory variables.

2.1 Cumulative Logit Models

Responses which are ordinal are very common in clinical, medical , epidemiological and social science studies.Mobility which is the main response in this research comes with a natural ordering. This is an ordinal response which has 3 levels such as no problems, some problems and extreme problems and these are coded as 1,2 and 3 respectively.
Logits can utilize the ordering of a response variable when the response categories are ordered naturally which results in simpler models with simple interpretation and potentially greater power than baseline category logit models (Agresti, 2007).

Latent Variable motivation for the proportional odds model

A motivation for the proportional odds model is that, due to the ordinal nature of the response variable, it can be associated with a model which assumes an underling continues variable since the different levels of the response were subjectively selected,(Agresti, 2007). According to (John and Jangman, 2009) ,the proportional odds model is often derived when a latent response variable ξ is assumed to follow a linear regression as given below,

$$\xi_i = \alpha + x_i'\beta + \varepsilon_i$$

$$\xi_i = \eta_i + \varepsilon_i$$

where x_i is the i_{th} row of the model matrix and β is a vector of regression coefficients, and α is the intercept parameter .It was further stated that although the latent response cannot be observed directly, a binned version of it, y, with n levels is available:

$$y_{=}\begin{cases} 1 & \text{for} \quad \xi_i \leq \alpha_1 \\ 2 & \text{for} \quad \alpha_1 < \xi_i \leq \alpha_2 \\ . & \\ . & \\ n-1 & \text{for} \quad \alpha_{n-2} < \xi_i \leq \alpha_{n-1} \\ n-2 & \text{for} \quad \alpha_{n-1} < \xi_i \end{cases}$$

where $\alpha_1 < \alpha_2 < ... < \alpha_{n-2}$, are estimates obtained from the data.The observed response now has a cumulative distribution given by;

$$\begin{aligned} \Pr(y_i \leq j) &= \Pr(\xi_i \leq \alpha_j) \\ &= \Pr(\alpha + \eta_i + \varepsilon_i \leq \alpha_j) \\ &= \Pr(\varepsilon_i \leq \alpha_j - \alpha - \eta_i) \end{aligned}$$

for $j = 1, 2, ...n-1$
If we assume the errors ε_i are logically distributed with a function

$$\Lambda(\varepsilon_i) = \frac{1}{1 + e^{-\varepsilon_i}}$$

leads to an ordered probit model known as the proportional odds model.

2.2 Proportional odds model

We first dichotomize the response variable Mobility incorporating the ordinal information,

$$P(Y \leq j) = \pi_1 + \pi_2 + ... + \pi_j = \sum_{k=1}^{j} \pi_k \quad for \quad j = 1, ...J-1,$$

$$log\left(\frac{P(Y \leq j)}{P(Y > j)}\right) = log\left(\frac{\pi_1 + \pi_2 + ... + \pi_j}{\pi_{j+1} + \pi_{j+2} + ... + \pi_J}\right) \quad for \quad j = 1, ...J-1,$$

$$log\left(\frac{P(Y \leq j)}{P(Y > j)}\right) = \gamma_j + X'\beta$$

Where Y is the response Variable ,J the levels of the response variable,γ_j is the constant, X' a matrix of the explanatory variables in the model and β representing estimates of the parameters corresponding to the explanatory variables in the model.Also π_j is the probability of response in category j.
Since the response variable is a three level variable, it produces two equations when we model in the direction of "No Problems in Walking" (Level 1) of the response variable.

$$log\left(P(Y \leq \text{"No Problems in Walking"})\right) = \gamma_j + X'\beta$$
$$log\left(P(Y \leq \text{"Some Problems in Walking"})\right) = \gamma_{j-1} + X'\beta$$

From the model above, proportional odds models assumes constant effect of each covariate across the different levels of the Mobility, which is the outcome of interest with the only difference being the constants γ_j and γ_{j-1}.Score test for proportional assumptions are carried out to check the assumption of proportionality before the model can be interpreted.If this assumption is violated,several other models in Multinomial family of models could be fitted but in this project, partial proportional odds model were adopted if the assumption of proportionality is violated.To begin with the model building, we need to select important covariates from the lots of available variables in the dataset.

2.3 Univariate Relationships

The variables in the "GRACE" data set are 212 with quite a lot of missing values in the explanatory variables.Univariate analysis were carried out for each covariate separately to determine its significance with a significance level of 0.05. Moreover, due the missing nature of the data, multiple imputation techniques were used to impute missing data before the univariate analysis were carried out. The SAS procedure PROC MI was used for this computations.
We looked at the the univariate model below.

$$log\left(\frac{P(Y \leq j)}{P(Y > j)}\right) = \gamma_j + x'\beta \quad for \quad j = 1, ...J-1,$$

Where Y is the response Variable ,J the levels of the response variable,γ_j is the constant, x' a vector of the explanatory variable in the model and β representing estimates of the parameter corresponding to the explanatory variable in the model. This model is used to and test for significance of each explanatory variable.

2.4 Interaction Relationships and confounders

The variables which were significant after the univariate analysis were used to test for interactions as represented in the model below.Tests for possible confounders were also carried out to determine if any other covariate interferes in the relationship between the outcome of interest known as Mobility and the covariate of interest known as Country.

$$log\left(\frac{P(Y \leq j)}{P(Y > j)}\right) = \gamma_j + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 \quad for \quad j = 1, ...J-1,$$

Where Y is the response Variable ,J the levels of the response variable,γ_j is the constant, x_1 and x_2 are the explanatory variables in the model and β_1 ,β_2 and β_3 representing estimates of the parameters corresponding to the explanatory variables in the model.

2.5 Stepwise regression to select determinants of the outcome

All the variables that were significant determinants of Mobility at the univariate analysis were subjected to stepwise regression or selection procedure.The Stepwise variable selection

method was used in the PROC LOGISTIC procedure of SAS software to obtain variables which improves the fit of the model.The stepwise selection method starts with model with only an intercept term and evaluates all the available predictors separately to find significant covariates which mostly improves the fit of the model.The selection criteria was done with a significance level of $\alpha = 10\%$ which was specified in "SLE" and "SLS" options of the model statement in SAS software. That is selection criteria to enter and selection criteria to stay were 10%. The variables selected from this stepwise selection process were used for the subsequent model building.Furthermore, other variables such as country known to be clinically important or seems to be a possible confounder will be included even if it is not statistically significant.

2.6 Partial Proportional odds model

The test for proportional odds assumption which assumes constant effect of the explanatory variables across the different level of the response variable is often rejected when we have a large number of explanatory variables as well as a mixture of continues and categorical explanatory variables in the model (Allison, 1999; Brant, 1990) . Score test was used in this analysis to test for this assumption.Once this test rejects the assumption of proportionality, we cannot interpret the model.Hence we fit a models of the Response to each covariate separately to determine which covariates satisfies the assumption and which covariates do not. This leads us to Partial proportional odds model which has a main feature of allowing non proportional and proportional covariates to still be kept in the model and still preserves the ordinal nature of the outcome variable.this finally result in more efficient models even though it often results in more estimates for the covariates.Paul J. Hilliard(2017).

