

# Heterogeneity in Subnational Mortality in The Context of The COVID-19 Pandemic: The Case of Belgian Districts

Benjamin-Samuel Schlüter (✉ [benjamin-samuel.schluter@uclouvain.be](mailto:benjamin-samuel.schluter@uclouvain.be))

Catholic University of Louvain: Universite Catholique de Louvain <https://orcid.org/0000-0001-6707-3448>

Bruno Masquelier

Catholic University of Louvain: Universite Catholique de Louvain

Carlo Giovanni Camarda

Institut National d'Etudes Demographiques: INED

---

## Research Article

**Keywords:** Covid19, Mortality, Subnational, Belgium, Bayesian models

**Posted Date:** December 17th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1160099/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

## RESEARCH

# Heterogeneity in subnational mortality in the context of the COVID-19 pandemic: The case of Belgian districts

Benjamin-Samuel Schlüter<sup>1\*</sup>, Bruno Masquelier<sup>1</sup> and Carlo Giovanni Camarda<sup>2</sup>

## Abstract

**Background:** The COVID-19 pandemic has caused major shocks in mortality trends in many countries. Yet few studies have evaluated the heterogeneity of the mortality shock at the sub-national level, rigorously accounting for the different sources of uncertainty.

**Methods:** Using death registration data from Belgium, we first assess the change in the heterogeneity of subnational standardized mortality ratios in 2020, when compared to previous years. We then measure the shock of the pandemic using district-level values of life expectancy, comparing the observed and projected districts life expectancy, accounting for all sources of uncertainty (related to the life-table construction at district level and to the projection methods at country and district level). The Bayesian modelling approach makes it easy to combine the different sources of uncertainty in the assessment of the shock. This is of particular interest at a finer geographical scale characterized by high stochastic variation in annual death counts.

**Results:** The heterogeneity in the impact of the pandemic on all-cause mortality across districts is substantial, with some districts barely showing any impact whereas the Bruxelles-Capital and Mons districts experienced a decrease in life expectancy at birth of 2.24 (95% CI:1.33-3.05) and 2.10 (95% CI:0.86-3.30) years, respectively. The year 2020 was associated with an increase in mortality levels' heterogeneity at a subnational level in comparison to past years measured by both the standardized mortality ratios and the life expectancies at birth. Decisions on uncertainty thresholds have a large bearing on the interpretation of the results.

**Conclusion:** Developing sub-national mortality estimates with their uncertainty is key to understanding why certain areas have been hard hit in comparison to others.

**Keywords:** Covid19; Mortality; Subnational; Belgium; Bayesian models

## Background

The COVID-19 pandemic has halted life expectancy progress and caused an abrupt drop in life expectancy in most Western countries, with reductions of more than a year documented in 11 countries in males and 8 countries in females in 2020 [1]. As of 25 October 2021, more than 4.9 million COVID-deaths had been reported to WHO [2] and the Institute for Health Metrics and Evaluation (IHME) estimated that the pandemic has resulted in more than 11.8 million COVID-deaths in total, after factoring in the unreported deaths [3]. Belgium has been hardly hit by the pandemic with

27,463 deaths according to Sciensano [4]. The country is ranked 17<sup>th</sup> worldwide according to its deaths per 100k population [5]. Tallies of deaths directly due to the virus, despite being informative, are influenced by variations in testing capacity and definitional inconsistencies in counting COVID-19 deaths [6, 7], making comparison across countries unreliable. As a consequence, all-cause mortality has been widely used to accurately measure the pandemic's impact on mortality using excess deaths [8, 9]. This indicator is defined as the difference between the observed number of deaths and what would have been expected, had the pandemic not happened. There is a long history of using excess deaths to assess the death toll of pandemics and other extreme events [10]. Despite its consensual use [8], this measure is not perfect as it mixes the direct effect and

\* Correspondence: [benjamin-samuel.schluter@uclouvain.be](mailto:benjamin-samuel.schluter@uclouvain.be)

<sup>1</sup>IACCHOS (DEMO), Catholic University of Louvain (UCLouvain), Louvain-la-Neuve, Belgium

Full list of author information is available at the end of the article

indirect effects of the pandemic on all-cause mortality, such as those due to a shortage of health services for other medical emergencies, a reduction in traffic accidents and influenza deaths in response to the lockdown or even the economic slow down. This has led some authors to consider both excess mortality and COVID-19 deaths [1, 11]. Other researchers used changes in life expectancy as a metric to compare countries, regions or time periods [1, 12, 13]. The main advantages of this synthetic indicator are that it is a commonly used metric, insensitive to variations in age structures and is expressed in years [14]. The magnitude of the shock can be assessed similarly as for excess mortality, comparing the observed levels of life expectancy in 2020 against estimates of life expectancy had the pandemic not happened, obtained from counterfactual scenario. It is however usually assessed by comparing life expectancy in 2020 with respect to previous years [1, 13]. Such comparisons often fail to account for the increasing trend in life expectancy observed in most western countries [15] resulting in underestimates of the impact of the pandemic on mortality. Accounting for the secular trend would require projecting mortality towards 2020 and propagating the uncertainty associated with the projection into the assessment of the shock introduced by the pandemic on life expectancy.

Considering the difficulties associated with assessing the mortality shock nationally, relatively few studies have evaluated the heterogeneity of the mortality shock at a sub-national level [12, 13, 16, 17]. Yet national-level analyses might conceal important heterogeneity in how various subnational areas have been impacted. Subnational estimation is required to identify districts that might be more vulnerable to potential next waves or other future pandemics. Assessing the mortality change on a finer geographical scale comes with at least two additional challenges. First, such assessment requires to work with smaller populations where annual death counts are subject to higher stochastic variation [18]. This directly impacts the uncertainty around computed life expectancy, but this is usually not accounted for at a subnational level (see [13] as an example). Second, obtaining the counterfactual scenario in 2020 requires to project life expectancy in 2020 at the district level while projections are commonly performed at the national level.

In this article, we propose a methodology to assess how various Belgian districts have been impacted in the context of the pandemic. Districts correspond to NUTS 3 in Belgium, the finest geographical scale used by Eurostat. With this method, we address two related research questions:

- How has the pandemic impacted the subnational heterogeneity in mortality levels across Belgian districts?

- What is the magnitude of the shock at the district level, accounting for the various sources of uncertainty?

First, we estimate Standardized Mortality Ratios (SMR) in a Bayesian Hierarchical Model (BHM). The hierarchy permits to estimate the change in SMRs heterogeneity over years by assuming that, on a yearly basis, they all come from the same distribution. In addition, SMRs are not affected by the population age structure of each district and only require total deaths instead of age-specific mortality rates at the district level which reduces the noise. Second, we use life expectancy at birth to assess the magnitude of the shock on district-level mortality in 2020. This indicator summarizes the mortality intensity across all ages, without being affected by the different districts' population structures. We propose a method, inspired by Ševčíková and Raftery (2021) [19], to project district life expectancy in 2020 while accounting for all sources of uncertainty. We then compare these estimates with what has been observed, while also accounting for uncertainty in life expectancy computation due to small population size at the district level. Finally, we highlight how important it is to use different uncertainty thresholds when performing analyses at a subnational scale.

