UHASSELT
KNOWLEDGE IN ACTION

Maastricht University

Faculty of Sciences
School for Information Technology
Master of Statistics and Data Science

Master's thesis

Statistical Modelling of COVID-19 Mortality and Excess Mortality in Belgium

Chinenye Innocent Okafor

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

SUPERVISOR :

Prof. dr. Geert MOLENBERGHS

De heer Johan VERBEECK

2020
2021

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

Maastricht University

Faculty of Sciences
School for Information Technology
Master of Statistics and Data Science

Master's thesis

Statistical Modelling of COVID-19 Mortality and Excess Mortality in Belgium

Chinenye Innocent Okafor

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

SUPERVISOR :

Prof. dr. Geert MOLENBERGHS

De heer Johan VERBEECK

Acknowledgement

First and foremost I would like to thank the almighty God for having made everything possible. Secondly, I would like to pay my special regards to my supervisors, Prof. Dr. Molenberghs and Mr. Johan Verbeeck for their guidance, and support during this thesis. It is also my privilege to thank all my family members, especially my mother Mrs. Christiana Osu for her prayers, unconditional love, patients, advice and being the source of my motivation in this journey. I wouldn't forget to say thank you to my most lovely father Mr. Innocent Osu, and siblings Ebere, Ikenna, Jideobi, Udoka, and Oluwaseun for their huge supports in one way or the other. Thanks to all professors at Censtat that aid me during my studies in Uhasselt.

Many thanks to my classmates, friends, and group members, most especially to Connie Musisi, Nur Ferdous, Michale Obimpeh, and Festus Okonofua for their constant support and encouragement.

Okafor Chinenye, Innocent
June 30, 2021.
Diepenbeek, Belgium

Abstract

Background: COVID-19 has increased the mortality worldwide, including in Belgium. Lethality of COVID-19 differs however between sexes and increases with age. Analysis based on reported COVID-19 deaths and on excess mortality both show a higher risk of death for males then for females. However, COVID-19 deaths may be subject to under-reporting and the model used for excess mortality ignores historical excess mortality and correlation of deaths between consecutive weeks.

Objective: To study age gender difference in COVID-19 mortality in the year 2020 in Belgium, by analyzing the reported COVID-19 mortality and the excess mortality estimated via an improved model. Additionally, differences between three all-cause mortality waves in 2020 are studied.

Methodology: A negative binomial regression model was fitted to the COVID-19 reported mortality to evaluate age and gender differences and their interaction. Excess mortality is obtained by taking the difference between the observed and predicted mortality. To predict the subject-specific all-cause mortality for the year 2020 (from week 11-week 52), a linear mixed model was fitted that allows for serial correlation and which reduces the influence of historical excess mortality caused by heat-waves and seasonal influenza, by two methodology. Adequacy of the linear mixed model prediction was assessed by taking the percentage of the root mean square error (RMSE%) and compared to the weekly average methodology.

Result: The COVID-19 reported mortality shows a higher mortality for females below 40 years compared to males during the spring, and winter waves, and on the entire year 2020, while the reverse is true above 40 years. Mortality in the summer wave does not show any age-gender differences. The weekly average and linear mixed model show more excess deaths in females than males in the age group above 80 years in all the three waves, and the entire year 2020. Reported COVID-19 mortality coincides with excess deaths in Belgium, except for the summer wave. This discrepancy can be largely explained by a heat wave in August 2020.

Conclusion: Based on the findings, there is a difference in mortality between age and gender, where more females died due to COVID-19 in Belgium than males and COVID-19 leads to more deaths in older individuals. The gender finding in this study is however opposite to findings reported in the literature, because we have not corrected for population size.

Keywords: COVID-19 Mortality, Excess Mortality, Negative Binomial Regression Model, Linear Mixed Model, Weekly Average

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Coronavirus (COVID-19)

Coronavirus disease (COVID-19) is an illness caused by the SARS-CoV-2 virus. Most people infected with COVID-19 experience symptoms like mild to moderate respiratory symptoms and might recover without requiring special treatment [15]. However, those with medical conditions such as diabetes, chronic respiratory disease, and the elderly are at a higher risk of becoming more seriously ill [15]. The SARS-CoV-2 virus can be spread when an infected person coughs or sneezes, via a droplet of saliva, and it can affect persons in different ways [15]. Measures to prevent or to reduce the transmission of COVID-19 include washing of hands regularly, ventilation and air filtering, social distancing, covering mouth when sneezing or coughing, wearing face masks in public, disinfecting surfaces, and monitoring and self-isolation for persons exposed or symptomatic [15].

Despite the measures to reduce the transmission of COVID-19, many people died due to COVID-19. Mortality is one of the measures to evaluate the gravity of COVID-19 within and between countries. In Belgium, the first deaths due to COVID-19 were reported on March 11, 2020, i.e the 11th week of the year [13]. From this date (week 11), the cumulative number of deaths due to COVID-19 gradually increased.

COVID-19 mortality can be determined in two ways, by the reported COVID-19 mortality or via excess mortality. The reported COVID-19 mortality however depends on the completeness and strategy of counting deaths on the testing procedure of COVID-19, or availability of testing materials. Hence, the reported COVID-19 mortality might not be accurate and may lead to variability in reporting completeness between countries [14]. In this case, excess mortality has been suggested to evaluate the impact of COVID-19 mortality [3] .

The baseline mortality, i.e the predicted all-cause mortality based on historical mortality data is the major part in determining the excess mortality. However, historical mortality data contains excess mortality due to heat waves and seasonal influenza [14]. To reduce the

influence of this historical excess mortality, the standardized residuals are used to down-weight the historical mortality [5]. However, other models such as the time series models, including ARIMA models can also be used, but these models are plausible only when there is stationarity in the time series.

Several authors have observed sex and age differences in COVID-19 mortality, either via reported mortality [1] or via excess mortality [7]. In 10 European countries, excluding Belgium, it was observed that in almost all age groups, more males have died from COVID-19 than females. Differences in sex increased until the ages of 60–69 years, but decreased thereafter with the smallest differences between sex at age 80+ years ([1]). In the 29 OECD countries, including Belgium, Islam et al.[7] showed and age and sex difference in excess mortality. In almost all countries, excess mortality rates were higher in men than women, and many countries had lower deaths than expected in children <15 years [7]. The model Islam et al. used to predict the all-cause mortality however ignores historical excess mortality by excluding periods with excess mortality due to a heat wave or seasonal influenza. Recently, Verbeeck et al.[14] suggested a linear mixed model to predict all-cause mortality that does not exclude any historical excess mortality, but down-weighs periods with excess mortality.

In this study, we will expand the LMM by Verbeeck et al.[14] to include age and gender effects to evaluate age and gender specific difference in COVID-19 mortality in Belgium during the year 2020. The results will be compared to a direct age and gender analysis of COVID-19 reported mortality via a negative binomial model and additionally mortality between three different waves are compared.

The data considered and methods applied in this study are discussed in chapter two and chapter three. The results of all the analyses are shown in chapter four. Finally, in chapter five the discussion and conclusion are described, and a brief suggestion for further research is provided.

Chapter 2

Data

2.1 COVID-19 Mortality

Sciensano, the Belgian institute for public health recorded the daily numbers of COVID-19 deaths [4]. These data were grouped in weeks according to international standard ISO 8601 definition, Monday is the first day of the week [8]. Week 53 in 2020 was excluded, since the week is not complete.

The age categories (in years) are 0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, 90+, and these ten categories are used throughout the analyses on COVID-19 data. Missing data were observed in the COVID-19 mortality data.

2.2 All-cause Mortality

All-cause mortality is provided daily by Statbel, the Belgian statistical office [12]. These data were grouped in weeks and used to form age categories for each of the sexes using the same week definition as for the COVID-19 mortality [8]. Data from January 2009-December 2020 are used. The first week of the year 2009 was excluded, since this week was incomplete. Additionally, as not every year has a week 53, these weeks were also excluded. The age (in years) is categorized into 3 groups (0-40, 41-80, 80+) in order to see the age effect below 40 and above 40 years, as well below and above 80 years.

2.3 Belgium Population

Belgium population is provided yearly by Statbel, the Belgian statistical office [12]. Age categories are formed for each of the sexes in 10 year bins (0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, 90+). Population data from 2009-2020 are used.

Chapter 3

Methodology

3.1 Missing Data Mechanisms

In order to take account for incompleteness we need to reflect on the nature of the missing value mechanism and its implications for statistical inference [11]. There are many statistical methods for handling missingness. One of these statistical methods involves imputing the missing observations. In this study, redistribution methods for missing data were used, where missing age and/or gender for reported COVID-19 deaths where randomly sampled to match with the age-sex distribution from historical mortality data [10].

3.2 Statistical Models

Two models are proposed to study the age and gender differences, one for the reported COVID-19 mortality and one for the excess mortality. This section describes the procedures used in building these models.

To select the most parsimonious model Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used. To evaluate age gender differences pairwise comparisons were performed between the multiple mean pairs, using the Bonferroni correction to both the count and continuous data to adjust the level of significance for multiple comparisons.

3.2.1 Negative Binomial Regression Model

A negative binomial regression model was considered to investigate potential differences in COVID-19 reported mortality between age groups and sex. The negative binomial is a Poisson distribution where the Poisson parameter follows a Gamma distribution and that

is usually used to model overdispersed count data.

For a fixed gamma, the negative binomial is a generalized linear model (GLM) [2]. Generalized linear models (GLMs) exists out of three parts: the random component or distributional assumption, systematic component or set of linear predictor, and a link function. The general form of the model is given below:

$$g(\mu_t) = \eta_t \tag{3.1}$$

Where g is the link function that links μ_t to the linear predictors (exploratory variables) η_t. Two negative binomial models were fitted in order to find the most parsimonious model. The full model is shown in Equation 3.2 while the reduced model is shown in Equation 3.3 .

$$\log(\mu_t) = \beta_0 + \sum_{k=1}^{8} \beta_k Age_t + \beta_9 Sex_t + \sum_{k=10}^{17} \beta_k Age_t * Sex_t + \sum_{k=18}^{n} \beta_k Time_t \tag{3.2}$$

$$\log(\mu_t) = \beta_0 + \sum_{k=1}^{8} \beta_k Age_t + \beta_9 Sex_t + \sum_{k=10}^{n} \beta_k Time_t \tag{3.3}$$

Where μ_t is the number of COVID-19 deaths in week t, and n the amount of weeks in a period summed to the k-value.

3.2.2 Linear Mixed Model

Often the weekly average methodology is used to predict all-cause mortality and calculate excess mortality. It is estimated by taking per week the average of the mortality in previous years (2009-2019). But this method ignores autocorrelation and can be influenced by the historical excess mortality. Therefore, a linear mixed model is suggested to estimate the COVID-19 excess mortality by predicting the subject-specific all-cause mortality for the year 2020. Since the mean of the weekly deaths is sufficiently large to invoke the central limit theorem [14], the mean will be approximately normally distributed. Within a linear mixed model, both the random effect structure and the serial correlation process are devices to capture association within units [9]. According to Harvey[6], the linear mixed models require less expertise to fit as compared to time series models (ARIMA), and also results in a smaller variance for the forecasted values.

Since a cyclic pattern was observed in the Belgium mortality (Figure A.7), and a correlogram [14] shows that a yearly cycle is strongly present with a less pronounced half-yearly cycle, both a yearly and half-yearly Fourier series are included into the linear mixed model. Due to the increase in population sizes or changes in the age distribution, the mortality in Belgium fluctuates year by year, therefore a random intercept is included into the model [14]. As there is more variability within the year (Figure A.7) than between years, a random slope was included to the yearly sine wave [14]. The linear mixed model is extended to include age, gender and age-gender interaction effects.

The general form of the model is

$$\begin{aligned} Y_{tj} = (\beta_0 + b_{0j}) + (\beta_1 + b_{1j}) \sin(\frac{2\pi t}{52}) + \beta_2 \cos(\frac{2\pi t}{52}) + \beta_3 \sin(\frac{2\pi t}{26}) + \\ \beta_4 \cos(\frac{2\pi t}{26}) + \sum_{k=5}^{6} \beta_k Age_{tj} + \beta_7 Sex_{tj} + \sum_{k=8}^{9} \beta_k Age_{tj} * Sex_{tj} + \varepsilon_{tj}, \end{aligned} \tag{3.4}$$

where Y_{tj} is the average weekly mortality with week t = 1,2,...,52 by year $j = 2009, ..., 2020$. b_{0j} and b_{1j} are the random effects (b_j), $\varepsilon_{tj} \sim$ N(0, $\sigma^2 I_{nj}$), $b_j \sim$ N(0, D) which are assumed to be independent. For the year 2020, only the first 10 weeks were used for the modelling, while the remaining weeks of year 2020 were forecasted.

Furthermore, to reduce the influence of historical excess mortality that might occur due to heat waves and seasonal influenza, two statistical methods are considered [14]. These methods are the weighted regression [5] and weighted observation [14]. These methods are applied after fitting model (3.4) for the first time.

Based on the standardized residuals (residual$_{tj}$) after fitting model (3.4), a weight is obtained as weight$_{(1)tj}$ = residual$_{tj}^{-2}$, for residual$_{tj} > 1$ [5]. Then, the weighted regression model is applied by fitting model (3.4) with the weight$_{(1)tj}$ [5].

For the second method, the standardized residuals (residual$_{tj}$) after fitting model (3.4) for the first time are also used, but the observations are multiplied with the weight (weight$_{(2)tj}$) to down-weigh the observations:

$$weight_{(2)tj} = 1 - (0.05 * (1 + residual_{tj})) \tag{3.5}$$

Thereafter, model (3.4) is fitted again on the down-weighted observations.

Finally, exponential serial correlation (ε_{2tj}) is added to the measurement error of model (3.4) after taking the influence of historical excess mortality into account. Resulting in the following model:

$$Y_{tj} = (\beta_0 + b_{0j}) + (\beta_1 + b_{1j})\sin(\frac{2\pi t}{52}) + \beta_2\cos(\frac{2\pi t}{52}) + \beta_3\sin(\frac{2\pi t}{26}) + \beta_4\cos(\frac{2\pi t}{26}) + \sum_{k=5}^{6}\beta_k Age_{tj} + \beta_7 Sex_{tj} + \sum_{k=8}^{9}\beta_k Age_{tj} * Sex_{tj} + \varepsilon_{(1)j} + \varepsilon_{(2)j}, \tag{3.6}$$

where $\varepsilon_{(1)j}$ is the measurement error while $\varepsilon_{(2)j}$ is the serial correlation component which follows N(0, $\tau^2 H_j$).

The restricted maximum likelihood (REML) was used to compare the covariance structure with the same mean structure for nested models [11].

To assess the prediction accuracy of the fitted models (3.6) predictions, percentage of the root mean square error (RMSE%) was used, where the predicted error (e_{tj}) is estimated as the difference between the observed mortality (y_{tj}) and the predicted mortality ($\hat{y}_{tj}$). The root mean square error percentage (RMSE%) is then calculated as:

$$RMSE\% = \frac{\frac{\sqrt{\sum e_{tj}^2}}{n}}{\frac{\sum y_{tj}}{n}} \times 100 \tag{3.7}$$

3.3 Mortality rate per million

The mortality rate per million is calculated as the number of deaths per sex and age group divided by the population for that sex and age group, and multiplied by 1000000.

3.4 Model Checking

Formal tests and graphical methods were used to evaluate the adequacy of the fitted models and to check whether necessary remedial measures should be applied. For GLMs, Scaled Deviance (χ^2_D) and Scaled Pearson Chi-Squared χ^2_P statistics were explored to determine whether there is lack of fit or overdispersion. A bar plot is used to compare the observed

and the predicted response variable.

For the linear mixed model, a likelihood ratio test based on a mixture of chi-squared distribution is used to to select which model is appropriate to provide a good fit to the data.

Statistical analyses were performed, and figures produced using SAS studio 3.81 and R studio 4.0.3. For statistical decision making, all hypotheses were tested at 5% level of significance except for the Bonferroni.

Chapter 4

Results

4.1 Data Description

In Belgium, for the first half-year of 2020, i.e the first wave (week 11-week 26), 9624 COVID-19 deaths were reported, 268 in the summer period, i.e second wave (week 27-week 35), and 9754 in the winter period, i.e third wave (week 36-week 52). However, missing data were observed in the COVID-19 mortality data for the first and third waves in Belgium. In the first wave, only for one man the age is unknown, ten deaths have their age but not their sex reported (all 60 years and above), and six deaths have neither age nor sex recorded. Then in the third wave, 5 male and 5 female ages are unknown. Due to the low amount of missingness, it is anticipated that the redistribution method does not influence the obtained results.

4.2 Exploring Reported COVID-19 Mortality For The First, Second, and Third Waves

The histogram plots show count data with overdispersion, indicating that a Poisson distribution cannot account for the variability in the COVID-19 mortality (Figure 4.1 and Figure 4.2). The mean and variance of the data indicates that there is more variability in the data than what can be explained from a Poisson distribution (Table 4.1). Although, in the second wave there seems less evidence for overdispersion (Table 4.1).