2.7 Statistical Software

SAS 9.4 and R version R-3.4.3 were used for all the statistical analysis and graphical illustrations of the dataset. All tests were also performed at a 5% significance level.

3 Results

3.1 Exploratory Data Analysis

As a way to discover patterns of systematic variation across groups of patients, as well as aspects of random variation that distinguish individual patients, an exploration of data was performed.

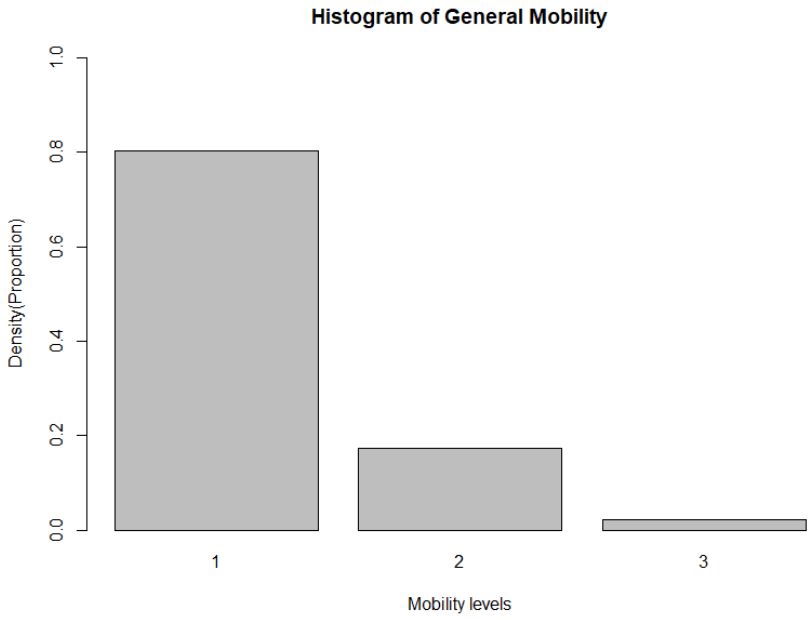


Figure 1: Overall distribution of Mobility question

From graph 1 above, a high proportion of responses irrespective of patient differences was recored as one which signifies no problems in walking or level 1 with regards to Mobility.

The graph in figure 2 shows the distribution of responses for patients in Flanders and for patients in Netherlands. Once again, both sides show high proportions of responses which corresponds to level 1 of the Mobility question as compared to levels 2 and 3.However since Mobility could have different meaning for patients from Flanders and patients from Belgium further analysis needs to be done to determine whether there is agreement in their scoring of this question.
The boxplot in figure 3 shows the distribution of the severity scores by the response category on the Mobility question.From the boxplots, we could see patients with high scores of disease severity tend to belong to levels 2 and 3 of the mobility and patients with low scores belong to the level 1 of the mobility outcome.

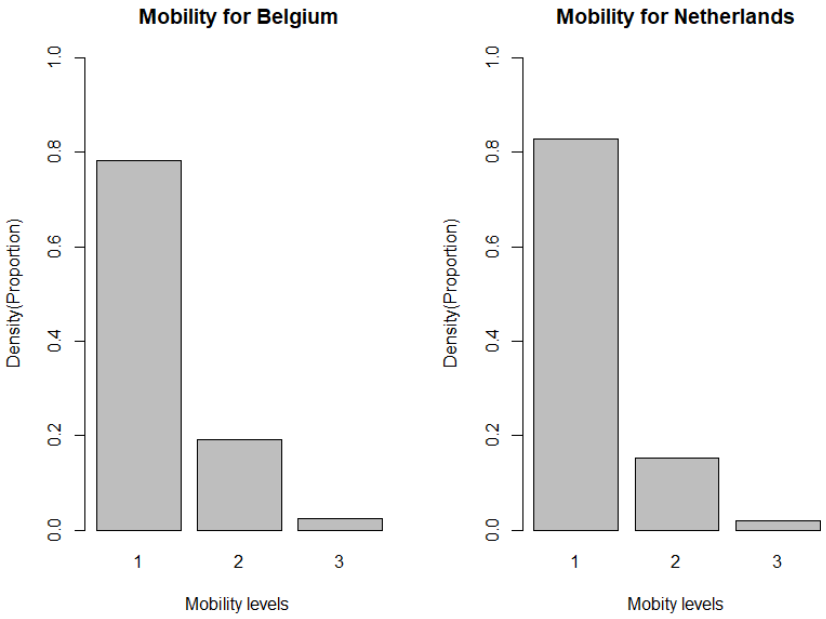


Figure 2: Distribution of Mobility by patients

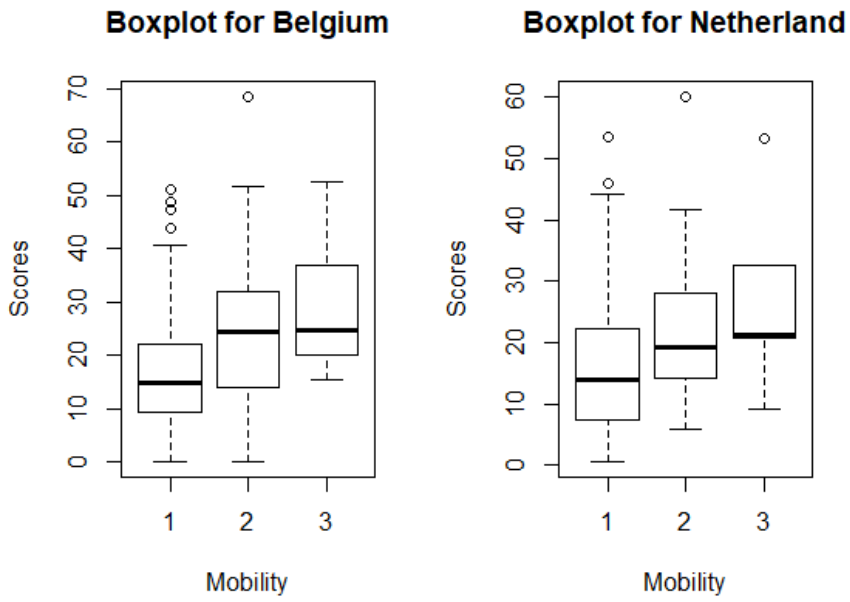


Figure 3: A boxplot of Mobility with respect to severity scores by patients

3.2 Univariate Analysis of Data

Table 1: Univariate Analysis to determine important covariates

Parameter Estimates (5 Imputations)			
Parameter	**Estimate**	**Std Error**	**Pr>|t|**
Country	0.1473	0.2304	0.5225
Breath_sev	0.708314	0.318706	0.0263
Feverb	1.747455	0.70055	0.0126
Wheeze_yn	0.869852	0.210541	<.0001
Fever_yn	-0.501307	0.227704	0.0277
Chest_pain_yn	-0.723368	0.212137	0.0006
Gen_tox_yn	-0.827953	0.288425	0.0041
symp_status	-1.18109	0.25675	<.0001
feverynb	0.501307	0.227704	0.0277
COPD_yn	-1.1854	0.2757	<.0001
Interf_actb	1.5932	0.3818	<.0001
or_ster_yn	-2.68	0.7739	0.0005
baselinem	-0.987302	0.220958	<.0001
symp_res_time	-0.059048	0.011631	<.0001
age	-0.0121	0.0066	0.00647
Severity_score	-0.0613	0.00977	<.0001