## Methods

### Data sources

We employ aggregate-level data from the Belgian national civil register data. Belgium has several administrative levels (3 regions, 10 provinces, 43 districts and 581 communes). Here we focus on districts, ranging in population size from approximately 38,000 to 1.2 millions. Our observations consists of deaths  $D_x^{d,t}$  and person-years of exposure  $N_x^{d,t}$  for district  $d$ , year  $t$  and age group  $x \in \{0, 1-4, 5-9, \dots, 90-94, 95+\}$ . Setting 95+ as the upper bound for age permits to reduce noise in the data in the oldest age range in the context of a subnational analysis. Exposure consists of mid-year populations. The mortality rate at age  $x$  for district  $d$  and year  $t$  is thus expressed as  $M_x^{d,t} = \frac{D_x^{d,t}}{N_x^{d,t}}$ .

### Standardized mortality ratios model

SMRs are widely used for indirect standardization [20] and consist of the ratio of district deaths to the number of deaths that would be expected if the district had faced the national mortality rates of the selected year. For district  $d$ , it can be expressed as follows:

$$\begin{aligned} SMR_d = \theta_d &= \frac{\sum_{x=0}^{95} M_x^d N_x^d}{\sum_{x=0}^{95} M_x^{Nat} N_x^d} = \frac{\sum_{x=0}^{95} D_x^d}{\sum_{x=0}^{95} M_x^{Nat} N_x^d} \\ &= \frac{Deaths_d}{Expected\ deaths_d}, \end{aligned}$$

where, for simplicity we removed indices for year and superscript  $Nat$  corresponds to national level. For a given year and district, deaths can then be modelled as follows

$$Deaths_{d,t}|\theta_{d,t} \sim Poisson(\theta_{d,t} \times Expected\ deaths_{d,t})$$

where  $\theta_{d,t}$  is the SMR we want to estimate for district  $d$  and year  $t$ . The Poisson distribution is of interest as it allows the uncertainty to vary with the size of the district's population. For each year in the period 2015-2020, we model  $\theta_d \forall d \in \{1, 2, 3, \dots, 43\}$  as coming from a single distribution allowing districts to borrow strength from each other and shrinking estimates for smaller districts towards the mean. We further assume that all yearly SMR's average,  $\mu_{\theta t}$ , share the same distribution, centered in  $M_{\theta}$ . The standard deviations of SMRs,  $\sigma_{\theta t}$ , are estimated independently for each year. We use the following weakly informative priors [21]

$$\begin{aligned}\theta_{d,t} &\sim Normal^+(\mu_{\theta t}, \sigma_{\theta t}) \\ \mu_{\theta t} &\sim Normal^+(M_{\theta}, 1) \\ M_{\theta} &\sim Normal^+(1, 1) \\ \sigma_{\theta t} &\sim Normal^+(1, 1)\end{aligned}$$

The posterior distribution of  $\sigma_{\theta t}$  for each year is of particular interest as it consists of the estimated standard deviation of the SMR distribution. Hence, comparing its median and credible interval permits to assess the heterogeneity in SMR values in 2020 in comparison to previous years.

The SMR being a relative measure of mortality level relatively to the national level, we also need an absolute measure of the shock in the context of the COVID-19 pandemic.

**Life expectancy as a measure of the shock on mortality**  
In order to assess how mortality rates have been impacted during the COVID-19 pandemic at the district level, we will use life expectancy at birth ( $e^0$ ) computed without stratifying by sexes, referred simply as life expectancy in the text. To do so, we need two pieces of information. First, a projection of  $e^0$  at the district level in 2020 is needed. This forecast value can be seen as a counterfactual scenario, had COVID-19 pandemic not happened. Second, we need the observed  $e^0$  at the district level in 2020 obtained with standard life table techniques. We account for the uncertainty around the observed life expectancy values using Chiang's method [22]. Computing the difference between the projected and observed estimates, combining their

associated uncertainty, allows us to adequately assess the magnitude of the shock on district-level mortality in 2020.

#### *Probabilistic projections of district life expectancy*

The methodology that we developed to obtain projections at the district level builds on recent work from Ševčíková and Raftery (2021) [19].

Let  $e_{Nat,t}^0$  national life expectancy at year  $t$ . We define life expectancy in district  $d$  at year  $t$ ,  $e_{d,t}^0$ , as a function of the national level

$$e_{d,t}^0 = e_{Nat,t}^0 + \alpha_{d,t},$$

where  $\alpha_{d,t}$  can be viewed as the gap between district and national life expectancy at year  $t$  and it will be treated as a stochastic variable.

We first project life expectancy in 2020 at the national level using the Lee-Carter model [23] estimated over the period 1991-2019. Estimating the model in a Bayesian framework allows to produce a set of trajectories from its posterior predictive distribution,  $e_{Nat,2020,i}^0$  for trajectory  $i$ .

District-specific deviations is model using a Bayesian hierarchical linear model:

$$\alpha_{d,t}|\beta_d, \sigma_{\alpha d} \sim Normal(\beta_{0,d} + \beta_{1,d} \cdot t, \sigma_{\alpha d}),$$

over the period 1991-2019. The linearity over time approximates quite well the data (Additional Fig. 1). This can be explained by the fact that over the period 1991-2019, the health performance of a district, in comparison to the national performance, is usually stable over time, or showing a constant trend. The hierarchical structure of the model allows that betas for smaller districts are informed by all other districts, leading to shrinkage in estimation. Weakly informative priors defined for intercepts ( $\beta_{0,d}$ ) and slopes ( $\beta_{1,d}$ ) can be found in the Additional File 1. We tested the sensitivity of our results to the priors.

Variance for  $\alpha_{d,t}$  denoted by  $\sigma_{\alpha d}$  is allowed to vary across districts because fluctuations in  $\alpha_{d,t}$  depend on district's population size (Additional Fig. 2). This uncertainty will thus be propagated into the projection of each district.

Finally, we simulate the  $i$ -th trajectory of  $\alpha$  for district  $d$  in year 2020,  $\alpha_{d,2020,i}$ , as follows:

$$\alpha_{d,2020,i} \sim Normal(\beta_{0,d,i} + \beta_{1,d,i} \cdot 2020, \sigma_{d,i}).$$

This step allows to account for both parameters uncertainty and variation in the data. The  $i$ -th trajectory from the posterior predictive distribution of district life expectancy in 2020 is given by

$$e_{d,2020,i}^0 = e_{Nat,2020,i}^0 + \alpha_{d,2020,i},$$

where  $e_{Nat,2020,i}^0$  is the  $i$ -th trajectory from the national life expectancy projection step. Consequently we account for the uncertainty in the national projection as well as in the estimation of the district-specific deviations.

#### Observed life expectancy and its uncertainty in 2020

In order to reflect uncertainty in life-table computation at the district scale, we applied Chiang method [22]. Let  ${}_nq_x$  and  ${}_nD_x$  be the probability of dying from the 2020 observed life-table and observed deaths in a given district, respectively. For each age-group  $[x, x + n]$ , we assume that deaths are realizations from a binomial distribution where the number of trials is equal to the people at risk computed as  $\frac{{}_nD_x}{{}_nq_x}$  and the success probability for each trial is the probability of dying  ${}_nq_x$ . We can thus simulate an  $i$ -th series of death counts as follows:

$${}_nD_{x,i} \sim \text{Binomial}\left(\frac{{}_nD_x}{{}_nq_x}, {}_nq_x\right).$$

We repeat this procedure 4,000 times for each district obtaining 4,000 associated life-tables and life expectancies. We combine these outcomes with 4,000 posterior trajectories from the projection exercise. This allows us to compute, for each district, a distribution of differences  $\Delta e^0$  between observed and counterfactual life expectancy in 2020 which accounts for uncertainty coming from both observed and forecast life expectancy in 2020. Quantities such as median and quantiles of these distributions provide point estimate and credible intervals for the difference between observed and projected districts'  $e^0$  and hence, an assessment of the shock on mortality.

