PUTTING SPAIN BACK IN SPANISH INFLUENZA: Quantifying the timing and mortality impact in Madrid of the 1918-1921 pandemic through spatial, demographic, and social lenses

Laura Ann Cilek

Universitat Autònoma
de Barcelona

PUTTING SPAIN BACK IN SPANISH INFLUENZA:

Quantifying the timing and mortality impact in Madrid of the 1918-1921 pandemic through spatial, demographic, and social lenses

Candidate:

Laura Ann Cilek

Supervisors:
Dr. Diego Ramiro Fariñas
Prof. Dr. Gerardo Chowell Puente

Tutor:
Dr. Albert Esteve Pálos

Instituto de Economía, Geografía y Demografía (IEGD)
Centro de Ciencias Humanas y Sociales (CCHS)
Consejo Superior de Investigaciones Científicas (CSIC)

Department of Geography, Universitat Autònoma de Barcelona
Centre d'Estudis Demogràfics (CED)

2019

DEVOLVIENDO ESPAÑA A LA GRIPE ESPAÑOLA:

Cuantificación del impacto temporal y de mortalidad en Madrid de la pandemia de 1918-1921 bajo una perspectiva espacial, demográfica y social

RETORNANT ESPANYA A LA GRIP ESPANYOLA:

Quantificació de l'impacte temporal i de la mortalitat a Madrid de la pandèmia de 1918-1921 mitjançant una perspectiva espacial, demogràfica i social.

Candidate:

Laura Ann Cilek

Supervisors:
Dr. Diego Ramiro Fariñas
Prof. Dr. Gerardo Chowell Puente

Tutor:
Dr. Albert Esteve Pálos

2019

Acknowledgements

I would like to thank several people and institutions for their role in the completion of this thesis. Even before I began this research, I have had the support of countless individuals who have encouraged, inspired, and supported me. My parents, Bob and Mary Cilek, have always been my number one fans, providing strength and guidance throughout my life. Other family and friends outside of the academic realm have also been there to listen to my problems and lend an ear of support.

As my studies progressed but before I began my PhD, professors and classmates became my new support system. During my master's degree at Florida State, my interest in Demography was spiked and the flames of learning stoked, fueling a fire to continue my education. I felt empowered by my professors and classmates as my eyes were opened to a new world of scientific inquiry and investigation. I continued to develop knowledge and (not only professional) friendships during the subsequent year, when I partook in the European Doctoral School of Demography. Encountering new demographic methods and concepts was substantially easier due to the camaraderie of my colleagues, with whom together we sought to succeed in the program. Days and nights of studying were not for want of frustration and feelings of helplessness, but they were also filled with laughter, happiness, and delicious Italian ~~food~~ wine.

During the last three years, I have gained even more sources of counsel. My supervisors and co-authors Diego Ramiro Fariñas and Gerardo Chowell have guided the development of myself as a student and my thesis, and they have spurred my understanding of not only social epidemiology, but also the more general world of academia. Beatriz Echeverri Dávila, a co-author on the paper associated with chapter five, was also always happy to learn about the progress made on individual research projects within the following document; she was always supportive of the statistical approaches taken to verify what she had discovered during her own years spent tirelessly researching Spanish Influenza in Madrid and Spain. I would also like to acknowledge my other colleagues at CSIC, most specifically Yolanda Casado Ruiz (another co-author on chapter five) and Pilar Aceituno Nieto, for being wonderful office mates, our shared coffee breaks, and of course, my free Spanish lessons. Acting as my tutor, Albert Esteve also provided the scholastic "home" for the thesis at the Autonomous

University of Barcelona (UAB) and the Center for Demography Studies (CED). Socorro Sancho and Ángela Martínez-Carrasco Martínez put forth every effort to help navigate the seemingly unending trail of Spanish bureaucracy.

In addition to the aforementioned members of the LONGPOP network, the other Early Stage Researchers provided a support system during and between trainings, meetings, and secondments–I hope our network of newly christened researchers continues to bear fruitful work in the future. Also within LONGPOP, Chris Dibben and Lee Williamson were vital to the completion of my studies, as they supervised and facilitated my research stays at the University of Edinburgh, ensuring a productive three months of work so that I could complete the International Doctoral Research Component of the degree.

The international mention would also not be complete without the external reviews of Michel Oris of the University of Geneva and Francesco Scalone from the University of Bologna, to whom I am thankful for taking the time to read and report on this body of work. The academic researchers making up my final dissertation defense committee must also be recognized for agreeing to and taking the time to read and evaluate the compilation of research articles presented here.

And of course, thanks to Mathias who has acted as a sounding board and a huge pillar of support for me both before and during the last three years, as we braved this lonesome journey together.

Finally, though I've mentioned some of the people involved above, I would like to thank the European Union and their Horizon 2020 research and innovation programme for my source of funding, the LONGPOP project, over the last three years, without which I would have been unable to dedicate myself to this research, as well as attended various trainings, workshops, and conferences that provided skills and feedback with which I was able to complete this dissertation.

The LONGPOP project received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 676060 and the project LONGHEALTH CSO2015-69834-R funded by the Spanish Ministry of Economy and Competitiveness. This publication reflects only the author's (Laura Ann Cilek) view, and the Agency is not responsible for any use that may be made of the information it contains.

Abstract

Just over one hundred years ago, a series of epidemic influenza outbreaks swept through the world. Dubbed the “Spanish flu," the entirety of these outbreaks contributed to millions of excess deaths in less than one year. While research has long been focused on several aspects of the pandemic, not limited to its virology, aspects of its morbidity and mortality impact, and social and spatial variation in its manifestation, questions remain unanswered. This dissertation attempts to address several unresolved questions persist today using newly digitized data sources, primarily the Madrid Civil Register death records from 1917-1922. As a body of work, the following chapters contextualize Madrid’s experience of Spanish flu to both the rest of Spain and other parts of the world. Better understanding of these historical issues can help contemporary epidemiologists and policy makers to better prepare for future outbreaks. A discussion of the broad themes noted below are woven through the totality of the work.

Timing and strength of individual and successive waves:

During the strong fall wave in Spain, is a geographic pattern of transmission visible?
In line with previous assessments of the spread of influenza in Spain in 1918, a clear pattern of movement from the northeast to the west and south was statistically found using sequence analysis. Rural areas appear to have a stronger (higher R) and longer lasting fall wave than the provincial capital cities. However, within the urban center of Madrid, while variation in timing and strength existed, any geographic pattern is difficult to ascertain. [Chapter 2 and Appendix D]

What is the impact of successive waves in a population?
In the case of Spain, cities and provinces known to have a herald wave in spring or summer 1918 appear to experience fall waves of lower Reproduction numbers than those areas in which a herald wave is not known to have occurred. Madrid was hit with a particularly strong herald wave, and the excess mortality in the city is lower relative to other large urban areas, especially in the fall wave. With regards to cause of death, it appears that earlier waves

were particularly deadly to those with pre-existing diseases, such as tuberculosis. These first outbreaks effectively harvested these deaths, which did not occur in later waves. [Chapters 2, 4, and 6]

Age-specific mortality patterns:

How much excess mortality occurred in each wave, and to what extent did excess mortality peak in young adult ages?
Calculated as the amount of observed mortality above the expected level, age-specific excess mortality patterns during the outbreak are known to have varied across the world, but often a particular pattern is visible of increased mortality excess in young-adult ages. In the city of Madrid, this pattern is less prominent than in other locations in the world, such as Scotland, but there is a small localized peak in absolute excess mortality during the fall-winter 1918-19 wave. Relative excess mortality is found to be the highest in young adults in all waves. Given the presence of a strong herald wave in Madrid, the extent to which a protective effect of earlier exposure may have tempered the amount of mortality and prevalence of a young adult mortality spike is unknown. In an echo wave in both Madrid (1919-20) and Scotland (1920-21), the youngest and oldest are most affected, implying both a return to seasonal influenza mortality patterns and a lack of immunity among the youngest children, conceived and born after the outbreaks. [Chapters 3, 4, and 6]

Mortality risk from a neighborhood perspective:

How did demographic, social, and spatial variation contribute to increased mortality?
At the time of the outbreaks, Madrid was a very heterogeneous city, consisting of the richest rich and poorest poor. Some neighborhoods consisted of similar groups of people, while others were incredibly diverse. This makes the analysis of an area's traits with respect to excess mortality difficult. Nonetheless, the thesis clearly finds that irrespective of other social and constructed characteristics, such as social class, literacy level, and population density, the amount of baseline mortality in an area is the best predictor of excess mortality during the outbreaks. That is to say, those areas with the highest rates of pre-pandemic mortality also faced devastating losses of life during each pandemic wave. This adds to other evidence in the thesis that suggests those most vulnerable before the influenza epidemic continued to experience higher levels of mortality during the outbreaks. [Chapters 5, and 6]

Resumen

Hace poco más de cien años, una pandemia de gripe se extendió por todo el mundo. Apodada la "gripe española", esta pandemia contribuyó a la muerte de millones de personas en menos de un año. Durante mucho tiempo el foco del análisis se ha centrado, no sólo en su virología, si no en aspectos como su impacto en la morbilidad y la mortalidad, y cómo afecto de forma diferente social y espacialmente, pero aún no se ha dado una respuesta exhaustiva a muchas de estas cuestiones. Esta tesis doctoral pretende abordar algunas cuestiones que todavía hoy se encuentran pendientes de resolver, usando para ello fuentes de datos recientemente digitalizadas, principalmente los registros de defunción del Registro Civil de Madrid de 1917-1922. En los siguientes capítulos se contextualiza la gripe española dentro de una gran urbe como Madrid, así como en el conjunto de España, comparándola con otras partes del mundo. Una mejor comprensión de la evolución histórica de la enfermedad puede ayudar a los epidemiólogos y responsables políticos contemporáneos a prepararse mejor para futuros brotes. A continuación, se mencionan los temas principales, transversales en la totalidad del trabajo:

Tiempo y fuerza de las olas individuales y sucesivas:

Durante la Gripe de 1918, la fuerte ola de otoño en España, tiene un patrón geográfico de transmisión visible y claro. En línea con las anteriores evaluaciones de la propagación de la gripe en España en 1918, se encontró estadísticamente un patrón claro de movimiento desde el noreste hacia el oeste y el sur mediante el método de análisis de secuencias. Las zonas rurales de las provincias parecen tener una ola más fuerte (más alta en R) y más duradera que la de las capitales de provincia. Sin embargo, dentro de Madrid, si bien existían variaciones en tiempo y fuerza, cualquier patrón geográfico es difícil de identificar. [capítulo 2 y capítulo D]

¿Cuál es el impacto de las olas sucesivas en una población?
En el caso de España, las ciudades y provincias que tenían una ola precursora en primavera o verano (de 1918) parecen experimentar una ola en otoño más atenuada que aquellas áreas en las que no se sabe que hayan sufrido esa ola precursora. Madrid fue golpeada por una primera ola precursora particularmente fuerte, y, sin embargo, el exceso de mortalidad es menor en comparación con otras grandes áreas urbanas, especialmente en la ola de otoño. Con respecto a las causas de muerte, parece que las olas anteriores fueron particularmente mortales para aquellos con enfermedades preexistentes, como la tuberculosis. Estos primeros brotes cosecharon efectivamente estas muertes, que no ocurrieron en oleadas posteriores.