From the table above, all variables underwent univariate analysis but only the significant ones were presented and would be used to conduct further analysis with respect to answering our objective.The p values of the variables are less than the stated significance level of 0.05 in the exception of Country of origin of the patient , this variable is however left in the analysis since it is our variable of main interest and has to be kept even if it is not important.

3.3 Missing data pattern

The graph below shows the proportion of missing observations in the selected explanatory variables and their pattern of missing observations. From these,we could see that approximately 99% of the data points are complete observations whereas less than 1% represent monotone and intermittent missing observations. Since the percentage of missing data is considerably low, the successive analysis were carried out without imputation

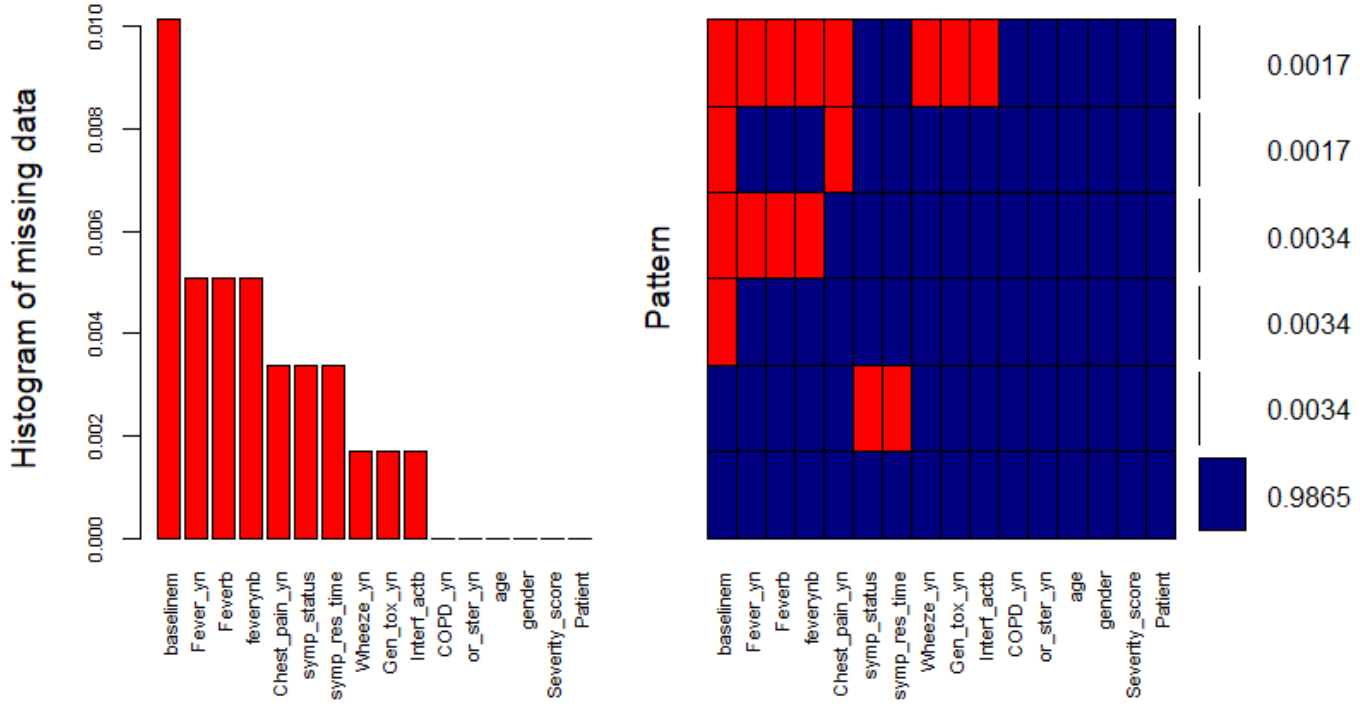


Figure 4: Missing data pattern and proportion in selected explanatory variables

3.4 Investigation of Interaction and confounding variables

The issue of confounding and interaction surfaces when the number of covariates in the model are quite many. A covariate confounds the relationship between another covariate of interest and an outcome variable if it is related to both the outcome and the covariate of interest. Most often , investigators determine whether there is confounding by estimating the measure of association before and after adjusting for a potential confounding covariate. When there is a change in the estimated measure of association by 10% or more , it would imply that confounding was present, but if the measure of association changes by less than 10%, there is likely to be little or no confounding by that variable (LaMorte and Sullivan 2013). The results are displayed in the successive tables below.

Table 2: Crude Odd ratio estimates for country only model

Effect	**Point Estimate**	**95% Wald Confidence Limits**	
country 0 vs 1	1.352	0.895	2.042

Table 3: Odd ratio estimates for country correcting for each covariate

Odds Ratio Estimates Adjusted for each covariate				
Adjusted Variable	**Effect**	**Point Estimate**	**95% Wald Confidence Limits**	
Coughb	**Country 0 vs 1**	1.343	0.888	2.032
Wheeze_yn	**Country 0 vs 1**	1.436	0.943	2.186
Chest_pain_yn	**Country 0 vs 1**	1.300	0.858	1.971
COPD_yn	**Country 0 vs 1**	1.331	0.876	2.024
Interf_actb	**Country 0 vs 1**	1.465	0.959	2.240
or_ster_yn	**Country 0 vs 1**	1.313	0.866	1.989
gender	**Country 0 vs 1**	1.354	0.897	2.046
age	**Country 0 vs 1**	1.385	0.915	2.097
symp_res_time	**Country 0 vs 1**	1.324	0.867	2.020
Severity_score	**Country 0 vs 1**	1.264	0.824	1.938
Breath_sev	**Country 0 vs 1**	1.455	0.895	2.365
Feverb	**Country 0 vs 1**	1.371	0.899	2.089
Fever_yn	**Country 0 vs 1**	1.374	0.907	2.082
Gen_tox_yn	**Country 0 vs 1**	1.231	0.807	1.877
symp_status	**Country 0 vs 1**	1.261	0.827	1.923
feverynb	**Country 0 vs 1**	1.374	0.907	2.082