## Results

### Standardized mortality ratios

Fig. 1 shows the annual estimates of SMR for each district over the period 2015-2020, with their associated 95% credible intervals. A SMR equal to one means that the observed deaths in the district equal the expected deaths if the population in the district had experienced the national mortality rates, while a SMR above (below) one reflects more (less) deaths than expected when facing national mortality. At the district level, there is substantial variation from year to year in SMR estimates. Heterogeneity in subnational SMR is present before the year 2020, reflecting disparity in mortality level within Belgium. However, the year 2020 seems to be characterized by higher variation in SMR with some district having higher SMR values than in previous years (Mons 1.34 (95% CI:1.30-1.39), Charleroi 1.30 (95% CI: 1.26-1.33), Liège 1.26 (95%

CI: 1.23-1.28), Tournai-Mouscron 1.23 (95% CI: 1.19-1.27) and Bruxelles-Capitale 1.09 (95% CI: 1.07-1.11)) and others having lower SMRs than in previous years (Bruges 0.82 (95% CI: 0.79-0.85), Maaseik 0.83 (95% CI: 0.79-0.86), Eeklo 0.84 (95% CI: 0.79-0.89), Louvain 0.87 (95% CI: 0.84-0.89), Hal-Vilvorde 0.87 (95% CI: 0.85-0.89), Malines 0.89 (95% CI: 0.86-0.92) and Anvers 0.93 (95% CI: 0.92-0.95)). We plot these estimates and their credible intervals width on two maps for the year 2020 (see Fig. 2 and Fig. 3). We see a clear North-South divide in mortality level within the country, not particular to the year 2020. Fig. 4 presents posterior distributions of  $\sigma_\theta$  for each year in the 2015-2020 period. It allows to assess the evolution in subnational SMR's standard deviation over the years. In 2020, the posterior distribution has clearly shifted to the right, reaching 0.15 (95% CI: 0.12-0.19). This indicates a higher heterogeneity in subnational mortality levels in comparison to previous years.

In the context of subnational analyses, ranking of units matters. Posterior ranking probabilities permit to rank districts according to their posterior expected rank while reflecting uncertainty in the ranking [24]. Fig. 5 presents posterior ranking probabilities for each district in 2020. Each square in the figure represents the posterior probability that district  $d$  has rank  $k \in \{1, 2, 3, \dots, 43\}$ . The darker the color the more certain position in the rank can be expected for a given district, with the sum over row equal to 1. The heat map shows a clear diagonal, allowing to conclude that ranking according to estimated SMR is rather certain, with Mons and Charleroi being most likely the districts experiencing the highest mortality in 2020, while Bruges, Maseik and Eeklo experienced the lowest rates with respect to the national mean.

Difference between projected and observed districts'  $e^0$  Moving on to our absolute measure of the shock, Fig. 6 shows the difference between the life expectancy observed in 2020 and the 2020 counterfactual projection with its associated credible intervals, for each district. Three credible intervals are presented: the red intervals correspond to taking the difference not accounting for uncertainty in life-table computation for the observed rates; the blue intervals account for uncertainty in life-table computation and set the credible interval at 80%; the yellow intervals account for uncertainty in life-table computation and set the credible interval at 95%. As an example, with respect to our counterfactual scenario, Bruxelles-Capitale, Mons, Arlon and Liège districts have experienced a drop in life expectancy of 2.24 (95% CI:1.33-3.05), 2.10 (95% CI:0.86-3.30), 2.0 (95% CI:0.42-3.6) and 2.0 (95% CI:1.0-2.9) years, respectively. However, the width of the credible intervals vary according to the district's population size,

the sources of uncertainty considered and the level of uncertainty chosen. For instance in the case of Arlon, we see that not accounting for uncertainty in life-table computation based on the observed rates overestimates the precision of the shock assessment (difference between blue and red intervals). This can be observed for all districts having relatively small population sizes (Ypres, Bastogne, Furnes, Dixmude, Marchen-Famenne, Neufchâteau and Virton). It is also worth noting that the width of the 95% credible intervals for districts such as Furnes or Bastogne (with population of 61,7 thousands and 49 thousands inhabitants, respectively) are close to three years of life expectancy. According to the most conservative credible interval (depicted in yellow), fourteen out of forty-three districts have had a statistically significant drop in their life expectancy in comparison to the projection. The magnitude of the drops differs across districts, however. When considering point estimates, all districts but one had a level of life expectancy in 2020 that was lower than expected based on continuing the secular decline in mortality had the pandemic not happened. We map the loss of life estimates and their 95% credible interval width on Fig. 7 and Fig. 8.

It is important to note that the amount of uncertainty we accept to tolerate is highly determinant. Accounting for both sources of uncertainty and setting 80% for credible intervals leads to conclude that only eleven districts did not experience a significant drop in their life expectancy in 2020 compared to the projections. When only considering the uncertainty around the projection, all but 10 districts experiences a significant drop. In addition, the width of the credible intervals vary substantially depending on the uncertainty scenario chosen and the district's population size. This calls for an uncertainty assessment when computing subnational life expectancy at this level of spatial disaggregation.

Ranking districts according to their posterior expected differences between observed and projected life expectancies,  $\Delta e^0$ , leads to high uncertainty in the ranking (see Fig. 9). This is a logical consequence of accounting for projection uncertainties in addition to uncertainty associated to life expectancy computation at the district level.

We also explored the life expectancy's heterogeneity in the year 2020 relatively to other years to consolidate our observation on the increased variation in subnational mortality levels for the year 2020. For each year and district combination, we obtained its associated 4,000 simulated life expectancies and computed across district the associated standard deviation. In each year, we ended up with 4,000 bootstrapped  $\sigma_{e^0}$ . We transformed that into a density of  $\sigma_{e^0}$  over the

years (Fig. 10). In line with previous observation on SMR, we see that the density of  $\sigma_{e^0}$  in 2020 has moved to the right, with its median equal to 1.72 (95% CI:1.63-1.81) years.

## Discussion

Our study proposed a methodology to assess how the mortality level of various subnational units have been impacted in the context of the COVID-19 pandemic in 2020 using vital registration data. Our application is on Belgian districts but the proposed method could be applied to other subnational contexts.

We first computed SMRs over the period 2015-2020 and explored how their variation changed in 2020 compared to previous years. In addition, we compared observed with projected district life expectancy in 2020, accounting for different sources of uncertainty. In 2020, three districts had deaths counts that were more than 20% higher than expected when assuming national mortality. On the other extreme, five districts had 10% less deaths than what would be expected if facing the national mortality rates. We found that heterogeneity in mortality level was present before the pandemic, with a clear North-South divide which has been documented earlier [25], [26], but the shock on mortality in 2020 has increased the SMRs variation with respect to previous years. This is also confirmed by a higher heterogeneity in districts' life expectancy at birth in the year 2020 relatively to previous years. When working at a subnational level, the goal of assessing which units have been the most affected is of policy relevance. Having that in mind, the SMR metric showed to be suitable to rank districts as it leads to greater certainty in the ranking.