[capítulos 2, 4, y 6]

Patrones de mortalidad específicos por edad:

¿Cuántas muertes añadidas trajo la Gripe en cada ola, y hasta qué punto el exceso de mortalidad alcanzó su punto máximo en la edad adulta joven?
Calculado como la cantidad de mortalidad observada por encima del nivel esperado, se sabe que los patrones de exceso de mortalidad por edad durante el brote han variado en todo el mundo, pero a menudo se observa un patrón particular de aumento del exceso de mortalidad en jóvenes y adultos. En la ciudad de Madrid, este patrón es menos prominente que en otros lugares del mundo, como Escocia, pero hay un exceso absoluto más elevado de mortalidad durante el otoño-invierno de 1918-19 en Madrid. El exceso relativo de la mortalidad es más alto entre los adultos jóvenes en todas las oleadas. Dada la presencia de una fuerte ola precursora en Madrid, se desconoce hasta qué punto un efecto protector de la exposición temprana pudo haber moderado la mortalidad y la prevalencia de un pico de mortalidad en adultos jóvenes en sucesivas oleadas. En una onda posterior, tanto en Madrid (1919-20) como en Escocia (1920-21), los más jóvenes y los mayores son los más afectados, lo que demuestra un retorno a los patrones de mortalidad por gripe estacional y la falta de inmunidad entre los ni´ nos más peque´ nos, concebidos y nacidos después de los brotes. [capítulos 3, 4, y 6]

Riesgo de mortalidad desde la perspectiva del barrio.

¿Cómo contribuyó la variación demográfica, social y espacial al aumento de la mortalidad?
En el momento de la pandemia, Madrid era una ciudad muy heterogénea. Algunos barrios los formaban grupos similares de personas, mientras que en otros los grupos de personas eran increíblemente diversos. Esto dificulta el análisis de los rasgos de un área respecto al exceso de mortalidad. No obstante, la tesis encuentra claramente que, independientemente de otras características sociales, como la clase social, el nivel de alfabetización y la densidad de población, la mortalidad inherente de cada zona es el mejor indicador del exceso de mortalidad durante cada uno de los brotes. Es decir, las zonas con las tasas más altas de mortalidad pre-pandémica también se enfrentan a pérdidas de vidas devastadoras durante cada ola pandémica. Esto se suma a otras conclusiones de la tesis que sugieren que los más vulnerables antes de la epidemia de gripe continuaron experimentando niveles más altos de mortalidad durante los brotes. [capítulos 5, y 6]

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Personal Demographic Beginnings

Until September 2016, when I began my studies at the Consejo Superior de Investigaciones Científicas (CSIC) in Madrid, my knowledge of the “Spanish Flu" epidemic largely remained limited to only understanding its high mortality impact. As I started my master’s degree in Demography and began my journey into the study of population, my very first lectures exposed me to the strength of the 1918 epidemic. Professors constantly use the drastic increase in mortality rates during this year to explain the impact age-specific mortality differences can have on aggregate population measures. Later, while attending the European Doctoral School of Demography (EDSD), Miguel Sanchez-Romero taught basic epidemic modeling and the calculation of the Reproduction Number (R) as part of a week-long section of the modeling module. This course, my first true exposure to quantifying and understanding the impact of epidemics, piqued my interest in this area of demography. Around this time, I was also searching for a PhD position and was referred to the LONGPOP project by several EDSD teachers. Of the listed available positions, I was immediately intrigued by the first, “Mapping epidemic diseases through time: influenza," and applied. Several months later, I was accepted to the project with 14 other Early Stage Researchers throughout Europe, and, after several more months of visa paperwork, I ultimately began the process of study, analyses, and review that has culminated in the following dissertation.

My first several months were spent reading and learning about themes associated with and evolution of the study of Spanish flu over the past one hundred years. As a novice on the subject, I had to learn about the different waves and their progression through the world, total and age-specific excess mortality (both absolute and relative) by cause, and how all of these differed by geography. Moreover, I learned about the continuing debate regarding acquired immunity and social gradients in mortality. Much work has already added to the

narrative and understanding of the Spanish flu in the city of Madrid; entire volumes have focused on the experience of the outbreaks in Madrid and Spain, though these books generally examined the events and their effects from a social perspective and qualitative lens. They contain calculations aggregated statistics and sought to understand how propaganda and the political situation before, during, and after the outbreaks led to changes in social and medical structures [84, 95, 112, 114]. As I continued to familiarize myself with the subject matter, I looked forward to how my own work might contribute to the existing body of knowledge on Spanish flu in Madrid and realized my quantitative skills would provide a missing statistical component to the story of the largest and most pervasive influenza epidemic in Madrid.

Recent technological developments and financial resources have led a resurgence in influenza research, both in the academic realm, through scientific articles and collections (e.g. *The Spanish influenza pandemic of 1918-1919: perspectives from the Iberian Peninsula and the America, The Spanish Influenza Pandemic of 1918-1919: New Perspectives*) and in population non-fiction (e.g. *The Great Influenza: The Story of the Deadliest Pandemic in History, Pale Rider: The Spanish Flu of 1918 and How it Changed the World*) [36, 115, 146, 260]. The ability to digitize historical death records means that the progression and mortality impact of the flu can be re-examined from quantitative perspectives, adding to our knowledge about how past epidemics spread and may manifest themselves in the future. These new analyses continue to both solve and call into question longstanding debates on the global outbreaks, while also reminding scientists, policy makers, and the general public alike of the importance of epidemic preparedness [286]. While nearly one hundred years have passed since the Spanish flu, a pandemic could similarly affect the world today should the global population be caught off guard. In this sense, the following dissertation is meant to complement and add to the re-emerging body of demographic and epidemiological analyses on the past pandemic, with the understanding that while other invaluable work on the topic has been completed, additional perspectives can continue to provide a wider insight into the ways pandemics have affected communities in the past and may do so in the future.

While my participation in the EDSD supplied an advanced knowledge of demographic analysis techniques, I needed to familiarize myself with additional methodologies and garner feedback from the scientific community through courses, conferences, and workshops. The LONGPOP grant, funded through the European Union's Horizon 2020 Program, afforded me the opportunity to attend many courses within our International Training Network (ITN) of 15 Early Stage Researchers (ESRs) as well as additional events throughout the world that helped in the overall development and completion of my work. In total, I have attended more than eight courses and eight conferences between December 2016 and June 2019.

The Spanish Influenza pandemic contains several defining characteristics, yet unique elements define its manifestation in individual locations. Thus, it seemed appropriate to chose another country with which I could compare the Madrid and Spain experience. Scotland seemed to be an ideal example from which to perform a comparison, especially given the extreme differences in climate compared to Spain, a relatively (to Madrid) rural population, and the country's direct involvement in World War I. The similarities and differences between the two places could highlight further peculiarities unique to either the pandemic *or* its manifestation in a specific geography.

Moreover, as a partner in the LONGPOP project, the University of Edinburgh allowed me to visit and perform two separate research stays through an associated project called "Digitising Scotland" at the National Records of Scotland. This project is part of a major undertaking to digitize and prepare 24 million birth, death, and marriage records in Scotland beginning from 1856 for statistical analysis. While to date, the dataset has not been completed and finalized, I was allowed the unique opportunity to work with all death records in Scotland between 1916 and 1923 for both my own analysis and to test the viability and quality of the data as it was prepared. I first spent two months in Edinburgh in September and October of 2018, then stayed an additional month in the summer of 2019. Some of the work I completed there is found in Chapter 6.

The rest of this introduction will present background information relating to the Spanish Flu in the world and in Spain, thus providing relevant information for the work presented in the succeeding chapters. While each chapter will include consider literature relating to the analysis, I have included in the introduction a small literature review to cover the most important topics of study in the flu. This includes overall timing and strength patterns, distinguishing age-specific mortality characteristics, and recent interest in the role of acquired immunity, environment, and sociodemographic status in influenza morbidity and mortality. I will also highlight the data sources used and provide an overview of their defining characteristics, but each chapter will also include more detailed relevant information to each specific analysis. As such, it should be noted that the while the culmination of this dissertation covers varying aspects of the Spanish Influenza Pandemic in 1918, it is meant to be read as a collection of individual research essays. Rather than a traditional PhD dissertation of successive chapters built upon each other, I feel the design employed here is more suitable to addressing the specific topics of note in each separate analysis.

1.2 Thesis Outline

1.2.1 Chapter II: Re-examining Strength and Timing Across Spain

The predominant focus of this thesis is the manifestation of Spanish Influenza within the city of Madrid, but it must be placed in the context of the world and regional neighbors within Spain. Thus, this chapter looks at the progression of the fall wave in Spain through all 49 provinces, their capitals, and rural populations between September and December 1918. Presented as a poster (for which an award was won) at the 2019 Annual Meeting of the Population Association of America, breakpoint analysis, or segmented regression, finds the start, peak, and end dates of the epidemic outbreak in each geographic area. From these values, the Reproduction number can be estimated, which serves as a proxy for wave strength. Using the trajectories of wave progression, the areas are grouped into distinct typologies using sequence analysis. Overall, the results show a clear north to south-west progression of the wave through the country as well as large differences between the waves' manifestation in urban/rural areas.

1.2.2 Chapter III: Estimating a Mortality Baseline from Limited Data

Although presented sequentially second in the thesis, the motivation for Chapter 3 arose *after* beginning the age-specific excess mortality calculations for each wave in Chapter 4. Because excess calculations are directly related to the underlying expected mortality during the examination period, it is imperative that the constructed seasonal baseline accurately reflects the timing and magnitude of mortality peaks and valleys. Generally, three years of data is considered the "gold standard" from which a baseline is calculated; the data I used only contained death records for one year, 1917, from which to create a baseline. Rather than simply acquiesce to using traditional estimation techniques with insufficient data, I took the opportunity to explore new methods of estimating seasonal mortality. Most generally, the contents of this chapter explore how aggregated monthly count data and adaptive methodologies can supplement a primary data source to construct a reasonable seasonal baseline. An early version of this analysis was presented in February 2018 at the 9th Demographic Conference of "Young Demographers" and a more complete work at the 2019 Annual Meeting of the Population Association of America and the 3rd European Society of Historical Demography Conference in 2019.

1.2.3 Chapter IV: Age-specific excess mortality patterns during the Spanish influenza pandemic in Madrid, Spain

Chapter four provides new calculations and estimates of excess mortality within the city. I first expand upon the background of the study of excess mortality in the context of Spanish flu and then briefly describe a method outlined in Chapter 2 to create a seasonal baseline. Some extra attention is given to the mortality records (from the Madrid Civil Register), particularly in the context of cause of death information, used in both this and the following chapter. From the calculated baseline, I present estimates of all-cause and respiratory-related absolute and relative excess mortality for three waves of Spanish flu in the city of Madrid between May 1918 and February 1920. These results are presented for several age groups, then compared to the 1889-90 influenza epidemic in Madrid, as well as other reports of excess mortality in Spain and cities throughout the world during the Spanish Flu pandemic. After oral presentations at both the IUSSP International Population Conference in Capetown and the IUSSP International Seminar on Pandemics in November 2017, a version of this chapter was published in an October 2018 special issue of the American Journal of Epidemiology.

1.2.4 Chapter V: Neighborhood variation in excess mortality across three waves

From the start of this project, I believed the analysis presented in Chapter 5 would become the most significant in terms of a contribution to the current study of epidemics, specifically how social, demographic, and economic inequalities contribute to mortality. However, I did not anticipate the challenges I would face in identifying and executing an appropriate approach to this analysis. The culmination of this undertaking is presented in this chapter. First, I briefly outline prior analyses on social inequalities in influenza morbidity and mortality. Then, I examine the case of Madrid in three ways: at the district level, at the neighborhood level, and, to a lesser extent, at the individual level in Madrid. I highlight the limitations to each approach and discuss the results in the context of the manifestation of the influenza epidemic in Madrid. As the work developed, I presented parts of this chapter at scientific meetings many times, including at the IUSSP International Seminar on Pandemics in November 2017, and in 2018 at the Annual Meeting of the Population Association of America (poster), the European Population Conference, the Social Impact of Epidemics workshop organized by the European Association for Population Studies (EAPS), and the VIII International Congress of the Latin American Population Association.