From table 3 above, estimates of odd ratios and their respective 95% confidence intervals

are reported.Mobility as a response variable was fitted against Country in table 2 to see the crude effect of Country on mobility without adjusting for any possible confounding variable. There was no significance effect of Country on the Mobility .
Further analysis were run by regressing Mobility outcome on country and adjusting for each variable separately to see if it confounds the relationship between Country and Mobility. Table 3 shows the odd ratio estimate for Country after adjusting for each of the possible confounding variable, the changes in this odds ratio after adjusting for potential confounding variable is less than 10% as compared to the crude estimate of the country variable in table 2.It was therefore concluded that there was no confounding by each of these variables. It is of interest to also note that after accounting for severity of sickness, the country variable's change in odds ratio estimate was still less than 10% ,implying severity though it is a significant determinant of Mobility does not confound the relationship between Country and Mobility.

From table A.2, country covariate and interaction of country and each covariate were tested separately to determine whether an interaction between the Country variable and any of the important covariate is significant, we could see from the output that among the different models that were fitted for country and each of the covariate, the last two models in the table which included gender and age separately had significant interaction with Country.

3.5 Variable selection process for model building

From table A.3, stepwise variable selection method was used in the PROC LOGISTIC procedure of SAS software to obtain variables which improves the fit of the model.The explanatory variables which were statistically significant from the univariate selection process were subjected to the stepwise selection procedure.The stepwise selection method starts with model with only an intercept term and evaluates all the available predictors separately to find significant covariates which mostly improves the fit of the model.The outcome of the stepwise selection process yielded the variables in table A.3. These are the variables which were selected by the stepwise procedure and used for the subsequent analysis.

3.6 Proportional odds model without interaction

In this model, we used the selected covariates from the stepwise regression and included country. Thus,the variables from the model selection process were used to fit this proportional model.Country of origin of the patients was not statistically selected by the stepwise procedure however, this variable is clinically relevant to this study and was included in the model and used for the subsequent analysis despite it being insignificant. Table A.4shows estimates for the proportional odds models which assumes constant effect of covariates across the different levels of the response variable. Furthermore, these estimates could not be interpreted since the proportionality assumption under which they were obtained has to be satisfied first.

Table 4: Score test for the proportional odds assumption for model without interaction term

Chi-Square	DF	Pr>ChiSq
47.6592	14	<.0001

Table 4 shows the score test for the proportional odds assumption and with a p value of 0.0001, the assumption of proportionality is violated. Hence other forms of multinomial logistic regression could be considered preferably partial proportional odds .

Table 5: Score test for proportional Assumption for each covariate

Variables	Chi-Square	DF	Pr>ChiSq
Coughb	10.745	3	0.0132
Wheeze_yn	3.6731	1	0.0553
Chest_pain_yn	4.6565	1	0.0309
COPD_yn	1.1418	1	0.2853
Interf_actb	19.2358	3	0.0002
or_ster_yn	0.3258	1	0.5681
symp_res_time	3.5201	1	0.0606
age	6.5725	1	0.0104
Severity_score	0.066	1	0.7972
Country	0.0001	1	0.9974

Since the proportional odds assumption was rejected, each of the covariates were tested for non proportionality to determine which covariates satisfies the assumption and which covariates do not in table 5. It was discovered that Coughb, Chest_pain_yn ,Interf_actb and age are the only covariates which do not satisfy the proportional odds assumption since their pvalues are all less than the 5% significance level. Partial proportional models were fitted as a result and the outputs shown in table 6

Model 1 formulation

$$
\begin{aligned}
log\left(P(Y \leq \text{"No Problems in Walking"})\right) &= \beta_0 + \beta_1 Country + \beta_2 age + \beta_3 Coughb \\
&\quad + \beta_4 Coughb + \beta_5 Coughb + \beta_6 Wheeze\_yn \\
&\quad + \beta_7 Chest\_pain\_yn + \beta_8 COPD\_yn \\
&\quad + \beta_9 Interf\_actb + \beta_{10} Interf\_actb + \beta_{11} Interf\_actb \\
&\quad + \beta_{12} Or\_ster\_yn + \beta_{13} Symp\_res\_time \\
&\quad + \beta_{14} Severity\_score
\end{aligned}
$$

$$
\begin{aligned}
log\left(P(Y \leq \text{"Some Problems"})\right) &= \beta_{15} + \beta_1 Country + \beta_{16} age + \beta_{17} Coughb \\
&\quad + \beta_{18} Coughb + \beta_{19} Coughb + \beta_6 Wheeze\_yn \\
&\quad + \beta_{20} Chest\_pain\_yn + \beta_8 COPD\_yn \\
&\quad + \beta_{21} Interf\_actb + \beta_{22} Interf\_actb + \beta_{23} Interf\_actb \\
&\quad + \beta_{12} Or\_ster\_yn + \beta_{13} Symp\_res\_time \\
&\quad + \beta_{14} Severity\_score
\end{aligned}
$$

Model 1

Table 6: Partial Proportional odds model without interaction

Analysis of Maximum Likelihood Estimates						
Parameter		**Mobility**	**DF**	**Estimate**	**Standard Error**	**Pr>ChiSq**
Intercept		**1**	1	3.3094	0.6675	<.0001
Intercept		**2**	1	3.057	1.7196	0.0754
Country	**0**		1	0.3853	0.24	0.1083
age		**1**	1	-0.0262	0.00831	0.0016
age		**2**	1	0.039	0.0277	0.159
Coughb	**1**	**1**	1	-1.3841	0.7388	0.061
Coughb	**1**	**2**	1	-3.7021	1.6383	0.0238
Coughb	**2**	**1**	1	-0.7379	0.4487	0.1001
Coughb	**2**	**2**	1	-1.5132	1.2075	0.2101
Coughb	**3**	**1**	1	-0.6594	0.2687	0.0141
Coughb	**3**	**2**	1	1.1181	0.8582	0.1926
Wheeze_yn	**0**		1	0.7482	0.2409	0.0019
Chest_pain_yn	**1**	**1**	1	-0.3357	0.2474	0.1747
Chest_pain_yn	**1**	**2**	1	-2.2824	1.1093	0.0396
COPD_yn	**1**		1	-0.766	0.3203	0.0168
Interf_actb	**1**	**1**	1	1.7625	0.485	0.0003
Interf_actb	**1**	**2**	1	2.3748	1.7113	0.1652
Interf_actb	**2**	**1**	1	0.8982	0.4286	0.0361
Interf_actb	**2**	**2**	1	16.4613	759.4	0.9827
Interf_actb	**3**	**1**	1	0.4288	0.3913	0.2731
Interf_actb	**3**	**2**	1	1.6159	0.7607	0.0336
or_ster_yn	**1**		1	-3.2335	1.07	0.0025
symp_res_time			1	-0.0284	0.0141	0.0433
Severity_score			1	-0.0394	0.0114	0.0006

The table 6 above shows the estimates of the effects of the partial proportional odds model which allows covariates which satisfies the proportional odds assumption to be included and with constant effects across difference levels of the response variable and also allows different effects of the covariate which do not satisfy the proportional odds assumption to be included all at a time.The odds ratio estimate for the patient variable is estimated to be 1.470 with confidence interval (0.918 ; 2.353).However this effect is not statistically significant.