For a given year, the SMR is expressed relatively to the national mortality level and is thus not suitable to quantify the magnitude of the shock on mortality in 2020. To reach that goal, comparing the projected life expectancy with what has been observed allows to assess the magnitude of the shock at the district level, while accounting for the increasing trend in life expectancy. This exercise requires to account for two sources of uncertainty; the uncertainty coming from projecting life expectancy at the district level in 2020, and the uncertainty in life table estimates based on age-specific mortality rates observed at the district level. The level of confidence we are willing to tolerate is thus determinant for interpreting our results. Indeed, the width of the intervals highly depends on the sources of uncertainty accounted for and the confidence level chosen. Choosing intervals of 95% and accounting for both the observed and projected uncertainties in life expectancies, we concluded that fourteen districts experienced a significant drop in their

life expectancy in comparison to their projections. The biggest drops were 2.24 and 2.10 years. However, when setting interval threshold at 80% and not accounting for life-table uncertainty, thirty-three districts out of forty-three experienced a significant drop. These order of magnitude are higher than the ones found by Aburto and colleagues (2021) [1] at the national level. In their study, they showed that Belgium saw a drop of approximately 1 year in life expectancy at birth for both male and female between 2019 and 2020. This difference in magnitude is expected as we are looking at the most affected districts within Belgium. Looking across districts, it is clear that the shock on mortality has been highly heterogeneous within the country, with some districts not showing any significant difference between observed and projected life expectancy. Other research also found important difference in the magnitude of the shock within Spain [16], [12], Sweden [17], France [27] and Mexico [13].

The observed subnational heterogeneity of the shock on mortality might come from an unequal distribution of specific characteristics across districts. An important indicator that might play a role is the number of care homes present within a district. Indeed, care homes have been strongly affected by the first wave [28] as they were not sufficiently prepared and host a majority of frail individuals. This number might substantially increase the pandemic's death toll at the district level. The intensity of mobility within a district might also lead to higher transmission of the virus [29], with district hosting more economic activities or international organizations being more at risk. This might be a factor explaining the fact that Bruxelles-Capitale has seen a big increase in its mortality level. The district's pre-pandemic health profile might also be an important factor [30]. Moreover, income at the individual level is a factor known to be closely related to health outcomes. Decoster and colleagues [31] showed that there was a significant and negative income gradient in excess mortality during the first wave in Belgium for the elderly. Poor housing conditions and higher occupational exposures have been shown to be the most likely mechanisms causing the higher burden of COVID-related mortality for the poor in France [32]. In Italy, Basellini & Camarda [33] found that regional differences could be explained by heterogeneity in intergenerational co-residence, number of ICU beds per capita, and timing of the outbreak. These factors might also have played a role in the Belgian context.

The estimated drops in life expectancy and their uncertainty can be used by other researchers to perform aggregate analysis on the association between the magnitude of the shock and various risk-factors at the district level. Our study also emphasizes how uncertainty

level that we define and are ready to tolerate impact the results and interpretations in the context of subnational analyses. Ideally, various uncertainty thresholds should be presented in the context of subnational research.

Our study is subject to several limitations. First, we did not stratify by gender for two reasons. On the one hand, we are already working with small deaths counts and hence, stratifying further by sex might increase the noisiness in the data. On the other hand, we were mostly interested in the magnitude of the shock on district-level mortality which can be measured by the overall life expectancy for both sexes. Further research could investigate how the sex ratios of mortality were affected by the pandemic at the subnational level. A second limitation is that we used a simple linear model for the temporal evolution of the difference across districts with national life expectancy over time. The explanatory data analysis showed that this linear relationship was appropriate for fitting the data. However, this assumption is only valid if we project over a really short time horizon of 1 or 2 years. If the goal is to obtain a projection over a longer time horizon, we do not advise to use our method and refer to the methodology developed by Ševčíková and Raftery (2021) [19]. Our method had the advantage to indirectly account for both the temporal trend with respect to the national life expectancy and the projection uncertainty at the district level. Third, period life expectancy as an indicator of the shock on mortality is not perfect. Indeed, it is built on the concept of a hypothetical cohort of individuals experiencing, throughout their entire life the age-specific mortality rates seen during 2020. But no real cohort will experience this mortality shock over a long period and it is likely that life expectancy in 2021 will bounce back, possibly reaching higher levels than those expected in 2021 based solely on the secular trend. This would then be a sign that we have experienced a harvesting effect [34] in 2020, indicating that the pandemic has advanced the deaths of frail individuals in time.

## Conclusions

Our study proposed a methodology to assess the impact of the COVID-19 pandemic at the Belgian subnational level. First, using SMR, we looked at the effect of the pandemic on the heterogeneity of mortality across districts. The year 2020 showed a widening of its SMR distribution, being a sign of a higher heterogeneity in mortality levels within Belgium compared to past years. Then, we compared projected and observed life expectancy at the district level. Using the projections as a counterfactual scenario permits to account for the increasing trend in life expectancy. When

estimating the difference between these two measures, we fully account for various sources of uncertainty. We showed that the shock has been highly heterogeneous within the country. Some districts experienced a drop of more than two years compared to their projection while others did not experienced any significant differences. The uncertainty we are ready to tolerate has a huge impact on what we consider to be a significant shock or not. This reinforces the need to carefully account for uncertainty and present various uncertainty levels in the context of subnational analyses.

## Abbreviations

WHO: World Health Organization; IHME: Institute for Health Metrics and Evaluation; NUTS: Nomenclature of territorial units for statistics; BHM: Bayesian hierarchical model; SMR: Standardized mortality ratio

## Declarations

### Ethics approval and consent to participate

Ethical concerns regarding the free and informed consent of participants are not relevant here as the study participants are deceased persons. Ethical clearance was not sought because the dataset used for this study does not contain any identifiable information of study participants.

### Consent for publication

All authors had access to the final manuscript and agreed to submit the final version.

### Availability of data and materials

The data are publicly available from Statbel, the Belgian statistical office and can be downloaded from <https://statbel.fgov.be/en/open-data>. The granularity of age groups for the death counts was higher in our data in comparison to what is found on STATBEL open data portal. This more disaggregated data was obtain in the context of the CAUSINEQ project, financed by BELSPO.

### Competing interests

The authors declare that they have no competing interests

### Funding

This work was supported by a FSR research grant from Université Catholique de Louvain (UCLouvain): IACS FSR19 MASQUE, Mr. Benjamin-Samuel Schlüter

### Authors' contributions

BSS analyzed the data and wrote the first of the manuscript. BM and CGC provided guidance on the analysis and revised the manuscript.

### Acknowledgements

Not applicable

### Author details

<sup>1</sup>IACCHOS (DEMO), Catholic University of Louvain (UCLouvain), Louvain-la-Neuve, Belgium. <sup>2</sup>Mortality, Health and Epidemiology, Institut National d'Etudes Démographiques (INED), Paris, France.