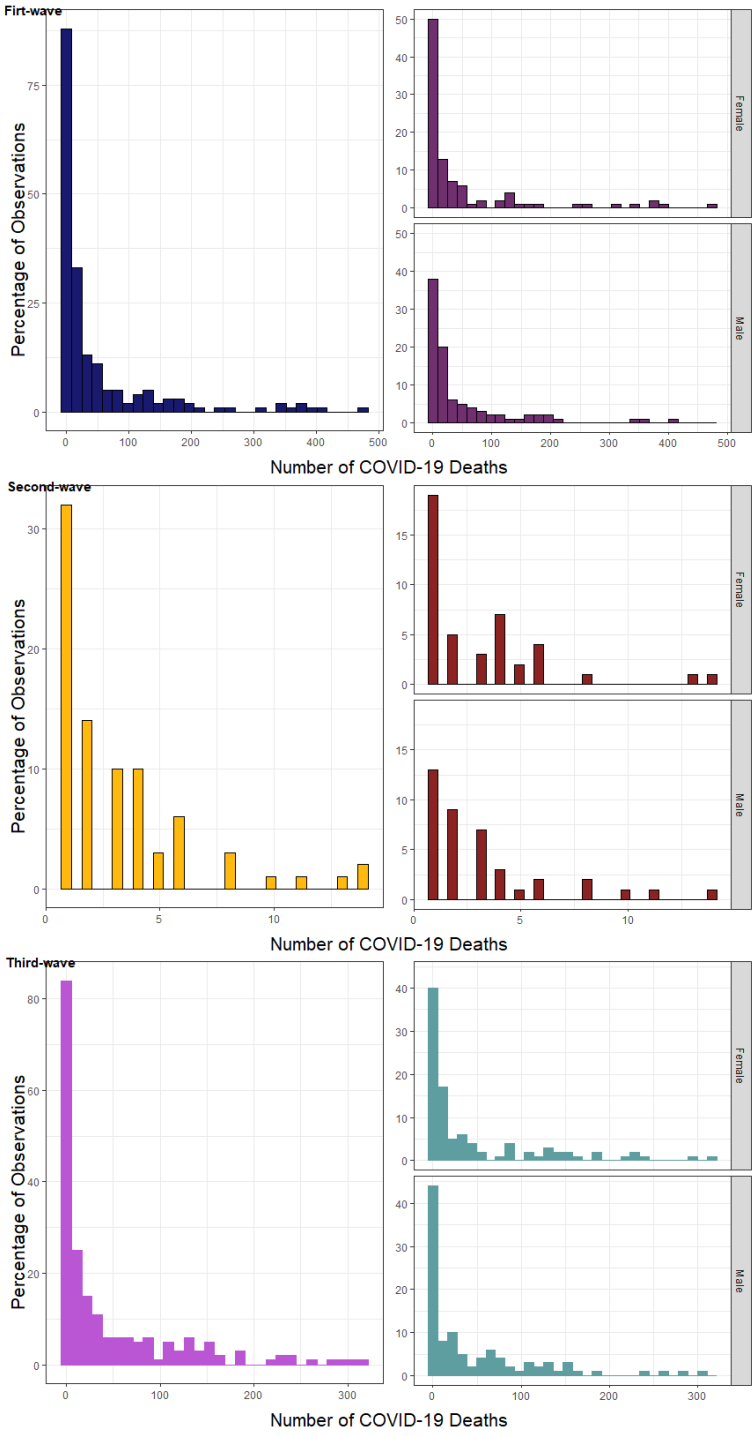

Figure 4.1: Histogram of the number of reported COVID-19 Deaths in Belgium for wave 1, 2, and 3

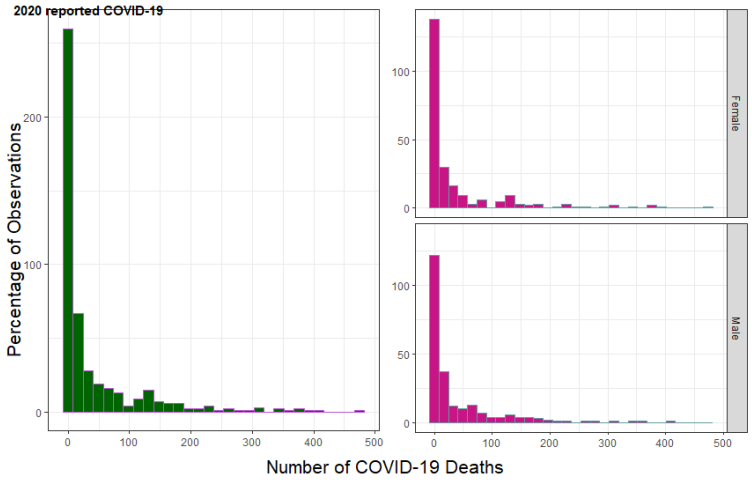

Figure 4.2: Histogram of the number of reported COVID-19 Deaths in Belgium for full year 2020

Table 4.1: Summary Statistics for the full year and weekly COVID-19 Reported Mortality for Wave 1, 2, and 3. "Std. means Standard deviation"

		First-wave (week11-week26)			Second-wave (week27-week35)			Third-wave (week36-week52)	
Variable	Mean	Variance	Std.	Mean	Variance	Std.	Mean	Variance	Std.
Deaths	51.19	8122.04	90.12	3.22	9.07	3.01	45.79	4514.16	67.19
Sex									
Male	49.30	6398.94	79.99	3.30	9.39	3.06	45.64	4125.32	64.23
Female	53.00	9851.33	99.25	3.14	8.98	2.99	45.95	4965.64	70.47

		Full year 2020	
Variable	Mean	Variance	Std.
Deaths	40.81	5523.54	74.32
Sex			
Male	40.21	4640.64	68.12
Female	41.41	6421.56	80.13

4.3 Modeling the reported COVID-19 Mortality Data Using Negative Binomial Model

Using the count data of the number of COVID-19 reported mortality, two negative binomial models were fitted and compared. The corresponding fit statistics of the proposed models is shown in Table 4.2. Taking model fit criteria into account, the model with the smallest Akaike Information Criterion(AIC) as well as Bayesian Information Criterion(BIC) was chosen, which is model 3.2 for the first and third wave and model 3.3 for the second wave.

Table 4.2: Criteria For Assessing Goodness of Fit

	First-wave		Second-wave		Third-wave	
Model	AIC	BIC	AIC	BIC	AIC	BIC
3.2	1102.71	1209.52	313.34	376.23	1183.91	1296.39
3.3	1244.96	1329.11	302.53	348.49	1320.53	1409.85

4.3.0.0.1 First-Wave COVID-19 Reported Mortality

Model (3.2) with age, sex, and the interaction between age and sex as covariates was found to fit well since the scaled Deviance and scaled Pearson Chi-Squared values are not larger than 1 (Table 4.3), and also the predicted values are close to the observed (Figure 4.3). This shows that the overdispersion problem has been taken into account.

Table 4.3: Negative Binomial Regression Model Goodness of Fit Statistics

Scaled Deviance	Value/DF	Scaled Pearson	Value/DF
160.6309	1.0297	183.9788	1.1794

The negative binomial model shows that age groups of 40 years and above are significantly different from the reference category 10-19 years (Table 4.4). Based on the fixed effects, mortality increases with age, but decreases again at 90+ years. There are more female deaths than male in the age groups below 40 and above 80 years old, since the sum of the gender effect (0.6937) with the interaction between age and sex is positive, while there are more male deaths then females in the age groups between 40 and 80 years, since the sum of the gender effect with the age-gender interaction effect is negative (Table 4.4).

Table 4.4: The Parameter estimates of the Negative Binomial Regression Model For the First-wave (week 11-week 26). "Age of the individual was in years, Age*Sex means interaction between Age and Sex, Std.Error means Standard Error and C.I means Confidence interval"

Effect	Group	Parameter	Estimate	Std.Error	Wald 95% C.I	P-value
Intercept	-	β_0	-5.0651	1.0317	[-7.0871;-3.0430]	$< .0001$
Age group (vs 10-19)						
Age	20-29	β_1	0.5035	1.4205	[-2.2806; 3.2875]	0.7230
Age	30-39	β_2	1.4073	1.0595	[-0.6693; 3.4838]	0.1841
Age	40-49	β_3	2.5848	1.0197	[0.5863; 4.5833]	0.0112
Age	50-59	β_4	3.9120	1.0089	[1.9346; 5.8894]	0.0001
Age	60-69	β_5	5.0358	1.0066	[3.0629; 7.0086]	$< .0001$
Age	70-79	β_6	5.8110	1.0060	[3.8394; 7.7826]	< 0.0001
Age	80-89	β_7	6.4681	1.0057	[4.4969; 8.4393]	< 0.0001
Age	90+	β_8	5.7298	1.0042	[3.7615; 7.6981]	< 0.0001
Sex (vs Male)						
Sex	Female	β_9	0.6937	0.0547	[0.5865; 0.8009]	$< .0001$
Age with Sex(vs 10-19 with Male)						
Age*Sex$_F$	20-29	β_{10}	1.1173	1.4214	[-1.6685; 3.9031]	0.4318
Age*Sex$_F$	30-39	β_{11}	-0.2741	0.4474	[-1.1510; 0.6027]	0.5401
Ag*Sex$_F$	40-49	β_{12}	-1.0059	0.2812	[-1.5571; -0.4546]	0.0003
Age*Sex$_F$	50-59	β_{13}	-1.3978	0.1585	[-1.7084; -1.0871]	$< .0001$
Age*Sex$_F$	60-69	β_{14}	-1.4568	0.1065	[-1.6656; -1.2480]	$< .0001$
Age*Sex$_F$	70-79	β_{15}	-1.0844	0.0828	[-1.2467; -0.9221]	$< .0001$
Age*Sex$_F$	80-89	β_{16}	-0.5633	0.0727	[-0.7058; -0.4207]	$< .0001$
Age*Sex$_F$	90+	β_{17}	-0.5828	0.0930	[-0.6861; -0.4704]	$< .0001$
Time (vs Time 11)						
Time	12	β_{18}	2.3029	0.2443	[1.8240; 2.7818]	$< .0001$
Time	13	β_{19}	3.6821	0.2364	[3.2187; 4.1454]	$< .0001$
Time	14	β_{20}	4.3713	0.2349	[3.9109; 4.8318]	$< .0001$
Time	15	β_{21}	4.5616	0.2347	[4.1016; 5.0216]	$< .0001$
Time	16	β_{22}	4.3964	0.2349	[3.9360; 4.8569]	$< .0001$
Time	17	β_{23}	4.0678	0.2354	[3.6063; 4.5292]	$< .0001$
Time	18	β_{24}	3.5443	0.2367	[3.0804; 4.0082]	$< .0001$
Time	19	β_{25}	3.2899	0.2376	[2.8242; 3.7556]	$< .0001$
Time	20	β_{26}	2.7506	0.2404	[2.2794; 3.2218]	$< .0001$
Time	21	β_{27}	2.4586	0.2427	[1.9828; 2.9343]	$< .0001$
Time	22	β_{28}	2.0843	0.2467	[1.6008; 2.5678]	$< .0001$
Time	23	β_{29}	1.7430	0.2520	[1.2491; 2.2370]	$< .0001$
Time	24	β_{30}	1.0713	0.2681	[0.5458; 1.5969]	$< .0001$
Time	25	β_{31}	0.7737	0.2791	[0.2266; 1.3209]	0.0056
Time	26	β_{32}	0.6970	0.2835	[0.1414; 1.2525]	0.0139

Furthermore, pairwise comparisons were performed between age groups above 50-59 years for both females and males. These comparisons were performed in order to see if there is any statistical difference between age groups in females and males. It is shown that, significant differences exist between age groups in females and males (Table 4.5), mortality increases with age in both males and females until 80-89 years and then decreases (Table 4.5).

Table 4.5: Bonferrori Pairwise Comparisons for Age group in Females and Males."Std.Error means Standard Error"

Label	Estimate	Std.Error	95% Confidence Limits	P-value
50-59 vs 60-69 Sex=Female	-1.0647	0.1409	[-1.515;-0.6143]	< .0001
50-59 vs 70-79 Sex=Female	-2.2124	0.1289	[-2.625;-1.8002]	< .0001
50-59 vs 80-89 Sex=Female	-3.3906	0.1247	[-3.789;-2.9919]	< .0001
50-59 vs 90+ Sex=Female	-3.2156	0.1251	[-3.615;-2.8157]	< .0001
60-69 vs 70-79 Sex=Female	-1.1477	0.0873	[-1.427;-0.8687]	< .0001
60-69 vs 80-89 Sex=Female	-2.3258	0.0808	[-2.584;-2.0674]	< .0001
60-69 vs 90+ Sex=Female	-2.1508	0.0814	[-2.411;-1.8905]	< .0001
70-79 vs 80-89 Sex=Female	-1.1782	0.0574	[-1.362;-0.9947]	< .0001
70-79 vs 90+ Sex=Female	-1.0032	0.0582	[-1.189;-0.8170]	< .0001
80-89 vs 90+ Sex=Female	0.7383	0.0543	[0.5650;0.9120]	< .0001
50-59 vs 60-69 Sex=Male	-1.1237	0.1024	[-1.451;-0.7962]	< .0001
50-59 vs 70-79 Sex=Male	-1.8990	0.0964	[-2.207;-1.5908]	< .0001
50-59 vs 80-89 Sex=Male	-2.5561	0.0937	[-2.856;-2.2565]	< .0001
50-59 vs 90+ Sex=Male	-1.8178	0.0969	[-2.127;-1.5081]	< .0001
60-69 vs 70-79 Sex=Male	-0.7752	0.0678	[-0.992;-0.5585]	< .0001
60-69 vs 80-89 Sex=Male	-1.4323	0.0639	[-1.637;-1.2282]	< .0001
60-69 vs 90+ Sex=Male	-0.6940	0.0684	[-0.913;-0.4752]	< .0001
70-79 vs 80-89 Sex=Male	-0.6571	0.0535	[-0.828;-0.4860]	< .0001
70-79 vs 90+ Sex=Male	0.0812	0.0590	[-0.107;0.2697]	0.9065
80-89 vs 90+ Sex=Male	0.1750	0.0479	[0.0220;0.3280]	0.0078

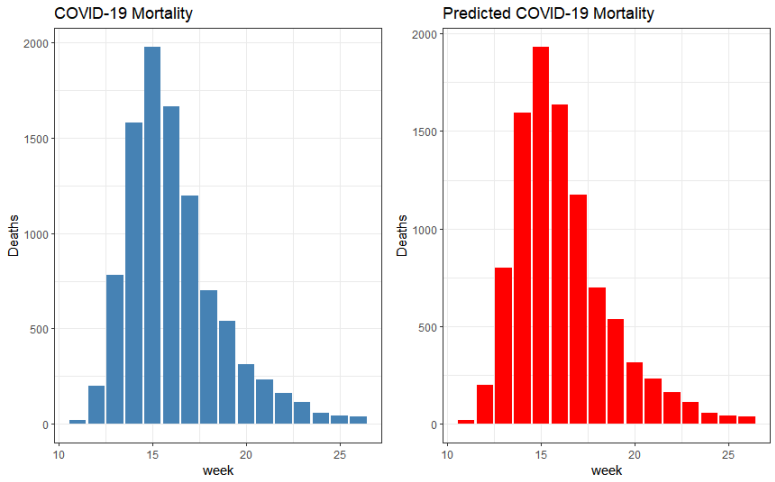

Figure 4.3: Comparing Observed COVID-19 Mortality with the Predicted COVID-19 For the First-wave

4.3.0.0.2 Second-Wave COVID-19 Reported Mortality

The negative binomial regression model fit the data well despite there is less evidence

for overdispersion (Table 4.6) and the predicted values coincide well with the observed (Figure 4.4).

Table 4.6: Negative Binomial Regression Model Goodness of Fit Statistics

Scaled Deviance	Value/DF	Scaled Pearson	Value/DF
32.1649	0.5546	31.7615	0.5476

None of the covariates are significant except the weeks during the heat wave (week 33-34), where mortality was significantly increased compared to the remaining weeks (Table 4.7). There is no age and gender effect (Table 4.7).