Table 7: Analysis of fixed effects in the model

Type 3 Analysis of Effects			
Effect	**DF**	**Wald Chi-Square**	**Pr>ChiSq**
Country	1	2.5783	0.1083
Age	2	13.6946	0.0011
Coughb	6	16.9974	0.0093
Wheeze_yn	1	9.6499	0.0019
Chest_pain_yn	2	5.3404	0.0692
COPD_yn	1	5.718	0.0168
Interf_actb	6	20.5984	0.0022
or_ster_yn	1	9.133	0.0025
symp_res_time	1	4.0822	0.0433
Severity_score	1	11.8681	0.0006

Table 7 above contains the test of hypothesis for each of the fixed effects in the model.We can conclude that all the fixed effects in the exception of Chest_pain_yn and Country of origin of the the patients are highly significant at the 5% significance level.

3.7 Model 2

Model 2 was formulated by adding a significant interaction term.This model however excludes some variables which were initially significant in the first model but insignificant in this model.Even though gender was not significant in any of the model building processes, its interaction with Country was significant and included in the model 2. Proportional odds was fitted and results are presented in table A.5. Moreover, the assumption of proportionality of the model is tested below.

Table 8: Score test for the proportional odds assumption for model with significant interaction

Chi-Square	**DF**	**Pr>ChiSq**
47.9216	13	<.0001

Table 8 above shows the score test for the proportional odds assumption and with a p value of 0.0001, the assumption of proportionality is violated. Hence we proceeded to check for the assumptions for each the covariates as displayed in table A.6 to check which covariates satisfies the proportionality assumption and which covariates do not. We therefore consider the partial proportional odds model below .

Partial proportional odds model fitting with significant interaction

Since the proportional odds assumptions were violated for some explanatory variables, we obtain the estimates of the effects of the partial proportional odds model which allows covariates which satisfies the proportional odds assumption to be included with constant effects across different levels of the response variable and also allows different effects of the covariate which do not satisfy the proportional odds assumption to be included all together.

3.8 Model 2 formulation

$$\begin{aligned} log\left(P(Y \leq \text{"No Problems in Walking"})\right) &= \beta_0 + \beta_1 Country + \beta_2 gender + \beta_3 age \\ &\quad + \beta_4 Coughb + \beta_5 Coughb + \beta_6 Coughb + \beta_7 Wheeze\_yn \\ &\quad + \beta_8 Interf\_actb + \beta_9 Interf\_actb + \beta_{10} Interf\_actb \\ &\quad + \beta_{11} Or\_ster\_yn + \beta_{12} Symp\_res\_time \\ &\quad + \beta_{13} Country * Gender \end{aligned}$$

$$
\begin{aligned}
log\left(P(Y \leq \text{"Some Problems"})\right) &= \beta_{14} + \beta_1 Country + \beta_2 gender + \beta_{15} age \\
&\quad + \beta_{16} Coughb + \beta_{17} Coughb + \beta_{18} Coughb + \beta_7 Wheeze\_yn \\
&\quad + \beta_{19} Interf\_actb + \beta_{20} Interf\_actb + \beta_{21} Interf\_actb \\
&\quad + \beta_{11} Or\_ster\_yn \beta_{12} symp\_res\_time \\
&\quad + \beta_{13} Countryt * Gender
\end{aligned}
$$

Table 9: Estimates with significant Country and Interaction terms

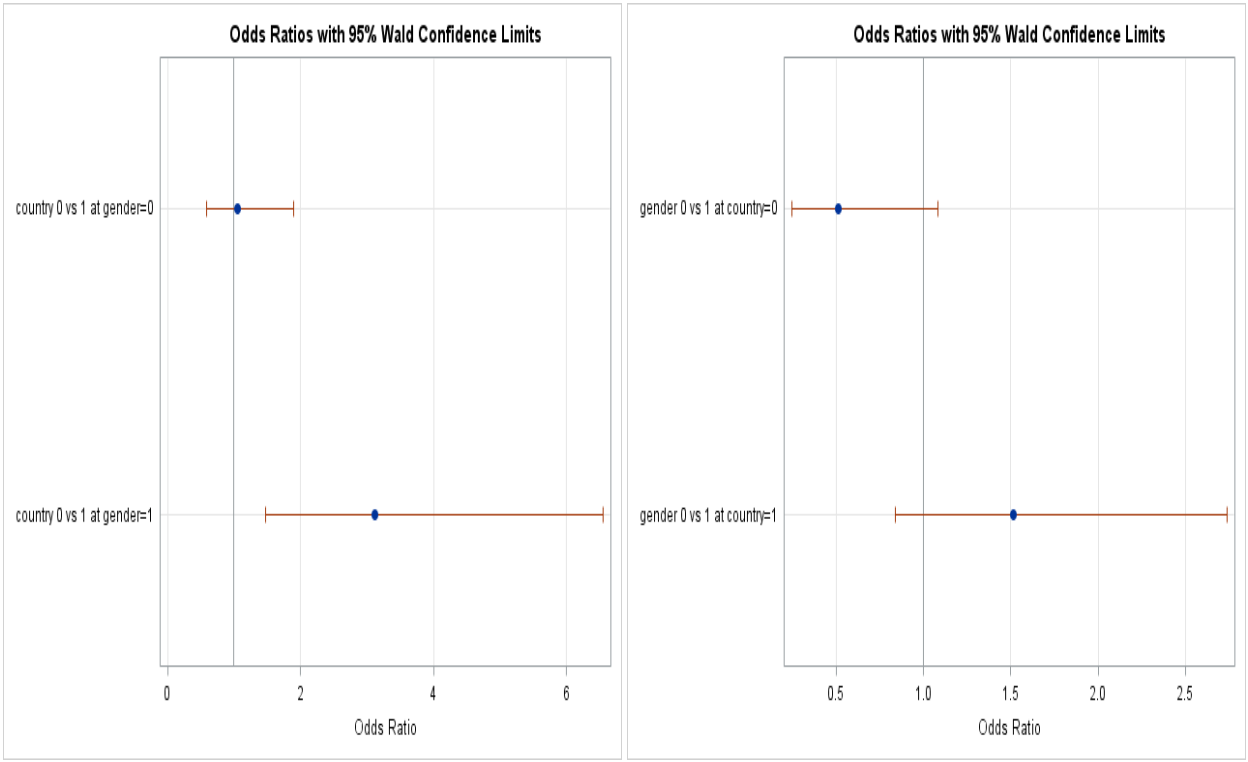
Analysis of Maximum Likelihood Estimates							
Parameter			**Mobility**	**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr>ChiSq**
Intercept			**1**	1	1.2424	0.7173	2.9996	0.0833
Intercept			**2**	1	3.721	210.8	0.0003	0.9859
country	**0**			1	0.297	0.1223	5.8998	**0.0151**
gender	**0**			1	-0.0628	0.1225	0.2625	0.6084
age			**1**	1	-0.0245	0.00779	9.9187	0.0016
age			**2**	1	0.0365	0.0209	3.0594	0.0803
Coughb	**1**		**1**	1	-0.9332	0.4986	3.5032	0.0613
Coughb	**1**		**2**	1	-2.9973	0.8081	13.7569	0.0002
Coughb	**2**		**1**	1	0.1355	0.3332	0.1654	0.6842
Coughb	**2**		**2**	1	0.5064	0.8255	0.3764	0.5396
Coughb	**3**		**1**	1	0.1281	0.2289	0.3133	0.5757
Coughb	**3**		**2**	1	1.7061	0.5933	8.2688	0.004
Wheeze_yn	**0**			1	0.4557	0.1162	15.3806	<.0001
Interf_actb	**1**		**1**	1	1.1802	0.2476	22.7146	<.0001
Interf_actb	**1**		**2**	1	-2.5518	210.8	0.0001	0.9903
Interf_actb	**2**		**1**	1	0.1707	0.2048	0.6947	0.4046
Interf_actb	**2**		**2**	1	11.2528	632.3	0.0003	0.9858
Interf_actb	**3**		**1**	1	-0.3668	0.1827	4.0315	0.0447
Interf_actb	**3**		**2**	1	-3.1934	210.8	0.0002	0.9879
or_ster_yn	**1**			1	-1.7565	0.5385	10.6388	0.0011
symp_res_time				1	-0.0504	0.0127	15.8585	<.0001
country*gender	**0**	**0**		1	-0.271	0.12	5.0969	**0.024**