1.2.5 Chapter VI: Excess Mortality and its effect on seasonal lifetable measures during the epidemic outbreaks in Scotland

The sixth and last chapter within this thesis expands the context of my work beyond Madrid and Spain to look at the effects of four influenza outbreaks in Scotland between 1918 and 1921. First, measures of excess mortality by age, sex, and wave are calculated using traditional methods. The rest of the chapter introduces “seasonal life tables," which are then decomposed according to Arriaga’s formula to understand how age-specific changes in mortality during the wave contributed to overall changes in life expectancy and life disparity during the outbreaks. Results are also shown by cause of death, which highlights the role a person’s underlying frailty (i.e. having tuberculosis prior to the outbreaks) had in their mortality risk during the fall wave. A preliminary version of this paper was presented in June 2019 at the Nordic Demographic Symposium in Iceland.

1.2.6 Thesis Conclusion

The conclusion summarizes the totality of the thesis, highlights its important contributions, then addresses some broad limitations of the work. Three main influenza pandemics have occurred in the century following the Spanish flu, and their spread and mortality impact is briefly summarized. With each pandemic, the world learns more about transmission, and with increased technology and preparedness plans, the ability to recognize and fight future pandemics is higher than ever. Nonetheless, several vulnerable populations still exist, particularly in countries and areas of the world with the fewest resources to prepare. Continued research in historical pandemics, such as that in 1918, can help better understand the ways in which future outbreaks may affect these areas. The conclusion calls attention to current preparedness plans, their continued evolution, and the need to remain vigilant to prevent another worldwide pandemic.

1.3 Origins and Characterization of Influenza

Human influenza outbreaks have been reported in history for hundreds of years, and the emergence of the influenza virus most certainly happened far before that. Their exact appearance is debated; some argue the transmission of the virus to humans required close contact with animals (i.e. domestication), however, influenza was not reported until well after this happened [161]. Generally, new virus strains jump to human and/or mammalian hosts after initially circulating among avians. While viruses can jump directly from birds to people, they often gain human transmissibility directly from swine [266]. The exact relationship

between new influenza strains and their transmission between species remains a debate today, but researchers understand that in general, viruses can transfer and mix within their hosts in a variety of ways [189]. The pandemic of 1918 could have been the first time the H1N1 virus transferred from avians to swine to humans and circulated around the world. This lack of immunity among humans was thereby a huge driver of the high mortality burden of the pandemic [228, 267]. It can be noted that during the 1918 human pandemic, simultaneous outbreaks of influenza occurred in swine [304].

While the virus can be divided into human types A, B, and C, most human epidemic outbreaks are directly related to Type A influenza, which was first isolated in 1933 [132, 254, 259]. Influenza A subtypes are defined by two types of proteins named haemagglutinin (HA) and neuraminidase (NA), which generally provide the nomenclature for each variety [132]. For example, the circulating virus that caused the Spanish influenza pandemic was a strain of H1N1 flu, and probably the first time this subtype circulated in humans, leading to the large epidemic [254, 304]. As person's immunity to the virus is derived from their immunity to both the HA and NA strands, the virus mutates quickly to ensure its survival and thereby continues to infect those non-immune to the newest strains.

In order to survive, influenza viruses mutate in two ways–antigenic drift and shift–so that their hosts are never fully immune [52]. Drift occurs much more subtly and frequently than shift, leading to a change in the virus over time. When the virus reproduces itself, stepwise mutations occur in the HA and NA proteins so that the host (human, swine, etc.) cannot properly fight against it [180, 301]. Shift occurs much less often and is overall much rarer than the first type of mutation, happening only in influenza A subtypes [52]. In this case, a new type of virus is created through a recombination of proteins, often when two separate virus strains are both present in the same host and mutate or mix together. [301]. Viruses caused by antigenic shift have the potential to become extremely virulent and cause pandemics, because hosts may lack the ability to fight against the disease.

1.3.1 Influenza Pandemics

The origins of influenza pandemics, both presumed and confirmed, can be traced back throughout history. Even as early as 412 BC, records of flu outbreaks exist, when Hippocrates noted an epidemic of influenza-like illness [116, 149]. However, whether or not these earliest outbreaks, including ones recorded in 1173 and 1510, existed on a global scale is unknown, due to a lack of records and evidence to confirm. Moreover, it would have been more difficult for these viruses to travel given that population settlements were less-connected and time needed to journey between them was longer [41]. Likely, the first global epidemic occurred in 1580, spreading from Asia to Africa, Europe, and America.

The dynamics of these pandemic-level outbreaks have changed slightly over time. Most generally, the pace at which the outbreaks spread increased in subsequent pandemics, reflecting the increased speed and communication within trade networks throughout the world. In 1729 an influenza pandemic began in Russia in the spring before spreading throughout Europe in only six months, and this pandemic was succeeded by one in 1781-2, which after starting in China, traveled to Russia and all of Europe in only eight months[213]. Another pandemic in 1830-3 also began in China, then spread through Europe and North America, and it featured successive echo waves in 1831-32 and 1832-33 in many locations [213].

1.3.2 Russian Flu Pandemic

While the spread of these pandemics from the place of initial outbreak to Europe and the rest of the world happened in less than a year, the so-called "Russian" flu in 1889 is considered the first pandemic in the modern, industrialized world, spreading throughout Europe in only six weeks and the world within 12 months [275]. Beginning in October 1889 in St. Petersburg, Russia, the first cases in Spain were reported in early December, and in Madrid, the flu peaked at the end of December 1899 [101]. The first true pandemic since that in 1848, the attack rate of this probable H3N8 virus was considered to be around 50%, and the mean peak mortality burden in 96 European and U.S. cities was 142%, meaning that mortality in the peak week was 142% higher than normal. While the spread of the flu was rapid and a large number of people became sick, the case fatality rate in most places was about 0.1%, on par with other pandemics in 1957 and 1968 [266, 275]. Conversely, data shows that the 1918 influenza pandemic case fatality rate was nearly one percent, ten times high than these other pandemics [179].

Similarly, a systematic review of surveys, vital statistics, and other primary sources revealed the age-specific mortality patterns of this pandemic was closer in shape to those of other pandemics in the 20th century [276]. Unlike the "w" shaped pattern often found in the Spanish flu (see section 1.4.3), age-specific mortality patterns typically followed a "j" shape, meaning that the youngest, and especially the oldest, experienced the highest excess mortality. Those around age twenty faced the lowest excess mortality.

1.3.3 The 1889 Pandemic in Madrid

The Russian flu was last major influenza pandemic to hit Madrid prior to the Spanish flu outbreaks, which occurred 28 years later. the first cases related to the epidemic already raging in much of Europe, were reported in the middle of December, although time-series of mortality information reveal the wave may have actually started a few weeks before

[101, 194, 218]. After peaking at the end of the year, mortality returned to expected winter levels in February 1890, but the outbreak produced a profound mortality impact in the city. During the wave, absolute excess mortality for all causes and respiratory-related deaths was 58.3 and 44.5 deaths per 10,000, with the highest age-specific rates occurring in those older than 50 (peaking at the oldest ages) and younger than 5. Meanwhile, the standardized mortality ratio (relative excess), peaked among 15-24 year olds, who experienced more than 10 times their expected normal mortality level across all causes of death [218].

Because the excess rates reported in Madrid deal with mortality only, it is difficult to quantify the amount of flu prevalence in these groups and the role an individual or age group's haleness may have played in their ability to fight the disease. While certainly during their lifetimes, many of the seniors who perished in this pandemic had been exposed to other strains of influenza virus, their high mortality rates suggest their bodies could not defend themselves after becoming sick. The high *relative* rate of mortality among youthful ages further reflects that this age-group did not carry immunity to this virus. The young-adult group generally has some of the lowest mortality rates over the life course and overall, their absolute mortality rates remained low compared to other age groups. However, the mortality impact of the Russian flu still produced nearly ten times as many expected deaths during the period, suggesting that even the healthiest groups could not fully fight the disease.

1.4 The Spanish Flu Pandemic in the World

1.4.1 Origins

The exact origin of the 1918 Spanish Influenza pandemic remains a debated topic, especially as to whether the virus began and circulated in avian species or if it was due to a re-assortment of proteins in swine [117, 267, 279]. In either case, the virus was likely circulating in swine and humans for at least several years prior to the start of the outbreak [258]. Researchers also debate about the specific geographic location from which the pandemic emerged. Due to the ongoing first World War, the initial reports of flu outbreak came from Spain, which, as a non-participant, did not face a press embargo and published news of the contagious illness in May 1918 [85]. Yet despite providing its name for the nomenclature of the pandemic, outbreaks began decidedly earlier. For some time, it was widely presumed human outbreaks began in military bases in Kansas (USA), home to many going, returning, and training to partake in the war [36, 79, 204]. However, continued investigations have revealed earlier possibilities of the first large-scale transmissions; documents have shown that even in the winter of 1916, a large-scale outbreak of influenza-like illness occurred in a crowded British

military base in Étaples [128, 204]. The illness reported high rates of mortality and nearly similar symptoms or severe bronchitis found in the ensuant pandemic. A few months later, a similar outbreak occurred in barracks in Aldershot, in the southern United Kingdom [15]. The crowded nature of these facilities and near-constant movement of soldiers during the war support the plausibility that initial transmission and outbreaks began in France, then spread to the UK and US before all over world [205]. However, tissue samples have not yet been tested to confirm these earlier origins, and overall uncertainty remains.

Still, other places in the world *have* noted higher than normal (though not necessarily of epidemic levels) influenza and respiratory related deaths in the several years preceding the pandemic. Spain itself may have experienced higher numbers of influenza and respiratory related deaths in 1915, 1916, and 1917 [112, 299].

Surveillance systems and vaccination programs continue to better understand and halt the transmission and occurrence of pandemics. Nonetheless, it remains impossible to predict the timing of a pandemic despite the relative frequency of occurrence and their study over the last several hundred years. Thus, ongoing studies of past pandemics, such as those presented here, continue to provide context and information so that scientists may understand the demographic and social mechanisms of influenza transmission, and perhaps one day eliminate the real threat of pandemics.

1.4.2 Waves

The timeline of the pandemic is broadly classified into three different waves–a short but intense “herald" wave in the Spring of 1918 in temperate regions, a longer wave in the late fall and winter of 1918-1919 which accounted for a majority of deaths, and a third, less lethal wave taking place in the following year. Both the first and third waves were not universally experienced; those locations that were more remote or practiced better preventative techniques often skipped the first wave but instead endured longer and more persistent second waves [141]. Yet the rapidity with which these waves reappeared was a new phenomenon compared to prior flu pandemics. These traditionally occurred on an annual basis, such as in the case of the Russian flu pandemic, which appeared in yearly waves between 1889 and 1893 [181]. Descriptions below focus on the herald and fall waves, but in addition to a third winter wave in some locations, “echo waves” in the years following the pandemic often occurred, generally featuring heightened excess mortality among those less than one year old [67, 68, 82, 198, 256].

Herald

Especially in the northern hemisphere, the epidemic outbreaks in different locations occurred during one or more of several wave periods [142, 229]. Following its initial appearance in Kansas recorded on the fifth of March, influenza spread through military bases in the midwest, south, and along the east coast of the United States, eventually transmitting to and appearing in civilian populations as well in March and April 1918 [79]. The wave reached New York before April, parts of Mexico in April and May [68, 198].