Table 4.7: The Parameter estimates of the Negative Binomial Regression Model For the Second-wave (week 27-week 35). "Age of the individual was in years, Age*Sex means interaction between Age and Sex, Std.Error means Standard Error and C.I means Confidence interval"

Effect	Group	Parameter	Estimate	Std.Error	Wald 95% C.I	P-value
Intercept	-	β_0	0.7015	1.0675	[-1.3908;2.7938]	0.5111
Age group (vs 10-19)						
Age	20-29	β_1	-1.2665	1.2635	[-3.7428; 1.2099]	0.3162
Age	30-39	β_2	-0.4033	1.2001	[-2.7555; 1.9489	0.7368
Age	40-49	β_3	-0.8873	1.1375	[-3.1168; 1.3422]	0.4354
Age	50-59	β_4	-0.9272	1.0824	[-3.0487; 1.1944]	0.3917
Age	60-69	β_5	-0.3876	1.0679	[-2.4807; 1.7055]	0.7166
Age	70-79	β_6	0.2533	1.0501	[-1.8048; 2.3114]	0.8094
Age	80-89	β_7	0.8661	1.0457	[-1.1834; 2.9156]	0.4075
Age	90+	β_8	0.5887	1.0472	[-1.4639; 2.6412]	0.5740
Sex (vs Male)						
SEX	Female	β_9	-0.0456	0.1229	[-0.2865; 0.1953]	0.7105
Time (vs Time 27)						
Time	28	β_{10}	-0.6559	0.3522	[-1.3462; 0.0345]	0.0626
Time	29	β_{11}	-0.1301	0.3060	[-0.7299 ; 0.4697]	0.6708
Time	30	β_{12}	-0.3009	0.3351	[-0.9578; 0.3559]	0.3692
Time	31	β_{13}	-0.1922	0.3002	[-0.7805 ; 0.3961]	0.5220
Time	32	β_{14}	0.2142	0.2732	[-0.3213; 0.7497]	0.4331
Time	33	β_{15}	1.0167	0.2320	[0.5620; 1.4713]	$< .0001$
Time	34	β_{16}	0.5851	0.2534	[0.0885 ; 1.0818]	0.0209
Time	35	β_{17}	0.0332	0.2767	[-0.5092; 0.5756]	0.9045

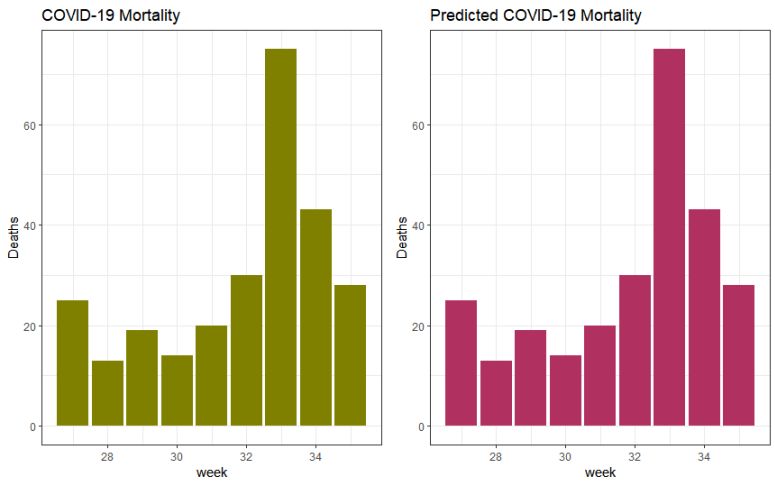

Figure 4.4: Comparing Observed COVID-19 Mortality with the Predicted COVID-19 For the Second-wave

4.3.0.0.3 Third-Wave COVID-19 Reported Mortality

The negative binomial regression model fits the data well (Table 4.8) and predicted mortality coincides with observed (Figure 4.5).

Table 4.8: Negative Binomial Regression Model Goodness of Fit Statistics

Scaled Deviance	Value/DF	Scaled Pearson	Value/DF
210.0017	1.2426	225.0492	1.3317

The negative binomial model shows that age groups of 40 years and above are significantly different from the reference category 10-19 years (Table 4.9). Based on the fixed effects, mortality increases with age, but decreases again at 90+ years. There are more female deaths than male in the age groups below 40 years, since the sum of the gender effect (0.5214) with the interaction between age and sex is positive, while above 40 years it is the reverse, since the sum of the gender effect with the interaction between age and sex is negative (Table 4.9).

Table 4.9: The Parameter estimates of the Negative Binomial Regression Model For the Third-wave (week 36-week 52). "Age of the individual was in years, Age*Sex means interaction between Age and Sex, Std.Error means Standard Error and C.I means Confidence interval"

Effect	Group	Parameter	Estimate	Std.Error	Wald 95% C.I	P-value
Intercept	-	β_0	-4.4433	0.7462	[-5.9059; -2.9807]	$< .0001$
Age group (vs 10-19)						
Age	20-29	β_1	0.5714	0.9155	[-1.2228; 2.3657]	0.5325
Age	30-39	β_2	0.9027	0.7843	[-0.6346; 2.4400]	0.2498
Age	40-49	β_3	2.0154	0.7277	[0.5891; 3.4417]	0.0056
Age	50-59	β_4	3.3397	0.7144	[1.9395; 4.7399]	$< .0001$
Age	60-69	β_5	4.5229	0.7113	[3.1288; 5.9169]	$< .0001$
Age	70-79	β_6	5.2843	0.7105	[3.8917; 6.6769]	$< .0001$
Age	80-89	β_7	5.9260	0.7102	[4.5340; 7.3180]	$< .0001$
Age	90+	β_8	5.1521	0.7088	[3.7628; 6.5415]	$< .0001$
Sex (vs Male)						
Sex	Female	β_9	0.5214	0.0456	[0.4320; 0.6108]	$< .0001$
Age with Sex(vs 10-19 with Male)						
Age*Sex$_F$	20-29	β_{10}	-0.4291	1.1576	[-2.6980; 1.8397]	0.7109
Age*Sex$_F$	30-39	β_{11}	-0.2388	0.4747	[-1.1692; 0.6916]	0.6149
Age*Sex$_F$	40-49	β_{12}	-1.0469	0.2677	[-1.5715; -0.5222]	$< .0001$
Age*Sex$_F$	50-59	β_{13}	-1.1702	0.1465	[-1.4573; -0.8832]	$< .0001$
Age*Sex$_F$	60-69	β_{14}	-1.2811	0.0931	[-1.4636; -1.0986]	$< .0001$
Age*Sex$_F$	70-79	β_{15}	-1.0234	0.0704	[-1.1613; -0.8855]	$< .0001$
Age*Sex$_F$	80-89	β_{16}	-0.5508	0.0587	[-0.6658; -0.4358]	$< .0001$
Age*Sex$_F$	90+	β_{17}	-0.8091	0.0664	[-1.0105; -0775]	$< .0001$
Time (vs Time 36)						
Time	37	β_{18}	0.1521	0.3257	[-0.4862; 0.7905]	0.6404
Time	38	β_{19}	0.3556	0.2984	[-0.2292; 0.9404]	0.2334
Time	39	β_{20}	0.7312	0.2799	[0.1826; 1.2797]	0.0090
Time	40	β_{21}	1.4480	0.2554	[0.9475; 1.9486]	$< .0001$
Time	41	β_{22}	1.9670	0.2456	[1.4857; 2.4483]	$< .0001$
Time	42	β_{23}	2.4710	0.2399	[2.0009; 2.9412]	$< .0001$
Time	43	β_{24}	3.2190	0.2352	[2.7581; 3.6799]	$< .0001$
Time	44	β_{25}	3.9644	0.2329	[3.5079; 4.4209]	$< .0001$
Time	45	β_{26}	4.2506	0.2324	[3.7951; 4.7061]	$< .0001$
Time	46	β_{27}	4.1886	0.2325	[3.7330; 4.6443]	$< .0001$
Time	47	β_{28}	4.0314	0.2328	[3.5751; 4.4876]	$< .0001$
Time	48	β_{29}	3.7796	0.2334	[3.3221; 4.2370]	$< .0001$
Time	49	β_{30}	3.5576	0.2340	[3.0991; 4.0162]	$< .0001$
Time	50	β_{31}	3.4993	0.2341	[3.0405; 3.9582]	$< .0001$
Time	51	β_{32}	3.4773	0.2343	[3.0182; 3.9364]	$< .0001$
Time	52	β_{33}	3.2777	0.2349	[2.8172; 3.7381]	$< .0001$

Furthermore, pairwise comparisons were performed between age groups above 50-59 years for both females and males. These comparisons were performed in order to see if there is any statistical difference between age groups in females and males. It is shown that, significant differences exist between age groups in females and males (Table 4.10). Mortality increases with increasing age in both genders until 80-89 years and then decreases (Table 4.10).

Table 4.10: Bonferrori Pairwise Comparisons for Age group in Females and Males."Std.Error means Standard Error"

Label	Estimate	Std.Error	95% Confidence Limits	P-value
50-59 vs 60-69 Sex=Female	-1.072	0.1310	[-1.4910 ; -0.654]	< .0001
50-59 vs 70-79 Sex=Female	-2.091	0.1202	[-2.4756; -1.707]	< .0001
50-59 vs 80-89 Sex=Female	-3.206	0.1158	[-3.5759 ; -2.835]	< .0001
50-59 vs 90+ Sex=Female	-2.983	0.1164	[-3.3547 ; -2.611]	< .0001
60-69 vs 70-79 Sex=Female	-1.019	0.0785	[-1.2700; -0.768]	< .0001
60-69 vs 80-89 Sex=Female	-2.133	0.0716	[-2.3624 ; -1.904]	< .0001
60-69 vs 90+ Sex=Female	-1.910	0.0725	[-2.1421; -1.679]	< .0001
70-79 vs 80-89 Sex=Female	-1.114	0.0492	[-1.2714 ; -0.957]	< .0001
70-79 vs 90+ Sex=Female	-0.891	0.0504	[-1.0524 ; -0.730]	< .0001
80-89 vs 90+ Sex=Female	0.774	0.0439	[0.6334 ; 0.9140]	< .0001
50-59 vs 60-69 Sex=Male	-1.183	0.0939	[-1.4835 ; -0.883]	< .0001
50-59 vs 70-79 Sex=Male	-1.945	0.0882	[-2.2265 ; -1.663]	< .0001
50-59 vs 80-89 Sex=Male	-2.586	0.0857	[-2.8601; -2.312]	< .0001
50-59 vs 90+ Sex=Male	-1.812	0.0889	[-2.0968 ; -1.528]	< .0001
60-69 vs 70-79 Sex=Male	-0.761	0.0573	[-0.9446 ; -0.578]	< .0001
60-69 vs 80-89 Sex=Male	-1.403	0.0533	[-1.5737 ; -1.233]	< .0001
60-69 vs 90+ Sex=Male	-0.629	0.0585	[-0.8162 ; -0.442]	< .0001
70-79 vs 80-89 Sex=Male	-0.642	0.0424	[-0.7772 ; -0.506]	< .0001
70-79 vs 90+ Sex=Male	0.132	0.0487	[-0.0235 ; 0.288]	0.1424
80-89 vs 90+ Sex=Male	0.223	0.0389	[0.0987 ; 0.3470]	< .0001

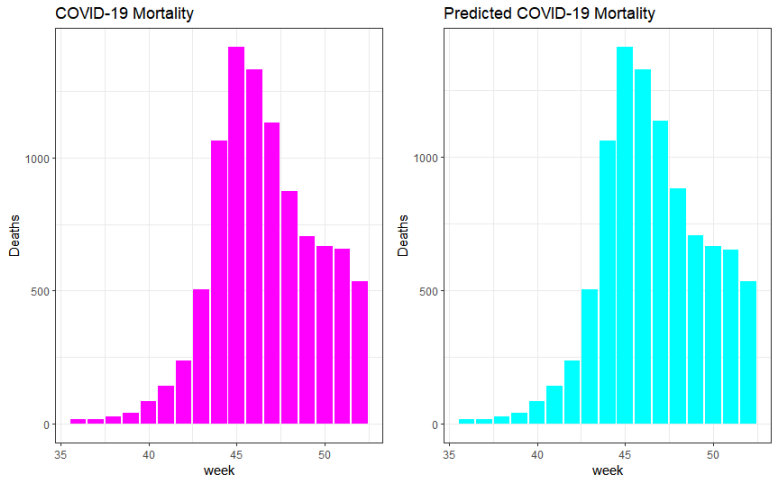

Figure 4.5: Comparing Observed COVID-19 Mortality with the Predicted COVID-19 For the Third-wave

4.3.0.0.4 Full year 2020 COVID-19 Reported Mortality

The negative binomial regression model fits the data well (Table 4.11) and predicted mor-

tality coincides with observed (Figure 4.6).

Table 4.11: Negative Binomial Regression Model Goodness of Fit Statistics

Scaled Deviance	Value/DF	Scaled Pearson	Value/DF
457.2628	1.1018	04.3624	1.2973

The negative binomial model shows that age groups of 40 years and above are significantly different from the reference category 10-19 years (Table 4.12). Based on the fixed effects, mortality increases with age, but decreases again at 90+ years. There are more female deaths than male in the age group below 40 and above 80 years old, since the sum of the gender effect (0.5920) with the interaction between age and sex is positive, while there are more male deaths then females in the age groups between 40 and 80 years, since the sum of the gender effect with the age-gender interaction effect is negative (Table 4.12).

Table 4.12: The Parameter estimates of the Negative Binomial Regression Model For the full year COVID-19 Reported Mortality. "Age of the individual was in years, Age*Sex means interaction between Age and Sex and Std.Error means Standard Error"

Effect	Group	Parameter	Estimate	Std.Error	Wald 95% C.I	P-value
Intercept	-	β_0	-4.3934	0.5553	[-5.4818 ; −3.3051]	< .0001
Age group (vs 10-19)						
Age	20-29	β_1	0.5245	0.6758	[-0.8000 ; 1.8489]	0.4377
Age	30-39	β_2	0.9046	0.5523	[-0.1778 ; 1.9870]	0.1014
Age	40-49	β_3	1.9841	0.5175	[0.9699 ; 2.9984]	0.0001
Age	50-59	β_4	3.2800	0.5079	[2.2844 ; 4.2755]	< .0001
Age	60-69	β_5	4.4305	0.5057	[3.4393 ; 5.4217]	< .0001
Age	70-79	β_6	5.1939	0.5052	[4.2038 ; 6.1841]	< .0001
Age	80-89	β_7	5.8423	0.5050	[4.8526 ; 6.8320]	< .0001
Age	90+	β_8	5.0928	0.5035	[4.1060 ; 6.0797]	< .0001
Sex(vs Male)						
Sex	Female	β_9	0.5920	0.0384	[0.5167 ; 0.6672]	< .0001
Age with Sex (vs 10-19 with Male)						
Age*Sex$_F$	20-29	β_{10}	0.2285	0.7362	[-1.2145 ; 1.6715]	0.7563
Age*Sex$_F$	30-39	β_{11}	-0.2954	0.3136	[-0.9102 ; 0.3193]	0.3463
Age*Sex$_F$	40-49	β_{12}	-1.0161	0.1915	[-1.3913; -0.6409]	< .0001
Age*Sex$_F$	50-59	β_{13}	-1.2454	0.1086	[-1.4582; -1.0325]	< .0001
Age*Sex$_F$	60-69	β_{14}	-1.3444	0.0733	;-1.4882 ; -1.2007]	< .0001
Age*Sex$_F$	70-79	β_{15}	-1.0352	0.0580	[-1.1488 ; -0.9216]	< .0001
Age*Sex$_F$	80-89	β_{16}	-0.5510	0.0507	[-0.6504 ; -0.4515]	< .0001
Age*Sex$_F$	90+	β_{17}	-0.5708	0.0486	[-0.6233; -0.4481]	< .0001
Time (vs Time 11)						
Time	12	β_{18}	2.2984	0.2447	[1.8188 ; 2.7779]	< .0001
Time	13	β_{19}	3.6710	0.2367	[3.2070 ; 4.1349]	< .0001
Time	14	β_{20}	4.3547	0.2352	[3.8937 ; 4.8158]	< .0001
Time	15	β_{21}	4.5459	0.2350	[4.0854 ; 5.0065]	< .0001
Time	16	β_{22}	4.3836	0.2352	[3.9226 ; 4.8446]	< .0001
Time	17	β_{23}	4.0577	0.2357	[3.5956 ; 4.5197]	< .0001
Time	18	β_{24}	3.5373	0.2370	[3.0727 ; 4.0018]	< .0001
Time	19	β_{25}	3.2832	0.2380	[2.8168 ; 3.7496]	< .0001
Time	20	β_{26}	2.7459	0.2408	[2.2740 ; 3.2178]	< .0001
Time	21	β_{27}	2.4535	0.2431	[1.9770 ; 2.9300]	< .0001
Time	22	β_{28}	2.0794	0.2470	[1.5952 ; 2.5636]	< .0001
Time	23	β_{29}	1.7398	0.2524	[1.2451 ; 2.2344]	< .0001
Time	24	β_{30}	1.0677	0.2685	[0.5415 ; 1.5940]	< .0001
Time	25	β_{31}	0.7713	0.2795	[0.2235 ; 1.3191]	0.0058
Time	26	β_{32}	0.6951	0.2838	[0.1389 ; 1.2513]	0.0143
Time	27	β_{33}	0.2431	0.3080	[-0.3607 ; 0.8468]	0.4301
Time	28	β_{34}	-0.3505	0.3632	[-1.0624 ; 0.3614]	0.3346
Time	29	β_{35}	0.1963	0.3281	[-0.4469 ; 0.8395]	0.5497
Time	30	β_{36}	−0.1404	0.3560	[-0.8380 ; 0.5573]	0.6933
Time	31	β_{37}	0.0244	0.3239	[-0.6104 ; 0.6592]	0.9400
Time	32	β_{38}	0.4856	0.2972	[-0.0970 ; 1.0681]	0.1023
Time	33	β_{39}	1.3206	0.2611	[0.8089 ; 1.8323]	< .0001
Time	34	β_{40}	0.8439	0.2797	[0.2957 ; 1.3922]	0.0026
Time	35	β_{41}	0.3391	0.3009	[-0.2507 ; 0.9290]	0.2598
Time	36	β_{42}	−0.0019	0.3280	[-0.6448 ; 0.6409]	0.9953
Time	37	β_{43}	0.1519	0.3284	[-0.4918 ; 0.7956]	0.6438
Time	38	β_{44}	0.3523	0.3010	[-0.2376 ; 0.9423]	0.2418
Time	39	β_{45}	0.7282	0.2827	[0.1741 ; 1.2822]	0.0100
Time	40	β_{46}	1.4487	0.2584	[0.9423 ; 1.9551]	< .0001
Time	41	β_{47}	1.9620	0.2486	[1.4747 ; 2.4493]	< .0001
Time	42	β_{48}	2.4664	0.2430	[1.9902 ; 2.9426]	< .0001
Time	43	β_{49}	3.2113	0.2382	[2.7444 ; 3.6781]	< .0001
Time	44	β_{50}	3.9648	0.2359	[3.5023 ; 4.4272]	< .0001
Time	45	β_{51}	4.2475	0.2354	[3.7861 ; 4.7088]	< .0001
Time	46	β_{52}	4.1858	0.2355	[3.7242 ; 4.6473]	< .0001
Time	47	β_{53}	4.0364	0.2358	[3.5742; 4.4986]	< .0001
Time	48	β_{54}	3.7887	0.2364	[3.3253 ; 4.2521]	< .0001
Time	49	β_{55}	3.5639	0.2370	[3.0994 ; 4.0284]	< .0001
Time	50	β_{56}	3.4941	0.2372	[3.0293 ; 3.9590]	< .0001
Time	51	β_{57}	3.4632	0.2372	[2.9982 ; 3.9282]	< .0001
Time	52	β_{58}	3.2653	0.2380	[2.7989 ; 3.7317]	< .0001

Furthermore, pairwise comparisons were performed between age groups above 50-59 years for both females and males. These comparisons were performed in order to see if there is any statistical difference between age groups in females and males. It is shown that, significant differences exist between age groups in females and male (Table 4.13). Mortality increases with increasing age in both genders until 80-89 years and then decreases (Table 4.13).