Finally the model estimates in table 9 was obtained and reports significant Country by gender variable .Once we have an interaction term which is significant, it means the estimate of country in this model cannot be interpreted alone,thus the effect of country on mobility according to the model above depends on the gender of the patients.Even though we have a significant interaction term,it is not straightforward to interpret its effect on Mobility of the patient.Therefore to fully understand the interaction between country and gender,we need to explore several odds ratios which depicts the level of association between patients from the Netherlands and patients from Belgium at the different levels of the gender variable.

Table 10: Odds ratio estimates adjusted for Country and gender

Adjusted Odds Ratio Estimates and Wald Confidence Intervals			
Odds Ratio	**Estimate**	**95% Confidence Limits**	
country 0 vs 1 at gender=0	1.053	0.584	1.901
country 0 vs 1 at gender=1	3.114	1.48	6.553
gender 0 vs 1 at country=0	0.513	0.244	1.08
gender 0 vs 1 at country=1	1.516	0.839	2.74

We can see that a male patient who comes from the Netherlands is 1.053 times likely to be in the lowest category as they are to be in the higher categories as compared to a male patient from Belgium holding other variables constant.Thus the probability to be in a lowest category is 0.053 higher for male patients from the Netherlands than for male patients from Belgium and it is statistically significant.However this difference is considerably low and we can safely assume there is no difference between these two groups of male patients from Belgium and Netherlands.
Secondly, we can see that a female patient who comes from the Netherlands is 3.114 times as likely to be in the lowest category of Mobility as they are to be in the higher categories of Mobility as compared to a female patient from Belgium holding other variables constant.Thus the probability to be in a lowest category is 3 times higher for female patients from the Netherlands than for female patients from Belgium.Hence we might have a problem assessing the patients' mobility when using the wrong EuroQol Q5D version for women.



(a) Odd ratios of Country vs gender (b) Odd ratios of gender vs country

Figure 5: Figure of Odd ratios

In order to further interpret the results, effect plot was obtained for country by gender interaction as shown below