## References

- Aburto, J.M., Schöley, J., Kashnitsky, I., Zhang, L., Rahal, C., Missov, T.I., Mills, M.C., Dowd, J.B., Kashyap, R.: Quantifying impacts of the covid-19 pandemic through life-expectancy losses: a population-level study of 29 countries. *International journal of epidemiology* (2021). <https://covid19.who.int> Accessed 5 Jul 2021
- World Health Organisation: WHO Coronavirus Dashboard (2020). <https://covid19.who.int> Accessed 5 Jul 2021
- Institute for Health and Metrics: COVID-19 Projections (2021). <https://covid19.healthdata.org/> Accessed 15 Nov 2021
- Sciensano: COVID-19 Epidemiological situation (2021). <https://datastudio.google.com/embed/reporting/c14a5cfc-cab7-4812-848c-0369173148ab/page/QTSKB> Accessed 10 Dec 2021
- Johns Hopkins Coronavirus Resource Center: Johns Hopkins University Mortality Analyses - COVID-19 (2020). <https://coronavirus.jhu.edu/data/mortality> Accessed 25 Oct. 2021
- Karlinsky, A., Kobak, D.: The world mortality dataset: Tracking excess mortality across countries during the covid-19 pandemic. *medRxiv* (2021)
- García, J., Torres, C., Barbieri, M., Camarda, C.G., Cambois, E., Caporali, A., Meslé, F., Poniakina, S., Robine, J.-M.: Differences in covid-19 mortality: Implications of imperfect and diverse data collection systems. *Population* **76**(1), 35–72 (2021)
- Beaney, T., Clarke, J.M., Jain, V., Golestaneh, A.K., Lyons, G., Salman, D., Majeed, A.: Excess mortality: the gold standard in measuring the impact of covid-19 worldwide? *Journal of the Royal Society of Medicine* **113**(9), 329–334 (2020)
- Leon, D.A., Shkolnikov, V.M., Smeeth, L., Magnus, P., Pechholdová, M., Jarvis, C.I.: Covid-19: a need for real-time monitoring of weekly excess deaths. *The Lancet* **395**(10234), 81 (2020)
- Karlinsky, A., Kobak, D.: Tracking excess mortality across countries during the covid-19 pandemic with the world mortality dataset. *Elife* **10**, 69336 (2021)
- i Arolas, H.P., Acosta, E., López-Casasnovas, G., Lo, A., Nicodemo, C., Riffe, T., Myrskylä, M.: Years of life lost to covid-19 in 81 countries. *Scientific reports* **11**(1), 1–6 (2021)
- Trias-Llimós, S., Riffe, T., Bilal, U.: Monitoring life expectancy levels during the covid-19 pandemic: Example of the unequal impact of the first wave on spanish regions. *PloS one* **15**(11), 0241952 (2020)
- García-Guerrero, V.M., Beltrán-Sánchez, H.: Heterogeneity in excess mortality and its impact on loss of life expectancy due to covid-19: Evidence from mexico. *Canadian studies in population*, 1–36 (2021)
- Luy, M., Di Giulio, P., Di Lego, V., Lazarevič, P., Sauerberg, M.: Life expectancy: frequently used, but hardly understood. *Gerontology* **66**(1), 95–104 (2020)
- Oeppen, J., Vaupel, J.W.: Broken limits to life expectancy. *American Association for the Advancement of Science* (2002)
- Saavedra, P., Santana, A., Bello, L., Pacheco, J.-M., Sanjuán, E.: A bayesian spatio-temporal analysis of mortality rates in spain: application to the covid-19 2020 outbreak. *Population Health Metrics* **19**(1), 1–10 (2021)
- Modig, K., Ahlbom, A., Ebeling, M.: Excess mortality from covid-19: weekly excess death rates by age and sex for sweden and its most affected region. *European journal of public health* **31**(1), 17–22 (2021)
- Alexander, M., Zagheni, E., Barbieri, M.: A flexible bayesian model for estimating subnational mortality. *Demography* **54**(6), 2025–2041 (2017)
- Ševčíková, H., Raftery, A.E.: Probabilistic projection of subnational life expectancy. *Journal of Official Statistics* **37**(3), 591–610 (2021)
- Preston, S., Heuveline, P., Guillot, M.: *Demography: measuring and modeling population processes*. 2001. Malden, MA: Blackwell Publishers (2000)
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B.: *Bayesian Data Analysis*, (1995)
- Chiang, C.L.: Life table and its applications. In: *Life Table and Its Applications*, pp. 316–316 (1984)
- Lee, R.D., Carter, L.R.: Modeling and forecasting us mortality. *Journal of the American statistical association* **87**(419), 659–671 (1992)
- Bryant, J., Zhang, J.L.: *Bayesian Demographic Estimation and Forecasting*. Chapman and Hall/CRC, ??? (2018)
- Renard, F., Tafforeau, J., Deboosere, P.: Mapping the cause-specific premature mortality reveals large between-districts disparity in belgium, 2003–2009. *Archives of public health* **73**(1), 1–18 (2015)
- Eggerickx, T., Sanderson, J.-P., Vandeschrick, C.: Mortality in belgium from nineteenth century to today. *Quetelet Journal* **8**(2), 7–59 (2020)
- Gaudart, J., Landier, J., Huiart, L., Legendre, E., Lehot, L., Bendiane, M.K., Chiche, L., Petitjean, A., Mosnier, E., Kirakoya-Samadoulougou, F., et al.: Factors associated with the spatial heterogeneity of the first wave of covid-19 in france: a nationwide geo-epidemiological study.

- The Lancet Public Health 6(4), 222–231 (2021)
28. Hardy, O.J., Dubourg, D., Bourguignon, M., Dellicour, S., Eggerickx, T., Gilbert, M., Sanderson, J.-P., Scohy, A., Vandael, E., Decroly, J.-M.: A world apart: Levels and determinants of excess mortality due to COVID-19 in care homes: The case of the Belgian region of Wallonia during the spring 2020 wave. *Demographic Research* 45(33), 1011–1040 (2021). doi:[10.4054/DemRes.2021.45.33](https://doi.org/10.4054/DemRes.2021.45.33)
  29. Basellini, U., Alburez-Gutierrez, D., Del Fava, E., Perrotta, D., Bonetti, M., Camarda, C.G., Zagheni, E.: Linking excess mortality to mobility data during the first wave of covid-19 in england and wales. *SSM-Population Health* 14, 100799 (2021)
  30. Dowd, J.B., Andriano, L., Brazel, D.M., Rotondi, V., Block, P., Ding, X., Liu, Y., Mills, M.C.: Demographic science aids in understanding the spread and fatality rates of covid-19. *Proceedings of the National Academy of Sciences* 117(18), 9696–9698 (2020)
  31. Decoster, A., Minten, T., Spinnewijn, J.: The income gradient in mortality during the covid-19 crisis: evidence from belgium. *The Journal of Economic Inequality*, 1–20 (2021)
  32. Brandily, P., Brébion, C., Briole, S., Khoury, L.: A poorly understood disease? the unequal distribution of excess mortality due to covid-19 across french municipalities. *NHH Dept. of Economics Discussion Paper* (15) (2020)
  33. Basellini, U., Camarda, C.G.: Explaining regional differences in mortality during the first wave of covid-19 in italy. *Population Studies*, 1–20 (2021)
  34. Islam, N., Shkolnikov, V.M., Acosta, R.J., Klimkin, I., Kawachi, I., Irizarry, R.A., Alicandro, G., Khunti, K., Yates, T., Jdanov, D.A., et al.: Excess deaths associated with covid-19 pandemic in 2020: age and sex disaggregated time series analysis in 29 high income countries. *bmj* 373 (2021)

#### Figures

**Figure 1** Estimated standardized mortality ratios with 95% credible interval by Belgian district, 2015-2020

**Figure 2** Estimated standardized mortality ratios in 2020

**Figure 3** Width of SMRs 95% credible intervals in 2020

**Figure 4** Density of the posterior draws for the standard mortality ratios' standard deviation, 2015-2020

**Figure 5** Posterior ranking probabilities of standardized mortality ratios, 2020

**Figure 6** Differences between observed and projected  $e^0$  at district level with uncertainty according to three scenarii, 2020

**Figure 7** Estimated loss of life expectancy in 2020

**Figure 8** Width of loss of life expectancy 95% credible intervals in 2020

**Figure 9** Posterior ranking probabilities of the differences between observed and projected  $e^0$  at district level, 2020

**Figure 10** Density of the simulated standard deviation of subnational life-expectancies, 2015-2020

#### Additional Files

Additional Fig. 1 — Linearity of  $\alpha_{d,t}$  over time

The figure shows the evolution of  $\alpha_{d,t} = e_{d,t}^0 - e_{Nat,t}^0$  over time for three districts of various population sizes, namely Liège, Nivelles and Marche-en-Famenne. We added the expected fit from a simple linear regression estimated individually for each district on the years 1991 to 2019.