Likely owing to the strong transportation flows of soldiers across the Atlantic, the virus appeared in France in April, spreading outward throughout Europe from there [209]. Flu came to Spain and Portugal in mid-May, whence the virus finally got the name "Spanish" flu. Around the same time, the virus had made it to south and central parts of the United Kingdom, Italy and Greece (through Mediterranean ports of entry), and southern Scandinavia [209]. By July, the flu had spread through the rest of Scandinavia and to parts of Poland and eastern Europe, but did not continue its easterly spread across land. Other reports indicate that the spring wave made it to India and other parts of Asia, but eventually, the spread of the harsh virus waned in the summer months. While it infected many individuals and was prevalent in the locations it affected, the spring herald wave had a much smaller mortality impact than that of the fall. However, some research has noted that as the wave progressed, it gained virulence and its manifestation began to more closely resemble that of the fall virus [280].

Fall

In late August, the fall wave appeared and began to spread through the world, seemingly first reported in Brest, a large port in western France [79, 209, 213]. Slightly after or nearly simultaneously, ships brought the virus to ports in Boston and Sierra Leone; when the crews were treated by doctors and nurses, the virus was transmitted onward, and from these three hubs spread across these landmasses [209]. In fact, large ports and railway lines proved to be a major source of transmission as the virus progressed through the world, as the beginning of outbreaks often began in or along ports and railways, then continued onwards in conjunction with large travel routes. As the virus continued to spread throughout the world, it became clear that it was much more virulent than its predecessor in the herald outbreaks. While both waves were were quite contagious and spread rapidly, the circulating strain in the fall had a case fatality rate of around 10 times higher than average flu outbreaks [107, 179].

Directionally, the fall wave spread in a similar manner to the herald wave. By the end of September, all of northern and western Europe was in the throes of the outbreaks, and

by October, the pandemic had reached the northernmost and easternmost parts of Europe, including Iceland, Lapland, Turkey and parts of Russia [209]. Most of the outbreaks in western, southern, and central, Europe peaked in October, but continued to rage through November or later. In North America, the pandemic spread via ports on the east, south, and west coasts and Chicago into the heart of the US via rivers such as the Mississippi, and from the east coast into Canada via the Saint Lawrence river [79, 172, 209]. As stressed above, ports served as entrance points in South America, Africa, and Asia, and from there, newly constructed railways carried the virus onwards [208, 209]. The pandemic even made it to New Zealand by October, but did not strike Australia until January 1919, when even a huge quarantine effort could no longer prevent the spread of the virus [171].

1.4.3 Mortality Impact

Several recent attempts to quantify the total number of individuals who perished to the pandemic place the total number of deaths in the millions worldwide, and some estimates argue the number is much closer to 50 or 100 million [141, 186, 261]. This number is far more than other past and contemporary pandemics; while the “Black Death" in the 14th century killed more than $\frac{1}{3}$ of the (smaller) population of Europe and AIDS has killed around 35 million throughout the world since 1990, these numbers pale in comparison to the number of deaths that happened in such a swift amount of time at the beginning of the 20^{th} century [274, 303].

The overall mortality impact remains highly debated for many reasons. The ongoing world war meant that several countries were careful to guard information about the outbreaks among civilian and especially their military populations. Moreover, destruction from subsequent wars in the twentieth century, such as the Spanish Civil War and World War II, has made the recovery and study of some historical documents nearly impossible.

Moreover, although the prevalence of the pandemic was nearly universal, the timing and severity of the outbreaks were highly variant, including differences in the onset, length, and number of individual waves, as well as demographic factors such as the total impact of morbidity and mortality by age and sex. At the time of the pandemic, much of the world lacked sufficient medical supplies, training, and funds, to help contain the spread and burden of the disease as it progressed [179]. While analyses in Europe, North and South America, parts of Asia and Oceania and other developed areas continue to update and refine estimates of influenza related mortality, rural and undeveloped populations lacked baseline and real-time information during the outbreaks. Thus, the true number of deceased is unknown, and general estimates still vary by millions today [141, 186, 209, 261].

Other aspects of the Spanish influenza pandemic and manifestation of its symptoms should be mentioned. The majority of influenza-related deaths resulted from secondary respiratory ailments following flu contraction (rather than influenza itself), most commonly bacterial pneumonia [80, 135, 181, 181, 182]. More related deaths came from additional types of bacterial infections and severe-acute respiratory distress, often evidenced by Cianosis, in which a lack of oxygen in the blood turns the skin blueish-gray shortly before death [181, 252]. Death also came quickly; generally, those who sought medical treatment died within two days of hospital or clinic admission [105]. However, while analyses studying flu-related deaths often focus, when possible, on the subset of influenza-like illnesses and other respiratory illnesses, other seemingly unrelated causes of death also tend to spike during each wave. Such is the case of the Russian influenza pandemic of 1889 in Madrid, when causes of death such as heart disease, meningitis, and brain disease spiked at the same time of the outbreaks in the city [101, 218]. Thus, the study of all causes of death are important in relation to total mortality impact.

Age-specific Mortality

While the flu's great strength is not doubted, a "normal" epidemic with similarly high virulence would result in a similar-to-normal mortality pattern in which the youngest and oldest die at the highest excess rates [229]. In many studies, this is not the case in the 1918 pandemic; rather, age-specific mortality in many places peaked in young-adulthood rather than among the elderly and statistically most impacted young adults between the ages of 25 and 30 [110, 233, 278, 284]. Supporting evidence can be found in analyses employing a variety of methods and different types of data; for example, a comparison of age specific mortality based on individual death records in Kentucky from 1911-1919 identified a peak mortality risk in 1918 (relative to the baseline years) for those aged between 24 and 26 [284]. Another study employed various sources of mortality data, including parish and civil registers, throughout the United States and Canada from September to December of 1918, finding mortality peaks around age 28 for both all cause and pandemic related mortality [110]. Despite the large number of these similar studies analyzing the age specific impact, still more continue–in 2018, another study found similar patterns in Copenhagen [278]. However, the authors argued against the proposal that exposure to the Russian flu pandemic may have played a role, instead highlighting that while the birth year of the mortality peak corresponded to the Russian flu, the hump of young adult excess was much wider, extending between birth years of 1878 and 1908.

These analyses have helped to develop and promote the idea of "Original Antigenic Sin" as a means of explaining the 1918 mortality peak among young adults [109, 298]. Most

generally, the “doctrine" states that young children experience a critical stage of development during which their immune systems learn to fight viruses. Individuals at or around age 28 during the Spanish flu (H1N1 strain) were exposed to the Russian flu (H2N2 or H3N8) in 1888-90, and thereby, their immune system´s developed the methods necessary to defend the body from the other strains [132, 275]. By “incorrectly" fighting the H1N1 strain, the immune system became compromised and weakened the body to the prolonged effects of the flu (i.e. pneumonia).

This possible explanation for the young-adult excess hump is not limited to the exposure directly surrounding the Russian flu pandemic. Other hypotheses include general ideas about the circulation of an H1 subtype in the population some time in the early to mid 1800s [229]. If the older adults with relatively less excess mortality than their younger counterparts had been exposed to an H1 variant prior to its departure as a circulating virus, perhaps this explains their disproportionate excess mortality to young adults [167]. Regardless, no firm consensus has been reached to directly explain the mid-adult hump.

When examining *old-age* influenza related mortality in 1918-1920, results vary. A few studies documenting excess age-specific mortality rates during the 1918-20 influenza pandemic reported either lower than expected elderly mortality in the US and European settings or little excess mortality [79]. For example an analysis of the three 1918-1919 pandemic waves in Copenhagen, using weekly surveillance data from 1910 to 1919, reported negligible excess mortality for those 65 and older [24]. These results mirrored another analysis of monthly mortality data from New York City that showed mortality risk of those 65 and older did not increase in pandemic periods between 1915 and 1918 [198].

Yet other analyses exist that contrast these findings. One such study uses single years of “normal" death count information (usually from 1917) to show elderly mortality remained constant [162], while another analysis, using three years of mortality archives as a baseline, found those 70+ faced two times the relative mortality risk in the fall and winter waves of 1918 and 1919 [68]. During the 1918 pandemic, in these cases of high excess mortality for both young adults and seniors, a true W-shaped pattern of excess mortality risk by age and a pattern occurs, where excess mortality rates peak in infants and young children, young adults, and the elderly population.

The appearance of this w-shaped mortality curve varies by location and perhaps exposure to previous strains of a familiar virus [156, 162]. Analyses conducted with census data and raw death counts during epidemic periods (with little to no baseline mortality information) reveal conflicting results as to a general mortality pattern by age; some evidence in rural and “geographically isolated" populations show a w-shaped mortality pattern [159, 249, 296]. However, similarly completed analyses in other remote areas instead find a v-shaped mortality

curve, in which the heightened mortality rates for adults does not decrease after the young-adult peak [79, 81, 162]. Other w-shaped mortality curves can be found in non-remote urban and rural locations throughout the world (for example, a study in the United States measuring country-wide annual excess mortality determined from a yearly five-year baseline [156] or analyses in Copenhagen using nine years of weekly surveillance and mortality data [24]).

Manifestation in Europe: Analyzing additional effects by lifetable decomposition

Estimates of excess mortality refer to the total number of extra deaths attributed to the flu while age-specific excess estimates can quantify the extent to which the epidemic generally disproportionately affected young adults. However, other analyses can provide additional insight into the impact of the flu on the total population and its mortality patterns during the time period of the epidemic. To demonstrate this and contextualize the mortality ramifications of the flu in a different manner, this section uses data from the Human Mortality Database (HMD) [136] to estimate the contribution of epidemic flu in several European countries to two classic demographic measures. Further methodology can be found in Appendix A.

Synonymous with demography and life tables, life expectancy is the average number of years a person born under specific age-specific mortality conditions lives [215]. While generally, this measure is best interpreted from a cohort perspective (so that it is represented as the average amount of years lived by those in a single birth cohort), this analysis uses period mortality rates, such that e_0 is the average number of years a person *would* live *if* mortality rates at the time they were born continue to stay the same as they age. E-dagger ($e_x^\dagger$) is another measure calculated via life tables that describes the disparity in age at death [281] and is the average life expectancy remaining at age of death. High values of e-dagger imply large variations throughout the population in the age at death.

Twelve countries from the HMD have period life tables for the time period during the pandemic, all of which (with the exception of New Zealand) are in Europe. The decomposition of these life tables shows the stark overall mortality impact of the Spanish flu had during 1918 and the way in which its unique characteristics, such as the age-specific mortality hump in young adults, impacted life expectancy and e-dagger. Graphics depicting the results for these countries by sex are found in A, and those for males in Spain are found here in figures 1.1 and 1.2.