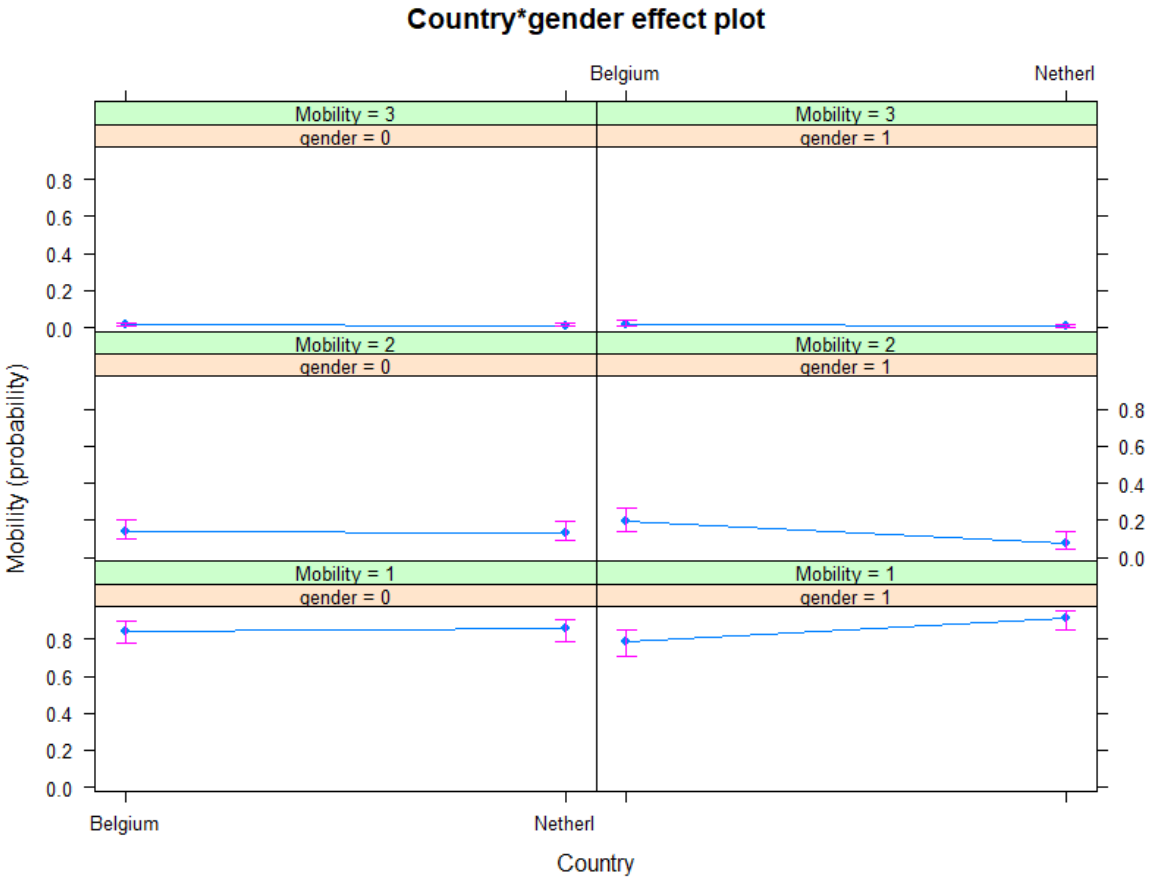


Figure 6: Effects plot of Country vs gender

We notice from the figure 6 above that, both male and female patients have equal probabilities of answering "Confined to bed" representing level three (3) of the mobility outcome in both countries. Also, we notice females in Belgium have slightly higher predicted probabilities of answering "some problems" than other patients who answered to "some problems". Finally, females from the Netherlands have the highest predicted probability of answering "No problems" as compared to other patients in the same category who also answered "No problems" which represents level 1 of the Mobility outcome.

Table 11: Comparison of Model with and without Interaction (Model 1 vs Model 2)

Model	**-2 Log L**	**AIC**
Model without Interaction	516.001	**564.001**
Model with significant Interaction	554.169	583.678

From table 11, above Model fit statistics were used to determine the fit of the model without interaction term and the model with interaction term. Comparing the AIC's of these models, we realized the Partial proportional odds model without interaction term provides a better fit since the model with the smaller AIC is preferred.
Moreover the covariates that explain the differences in scoring mobility is best described by model 1. Model 2 focuses more on country of origin and differences in scoring between patients from Belgium and patients from the Netherlands.Hence irrespective of their AIC values ,both models are useful.

4 Conclusion and Discussion

The main objective of this project is determine the impact on the quality of life assessment of the difference in meaning of words between dutch speaking patients from Flanders, Belgium and the Netherlands. Univariate analysis were carried out to select important variables which could determine this impact.There are several methodologies which could have been used to address this particular problem.However, due to certain limitations of these methodologies, Multinomial logits models were settled upon to answer the research question.Particularly proportional and partial proportion odds models were fitted to the data in a bid to answer the research question.

Two different models were fitted in a bid to answer the research question.After selecting important explanatory variables and checking for confounding variables as well as significant interaction terms,two models were fitted.Model 1 was without an interaction term and model 2 two was with a significant interaction term.Hence,covariates that explain the differences in scoring mobility is best described by model 1 and also, conditional on the differences in scoring mobility, model 2 looks at the impact of country of origin of the patients,however, model 1 is needed in order to obtain model 2.

Hence from the model without interaction, we can conclude that a patient who comes from the Netherlands is approximately 1.5 times as likely to be in the lowest category as they are to be in the higher categories as compared to a patient from Belgium holding other variables constant. Thus in model 1,the probability to be in a lowest category is 50% higher for patients from the Netherlands than for patients from Belgium however, this is statistically not significant.Also in model 2, female patients from the Netherlands are approximately 3 times as likely to respond in the lowest category of Mobility as compared to female patients from Belgium and this is statistically significant.Thus based on model 1, we don't need to know what country a patient is from to predict his or her mobility score.However, based on model 2, conditional on this optimal prediction of the mobility score, scores from patients in Flanders, Belgium and the Netherlands differ among female patients. We thereby conclude that there is a logic in the fact that Flemish responders have more problems with 'running' than Dutch responders have with 'walking', and this is true only for female responders.

5 References

1. Agresti A (1988). A model for agreement between ratings on an ordinal scale. Biometrics; 44: 539–48.

2. Agresti, A. (2007), An Introduction to Categorical Data Analysis, Second Edition, New York: John Wiley & Sons

3. Allison P D (1999). "Logistic Regression Using the SAS System: Theory and Application," Cary, NC.: SAS Institute.

4. Bagheban A A and Zayeri F (2010).A generalization of the uniform association model for assessing rater agreement in ordinal scales.Journal of Applied Statistics Vol. 37, No. 8, 1265–1273.

5. Brant R (1990). Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression," Biometrics, 46, 1171-1178.

6. Bob Derr,(2013).Ordinal Response Modeling with the LOGISTIC Procedure ,SAS Institute Inc.Paper 446-2013.

7. Carr A J and Higginson I J. "Are Quality of Life Measures Patient Centred"?.BMJ: British Medical Journal 322.7298(2001):1357-1360.

8. Hilliard P J (2017).Using New SAS 9.4 Features for Cumulative Logit Models with Partial Proportional Odds , Educational Testing Service (ETS).Paper Accompaniment for E-Poster 406-2017.

9. Hunt P R, Friesen M C, Sama S, Ryan L and Milton D,(2015) .Log-Linear Modeling of Agreement among Expert Exposure Assessors.

10. Jangman H J and John F (2009).Effect Displays in R for Multinomial and Proportional-Odds Logit Models:Extensions to the effects Package. McMaster University.Journal of Statistical Software October 2009, Volume 32, Issue 1.

11. Kerrie P N and Edwards D.Measures of agreement between many raters for ordinal classifications.Received 22 October 2014, Accepted 16 May 2015 Published online 21 June 2015 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002sim.6546

12. Lin C Y, Strong C, Tsai M C,Chih-Ting Lee (2017).Raters Interpret Positively and Negatively Worded Items Similarly in a Quality of Life Instrument for Children: Kid-KINDL

13. Marc J. Gameroff.Using the Proportional Odds Model for Health-Related Outcomes: Why, When, and How with Various SAS® Procedures.New York State Psychiatric Institute, New York, NY

14. Muldoon M F, Barger D S, Flory J D, Manuck S B,(1998).What are quality of life measurements measuring?

15. Nelson J C and Pepe M S.Statistical description of inter rater variability in ordinal ratings (2000).Statistical Methods in Medical Research: 475–496.

16. Reenen M V ,Oppe M, (2015). EQ-5D-3L User Guide.Basic information on how to use the EQ-5D-3L instrument Prepared.

17. Shana K (2017). Fitting a Cumulative Logistic Regression Model , Spectrum Health: Healthier Communities, Grand Rapids, MI. Paper 1108-2017.

18. Wayne W L and Lisa Sullivan(2013).Confounding and Effect Measure Modification,Boston University School of Publich Health.