**Figure 11**  $\alpha_{d,t}$  over time, Liège

**Figure 12**  $\alpha_{d,t}$  over time, Nivelles

**Figure 13**  $\alpha_{d,t}$  over time, Marche-en-Famenne

Additional File 1 — Priors in BHM for  $\alpha_{d,t}$

Here below are the priors set to model  $\alpha_{d,t} \sim Normal(\beta_{0,d} + \beta_{1,d} \cdot t, \sigma_d)$ .

$$\beta_{0,d} \sim Normal(\mu_{\beta_0}, \sigma_{\beta_0})$$

$$\beta_{1,d} \sim Normal(\mu_{\beta_1}, \sigma_{\beta_1})$$

$$\sigma_d \sim Normal^+(\mu_\sigma, \sigma_\sigma)$$

where priors for  $\mu_{\beta_0}$  and  $\mu_{\beta_1}$  are  $Normal(0, 1)$  and for  $\mu_\sigma, \sigma_{\beta_0}, \sigma_{\beta_1}, \sigma_\sigma$  are  $Normal^+(0, 1)$ .

Additional Fig. 3 — Prior predictive checks for  $\alpha_{d,t}$

In the figure below is the prior predictive distribution of possible  $\alpha_{d,t}$  values according to our priors. The range is much higher than what is observed in the data and hence is not informative.

**Figure 14** Prior predictive distribution of  $\alpha_{d,t}$

Additional Fig. 4 — Posterior predictive checks for  $\alpha_{d,t}$

In the figure below is the posterior predictive distribution of possible  $\alpha_{d,t}$  obtain for 11 trajectories. The simulated data does a good job in reproducing the data.

**Figure 15** 11 simulations of  $\alpha_{d,t}$  from its posterior predictive distribution

# Figures

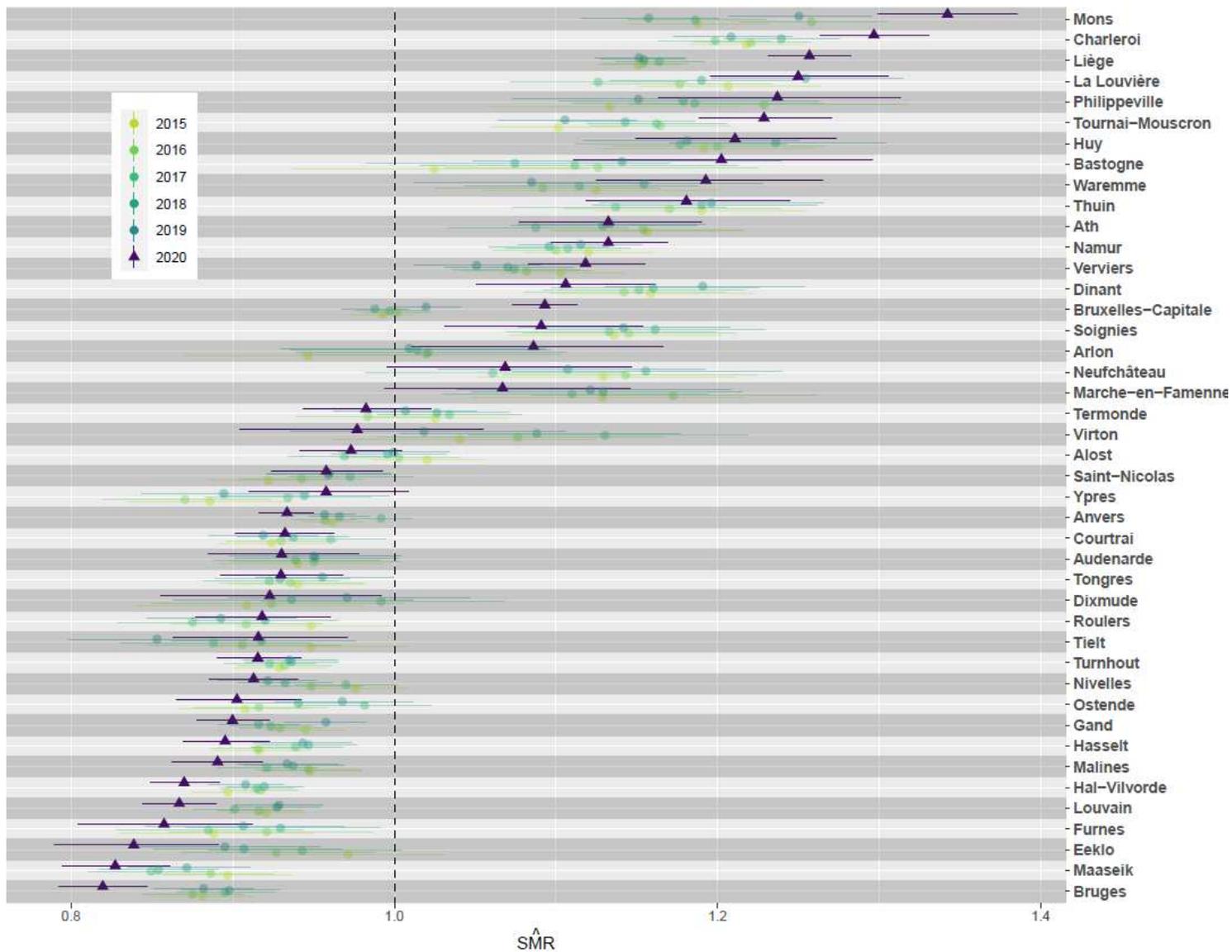


Figure 1

Estimated standardized mortality ratios with 95% credible interval by Belgian district, 2015-2020