Table 4.13: Bonferrori Pairwise Comparisons for Age group in Females and Males."Std.Error means Standard Error

Label	Estimate	Std.Error	95% Confidence Limits	P-value
50-59 vs 60-69 Sex=Female	-1.051	0.0958	[-1.3577 ; -0.7452]	< 0.0001
50-59 vs 70-79 Sex=Female	-2.124	0.0879	[-2.4052; -1.8430]	< 0.0001
50-59 vs 80-89 Sex=Female	-3.257	0.0849	[-3.5281 ; -2.9853]	< 0.0001
50-59 vs 90+ Sex=Female	-3.058	0.0852	[-3.3307 ; -2.7858]	< 0.0001
60-69 vs 70-79 Sex=Female	-1.073	0.0601	[-1.2648 ; -0.8806]	< 0.0001
60-69 vs 80-89 Sex=Female	-2.205	0.0555	[-2.3828 ; -2.0277]	< 0.0001
60-69 vs 90+ Sex=Female	-2.007	0.0560	[-2.1859 ; -1.8276]	< 0.0001
70-79 vs 80-89 Sex=Female	-1.133	0.0404	[-1.2617 ; -1.0035]	< 0.0001
70-79 vs 90+ Sex=Female	-0.934	0.0411	[-1.0654 ; -0.8028]	< 0.0001
80-89 vs 90+ Sex=Female	0.749	0.0377	[0.6290 ; 0.8699]	< 0.0001
50-59 vs 60-69 Sex=Male	-1.150	0.0707	[-1.3765 ; -0.9245]	< 0.0001
50-59 vs 70-79 Sex=Male	-1.914	0.0666	[-2.1269 ; -1.7011]	< 0.0001
50-59 vs 80-89 Sex=Male	-2.562	0.0648	[-2.7695; -2.3552]	< 0.0001
50-59 vs 90+ Sex=Male	-1.813	0.0670	[-2.0271 ; -1.5986]	< 0.0001
60-69 vs 70-79 Sex=Male	-0.763	0.0466	[-0.9125 ; -0.6144]	< 0.0001
60-69 vs 80-89 Sex=Male	-1.412	0.0440	[-1.5524 ; -1.2712]	< 0.0001
60-69 vs 90+ Sex=Male	-0.662	0.0472	[-0.8133 ; -0.5114]	< 0.0001
70-79 vs 80-89 Sex=Male	-0.648	0.0369	[-0.7664 ; -0.5303]	< 0.0001
70-79 vs 90+ Sex=Male	0.101	0.0408	[-0.0292 ; 0.2314]	0.2414
80-89 vs 90+ Sex=Male	0.198	0.0339	[0.0900 ; 0.3070]	< 0.0001

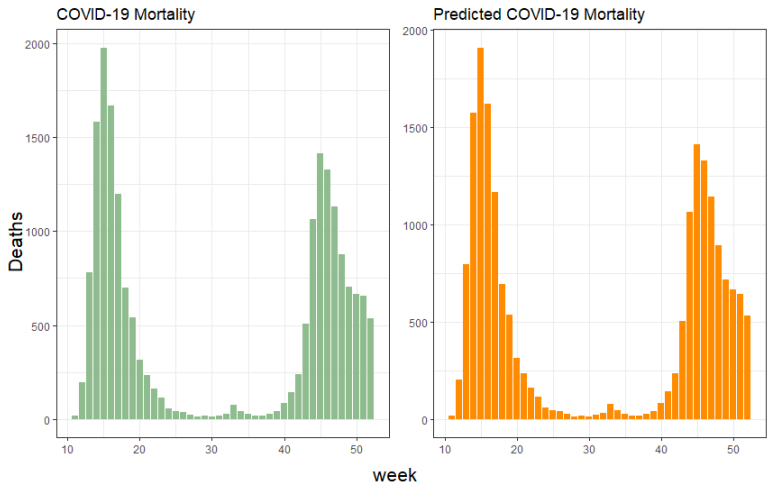

Figure 4.6: Comparing Observed COVID-19 Mortality with the Predicted COVID-19 For the Full year 2020

4.4 Exploring All-cause Mortality For The First, Second, and Third Waves

For each wave, the COVID-19 reported mortality is compared to the excess mortality based on the weekly average predictions of all-cause mortality. When the number of all-cause deaths in 2020 (blue line) exceeds the upper or lower predicted mortality limits (black line), there is respectively excess mortality or lower than average mortality (Figure 4.7). The observed excess mortality coincides with the reported COVID-19 mortality in the first and third waves, but not in the second-wave. This is the result of the heat wave that occurred in the weeks 33-34 in 2020 (Figure 4.7). Also, it can be seen that the peak of excess mortality occurred mostly in the older age groups (Figure A.1, Figure A.2, Figure A.3, Figure A.4, Figure A.5, and Figure A.6).

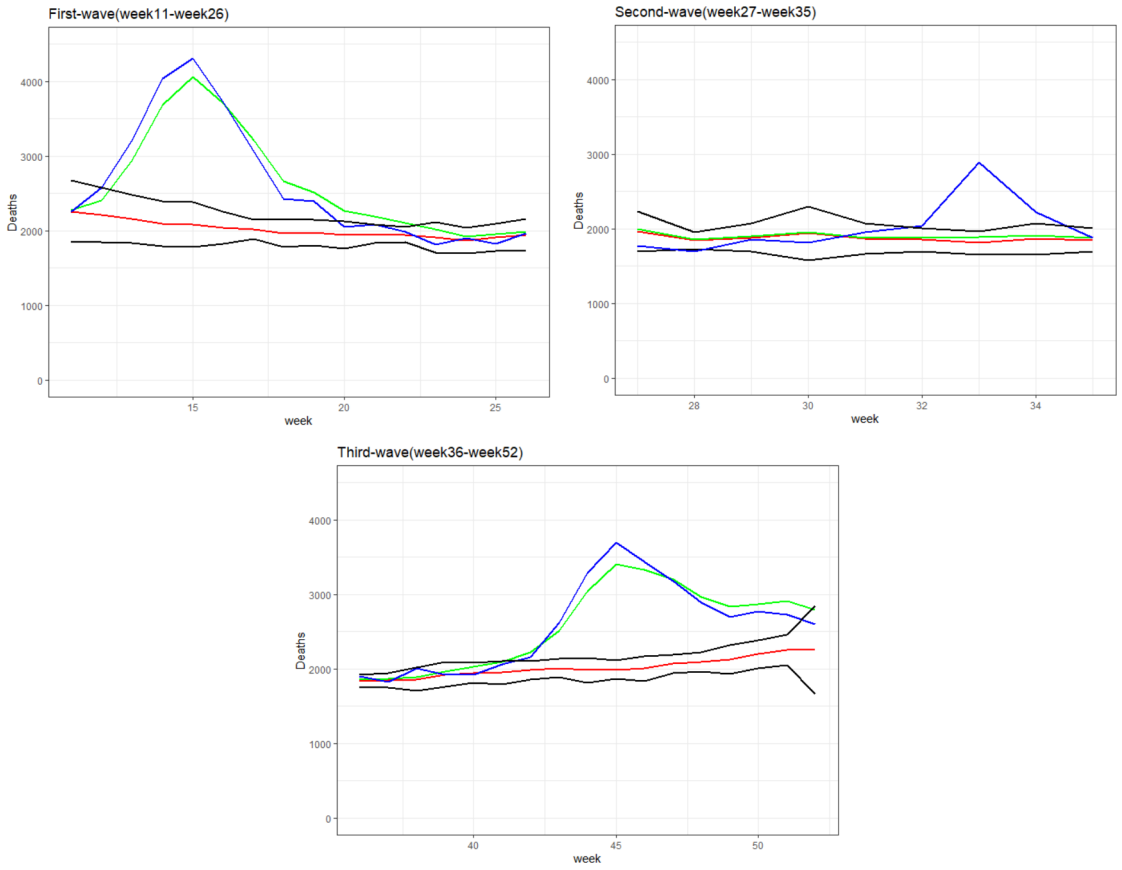

Figure 4.7: Weekly observed all-cause mortality (blue), predicted all-cause mortality (red), with lower and upper bound (black) and reported COVID-19 mortality combined (green) for the three waves in 2020 in Belgium.

4.5 Age Gender Difference Based on Excess Mortality

4.5.1 Predicting Excess Mortality Using Linear Mixed Model

Since the mean of the weekly all-cause deaths is sufficiently large to invoke the central limit theorem, a linear mixed model was fitted to predict mortality in the year 2020. The model with both random intercept and random slope was fitted and was compared to the model with only a random intercept, a model with only a random slope and a model without random intercept and random slope.

The model with both random intercept and random slope is most appropriate at 5% level of significance (Table 4.14).

Table 4.14: Likelihood ratio test for the need of random effect model

Hypothesis	$-2ln(\lambda_N)$	Asymptotic H_0	P-value
random intercept and slope vs Only random intercept	53.8	$\chi^2_{1:2}$	< 0.0001
random intercept and slope vs.Only random slope	76.7	$\chi^2_{1:2}$	< 0.0001
random intercept and slope vs. No random effect	124.6	$\chi^2_{0:2}$	< 0.0001

Model 3.4 with the weighted regression and weighted observation approach was fitted to the data and it was tested if adding serial correlation to the measurement error improved the fit (model 3.6). It can be seen that adding exponential serial correlation to both strategies; weighted regression and weighted observations, is important and also improves the models significantly (Table 4.15). Other serial correlation structure either did not converge or did not improve the fit.

The prediction accuracy was evaluated only in the years with little or no excess mortality, since the fitted models down-weigh the past excess deaths. In the year 2014 in Belgium, no excess deaths were observed due to heat wave or influenza, while in the year 2016 there was little excess mortality due to a heat wave and in the year 2018 excess mortality due to heat wave was very low (Figure A.8, Figure A.9, Figure A.10). The model prediction accuracy is evaluated by not including mortality from week 11 to week 52 for the years 2014, 2016 and 2018 each in turn.

Based on the 5 years weekly average mortality, the year 2014 was predicted by using the years 2009-2013, the years 2011-2015 to predict year 2016, and the years 2013-2017 to predict year 2018. There are very small differences in accuracy of the prediction. The weighted observations model shows a slightly better predicting accuracy for the year 2014, while the weighted regression model and the weekly average shows a slightly better predicting accuracy for the years 2016, and 2018 (Table 4.15). By comparing prediction based on the linear mixed model and the weekly average, it seems the linear mixed model improves

the predication accuracy only in 2014 (Table 4.15).

Table 4.15: REML log-likelihood for the fitted models, forecasting accuracy, excess mortality estimation (95% CI) and reported COVID-19 mortality. "ll means log-likelihood and RMSE means root mean square error"

			Weighted regression					Weighted observations		
Model	-2*ll*	RMSE% 2014	RMSE% 2016	RMSE% 2018	Excess Mortality (2020)	-2*ll*	RMSE% 2014	RMSE% 2016	RMS% 2018	Excess Mortality (2020)
3.4	31989.7	0.594	0.698	0.887	22586[4011;41162]	30835.3	0.596	0.708	0.881	23543[7324;39761]
3.6	31352.5	0.630	0.711	0.909	25441[8081;42801]	30814.7	0.589	0.704	0.878	22707[6388;39027]
			Weekly Average					COVID-19 Mortality		
		RMSE% 2014	RMSE% 2016	RMSE% 2018	Excess Mortality (2020)			19345[13783;24907]		
		0.680	0.707	0.861	19957[11638;28275]					

Based on the fixed effects, there is a positive association of age on the all-cause mortality (Table 4.16). There are less female deaths than male from age groups below 80 years old, since the sum of the gender effect (-15.52) with the interaction between age and sex are negative, while for the age group above 80 years the gender effect is the reverse (Table 4.16).

The random intercept variance is higher than the random slope variance in both models, meaning that the all-cause mortality was highly different between years. The estimated residual variance in both models (1069.93 and 1009.64) is very large compared to the random intercept variance (d_{11}) implying that there is more variability within than between a year (Table 4.16).

Table 4.16: Parameter Estimates and associated standard errors for the Linear Mixed Model with weighted regression and weighted observations and Estimates for the Variance Components. "Age are in years".

			Weighted Regression		Weighted Observations	
Effect	Group	Parameter	Estimate(Std. error)	P-value	Estimate(Std. error)	P-value
Intercept		β_0	-107.80 (10.74)	< 0.0001	32.37(1.56)	< 0.0001
sine	Full year	β_1	19.70(2.26)	< 0.0001	18.83(1.67)	< 0.0001
cosine	Full year	β_2	23.90 (1.08)	< 0.0001	24.65(1.25)	< 0.0001
sine	Half year	β_3	5.5657 (1.20)	< 0.0001	6.65(0.98)	< 0.0001
cosine	Half year	β_4	1.0812 (1.33)	0.4183	1.35(1.44)	0.3473
Age group (vs 0-40)						
Age	41-80	β_5	512.74 (3.54)	< 0.0001	511.07(3.02)	< 0.0001
Age	80+	β_6	384.46 (8.97)	< 0.0001	386.61(9.01)	< 0.0001
Sex (vs Male)						
Sex	Female	β_7	-15.52 (0.43)	< 0.0001	-15.78(0.50)	< 0.0001
Age with Sex (vs 0-40 with Male)						
Age*Sex$_F$	41-80	β_8	-167.80 (1.95)	< 0.0001	-166.06(1.69)	< 0.0001
Age*Sex$_F$	80+	β_9	231.60 (2.07)	< 0.0001	231.17(2.62)	< 0.0001
Covariance of b_j:						
var(b_{1j})		d_{11}	1013.35(668.09)	0.0647	58.4064(25.18)	0.0637
cov(b_{1j}, b_{2j})		$d_{12} = d_{21}$	80.4642(89.05)	0.3662	30.5102 (16.69)	0.0677
var(b_{2j})		d_{22}	15.4022(20.73)	0.2288	25.7486(16.19)	0.2554
Residual Variance:						
var($\varepsilon_{(1)j}$)		σ^2	1069.93(27.89)	< 0.0001	1009.64(26.65)	< 0.0001
var($\varepsilon_{(2)j}$)		σ^2	9110.86(6798.98)	0.0901	4.7922(2.29)	0.0182

Furthermore, pairwise comparisons were performed between all age groups for females and males. These comparisons were performed in order to see if there is any statistical differences between age groups in females and males. It is shown that a significant difference exist between age groups in females and males (Table 4.17). Historical mortality for both genders differs by age group. Also, it can be seen that females in the age group above 80 years have a higher mortality rate compared to females in the age groups below 80 years, while males in the age groups below 80 years (Table 4.17) have a higher mortality rate compared to males in the age group above 80 years.

Table 4.17: Bonferrori Pairwise Comparisons for Age group in females and males. "Std.Error means Standard Error"

	Weighted Regression		Weighted Observations	
Label	Estimate (Std.Error)	95% Confidence Limits	Estimate (Std.Error)	95% Confidence Limits
41-80 vs 80+ Sex=Female	-304*(4.57)	[-311 ; -296]	-273* (2.04)	[-278 ; -268]
41-80 vs 80+ Sex=Male	123* (3.12)	[116 ; 131]	124* (1.54)	[120; 129]

**Significant at 0.025% alpha level*

The excess mortality from the linear mixed model for each wave is higher than the COVID-19 reported mortality and the excess mortality estimated with the weekly average methodology (Table 4.18). Due to the heat wave in the year 2020, there is a large differ-

ence between the reported COVID-19 mortality and excess mortality in the second wave (Table 4.18). The excess mortality was mostly present in the older age group (Figure A.11, Figure A.12, Figure A.13, Figure A.14, Figure A.15, Figure A.16, and Table 4.19).

Table 4.18: Reported COVID-19 and expected excess mortality based on predictions from the weekly average methodology and the linear mixed model (LMM)

		Weekly Average	LMM	
			Weighted Regression	Weighted Observations
Wave	COVID-19 Mortality	Excess Mortality(95%)	Excess Mortality(95%)	Excess Mortality(95%)
First	9624	9338[5741;12935]	11395[4953;17836]	10875[4609;17141]
Second	268	1252[-549;3053]	2302[-1414;6018]	1556[-1926;5039]
Third	9453	9367[6446;12287]	11745[4541;189481]	10276[3705;16847]

The weekly average and the linear mixed model show that there are more female excess deaths than male in the age group above 80 years in all three waves, and in the full year 2020, while there are more male excess deaths than female in the age group between 41-80 years in the first wave, third wave, and the full year 2020. For the reported COVID-19 mortality, there are more female excess deaths than male in the age group above 80 years in all three waves, and in the full year 2020 (Table 4.19). However, taking the population size into account, more males than females died due to COVID-19 in the age group of 40 years and above, in the first and third waves, and in the full year 2020 (Table A.1, Table A.2, Table A.3, and Table A.4). For the weekly average and the linear mixed model excess mortality rate per million inhabitants, more males than females died above 50 years in the first wave, third wave, and the full year 2020 (Table A.5, Table A.6, Table A.7, Table A.8, Table A.9, Table A.10, Table A.11, and Table A.12).