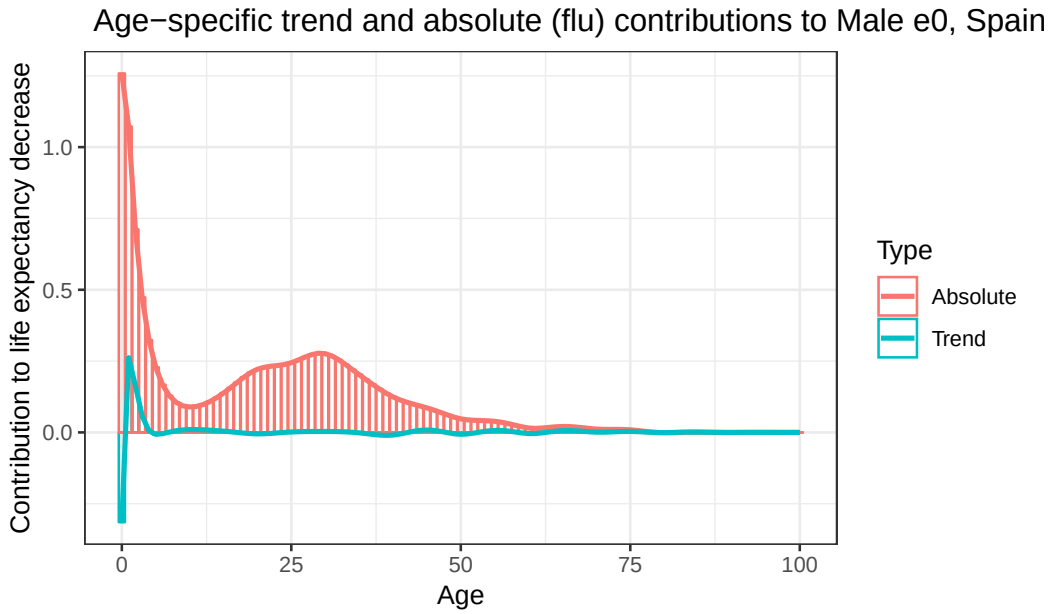

Fig. 1.1 Contour Decomposition for Life Expectancy of Males in Spain

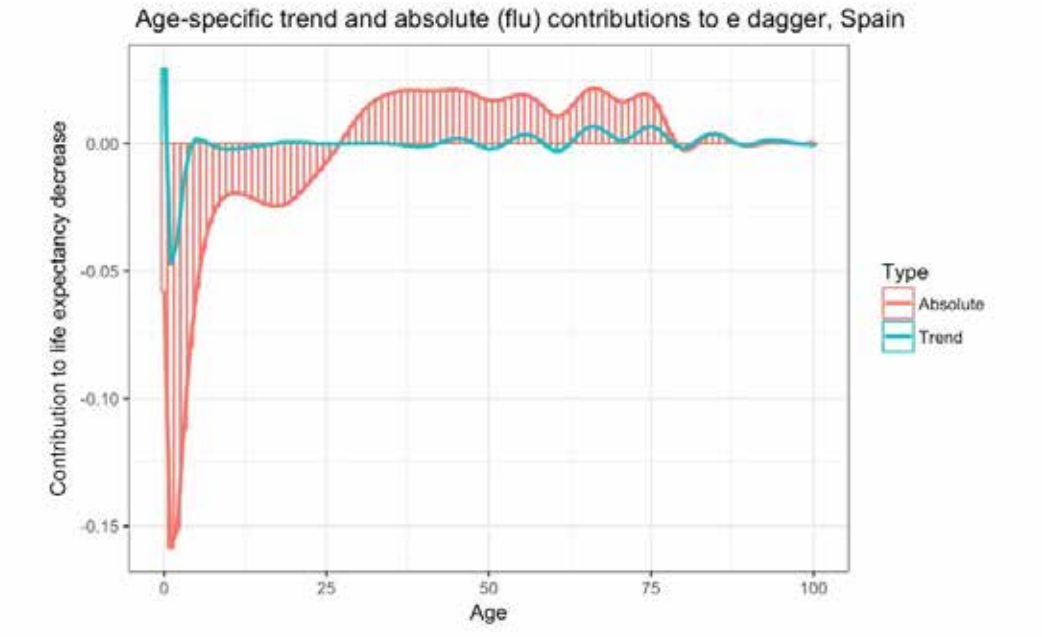

Fig. 1.2 Contour Decomposition for E Dagger of Males in Spain

As this thesis focuses predominantly within the country of Spain, some additional information related to the decomposition of Spanish males is noted here. Aggregate (smoothed)

life expectancy at birth in 1917 was 42.03 and when trends in life expectancy were extrapolated, the value was expected to be 41.70 in 1918. Notably, in a time of generally declining mortality in Europe, the life expectancy (perhaps due to changing population distribution by age, though not a part of this analysis) in Spain was expected to decrease slightly in 1918 given the smoothed trends in age-specific mortality from 1908 to 1917. Despite this expected value, real life expectancy of Spanish males was only 30.02 in 1918. With these total differences in average life expectancy in mind, the decomposition explains how each age group contributed temporally and absolutely to the total change between 1917 and 1918.

As expected, there is little temporal change in the contributions to both life expectancy and e-dagger. This temporal change is greatest at age 0, which reflects improvements in infant mortality in the preceding ten years that, despite the outbreak of the Spanish flu, continued to decrease in Spain in 1918. These lower rates contributed to an *increase* in life expectancy, despite the more than 10 year decrease in net average life expectancy. The largest absolute changes occur around age 28, which is expected given known age distribution of influenza-related excess mortality in many countries in 1918. Several countries show odd patterns in terms of the temporal trends at middle ages (France, Great Britain, Italy). This is perhaps a relic of mortality trends due to the First World War, as the focus is on men.[1] High mortality among soldiers in the four preceding years of the war may influence the smoothing parameters and therefore the extrapolated expected rates in 1918.

Trends of decreasing life expectancy and e-dagger do not exist in the female population during this time, likely, as mentioned, due to women *not* being on the front lines in the First World War. The trend component does exist, especially at the youngest and, to a lesser extent, in older ages. However, as expected, the difference in both life expectancy and e-dagger primarily exists due to the absolute differences between expected and observed mortality in 1918. That is to say the likely effect of influenza played a predominant role in the decrease in life expectancy and increase in lifespan disparity in the year 1918. Moreover, the change in these two elements comes from contributions at the youngest ages and in the mid-twenties, corresponding with both high infant mortality and age-specific mortality patterns during the flu.

[1] Spain was *not* involved in World War I.

1.5 The evolution of Madrid: small village to large capital

1.5.1 A short history of the city: early expansion and growth

As a city, the importance of Madrid grew significantly beginning in 1561, when Philip II moved the Habsburg court to the city, effectively making it the center of the vast Spanish empire [152]. At the time, the city had a modest population compared to other court cities in Spain, only 18,000, but by 1598 and the death of Philip II, the city had grown to more than 80,000 [50]. Still yet, relative to other European capitals, the city maintained an identity as a small town until the 1800s [241].

Following a time of unstable monarchy, the Napoleonic years, and finally the death of Ferdinand VII, the growth of the city began in earnest as new policies lead to an influx of immigrants from all over Spain. In addition to being the capital of Spain, the location of the city, in roughly the center of the Iberian peninsula, likely also contributed to its increased growth [71, 108]. Unfortunately, the physical size of the city remained small, and crowding contributed to epidemics that often occurred, such as a large outbreak of cholera in 1834, though this was only one of many to take place [152].

Extreme population increase and growing inequality

In the mid 19^{th} century, Madrid began to grow at an astounding rate, and construction and renovation projects attempted to both allow room for growth and modernize the city [287]. Railroad lines began to be constructed, such as the first route to Aranjuez in 1851, and additional water sources were brought to the city via the construction of the Canal de Isabel II and its associated projects (e.g. the Acueducto de Amaniel) [56, 76]. Despite these improvements, inequality between the social and economic classes was a large issue. In 1857, Carlos María's plan of city enlargement, the *Ensanche de Madrid*, was commissioned [51, 287]. The urban planner and architect's design involved an expansion of the city's footprint by about three times its size, roughly divided into three new parts. In the north and east, new the new neighborhoods of Argüelles, Chamberí, Salamanca, and Retiro primarily housed those middle and upper class, while in the south, areas such as Arganzuela grew to become home to many of the poorer and unqualified migrants from the south of Spain [35]. This pattern of organized segregation was not unheard of in Europe at the time, as wealthy and upper class individuals sought to isolate themselves from foreign and inhabitants of lower socioeconomic status [203]. While ultimately, the plan came to fruition, construction of the Ensanche occurred slower than the initially proposed, meaning that crowding in the city continued to be an issue through the end of the 19^{th} and early 20^{th} centuries.

Nonetheless, the ultimate effect of these newly constructed neighborhoods and the expansion of the city led to an increased separation by social class; where before, classes had intermingled within the same neighborhoods, they soon lived apart and interacted only within some of the city's public spaces (i.e. parks such as Retiro, which became a public space in 1868) [32, 152]. Rents in the enlarged areas of the city to the east were on average much more than four times as much as those in the south [17]. Within both of these neighborhoods, immigrants settled close to main transportation routes and paid less, in houses of poorer quality, and the most recent im-migrants paid even smaller rents. This is a relic of the un- and under-employment of those both new to the city and those residing in the southern parts [17, 287].

During this time of transformation and physical growth, the population also grew rapidly. What had until the 1800s remained largely a small court capital of modest numbers of people, the population rose from around 200,000 in the early part of the 19^{th} century to nearly half a million by the end of the century. Internal migration drove the majority of this change, and the bulk of these new inhabitants were of lower social class, crowded into areas in the southern part of the city, such as the Latina and Inclusa districts, or in the very periphery of the city, far away from the newly planned and built neighborhoods for the wealthy [211]. The cramped and low-quality living quarters of these areas meant that high mortality, especially among young children, and disease outbreaks such as cholera (1885 and 1890) were common. Yellow fever and small pox also contributed to excess mortality, as more than half a million individuals succumbed to these infectious diseases in several outbreaks over the course of the 19^{th} century [244]. Tuberculosis was another disease to hit the area hard. While general mortality declines had begun by the end of the 19^{th} century and tempered some of the extreme levels of deaths, from the end of the Russian Influenza epidemic until at least 1905, frequent outbreaks of smallpox and cholera continued to plague the city [44].

Expanding infrastructure and migrant settlement

The extreme growth of Madrid during this time not only changed the demographic makeup of the city, it dramatically altered the built urban environment. More people led to a higher demand in infrastructure such as housing and transportation options, and construction in the city during the Ensanche reflected that. The first tram line linked the new middle and upper class neighborhoods of Argüelles and Salamanca to Sol, perhaps the largest commercial and shopping hub, in 1871, and construction on these new forms of transportation continued [119]. At the turn of the century in 1902, the tram lines had all been electrified, and less than twenty years later in October 1919, just before the large fourth wave of Spanish flu outbreak in the city, the first line of the Madrid metro was opened by King Alfonso XIII and

immediately followed by the continued construction and opening of more lines and stations in the following years [152, 240]. As these lines of transport within the city grew, so did transportation to nearby municipalities in the province, whose population was growing at an even higher rate than the city itself [87].

One of the now-most iconic streets of the city, running through the center of Madrid, Gran Vía, was the result of decades of planning (beginning in the mid nineteenth century) and construction (beginning in 1910) [166]. Several neighborhoods and streets were destroyed in its making, particularly those of urban slums filled with not just migrants, but also long-term residents with lower social class, which forced the removal of poor individuals from the city center to the outskirts [152]. While the construction of this now iconic main avenue led to the creation of a "modern" street front to connect two important areas of the city, its synthesis can be seen as an analogy for the contradiction of the city at the time; as the city grew into the large capital it is today, many parts were left to struggle unnoticed. Moreover, the new areas to which many of the former Gran Vía area inhabitants moved to did not provide improved living standards. The crowdedness and lack of modern utilities in these areas contributed to their elevated mortality rates in comparison with central and "richer" areas.

Also, new technology contributed to modernization within the city as it grew during this time period. Electric lights first lit Sol in 1878 and then illuminated city streets by the end of the 19^{th} century [143]. However, while several private companies also installed electricity in homes and other private places, these were largely limited to those who could afford it [152]. But the while addition of these technologies increased the facility of life, they continued to reinforce the persistent inequality within the city as the divide between those able and unable to afford modern conveniences increased.