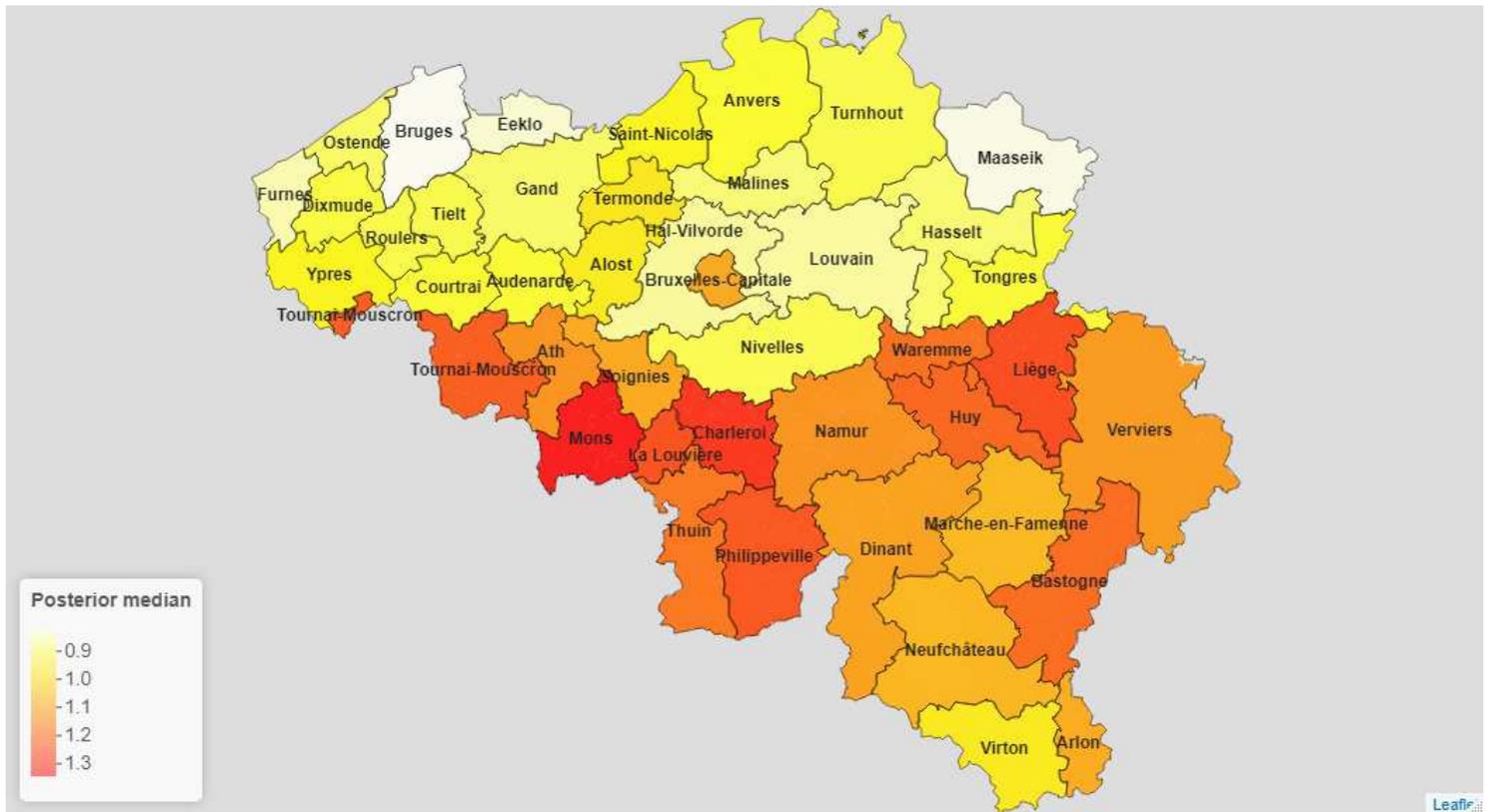


Figure 2

Estimated standardized mortality ratios in 2020

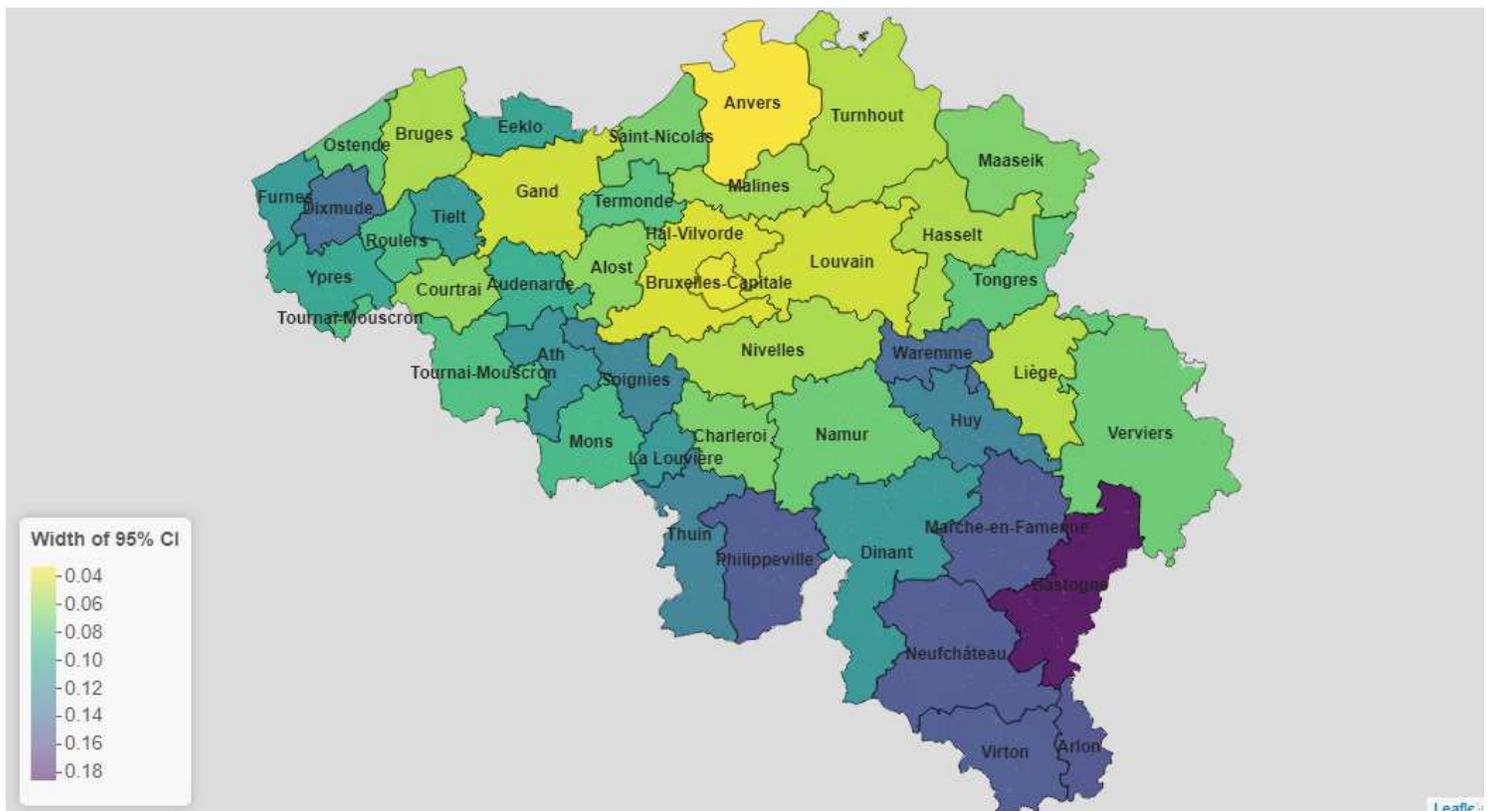


Figure 3

Width of SMRs 95% credible intervals in 2020

Figure 4

Density of the posterior draws for the standard mortality ratios' standard deviation, 2015-2020