Table 4.19: Expected excess mortality based on the weekly average and the linear mixed model (LMM) for the three waves and the full year 2020

First-wave

			Weekly Average		Weighted Regression		Weighted Observations
Age	Sex	COVID-19 Mortality (95%)	Excess Mortality (95%)	Sex	Excess Mortality (95%)	Sex	Excess Mortality (95%)
0-40	Female	16[14;17]	-57[-109;-4]	Female	265[-798;1329]	Female	187[-847;1222]
41-80	Female	1119[969;1268]	777[384;1169]	Female	1076[11;2141]	Female	997[-39;2033]
80+	Female	3965[3425;4504]	4479[3129;5828]	Female	4951[3860;6043]	Female	4845[3787;5904]
0-40	Male	12[9;14]	-95[-213;23]	Male	237[-826;1301]	Male	154[-880;1190]
41-80	Male	1784[1555;2012]	1115[476;1753]	Male	1511[443;2580]	Male	1455[416;2495]
80+	Male	2728[2353;3102]	3119[2012;4225]	Male	3351[2263;4438]	Male	3234[2172;4295]

Second-wave

			Weekly Average		Weighted Regression		Weighted Observations
Age	Sex	COVID-19 Mortality (95%)	Excess Mortality (95%)	Sex	Excess Mortality (95%)	Sex	Excess Mortality (95%)
0-40	Female	4[4;4]	-21[-48;6]	Female	351[-264;968]	Female	232[-344;808]
41-80	Female	45[38;51]	1[-208;210]	Female	210[-406;827]	Female	90[-486;666]
80+	Female	87[73;100]	840[177;1502]	Female	785[161;1409]	Female	650[64;1236]
0-40	Male	4[2;5]	-56[-119;7]	Male	332[-284;948]	Male	210[-366;786]
41-80	Male	51[44;57]	-93[-385;199]	Male	74[-543;692]	Male	-32[-610;544]
80+	Male	77[63;90]	581[26;1135]	Male	547[-74;1170]	Male	406[-182;995]

Third-wave

			Weekly Average		Weighted Regression		Weighted Observations
Age	Sex	COVID-19 Mortality (95%)	Excess Mortality (95%)	Sex	Excess Mortality (95%)	Sex	Excess Mortality (95%)
0-40	Female	20[18;21]	-69[-129;-8]	Female	301[-895;1497]	Female	65[-1021;1151]
41-80	Female	1794[1615;1972]	756[259;1252]	Female	1061[-137;2259]	Female	823[-264;1912]
80+	Female	2818[2530;3105]	3796[2707;4884]	Female	4399[3194;5605]	Female	4134[3029;5240]
0-40	Male	24[21;26]	-109[-203;-15]	Male	266[-930;1463]	Male	25[-1061;1113]
41-80	Male	2738[2461;3014]	1688[858;2517]	Male	2028[826;3230]	Male	1816[725;2907]
80+	Male	2059[1860;2257]	3304[2242;4365]	Male	3687[2483;4890]	Male	3410[2299;4521]

Full-year

			Weekly Average		Weighted Regression		Weighted Observations
Age	Sex	COVID-19 Mortality (95%)	Excess Mortality (95%)	Sex	Excess Mortality (95%)	Sex	Excess Mortality (95%)
0-40	Female	40[38;41]	-147[-268;-25]	Female	918[-1959;3796]	Female	484[-2213;3182]
41-80	Female	2958[2796;3119]	1534[577;2490]	Female	2348[-532;5228]	Female	1911[-790;4612]
80+	Female	6870[6458;7281]	9115[6271;11959]	Female	10136[7215;13058]	Female	9630[6880;12381]
0-40	Male	40[37;42]	-260[-504;-15]	Male	835[-2042;3713]	Male	390[-2308;3089]
41-80	Male	4573[4323;4822]	2710[1105;4314]	Male	3614[726;6503]	Male	3239[531;5947]
80+	Male	4864[4579;5148]	7003[4370;9635]	Male	7586[4672;10500]	Male	7050[4289;9812]

4.5.1.1 Diagnostics Checking

The residuals from both models fitted (weighted regression and weighted observation model) are scattered around the horizontal line centered around zero in the scatter plots against the predicted values. These plots indicate that the linear model is appropriate and also show that the error variance is constant. Based on the QQ plots, there seems to be a linear

relationship between the residuals and their corresponding expected values which suggests normality of the error terms (Figure 4.8 and Figure 4.9).

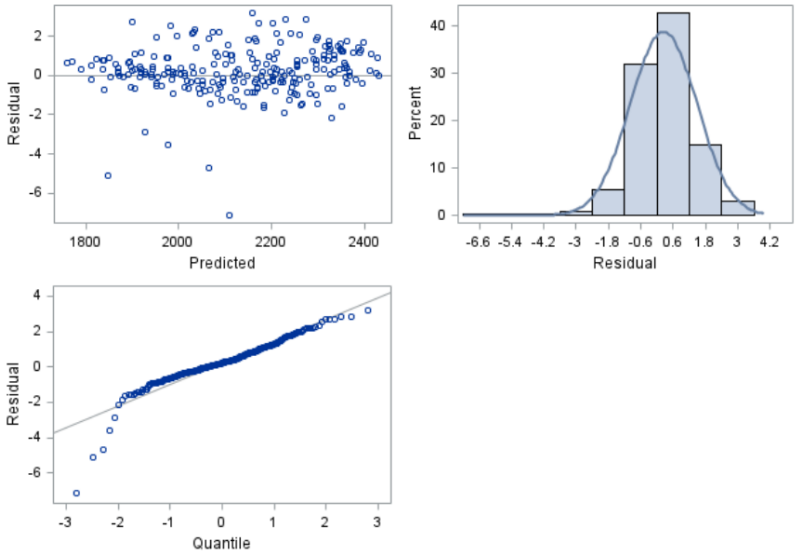

Figure 4.8: Scatterplot (Upper left), histogram plot(Upper right), with QQ plot (Bottom left) for the Weighted Regression model

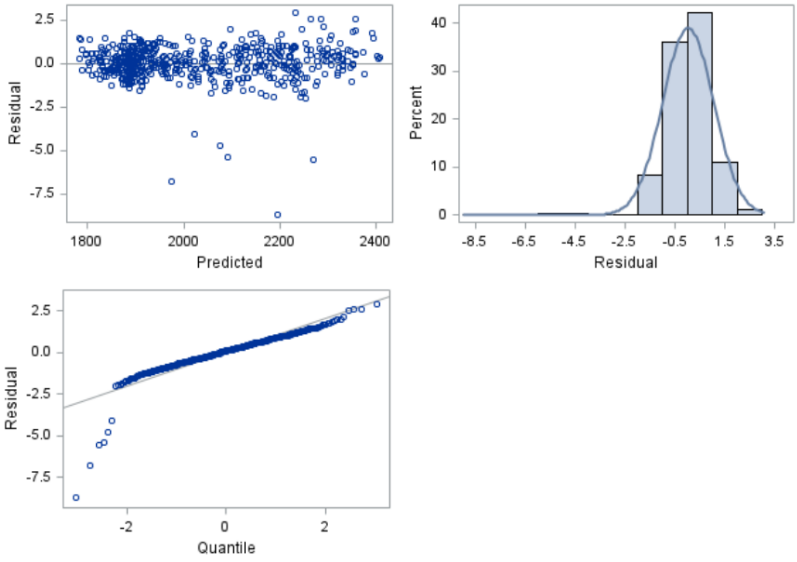

Figure 4.9: Scatterplot (Upper left), histogram plot(Upper right), with QQ plot (Bottom left) for the Weighted Observations

Chapter 5

Conclusion

To evaluate the gravity of COVID-19 within and between countries, COVID-19 reported deaths are used, where Belgium has been among the countries with a high rate of COVID-19 deaths in 2020. However, the completeness of reported COVID-19 mortality is heterogeneous between countries, therefore excess mortality has been suggested instead. Age and sex differences in COVID-19 mortality was investigated in Belgium to see whether the reported COVID-19 mortality differs across age and sex. Indeed, Ahrenfeldt et al.[1] has showed in Europe that more males died from COVID-19 than females in all age groups. Also, Islam et al.[7] found more deaths in males than females based on excess mortality.

Based on our findings, below 40 years and above 80 years there are more female than male deaths due to COVID-19 in the first wave, while between 40 and 80 years it reversed. For the third wave, there are more female deaths than male in the age group below 40 years, while above 40 years it is the reverse. There are no age and gender difference found in the summer peak, which may be the result of low COVID-19 death counts during this period. For the entire year 2020, there are more female deaths than male in the age groups below 40 and above 80 years old, and more male deaths than female in the age groups between 40 and 80 years, which is exactly what is also found based on the excess mortality from the linear mixed model. Both the reported COVID-19 mortality and excess mortality analyses show that mortality due to COVID-19 increases with age, where the age group of 80-89 years has the highest mortality rate. Both reported COVID-19 mortality and excess mortality analyses can thus be used to evaluate age gender difference for Belgium. This is a direct consequence of the rather complete reporting of COVID-19 related mortality in Belgium.

Correcting the reported COVID-19 mortality and excess mortality predictions from the weekly average method for population size, more males died than females in all age categories above 40, respectively 50 years. These results coincide with the results from

Ahrenfeldt et al.[1] and Islam et al.[7].

The excess mortality from the linear mixed model, for each wave is higher than the COVID-19 reported mortality, but in the second wave there is a large difference between reported COVID-19 mortality and excess deaths due to the heat wave that occurred in August 2020. However, the excess mortality based on the weekly average method was closer to the COVID-19 reported mortality than the excess mortality based on the linear mixed models.

The linear mixed model improves predication accuracy only for the year 2014, and is worse for 2018, while the weekly average methodology improves predication accuracy for the year 2018, and is worse for 2014. In 2016, both methods (weekly average and linear mixed model) gave similar predicting accuracy. The linear mixed model predicts a conditional mean, allows for serial correlation, and down-weighs the influence of historical excess mortality, while the weekly average predicts a marginal mean, ignores autocorrelation, and can be influenced by historical excess mortality. In this case, the results from the linear mixed model should be considered and trusted.

In conclusion, this study showed that there is a difference between the second wave and the other two waves indicating that excess mortality that occurred during the summer of 2020 was not caused by COVID-19 but rather due to the heat wave. Statistical differences were shown between age groups and sex, where females below 40 years died more of COVID-19 than males during the spring, and winter periods and the entire year 2020. However, taking the population size into account, the results change and we find more deaths in males above 40 years. Elderly people died more from COVID-19 in Belgium than younger people.

In this study, the effect of age and gender on COVID-19 mortality was studied, which cannot completely tale the effect of mortality due to COVID-19. For further research, regions, and nursing home differences within Belgium should be considered as they are associated with COVID-19 mortality as seen in Molenberghs et al.[10] and Ahrenfeldt et al.[1]. Also, this modelling exercise should be repeated with population size corrections.

References

[1] Ahrenfeldt, Linda Juel, Otavova, Martina, Christensen, Kaare, Lindahl-Jacobsen, and Rune. "Sex and age differences in COVID-19 mortality in Europe". In: *Wiener klinische Wochenschrift* 133.7 (2021), pp. 393–398.

[2] Agresti Alan. *Categorical data analysis*. Vol. 482. John Wiley & Sons, 2003.

[3] Aron, Janine, Giattino, C, Muellbauer, J, Ritchie, and Hannah. "A pandemic primer on excess mortality statistics and their comparability across countries". In: *Our World in Data* (2020).

[4] EPISTAT:COVID-19. *Sciensano, 2020*. Sept. 23, 2020. URL: `https://epistat.wiv-isp.be/covid/`.

[5] Farrington, CP, Andrews, Nick J, Beale, AD, Catchpole, and MA. "A statistical algorithm for the early detection of outbreaks of infectious disease". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 159.3 (1996), pp. 547–563.

[6] Andrew C Harvey. "Forecasting, structural time series models and the Kalman filter". In: (1990).

[7] Islam, Nazrul, Shkolnikov, Vladimir M, Acosta, Rolando J, Klimkin, Ilya, Kawachi, Ichiro, Irizarry, Rafael A, Alicandro, Gianfranco, Khunti, Kamlesh, Yates, Tom, Jdanov, and Dmitri A. "Excess deaths associated with covid-19 pandemic in 2020: age and sex disaggregated time series analysis in 29 high income countries". In: *bmj* 373 (2021).

[8] ISO8601. *Date and time — Representations for information interchange*. Feb. 2019. URL: `https://www.iso.org/obp/ui/#iso:std:iso:8601:-1:ed-1:v1:en`.

[9] Molenberghs, G, Verbeke, and G. "Models for Discrete Longitudinal Data. Springer-Verlag". In: *New York* (2005).

[10] Molenberghs, Geert, Faes, Christel, Verbeeck, Johan, Deboosere, Patrick, Abrams, Steven, Willem, Lander, Aerts, Jan, Theeten, Heidi, Devleesschauwer, Brecht, Sierra, and Natalia Bustos. "Belgian COVID-19 Mortality, Excess Deaths, Number of Deaths

per Million, and Infection Fatality Rates (9 March—28 June 2020)". In: *medRxiv* (2020).

[11] Molenberghs, Geert, and Verbeke Geert. *Linear mixed models for longitudinal data*. Springer, 2000.

[12] STATBEL. *Number of deaths per day, sex, age, region, province, district, 2020*. Sept. 23, 2020. URL: https://statbel.fgov.be/en/open-data/number-deaths-day-sex-district-age.

[13] Statista. *Death toll in Belgium*. 2020. URL: https://www.statista.com/statistics/1101080/coronavirus-cases-in-belgium/.

[14] Verbeeck, Johan, Faes, Christel, Neyens, Thomas, Hens, Niel, Verbeke, Geert, Deboosere, Patrick, Molenberghs, and Geert. "A linear Mixed Model to Estimate COVID-19-induced Excess Mortality". In: *medRxiv* (2021).

[15] WHO. *Coronavirus*. 2020. URL: https://www.who.int/health-topics/coronavirus#tab=tab_1.

Appendix A

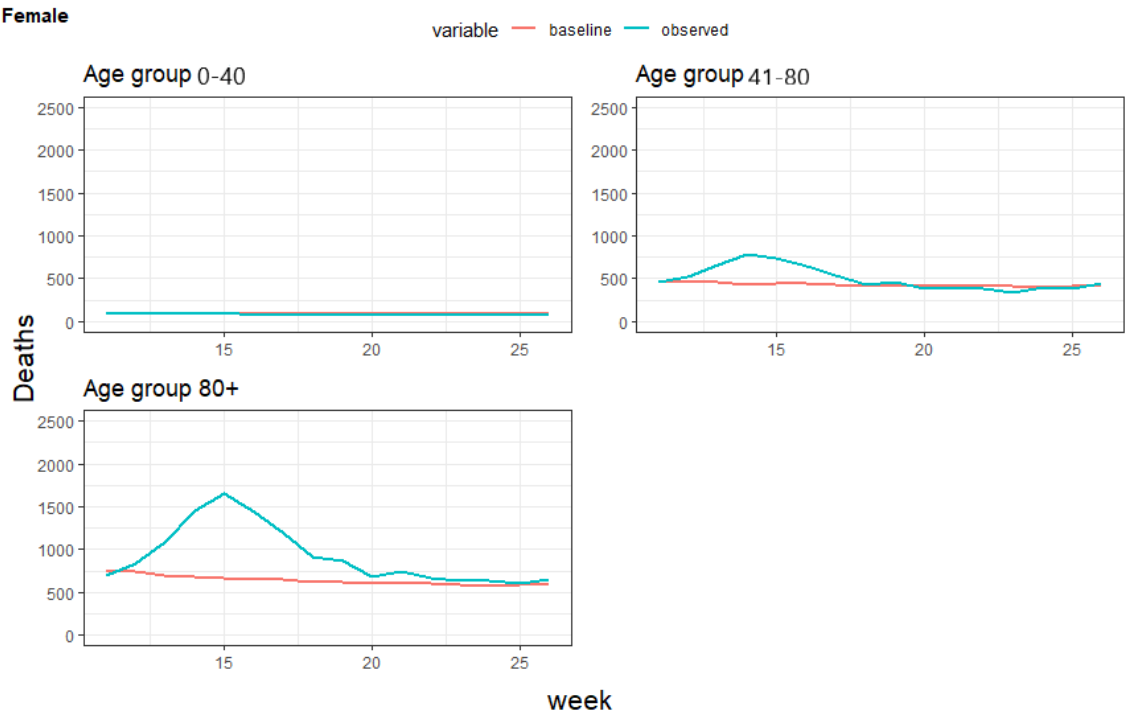

Figure A.1: Weekly observed all-cause mortality (observed) and predicted all-cause mortality (baseline) by age group and gender in Belgium for the first-wave (week 11-week 26) by the weekly average method