Because many migrants from similar social groups and provinces of Spain tended to settle in the same neighborhoods of Madrid upon their arrival, often, lower socioeconomic classes in Madrid consisted of entire migrant groups. That is to say, especially in the case of migrants from the south and areas within 100 kilometers of the city, little differences in group status existed. For example, a large crowded building housing around 1,000 people near the Puerta de Toledo was almost exclusively occupied by people from Toledo [287]. Other similarly crowded parts of the city, particularly in the south, were primarily composed of migrants from nearby towns. That is to say, in the initial parts of large migration to the city, even those born in Madrid to migrant families were often surrounded by the culture of their parents' home, limiting their overall ability to interact with other groups and areas of the city [16, 34].

However, as time wore on, the poor living conditions actually pushed people to seek other, less expensive living arrangements, regardless of their familial hometown. This integration

into the city is also apparent in the percentage of married couples from the same region, which drastically reduced between 1860 and 1930 [287]. However, the extent to which these trends prevail also depended on one's social status; those of highly regarded professions most often coupled with those from outside their home province, while those without any qualifications, who often worked odd jobs and preformed manual labor, most often partnered with someone with their home province [272]. Thus, one could argue economic hardship among migrants led to an integration of the population but reinforced patterns of financial inequality [287].

Yet, this pattern could also be observed from an opposite direction. In the southern part of the city, the example of a large tobacco factory can be used as the embodiment of the migratory experience for many. The workers in the factory were overwhelmingly women, from Alicante, and extremely poor. The daily toil at the factory rolling cigarettes was also long, demanding, and exhausting. But the integration found between many other migrants existed to a much lesser extent, as jobs predominantly passed down from mother to daughter. The result of this environment led to less integration between these migrants and others, but fostered a sense of community within themselves [17].

1.5.2 Population changes and demographic shifts

As noted, during the early 20^{th} century, the city of Madrid experienced vast change as rural-urban migration brought Spaniards to the city and new technology modernized the growing area. In 1900, the population was roughly 570,000, but by the end of the Spanish flu outbreaks in 1920, the population had already grown to 728,937, or more than 30% in the twenty year period [29]. During the time period between 1887 and 1930, the rate at which the population grew increased from 15% in the period of 1887-90 to 27% between 1920-30, indicative of the increased momentum of growth in the capital [87].

Yet while the overall population increased substantially in the late 19^{th} and early 20^{th} centuries, demographic indicators reveal the ongoing process of the demographic and epidemiologic transitions during this time [73, 199]. Especially in the Spanish context, urban areas, and particularly their male inhabitants, faced a higher mortality risk [225]. But broadly speaking, despite the associated increased penalties of living in a densely populated area, fertility, mortality, and the prevalence of some diseases in Madrid were in the middle of a steep decline during the period prior to the Spanish flu outbreaks [87].

Mortality decline and improvement of sociodemographic indicators

In general, the country of Spain experienced the demographic transition quite late, compared to other European countries. Prior to the transition, without migration, the city's growth would have been stunted due to extremely high rates of mortality, particularly among infants and children [17]. However, mortality improvement had slowly begun in the beginning of the twentieth century, particularly in indicators of infant and early child mortality, which steadily declined during this time. Madrid, which had one of the highest levels of infant mortality ever recorded in Spain in 1882 (358 per 1,000), experienced a drop by nearly half in 0 - 1 year old mortality between 1900 and 1930 [87]. In fact, in the city of Madrid, the probability of death between birth and age 1 ($_0q_1$) fell drastically from between around 225-250 in the 1860s to normal rates of roughly 170-200 in the first decade of the 1900s [245]. These rates continued to fall, and by 1960 child mortality was just above 30 per 10,000 [221]. [2]

In other locations in central Spain, infant and child mortality due to infectious diseases steadily declined beginning at the turn of the 20^{th} century and reached levels nearly 40% lower at the beginning of the Spanish flu epidemic than in 1885 [246]. Prior to this steep decline, many areas actually faced higher mortality between ages one and four than from zero to one. This is indicative of the presence of many diseases from which breastfeeding provides immunity to a child, but when weaned, made the child susceptible to contracting and dying from the disease [227, 246].

Some discussion has questioned to what extent an "urban penalty" may actually exist during this time and in contemporary society not only in Madrid and Europe, but also throughout the developing world [203, 220]. During the first third of the 20^{th} century in Madrid, infant and child mortality fell across both urban and rural areas, but the gap between the higher urban and lower rural mortality decreased so that by the mid 1920s, both population types had similar young child mortality (and rates of decline), and infant mortality in urban areas fell below that of rural regions. Before this change, perhaps sick children were brought to the city to die in hospitals, inflating the urban rates. Or, deaths were mis- or under-registered in rural areas. Other evidence using data from the charity Foundling Hospital of Madrid, which normally housed pregnant women facing adverse economic and health problems and orphans or abandoned infants, found no clear evidence of an urban penalty in infant mortality in the early 20^{th} century in Madrid [231]. Overall, the high mortality in Madrid and its quick decline between 1900 and 1933 reflected, despite the afore-mentioned crowded living situations and large inequality, the overall improvement of health in the city during the time.

[2] It should be noted that these mortality rates returned to high levels during the years associated with Spanish flu.

In the years prior to the outbreaks, some attempts had been made by medical professionals to call attention to the need for precautionary measures in the event of epidemic outbreaks, especially given high mortality rates to preventative diseases in the country [111]. Despite overall improvements in mortality in the city and across the country, many doctors and medical professionals in Spain recognized the extent to which medical institutions and governmental policies lacked the ability to provide effective support in the case of large scale outbreaks [130, 243]. Although mortality rates were falling, both air and waterborne diseases continued to play a large role in the high mortality of the city [115]. However, these calls for more authority and regulations often fell on deaf ears; the social position and overall authority of doctors and medical professionals at the time was low, despite recent advancements in diagnosis and treatment of diseases [112, 288]. In fact, when the first outbreaks of influenza began, there was no government ministry related directly to public health [114]. Instead, any general medical direction or regulation fell under the Ministerio de Gobernación, or Ministry of Government.

Fertility decline and changing marriage patterns

In line with the first demographic transition, as mortality declined, so did fertility rates. In Madrid, births to women of childbearing age fell from 129 per 1,000 in 1877 to 103 in 1900. After stagnating a bit, they fell again to 91 per 1,000 by 1920, a drop of more than 30% in 50 years. By contrast, the levels across all of Spain fell from 143 per 1,000 in 1877 to 114 in 1920, nearly 10% less than the fall observed in Madrid [87].

In all of Spain, some of this decline in births may also have been due to a delay in marriage. Between 1887 and 1930, the percentage of women married by age 30 fell from nearly 42% to under 30% [87]. However, during this same period, Madrid did not experience the same decline. Rather, in 1887, only 27.6% of women under 30 were married, falling only slightly to 26.7% in 1930. Yet despite this seeming delay in first marriage age, among both men and women, the percentage of widowers and widows decreased during this time, theoretically meaning several things. First, some of the social pressure (for women especially) to *not* remarry dissipated. But also, decreasing mortality rates were leading to longer lives lived together (expressed demographically as "l-bar" ($\bar{l}$)) [87]. Finally, perhaps the extremely high migration to the city from elsewhere in Spain led to a lack of familial support, resulting in a perceived need to take financial support and emotional comfort from marriage.

Other factors in the dynamics of the higher age at marriage in Madrid are likely due to the large role migration played in the development of the population and the need (or desire) for young migrant women and men to establish themselves in new surroundings before marriage. With respect to this, Madrid had a disproportionately high population of young women who

came from other parts of Spain to work as domestic helpers, which often prevented their marriage possibilities [18].

When considering all migrants, most (men and women) came from within 100 kilometers of the city, but large numbers of single women also arrived from northern provinces [87, 287]. Men who migrated looking for industrial work also came from unindustrialized areas in the South, because other urban poles in the north of the country (Barcelona and Bilbao) attracted migrants from those areas [17]. Those male migrants from the north would often be professionals of some sort whose work required their presence in Madrid.

Initially, most settled outside of the city center (in 1860, only 11% of migrants in the province lived in the center), but this percentage grew over time, so that by 1930, 21.3% of all migrants in the province of Madrid lived in the capital [287]. These migrants, who initially came on their own in search of work (with a temporary stay in mind), gradually started to both become permanent residents and attempt to reunite with their families in their new home city. When multiple generations of extended families lived in the city, they often settled in the same neighborhoods, even blocks, to remain close [155].

With regard to the evolution of economic activity and employment within the population, large differences existed between men and women. As noted above, women dominated the large domestic labor industry in Madrid, and between 1900 and 1930, there was hardly any change in the percentage of women employed as maids, nannys, etc. The changing employment status of men, however, drove a lot of the economic sector changes in the first third of the 20^{th} century. Overall, the percent of men employed in the "modern" workplace (industrial, government, and other professional jobs) nearly doubled to 43% in 1930. The percentage of those in "traditional" industries (agricultural, commercial, etc.) declined by half to nearly 15%, and an increased amount of men and women (though more men) went to school [87]. It should also be noted that as the capital city, Madrid also had a large presence of soldiers, and that over the several decades before the Spanish Influenza outbreaks, their numbers in the city at any given time had increased due to a general decrease in Spanish military conflicts.

Thus, on the eve of the Spanish Influenza outbreaks, the state of Spain and especially Madrid can be summarized as follows. While a large world war was fought one country away, Spain was in the middle of a large demographic shift; long after many other European countries, the demographic transition was finally in full swing as birth and death rates plummeted. Migration to the city from all over Spain, but particularly the non-industrialized south, drove the increase in both population and inequality. New technologies spread through the city, bringing electricity to the masses and making transportation within the city easier. And no one suspected that the widely popular and catchy song, "Naples Soldier" from the

opera The Song of Forgetting (La Canción del Olvido) would lend its name to a massive influenza outbreak in the city, later known worldwide as the Spanish Influenza pandemic [85].

Chapter 2

Re-examining strength and timing across Spain

Abstract

Transmission and spread of the Spanish flu through Spain in the fall of 1918 has been theorized, but not statistically proven. This work uses newly digitized sources of daily influenza death counts for 49 provinces in Spain, their capitals, and their remaining area to estimate the start, peak, and end dates of the wave with segmented regression, as well as the Reproduction Number of the outbreak. Then, sequence analysis groups the trajectories of each geography, identifying distinct patterns in the manifestation of the wave. The results find a distinct north to south west pattern of influenza spread in the fall of 1918. Clear differences in the Reproduction number between the provincial capitals and other areas also exist, perhaps indicating a small herald wave may have provided some immunity to those living in urban areas. The analysis highlights the importance of understanding the progression of influenza through Spain and the role consecutive waves may have played in the mortality during the fall wave.

Resumen

El capítulo 2, trata la transmisión y propagación de la gripe española a través de España en el otoño de 1918, que hasta ahora había sido descrita a través de fuentes secundarias pero que ahora se ha probado estadísticamente el proceso de difusión. Este trabajo utiliza nuevas fuentes digitalizadas de recuentos diarios de defunciones por gripe en 49 provincias de España, sus capitales y el resto de las provincias para estimar las fechas de inicio, ápice y fin de la ola, usando para ello la metodología de regresión segmentada, así como la tasa de reproducción del brote. Posteriormente, la metodología del análisis de secuencias, agrupa las trayectorias de cada zona geográfica, identificando distintos patrones en la manifestación de la pandemia. Los resultados encuentran un patrón distinto de diseminación de la gripe de norte a suroeste en el otoño de 1918. También existen claras diferencias en la tasa de reproducción entre las capitales de provincia y otras áreas, lo que tal vez indica que una pequeña ola precursora podría haber proporcionado cierta inmunidad a quienes viven en áreas urbanas. El análisis destaca la importancia de entender la progresión de la gripe en España y el papel que las olas consecutivas pueden haber desempeñado en la mortalidad durante la ola de otoño.