Figure 5

Posterior ranking probabilities of standardized mortality ratios, 2020

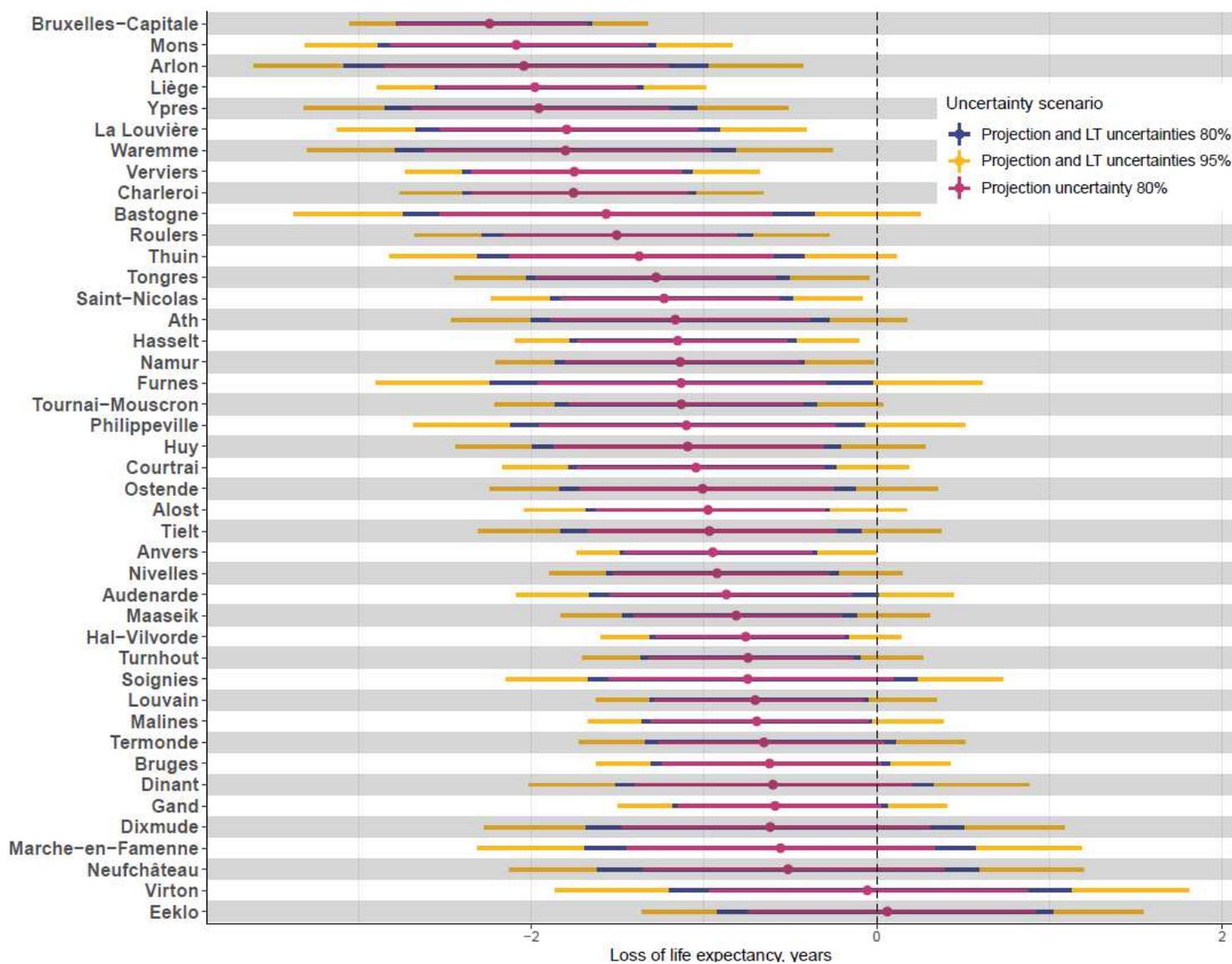
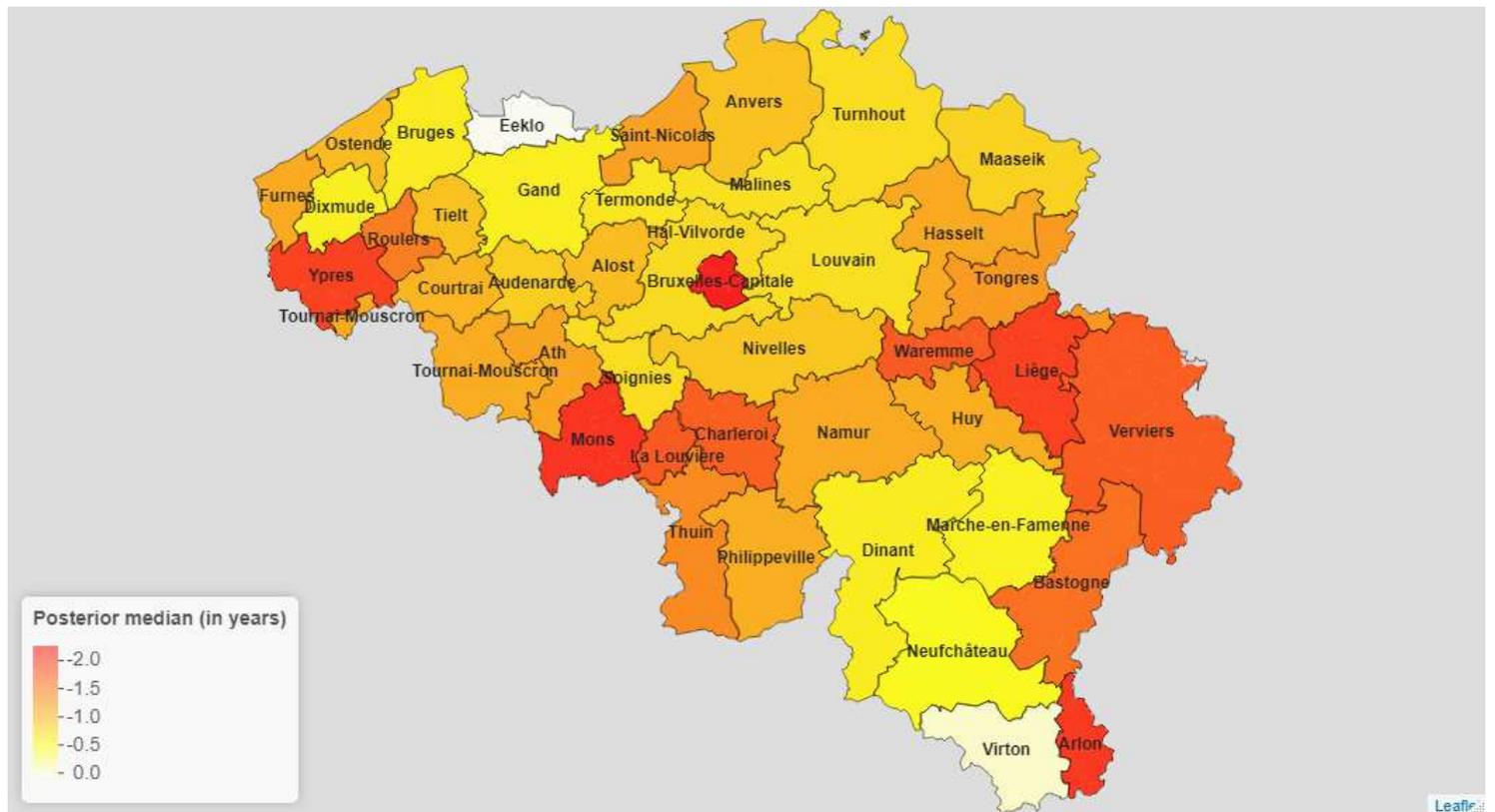


Figure 6

Differences between observed and projected  $e^0$  at district level with uncertainty according to three scenarii, 2020



**Figure 7**

Estimated loss of life expectancy in 2020

**Figure 8**

Width of loss of life expectancy 95% credible intervals in 2020

**Figure 9**

Posterior ranking probabilities of the differences between observed and projected  $e^0$  at district level, 2020

**Figure 10**

Density of the simulated standard deviation of subnational life-expectancies, 2015-2020

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [fig11.png](#)
- [fig12.png](#)
- [fig13.png](#)
- [fig14.png](#)
- [fig15.png](#)