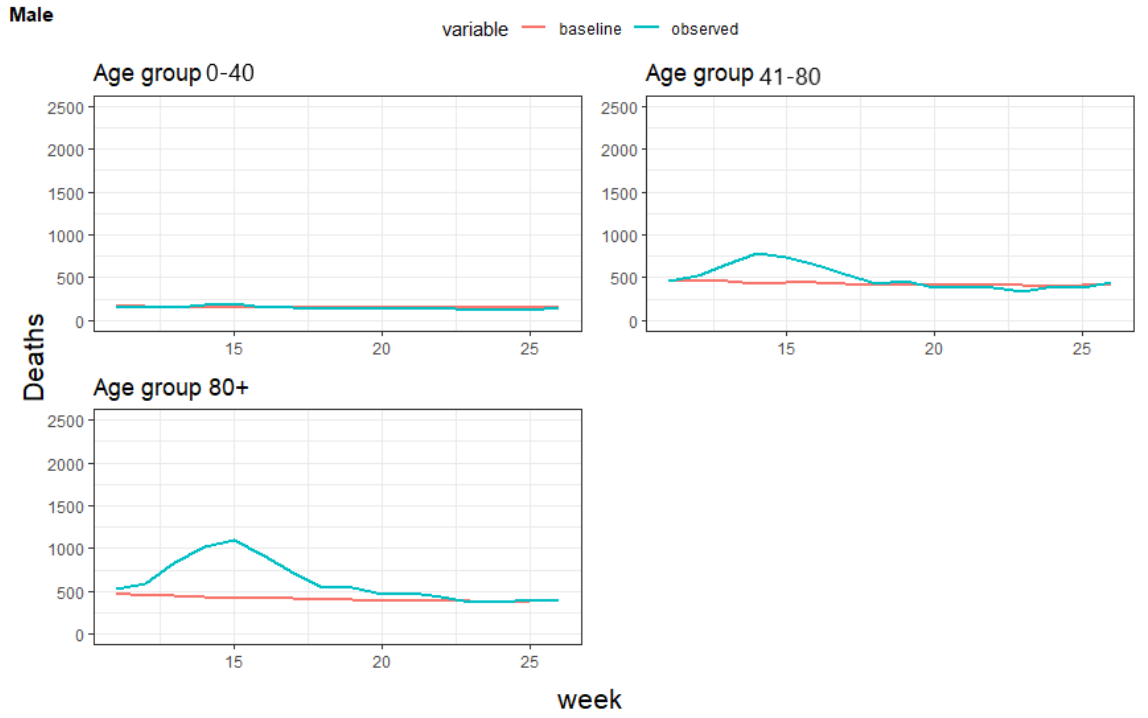

Figure A.2: Weekly observed all-cause mortality (observed) and predicted all-cause mortality (baseline) by age group and gender in Belgium for the first-wave (week 11-week 26) by the weekly average method

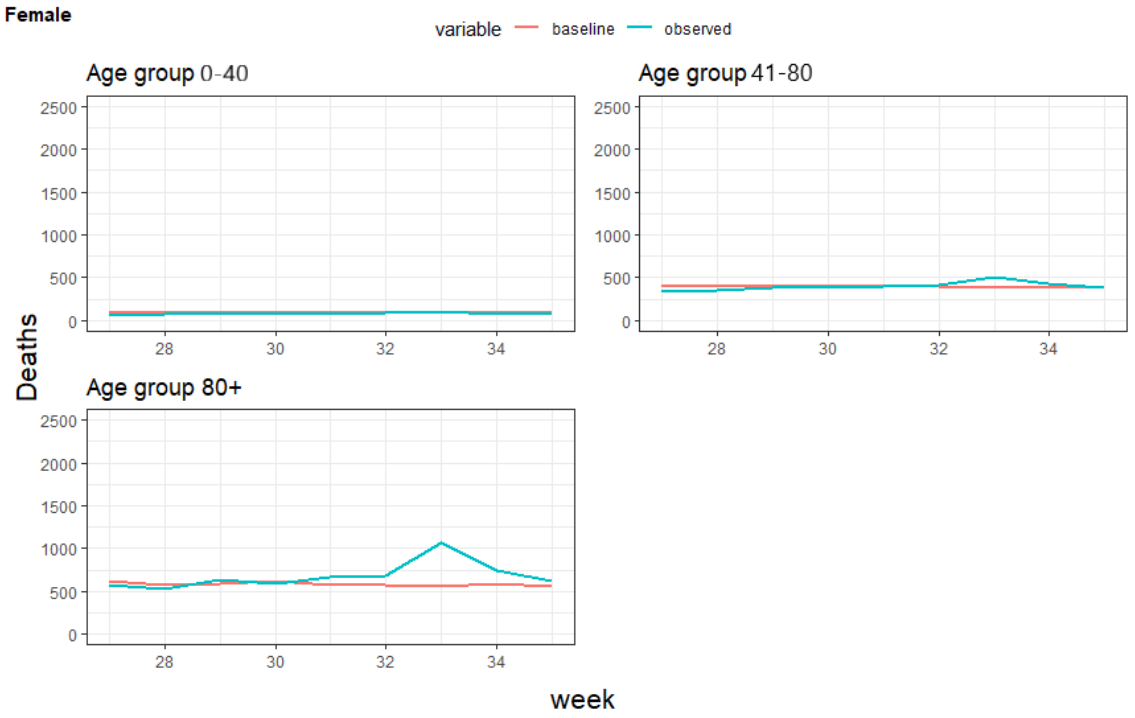

Figure A.3: Weekly observed all-cause mortality (observed) and predicted all-cause mortality (baseline) by age group and gender in Belgium for the second-wave (week 27-week 35) by the weekly average method

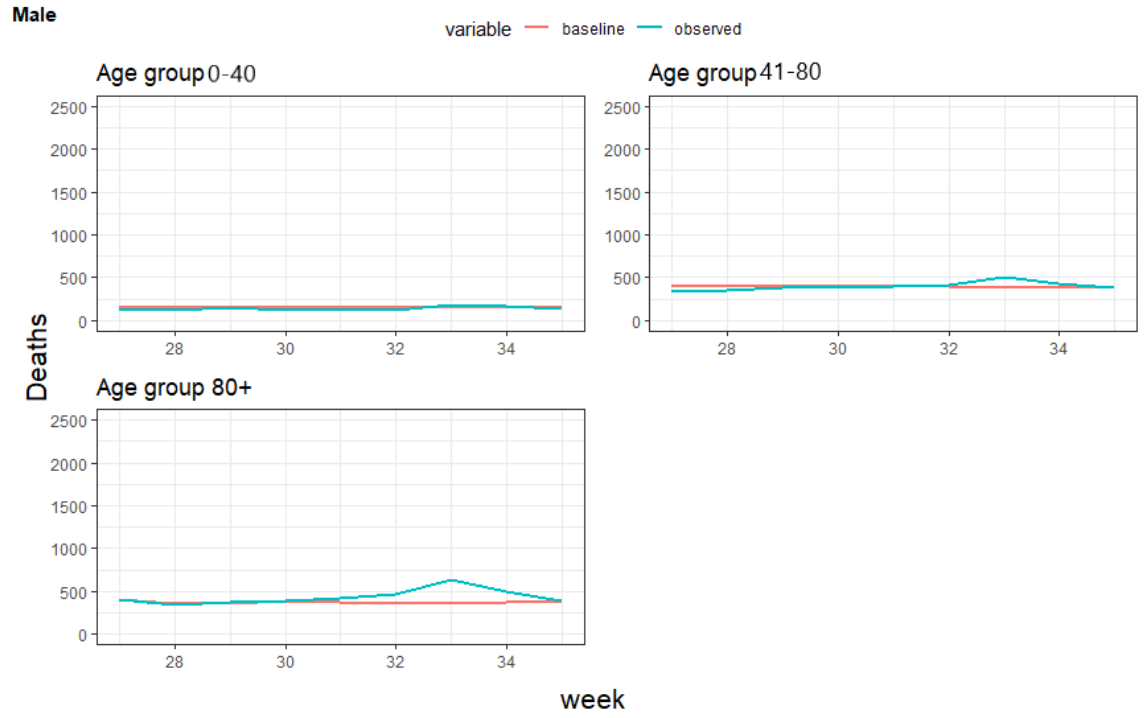

Figure A.4: Weekly observed all-cause mortality (observed) and predicted all-cause mortality (baseline) by age group and gender in Belgium for the second-wave (week 27-week 35) by the weekly average method

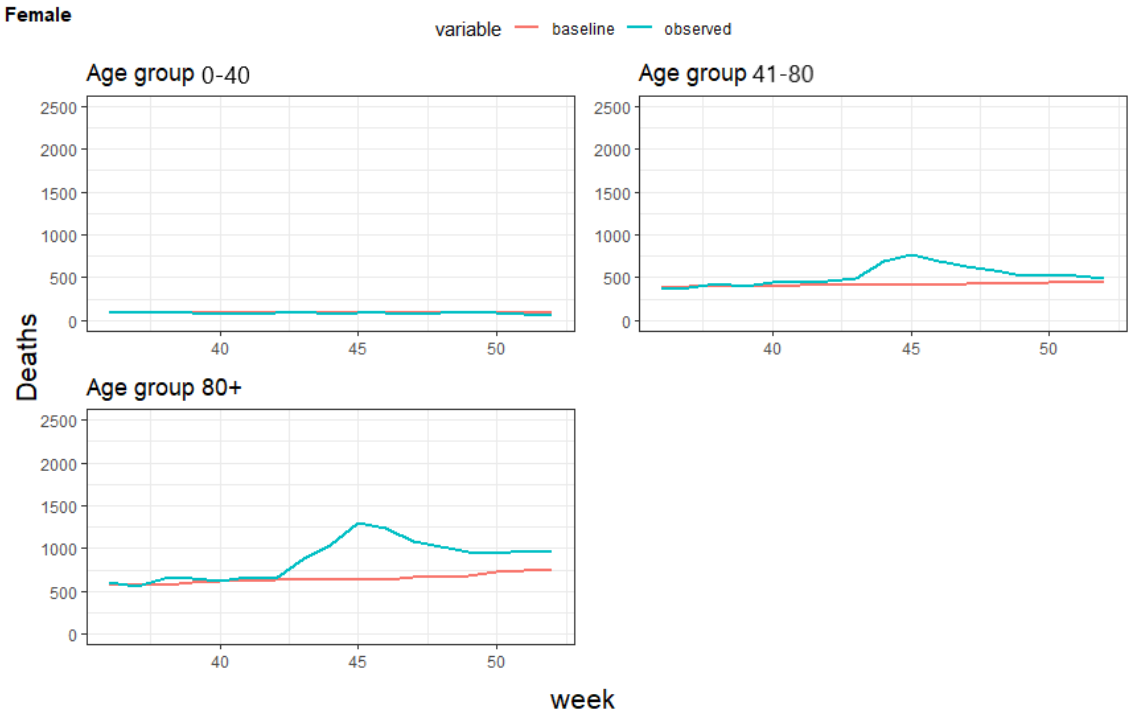

Figure A.5: Weekly observed all-cause mortality (observed) and predicted all-cause mortality (baseline) by age group and gender in Belgium for the third-wave (week 36-week 52) by the weekly average method

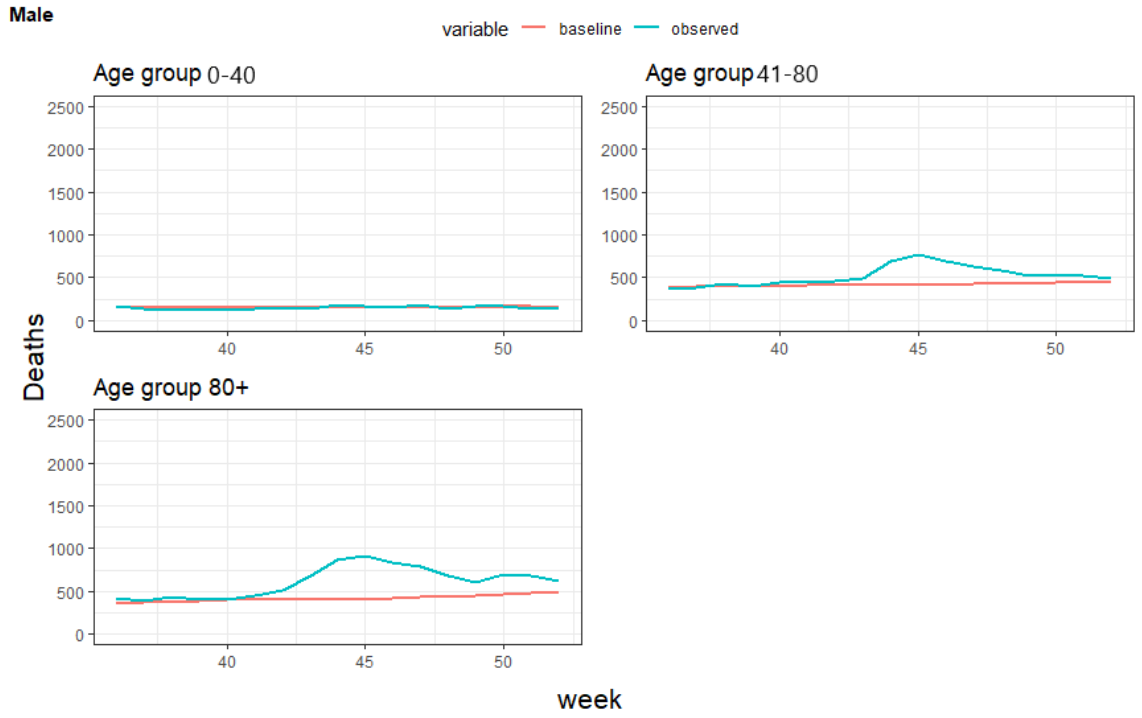

Figure A.6: Weekly observed all-cause mortality (observed) and predicted all-cause mortality (baseline) by age group and gender in Belgium for the third-wave (week 36-week 52) by the weekly average method

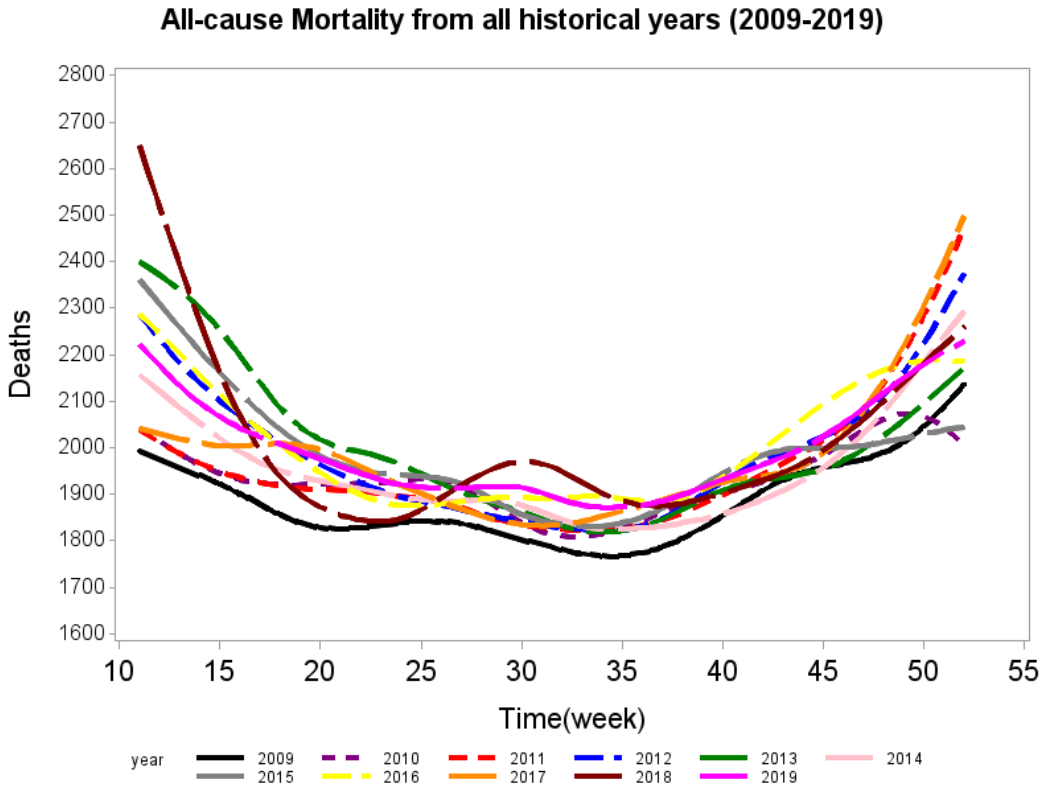

Figure A.7: All-cause Mortality of historical years (2009-2019)