2.1 Background: Influenza transmission and manifestation in Spain

Considerable research has examined how spatial differences and temporal patterns affect flu mortality. In seasonal influenza outbreaks, climatic differences between tropical and northern countries appear play a role in seasonal influenza activity [90]. However, while the country of Spain contains diverse climate biomes that may affect transmission at a larger scale, the rapidity and strength at which the outbreaks spread does not appear to have played a large role in its movement through the Iberian peninsula [141, 265].

Moreover, the geographic route of transmission also can vary considerably, dependent on the amount of human interaction and movement from place to place. Very often, research has found a correlation between the route of transmission and high volume transport networks, including during the 1918 Spanish flu [232]. Historically, the transmission network that defines the mechanisms of disease spread is widely used and connects large nodes of population, whether on foot, by train, or in contemporary society, by airplane [147, 300, 302]. Another study has also looked at transportation networks at the time of the fall 1918 Spanish flu outbreak in Newfoundland, trying to understand the connection between when cases were first identified in conjunction with nearby transport nodes [207], and a recent analysis found that in India, the disease spread along rail networks [232].

Some parts of Spain experienced a herald wave in May and June of 1918, which was thought to have begun in Madrid before moving outward on main transportation routes in the country [112, 273]. This certainly may have affect the absolute mortality impact between and within provinces during the large fall wave. For example, a virus will spread more slowly in a population with some immunity (i.e. the reproduction number $R_{effective}$ in a partially-immune population will be lower than R_0),[1] assuming the virus has not yet mutated and evolved to be significantly different from the initial strain [168]. Additionally, the total length of pandemic wave is important, as it describes the heightened risk of others to contract the virus, allowing it to continue to spread. Several calculations for the R values of the Spanish Influenza outbreaks have found that despite its great mortality burden, the Reproduction Number tended to fall in line with other epidemics and pandemics in the 20^{th} century [179]. Rather, the high case fatality rate, roughly ten times higher than a normal influenza epidemic, drove much of the high numbers of death [107, 179].

[1] The Reproduction Number (R_0), can be interpreted as the number of additional cases a single case will cause [60]. In an epidemic period, the Reproduction number will be greater than one, meaning that the number of infected individuals increases as time progresses.

Calculations of 45 cities in the United States found that an “extreme" R would still have been less than 4, additional estimates in Winnipeg, Canada were found to be 2 [60, 179]. These values are even lower in some South American estimates; depending on the generation interval, Boyacá, Colombia estimates ranged between 1.4 and 1.7, and in Lima, Peru, the number was revealed to likely be between 1.3 and 1.5 [67, 69]. In Europe, studies calculated Copenhagen, Oslo, and Gothenburg had higher values in the spring wave, and in Geneva, the opposite was found, with a mean estimated spring value of 1.49 and fall of 3.75 [24, 58]. A review across several studies found that among 51 Reproduction Numbers reported from studies on the 1918 pandemic, the median value was 1.8 [42].

Several research articles explain and quantify the outbreaks in individual provinces and regions (for example in Palencia and Logroño [33, 88]), but in order to fully understand the totality of its transmission throughout Spain, the entire country should be considered together. A comprehensive background on the Spanish Influenza virus in Spain qualitatively describes the four unique, but not universally experienced waves of influenza in Spain beginning with a herald wave in May 1918 and ending with a large echo wave in the winter of 1919-1920 [95, 273]. The brunt of the spring herald outbreak was found in the city of Madrid [64, 72]. Attempts to quantify the spatial-temporal influenza patterns throughout Spain found significant variations in excess mortality among 49 provinces of Spain, finding cumulative (across all waves) excess rates as high as 212.2 (per 10,000) in Zamora and 6.2 in the Canary Islands [64].

Slightly over half of the provinces experienced herald and or winter waves in the spring and summer of 1918 and winter of 1919, but with the exception of Madrid, this mortality paled in comparison to the dramatic fall wave of 1918. As the virus spread through the peninsula, seemingly nothing could stop it. As the virus began to rage, there was some hesitancy to declare an epidemiologic emergency, due to the perceived negative implications of having to cancel planned festivities, temporarily close offices, and provide additional resources. However, the epidemic was first officially recognized in Valladolid, and other entities soon followed suit [217]. Nonetheless, while often official events or offices were closed to prevent transmission, such non-shuttered gathering places as church and theaters still allowed for continued transmission of the virus between individuals [40, 95]. In fact, people often flocked to church in order to pray for the end of the pandemic, even further allowing the dissemination of the virus [95].

When the epidemic struck, doctors and other medical officials, especially in smaller cities where their numbers were few, were overwhelmed with the number of cases, both stumped by the virulence of the disease and the lack sufficient supplies and supplies to care properly for each patient [33, 40, 88, 217]. Efforts were made to educate and encourage the public

in methods to individually reduce individual transmission via disinfecting and avoidance of individual contact. Large-scale attempts to contain the virus were also employed via use of disinfectant on streets, in railway stations and on trains, and in other public and private spaces [40, 95, 236, 273]. However, its overall effectiveness in preventing spread it not known.

Two more waves of influenza hit the country in the first several months of 1919 and later that year in December 1919 to January 1920, when an echo wave occurred. Considering the entire country, the total mortality impact of the third wave pales in comparison to the fourth [273]. The echo wave in 1919-1920 is well documented in qualitative accounts, but the impact of this outbreak in contemporary literature is muted [26, 95, 120]. In this last echo wave, most reports note mortality was concentrated in 0 to 1 year olds, who would not have been alive during the previous outbreaks.

In the work quantifying provincial excess in Spain, the data used contains several issues limiting the overall interpretation of the results. Not only does the “Boletín mensual de estadística demográfica sanitaria" data end in 1919, making the study of the aforementioned echo wave impossible, the death counts are presented as monthly statistics, which make it difficult to pinpoint the onset timing of each wave. The quality of the data in these monthly bulletins also varies by each Spanish province, as the data was aggregated based on reported statistics from each town in the province and do not have universal coverage based on various response rates [219].[2]

A large debate also revolves around the role of social status and mortality risk. At the individual and community level, those with lower social status, especially among youth and elderly, tend to experience a higher mortality risk from seasonal influenza [77, 187], but questions remain about this pattern during epidemic outbreaks. Some analyses point to the similar gradients [123, 160, 163, 271], while others find no relationship or conflicting results between class and mortality, perhaps a result of extreme virulence of the strain [78, 271, 277]. In the context of Spain during the Spanish Influenza outbreaks, provinces and their capitals can be examined as individual entities with varying population structures, main occupational sectors, median incomes, population density, etc.

Disentangling the role immunity and environment (and thereby transmission processes) played in total and wave-specific mortality risk has been and remains a complex process. This chapter first attempts to quantify the progression and strength of each epidemic Spanish influenza wave in Spain, using newly uncovered daily mortality logs for each province and provincial capital of Spain in 1918. Using linearizion techniques to estimate breakpoints and linear slopes of the start, peak, and end dates of each wave, one can estimate to the

[2]For example, in the province of Cuenca, the median percentage of reported statistics (based on a weighted average of the total population in the province) is only 38 percent (low of 21% in July/September and a high of 53% in August), but in the province of Toledo, generally, there is 89% coverage of statistics.

day when each wave began and map its progression through Spain, as well as the wave strength (via the R). After these initial estimations, the relationship between wave strength and province-specific measures can be analyzed to examine to what extent preceding waves (immunity) and other socio-demographic variables may have played a role in overall excess mortality due to these waves.

2.2 Data: monthly counts and death records

From May through July and September through December 1918, the Spanish Institute of Geography and Statistics collected daily mortality death counts for each provincial capital and entire province in Spain, thus capturing with detail the entirety of the spring wave and most of the immense fall wave throughout the country [299]. Called the “Resumen (Mensual) del Movimiento Natural de la Población de España y de las Capitales de provincia," or Monthly Spanish Statistics Bulletin, these reports include daily death counts of those who succumbed to influenza.

In the provincial capitals of Spain between September and December 1918, 21,048 deaths were recorded attributed to “Gripe" (influenza). This does not strictly take into account deaths according to respiratory and other diseases normally associated with influenza mortality, such as pneumonia, broncopneumonia, and bronchitis [181, 182]. Figure 2.1 displays the daily death rates for influenza in the provincial capitals of Spain for the available data, highlighting selected cities. In both graphs, the figure shows clear differences in the rate of growth in the initial phase of the fall wave, as well as differences in timing. However, the spring wave shows greater noise and much smaller mortality rates. Many provinces experience few deaths, making the following analysis difficult to preform. Thus, the spring wave is meant to provide a visual understanding of the spring wave to complement and contextualize the results, but will not be analyzed here.

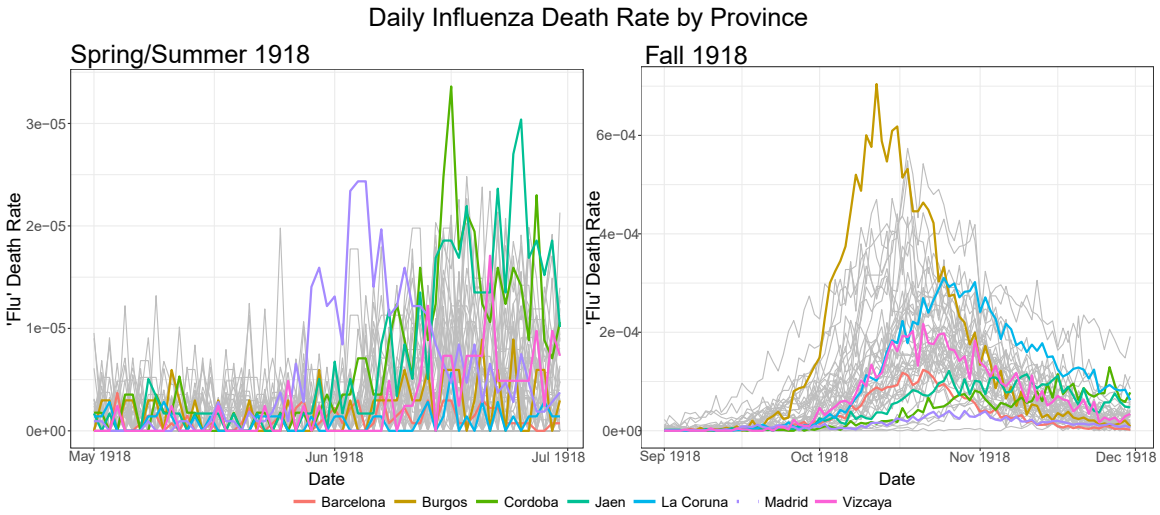

Fig. 2.1 Daily death counts for each of the 49 provinces in Spain are plotted for the period May 1st through July 31st and September 1st through December 31th in gray. Notable provinces are plotted in distinct colors.

With respect to the rest of Spain, the experience of Spanish flu in Madrid was extremely different, and so here, differences in strength and timing by wave in the city are also examined by aggregating weekly deaths from death records from the Civil Register between 1917-1922 [72, 224]. This source contains more than 103,500 mortality records, which allows the examination of the differences in the progression of the Spanish flu and each wave's (Spring 1918, Fall 1918, Winter 1919, and Winter 1918-19) strength within the city, to see to what extent differences exist within a singular urban environment. Figure 2.2 shows the raw mortality rate during the period immediately before, during, and after the epidemic outbreaks.[3]

[3]For more information regarding the death records in the Madrid Civil Registry between 1917 and 1922, refer to section 4.2.1 in chapter 4.