A Appendix

Table A.1: Overview of selected variables in the data

Variable	Description
Endpoint	
Mobility	Inference with mobility (i.e. walking) (3-point scale)
Explanatory variables	
Country	Country of patients(0-Netherlands;1-Belgium)
Gender	Gender of patients(0-male,1-female)
Breath_sev	Severity of breathing
Feverb	severity of Fever at baseline
Coughb	severity of cough at baseline(4-point scale)
Wheeze_yn	Presence of wheezing(yes/no)
Fever_yn	Presence of Fever(yes/no)
Chest_pain_yn	Presence of chest pain (yes/no)
Gen_tox_yn	Sick impression(yes/no)
symp_status	status across all symptoms
COPD_yn	Presence of chronic obstructive pulmonary disorder (yes/no)
Interf_actb	Severity of interference with daily activities (5-point scale)
or_ster_yn	Use or oral steroids (yes/no)
baselinem	mean baseline symptom
symp_res_time	time to resolution of symptoms or last day of diary entry
age	Age of patient
Severity_score	Average severity score of patients corresponding to Mobility

Table A.2: Tests for significance of interaction

Joint Tests				
Model	**Effect**	**DF**	**Wald Chi-Square**	**Pr>ChiSq**
1	**Country**	1	0.0107	0.9175
	Coughb	3	0.5897	0.8988
	Country*Coughb	3	4.2075	0.2399
2	**Patient**	1	1.6407	0.2002
	Chest_pain_yn	1	11.3047	0.0008
	Country*Chest_pain_y	1	0.1523	0.6964
3	**Country**	1	2.7255	0.0988
	COPD_yn	1	14.7339	0.0001
	Country*COPD_yn	1	1.0079	0.3154
4	**Country**	1	3.1529	0.0758
	Interf_actb	3	23.5541	<.0001
	Patient*Interf_actb	3	5.9805	0.1126
5	**Country**	1	0.9842	0.3212
	or_ster_yn	1	12.8537	0.0003
	Country*or_ster_yn	1	1.8449	0.1744
6	**Country**	1	0.584	0.4447
	symp_res_time	1	19.8889	<.0001
	symp_res_tim*Country	1	3.4614	0.0628
7	**Country**	1	0.0668	0.7961
	Severity_score	1	40.4814	<.0001
	Severity_sco*Country	1	0.7704	0.3801
8	**Country**	1	0.1115	0.7385
	gender	1	0.5056	0.4771
	gender*Country	1	4.7929	**0.0286**
9	**Country**	1	2.2479	0.1338
	age	1	1.5757	0.2094
	age*Country	1	3.9271	**0.0475**

Table A.3: Selected variables by the stepwise regression

Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr>ChiSq**
Intercept	**1**	1	1.5544	0.7194	4.6684	0.0307
Intercept	**2**	1	4.2668	0.7524	32.1602	<.0001
Coughb	**1**	1	-1.2556	0.5108	6.0435	0.014
Coughb	**2**	1	0.0811	0.3354	0.0585	0.809
Coughb	**3**	1	0.2925	0.2341	1.5612	0.2115
Wheeze_yn	**0**	1	0.3538	0.1202	8.6569	0.0033
Chest_pain_yn	**1**	1	-0.2064	0.1226	2.8347	0.0922
COPD_yn	**1**	1	-0.4136	0.1625	6.476	0.0109
Interf_actb	**1**	1	1.0317	0.2577	16.029	<.0001
Interf_actb	**2**	1	0.1464	0.2076	0.4973	0.4807
Interf_actb	**3**	1	-0.2365	0.1812	1.7045	0.1917
or_ster_yn	**1**	1	-1.3964	0.4501	9.6258	0.0019
symp_res_time		1	-0.0282	0.0142	3.9198	0.0477
age		1	-0.0213	0.00815	6.8021	0.0091
Severity_score		1	-0.0438	0.0116	14.2134	0.0002

Table A.4: Estimates of Proportional odds model after adding Country to the model

Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr>ChiSq**
Intercept	**1**	1	1.9089	0.7335	6.7722	0.0093
Intercept	**2**	1	4.6345	0.7693	36.2888	<.0001
Country	0	1	0.1703	0.1190	2.0465	0.1526
Coughb	**1**	1	-1.0011	0.5098	3.8564	0.0496
Coughb	**2**	1	0.0275	0.3346	0.0067	0.9346
Coughb	**3**	1	0.1649	0.2331	0.5006	0.4792
Wheeze_yn	**0**	1	0.3668	0.1198	9.3714	0.0022
Chest_pain_yn	**1**	1	-0.1901	0.1219	2.4327	0.1188
COPD_yn	**1**	1	-0.371	0.1611	5.3049	0.0213
Interf_actb	**1**	1	0.9791	0.2516	15.1488	<.0001
Interf_actb	**2**	1	0.2113	0.2077	1.0347	0.3091
Interf_actb	**3**	1	-0.2435	0.1812	1.8059	0.179
or_ster_yn	**1**	1	-1.3701	0.4508	9.2392	0.0024
symp_res_time		1	-0.0287	0.0141	4.1191	0.0424
age		1	-0.0222	0.00806	7.5909	0.0059
Severity_score		1	-0.043	0.0115	13.9991	0.0002

Table A.5: Proportional odds estimates with significant interaction term

Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr>ChiSq**
Intercept	**1**	1	1.2501	0.6477	3.7251	0.0536
Intercept	**2**	1	3.8755	0.6831	32.1857	<.0001
country	**0**	1	0.2796	0.1211	5.3297	0.021
gender	**0**	1	-0.0469	0.1213	0.1496	0.6989
age		1	-0.0212	0.0076	7.7604	0.0053
Coughb	**1**	1	-1.1797	0.4895	5.8084	0.0159
Coughb	**2**	1	0.2007	0.3293	0.3713	0.5423
Coughb	**3**	1	0.236	0.2262	1.0889	0.2967
Wheeze_yn	**0**	1	0.4578	0.1158	15.6195	<.0001
Interf_actb	**1**	1	1.1844	0.2449	23.3943	<.0001
Interf_actb	**2**	1	0.2369	0.2036	1.3543	0.2445
Interf_actb	**3**	1	-0.289	0.1799	2.5812	0.1081
or_ster_yn	**1**	1	-1.4495	0.4477	10.4818	0.0012
symp_res_time		1	-0.0514	0.0127	16.411	<.0001
country*gender	**0 0**	1	-0.2466	0.1189	4.3024	0.0381

Table A.6: Score test for proportional Assumption for each covariate with interaction term

Variables	**Chi-Square**	**DF**	**Pr>ChiSq**
Country	0.0001	1	0.9974
gender	0.1204	1	0.7286
age	6.5725	1	0.0104
Coughb	10.745	3	0.0132
Wheeze_yn	3.6731	1	0.0553
Interf_actb	19.2358	3	0.0002
or_ster_yn	0.3258	1	0.5681
symp_res_time	3.5201	1	0.0606
Country*gender	1.2421	1	0.2651