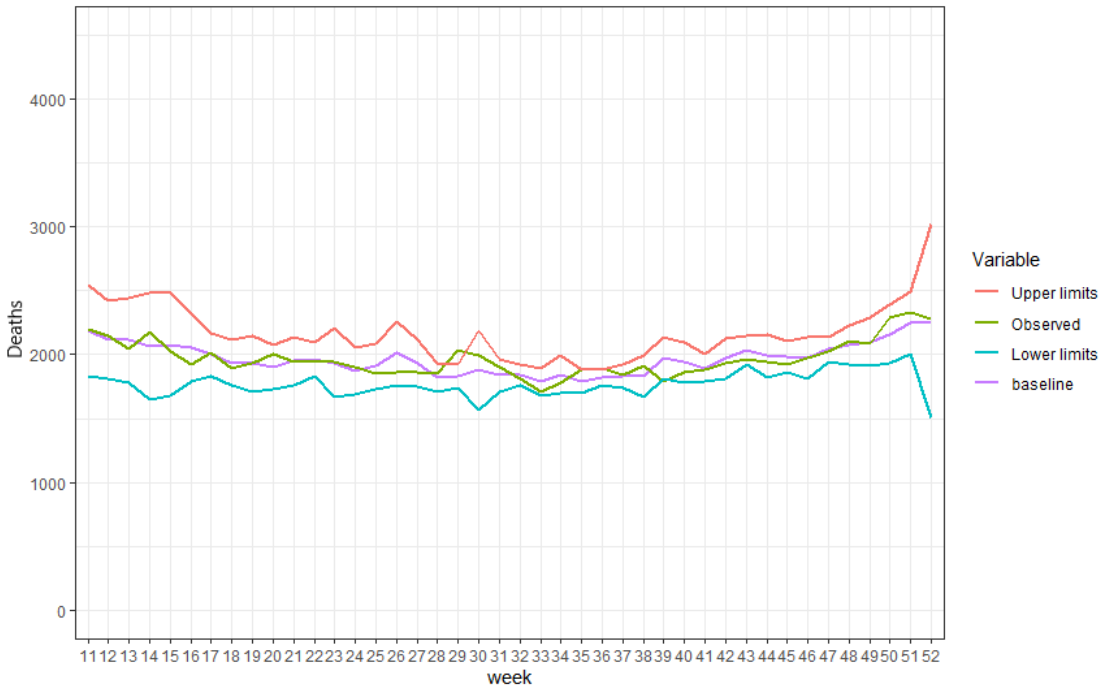

Figure A.8: Weekly observed all-cause mortality (observed) and predicted all-cause mortality (baseline) plot of the year 2014 by the weekly average method

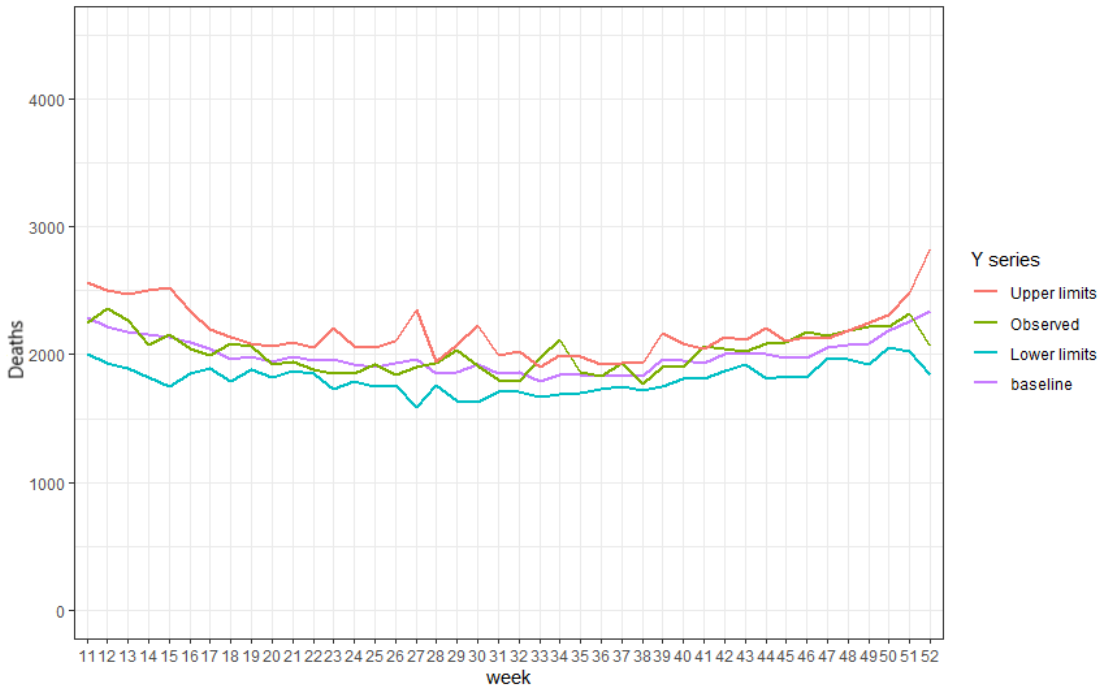

Figure A.9: Weekly observed all-cause mortality (observed) and predicted all-cause mortality (baseline) plot of the year 2016 by the weekly average method

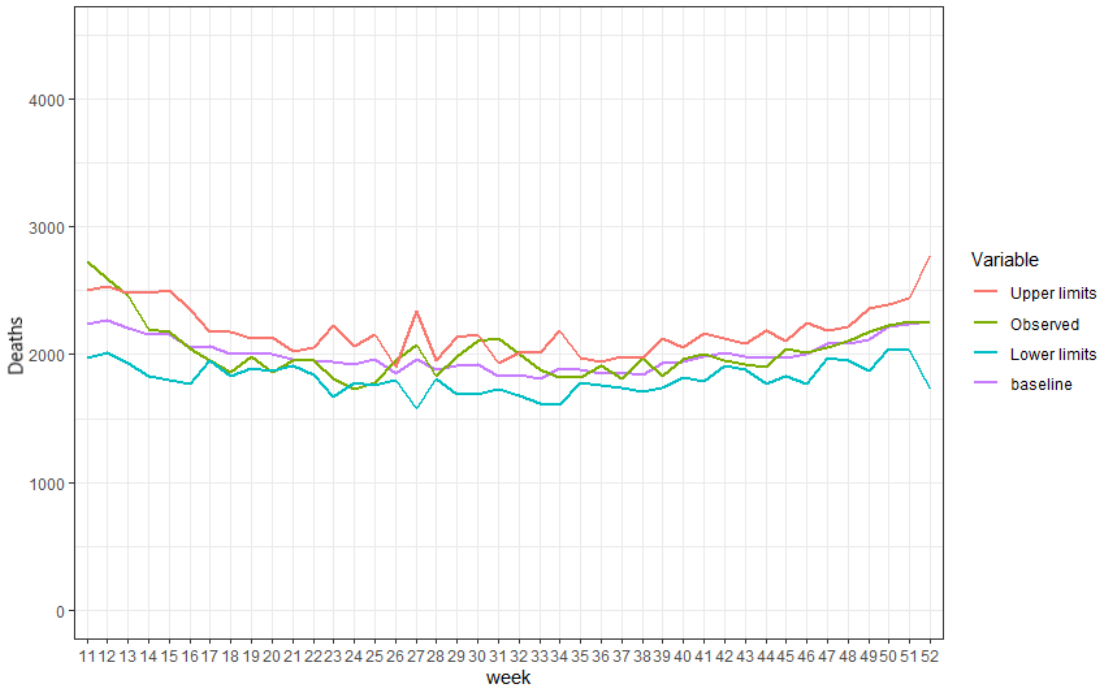

Figure A.10: Weekly observed all-cause mortality (observed) and predicted all-cause mortality (baseline) plot of the year 2018 by the weekly average method

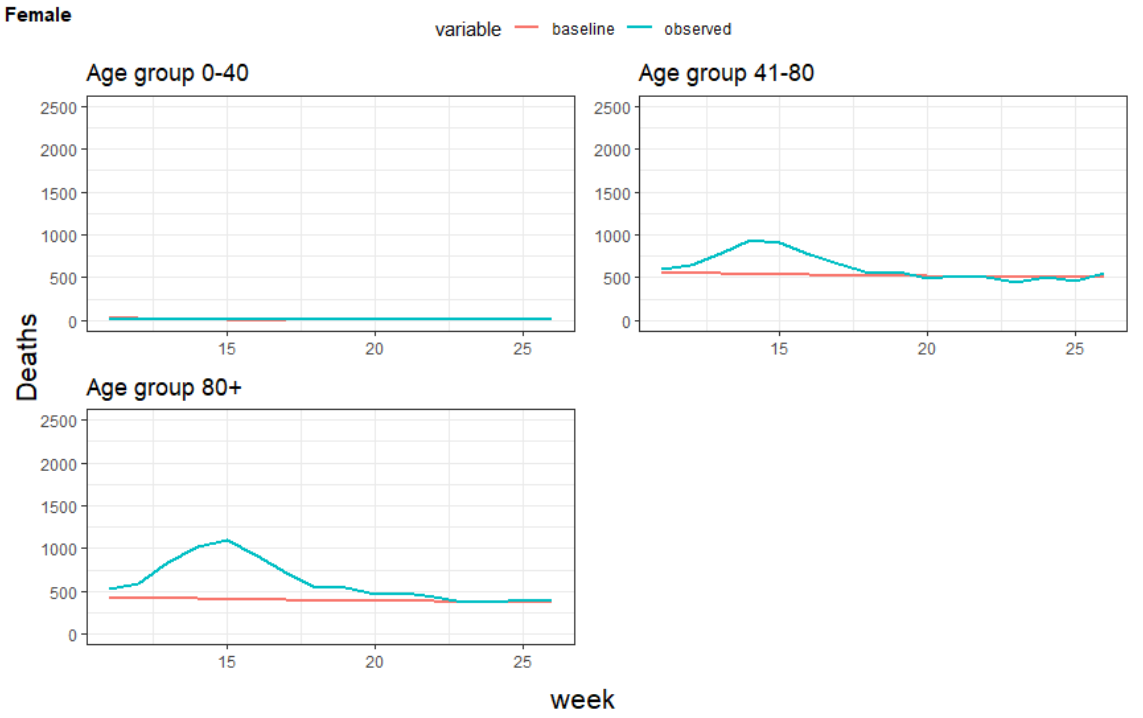

Figure A.11: Weekly observed all-cause mortality (observed) by age group and gender in Belgium for the first-wave (week 11-week 26) and the all-cause mortality predicted (baseline) from the linear mixed model

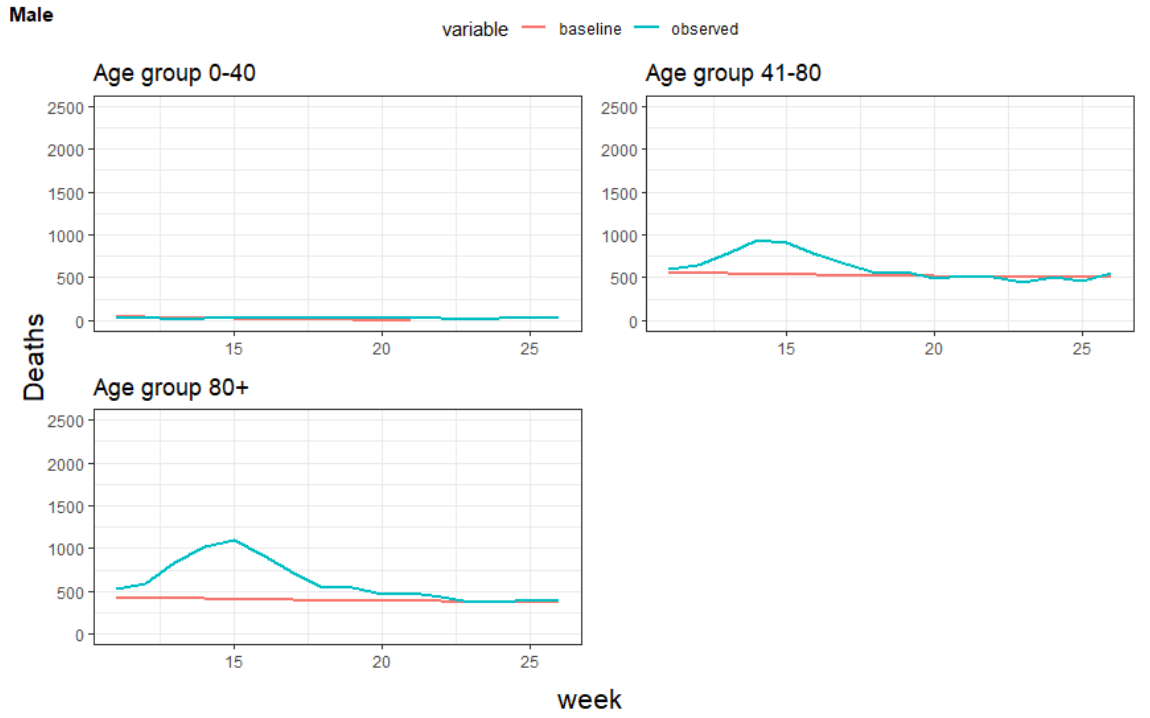

Figure A.12: Weekly observed all-cause mortality (observed) by age group and gender in Belgium for the first-wave (week 11-week 26) and the all-cause mortality predicted (baseline) from the linear mixed model

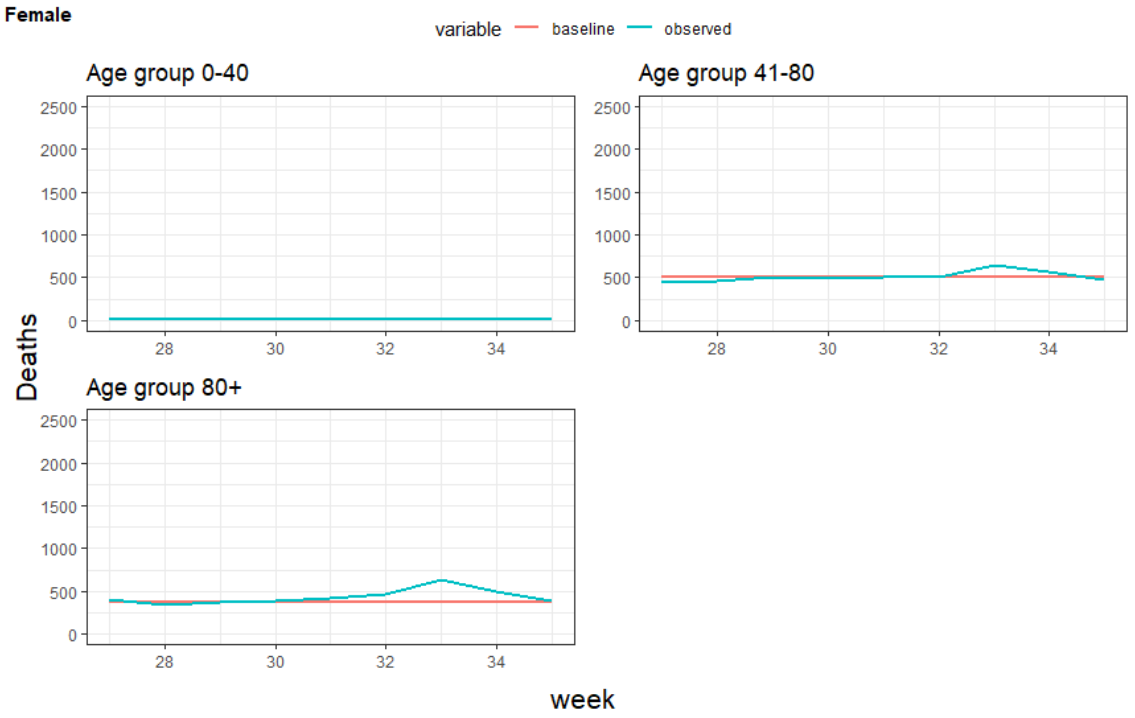

Figure A.13: Weekly observed all-cause mortality (observed) by age group and gender in Belgium for the second-wave (week 27-week 35) and the all-cause mortality predicted (baseline) from the linear mixed model

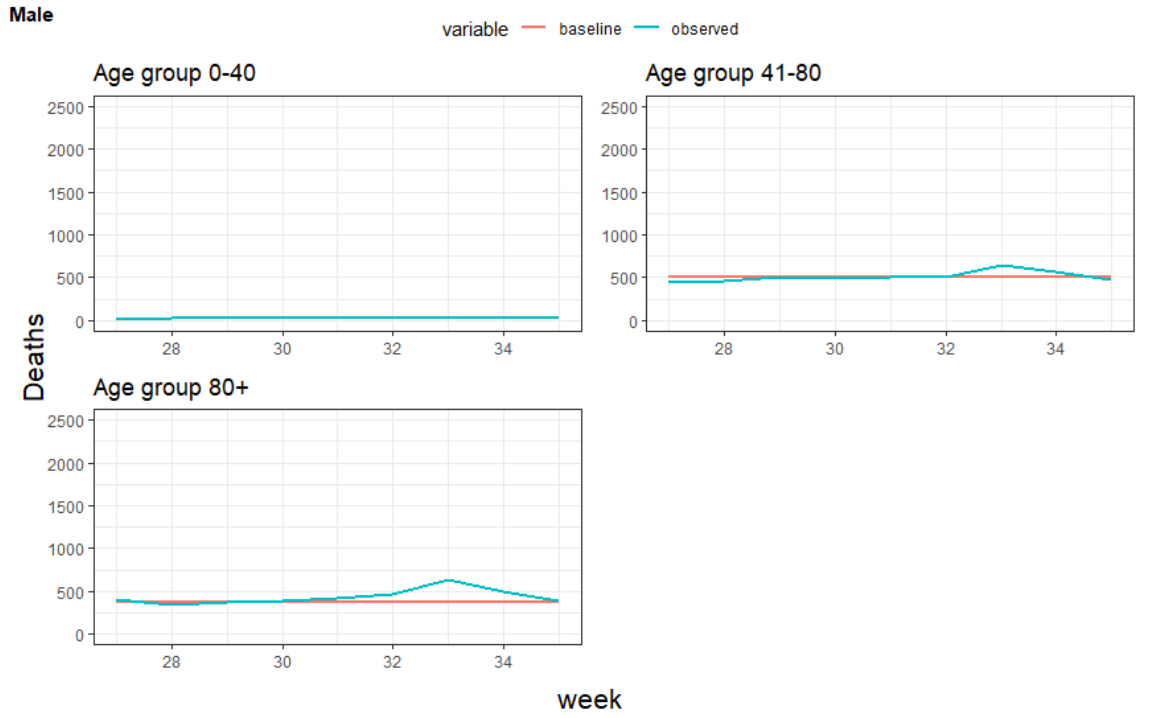

Figure A.14: Weekly observed all-cause mortality (observed) by age group and gender in Belgium for the second-wave (week 27-week 35) and the all-cause mortality predicted (baseline) from the linear mixed model

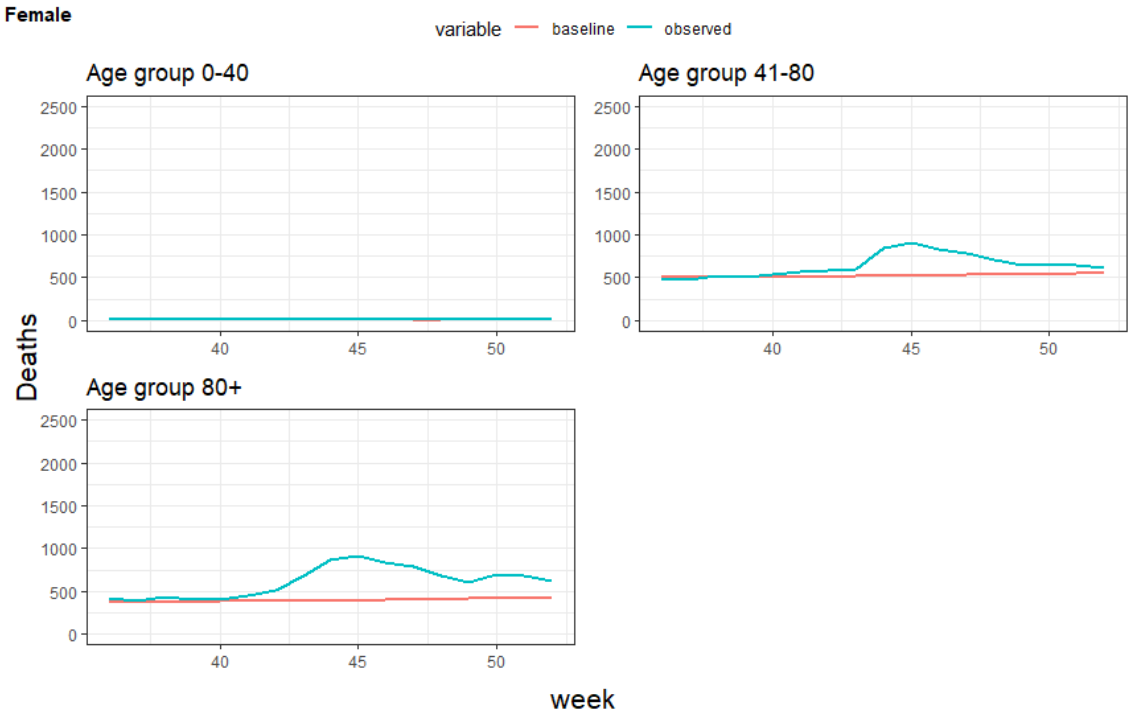

Figure A.15: Weekly observed all-cause mortality (observed) by age group and gender in Belgium for the third-wave (week 36-week 52) and the all-cause mortality predicted (baseline) from the linear mixed model

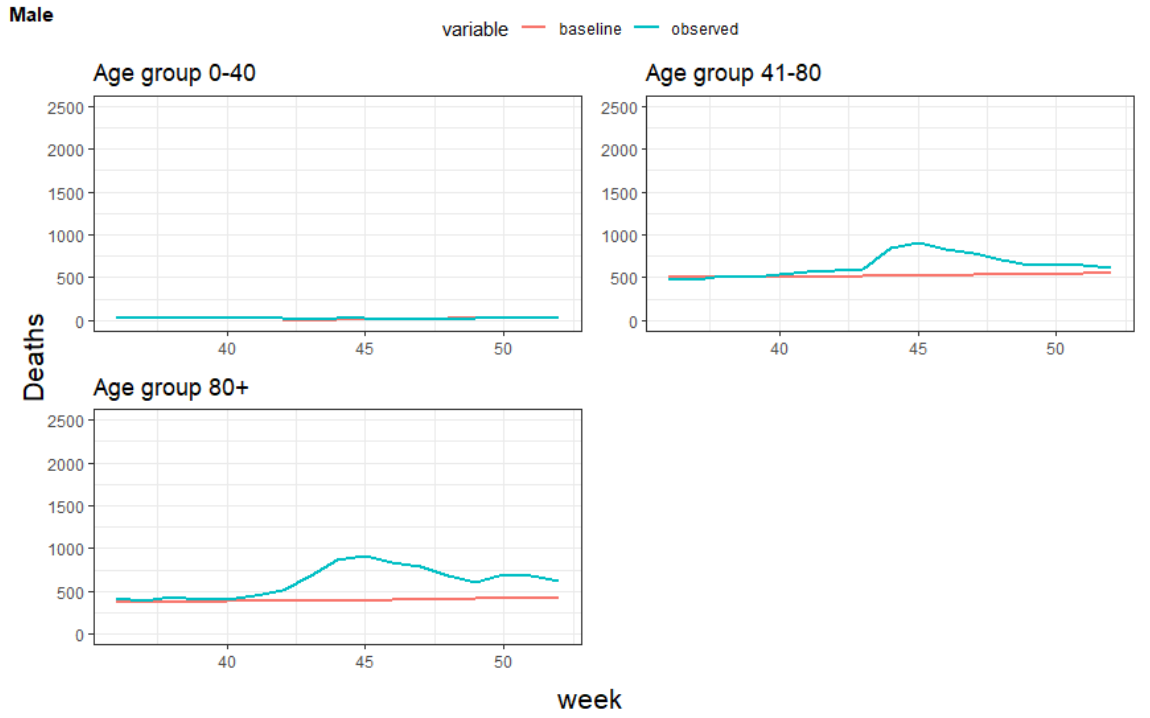

Figure A.16: Weekly observed all-cause mortality (observed) by age group and gender in Belgium for the third-wave (week 36-week 52) and the all-cause mortality predicted (baseline) from the linear mixed model

Table A.1: COVID-19 reported mortality rate per million inhabitants over sex and age group in Belgium for the First-wave.

Sex	Age group	Number of COVID-19 Deaths	Number of COVID-19 per million
Male(M)	30-39	9	5
	40-49	35	19
	50-59	146	182
	60-69	453	687
	70-79	1001	2343
	80-89	1968	9305
	90+	923	28201
	Total Deaths	**4535**	
Female(F)	30-39	12	8
	40-49	22	12
	50-59	73	92
	60-69	214	311
	70-79	682	1372
	80-89	2211	6743
	90+	1872	22110
	Total Deaths	**5086**	

Table A.2: COVID-19 reported mortality rate per million inhabitants over sex and age group in Belgium for the Second-wave

Sex	Age group	Number of COVID-19 Deaths	Number of COVID-19 per million
Male	30-39	2	1
	40-49	3	2
	50-59	5	6
	60-69	10	15
	70-79	31	73
	80-89	48	227
	90+	32	978
	Total Deaths	**131**	
Female	30-39	1	1
	40-49	2	1
	50-59	6	8
	60-69	11	16
	70-79	22	44
	80-89	49	149
	90+	42	496
	Total Deaths	**133.00**	

Table A.3: COVID-19 reported mortality rate per million inhabitants over sex and age group in Belgium for the Third-wave

Sex	Age group	Number of COVID-19 Deaths	Number of COVID-19 per million
Male	30-39	9	5
	40-49	39	21
	50-59	155	193
	60-69	508	771
	70-79	1093	2559
	80-89	2064	9759
	90+	950	29026
	Total Deaths	**4818**	
Female	30-39	9	6
	40-49	23	12
	50-59	80	101
	60-69	237	344
	70-79	657	1322
	80-89	2014	6142
	90+	1609	19004
	Total Deaths	**4629**	

Table A.4: COVID-19 reported mortality rate per million inhabitants over sex and age group in Belgium for the Full year 2020

Sex	Age group	Number of COVID-19 Deaths	Number of COVID-19 per million
Male	30-39	20	12
	40-49	77	41
	50-59	306	382
	60-69	971	1473
	70-79	2125	4975
	80-89	4080	19291
	90+	1905	58205
	Total Deaths	**9484**	
Female	30-39	22	14
	40-49	47	25
	50-59	159	201
	60-69	462	671
	70-79	1361	2738
	80-89	4274	13035
	90+	3523	41610
	Total Deaths	**9848**	

Table A.5: Weekly average excess mortality rate per million inhabitants over sex and age group in Belgium for the First-wave

Sex	Age group	Number of Excess Deaths	Number of Excess Deaths per million
Male	30-39	6	4
	40-49	-71	-38
	50-59	-11	-14
	60-69	411	624
	70-79	708	1657
	80-89	1737	8213
	90+	1460	44609
	Total Deaths	**4240**	
Female	30-39	1	1
	40-49	-62	-33
	50-59	-23	-29
	60-69	222	323
	70-79	531	1068
	80-89	1709	5212
	90+	2876	33968
	Total Deaths	**5254**	

Table A.6: Weekly average excess mortality rate per million inhabitants over sex and age group in Belgium for the Second-wave

Sex	Age group	Number of Excess Deaths	Number of Excess Deaths per million
Male	30-39	-13	-8
	40-49	-38	-20
	50-59	-53	-66
	60-69	53	80
	70-79	-29	-68
	80-89	142	671
	90+	416	12710
	Total Deaths	**478**	
Female	30-39	-8	-5
	40-49	-27	-15
	50-59	-25	-32
	60-69	41	60
	70-79	24	48
	80-89	89	271
	90+	740	8740
	Total Deaths	**834**	

Table A.7: Weekly average excess mortality rate per million inhabitants over sex and age group in Belgium for the Third-wave

Sex	Age group	Number of Excess Deaths	Number of Excess Deaths per million
Male	30-39	3	2
	40-49	-23	-12
	50-59	14	17
	60-69	418	634
	70-79	1183	2769
	80-89	1840	8700
	90+	1555	47511
	Total Deaths	**4990**	
Female	30-39	-18	-11
	40-49	-19	-10
	50-59	-48	-61
	60-69	173	251
	70-79	658	1324
	80-89	1243	3791
	90+	2546	30070
	Total Deaths	**4535**	

Table A.8: Weekly average excess mortality rate per million inhabitants over sex and age group in Belgium for the Full year 2020

Sex	Age group	Number of Excess Deaths	Number of Excess Deaths per million
Male	30-39	-4	-2
	40-49	-132	-70
	50-59	-51	-64
	60-69	882	1338
	70-79	1862	4359
	80-89	3718	17579
	90+	3431	104831
	Total Deaths	**9706**	
Female	30-39	-25	-16
	40-49	-108	-58
	50-59	-97	-123
	60-69	437	635
	70-79	1213	2440
	80-89	3041	9274
	90+	6162	72778
	Total Deaths	**10623**	

Table A.9: Weighted observations excess mortality rate per million inhabitants over sex and age group in Belgium for the First-wave

Sex	Age group	Number of Excess Deaths	Number of Excess Deaths per million
Male	30-39	45	27
	40-49	-29	-16
	50-59	30	38
	60-69	432	656
	70-79	819	1919
	80-89	1875	8867
	90+	1415	43252
	Total Deaths	**4588**	
Female	30-39	39	25
	40-49	-26	-14
	50-59	16	21
	60-69	234	341
	70-79	583	1174
	80-89	2033	6201
	90+	2939	34720
	Total Deaths	**5820**	

Table A.10: Weighted observations excess mortality rate per million inhabitants over sex and age group in Belgium for the Second-wave

Sex	Age group	Number of Excess Deaths	Number of Excess Deaths per million
Male	30-39	62	38
	40-49	46	25
	50-59	17	22
	60-69	45	70
	70-79	-73	-172
	80-89	-25	-121
	90+	350	10695
	Total Deaths	**423**	
Female	30-39	65	41
	40-49	45	25
	50-59	32	41
	60-69	69	101
	70-79	11	23
	80-89	-21	-66
	90+	638	7544
	Total Deaths	**841**	

Table A.11: Weighted observations excess mortality rate per million inhabitants over sex and age group in Belgium for the Third-wave

Sex	Age group	Number of Excess Deaths	Number of Excess Deaths per million
Male	30-39	29	18
	40-49	-5	-3
	50-59	42	53
	60-69	425	645
	70-79	1222	2862
	80-89	1974	9335
	90+	1593	48690
	Total Deaths	**5281**	
Female	30-39	15	10
	40-49	11	6
	50-59	-16	-20
	60-69	176	257
	70-79	665	1339
	80-89	1513	4617
	90+	2733	32289
	Total Deaths	**5100**	

Table A.12: Weighted observations excess mortality rate per million inhabitants over sex and age group in Belgium for the Full year 2020

Sex	Age group	Number of Excess Deaths	Number of Excess Deaths per million
Male	30-39	137	83
	40-49	11	6
	50-59	90	113
	60-69	903	1371
	70-79	1968	4609
	80-89	3824	18081
	90+	3359	102637
	Total Deaths	**10294**	
Female	30-39	119	76
	40-49	30	17
	50-59	32	42
	60-69	480	698
	70-79	1260	2536
	80-89	3525	10752
	90+	6312	74552
	Total Deaths	**11762**	

IMPORTANT SAS AND R CODES ONLY

```
##Negative Binomial Modelling For the ##
##Reported COVID-19 Mortality ##



summary(First-wave <- glm.nb(Deaths ~ as.factor(Age)
         + as.factor(Sex) + as.factor(Sex)*as.factor(Age)
         + as.factor(Time),data = First-wave))

means <- emmeans(First-wave,  "Age",  by="Sex",
                  adjust = "bonferroni")
pairwise=pairs(means)
confint(pairwise, adjust = "bonferroni")


summary(Second-wave <- glm.nb(Deaths ~ as.factor(Age)
+ as.factor(Sex) + as.factor(Time),data = Second-wave))


summary(Third-wave <- glm.nb(Deaths ~ as.factor(Age)
         + as.factor(Sex) + as.factor(Sex)*as.factor(Age)
         + as.factor(Time),data = Third-wave))

means <- emmeans(Third-wave,  "Age",  by="Sex",
adjust = "bonferroni")
pairwise=pairs(means)
confint(pairwise, adjust = "bonferroni")
```

```
summary(Full-year <- glm.nb(Deaths ~ as.factor(Age)
         + as.factor(Sex) + as.factor(Sex)*as.factor(Age)
         + as.factor(Time),data = Full-year))

means <- emmeans(Full-year,  "Age",  by="Sex",
adjust = "bonferroni")
pairwise=pairs(means)
confint(pairwise, adjust = "bonferroni")
```

```
/*Liner Mixed Model For the All-Cause Mortality*/

/*data preparation*/

data analysis1;
set analysis (where=(week^=53));/*remove week 53*/
/*add fourier terms to model seasonality*/
sine_full_year=sin(2*constant("pi")*week/52);
consine_full_year=cos(2*constant("pi")*week/52);
sine_half_year=sin(2*constant("pi")*week/26);
consine_half_year=cos(2*constant("pi")*week/26);
run;

/*fit model first time (model 3.4)*/

proc mixed data=analysis1 method = reml  empirical covtest plots=none ;
class year AgeP(ref='0-40') SEX(ref='Male');
model Deaths_avg = Age Sex Age*Sex sine_full_year
consine_full_year sine_half_year consine_half_yea /
solution residual outp=pred1; /*outp prints conditional residuals*/
random  intercept sine_full_year / subject=year type=un  ;
repeated / subject=year;
run;

/* downweighing observations*/
```

```
data analysis2;
set pred1;
weight_(2)=Deaths_avg;
if Pearsonresid > 1 then
weight_(2)=Deaths_avg*(1-(0.05*(Pearsonresid+1)));run;

/*fit model second time (Model 3.6)*/

proc mixed data=analysis2 covtest method = reml empirical;
class year Age(ref='0-40') Sex(ref='Male');
model weight_(2) = Age Sex Age*Sex sine_full_year
consine_full_year sine_half_year consine_half_yea /
solution residual outp=pred2 ;
random intercept sine_full_year  / subject=year type=un;
repeated / subject=year type=sp(exp)(week) local;/*fit serial correlation*/
estimate 'Age 41-80 vs Age 80+ Sex=Female' Age 0 1 -1   Sex -1 1
/  cl alpha=0.025 ;
estimate 'Age 41-80 vs Age 80+ Sex=Male' Age 0 1 -1   SEX 1 -1
/  cl alpha=0.025 ;
run;


/*weighted regression*/

data analysis3;
set pred1;
weight_(1)=1;
if Pearsonresid > 1 then w2=1/(Pearsonresid**2);
run;

/*fit model second time (model 3.6)*/

proc mixed data=analysis3 covtest method=reml empirical;
class year Age(ref='0-40') Sex(ref='Male');
model Deaths_avg = Age Sex Age*Sex sine_full_year
consine_full_year sine_half_year consine_half_yea /
solution residual outp=pred3 ;
random intercept sine_full_year / subject=year type=un;
repeated / subject=year type=sp(exp)(week) local;/*fit serial correlation*/
estimate 'Age 41-80 vs Age 80+ Sex=Female' Age 0 1 -1   Sex -1 1
```

```
/  cl alpha=0.025 ;
estimate 'Age 41-80 vs Age 80+ Sex=Male' Age 0 1 -1   SEX 1 -1
/  cl alpha=0.025 ;
weight weight_(1);
run;
```