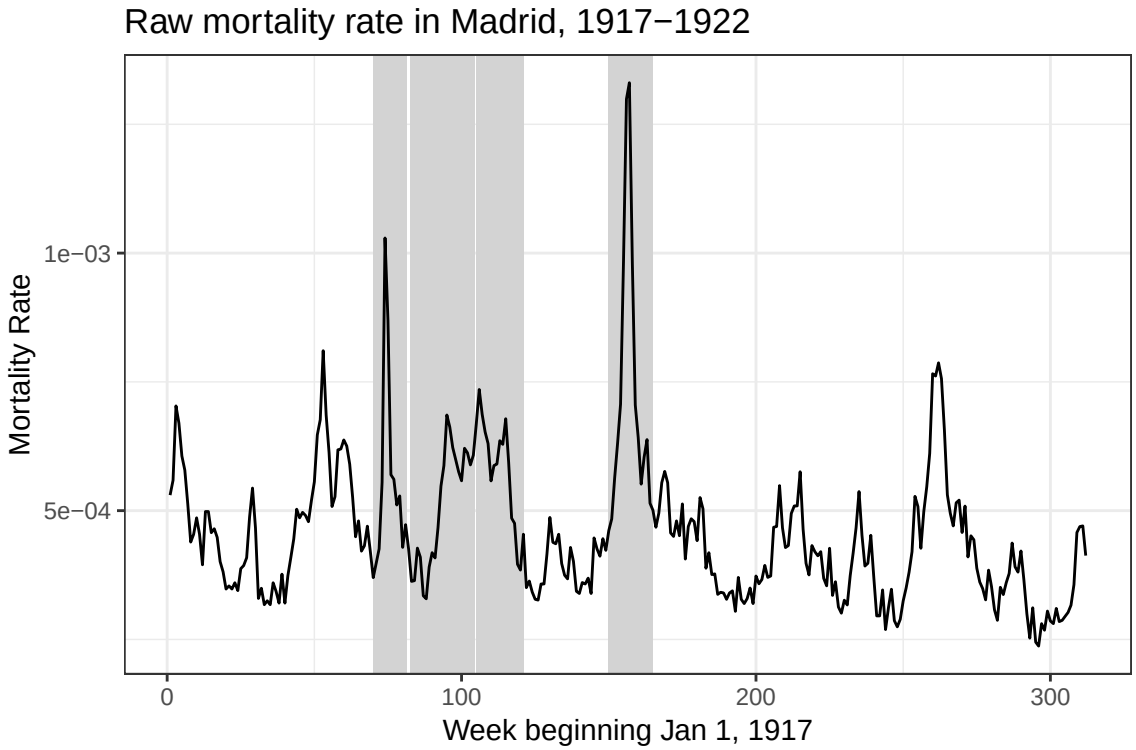

Fig. 2.2 Weekly raw mortality rate for the city of Madrid from 1917-1922 based on the civil register of death records. Defined epidemic periods are shaded in gray.

2.3 Methods

2.3.1 Breakpoint analysis

Mathematically estimating the starting, peak, and ending dates of the wave allows for a precise understanding of the timing of each attribute of an outbreak; simply using secondary sources and visual analysis of the timeline of deaths can lead to an observation bias and subsequently estimate incorrect values for the reproduction number. These dates are calculated according to a segmented method that estimates the break points of each wave based on a traditional model of an epidemic outbreak of influenza [184, 185, 196]. The typical outbreak may be generalized into four roughly linear parts: the initial seasonal baseline mortality (pre-epidemic), the period of increasing deaths as the epidemic breaks out and spreads, the decline of new daily cases following the peak outbreak period, and finally a return to baseline mortality. Due to the expected exponential increase in the number of deaths during the ascending phase, the log number of deaths is the response variable. Each phase can then be considered a log-linear component of a piecewise regression.

This linearization technique allows the estimation of breakpoints in the data, thereby identifying likely dates and associated standard errors of the beginning, peak, and end of each wave [184]. First, a general linear regression is fit predicting the log number of deaths based on the cumulation of time passed during a period of time that fully encompasses the potential dates of the wave of study. Breakpoints are estimated from the linear regression model through the R `Segmented` package [185]. The outcome provides estimates and standard errors for the breakpoints that delineate the change in phases according to the general epidemic model. Using the equation below, the first and second breakpoint are used to calculate the Reproduction number, such that:
$R_0 = e^{r*t}$
such that R is equal to the slope of the increase in deaths during the ascending period of the epidemic.

2.3.2 Sequence analysis

In order to better contextualize the progression of the fall wave of the epidemic throughout Spain, the results of the breakpoint analysis are further explored through sequence analysis. Most generally, sequence analysis allows individual trajectories to be analyzed as a whole so that general patterns and clusters within the entire group can be observed [238]. In the case of the fall wave in Spain, this analysis groups provinces or one of their sub-populations together based on the length and timing of each phase of the epidemic cycle. Normally, the general grouping and comparison of trajectories requires a large number of individual sequences to identify and group from which overarching patterns. However, while only 49 total observations are analyzed in any grouping (for each province), the clustered results can still provide some insight into the geographic patterns of wave progression through Spain.

More specifically, the sequences for each province are created by assigning an epidemic phase (pre-epidemic baseline, ascending, descending, and post-epidemic mortality) for each of the 91 days of the observation period. The change from one state to the next is defined from the breakpoints found in the segmented regression. Each of the provincial capitals for which no wave could be statistically detected were assigned an additional “no wave" value for all 91 days. After defining the “distances," or general differences, between each trajectory and the cost of changing one sequence to another, the provinces are grouped according to the calculated similarities and differences between each trajectory [238].

While different methods exist to determine the extent of dissimilarity between sequences (provinces), many are not optimal for this analysis, as they take into account changes in the progression of one state to the next [263]. The primary interest lies in the timing and length of each spell, as the inherent structure of an epidemic wave infers that each sequence

progresses through states in the same order. Thus, we estimate the dissimilarities using the Euclidean and Chi-squared methods, which are better suited to examining when and for how long each province remained in each state [263]. Separate analyses of wave progression are completed for each provincial, capital, and “rural" population.

2.4 Results

2.4.1 Breakpoint analysis

The segmented regression reveals that most provinces experienced the entirety of the fall wave between September and November 1918, though there are some exceptions. Statistically speaking, a few provinces, including Sevilla, Huelva, Guadalajara, and Cordoba, have not yet returned to pre-epidemic mortality levels by the end of November. In the case of the Canary Islands, to which the outbreak arrived the latest, the results show that the epidemic is still in the ascending phase when the availability of the daily data ceases. Despite the variations in the number of deaths and overall mortality rate during this time, figures 2.3, 2.4, and 2.5 reveal that nearly all provinces, their capitals, and their rural populations faced the same typical triangle shape of influenza epidemics mortality characterized by a rise, peak, and fall of mortality rates as they return to initial levels. A summary of the numerical values of the analysis is presented in table 2.1.

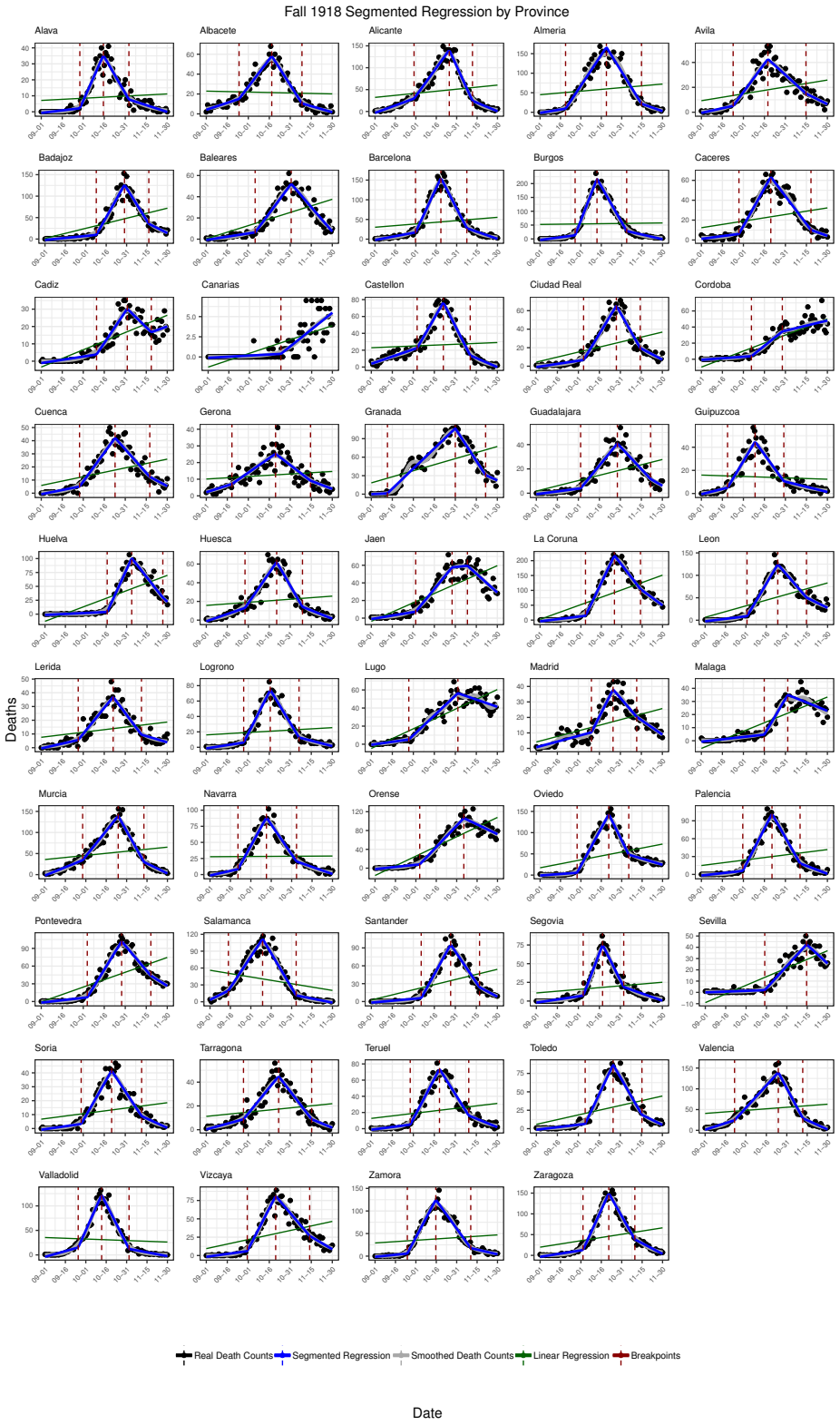

Fig. 2.3 Segmented Regression for all-province population in the fall wave 1918 shows estimated breakpoints (red dotted lines) for start, peak, and end dates of the wave in 49 provinces of Madrid. Black dots represent raw mortality rates, gray lines represent smoothed mortality, green line represents linear regression, and blue lines show segmented linear regression prediction.

Table 2.1 Summary of Numerical Results from Segmented Regression

	Min.	Median	Mean	Max.	NA's
Rural R	2.21	4.00	3.94	5.23	2
Capitals R	0.67	1.84	1.63	4.00	13
Provincial R	2.28	4.15	4.06	5.27	2
Rural Length	31.15	42.21	43.47	58.73	10
Capitals Length	13.61	36.60	36.69	51.75	14
Provincial Length	29.58	41.75	43.47	70.08	10
Rural Start	13.52	30.24	31.39	54.71	0
Capitals Start	16.64	31.37	33.65	72.35	5
Provincial Start	12.52	30.07	31.17	53.71	0

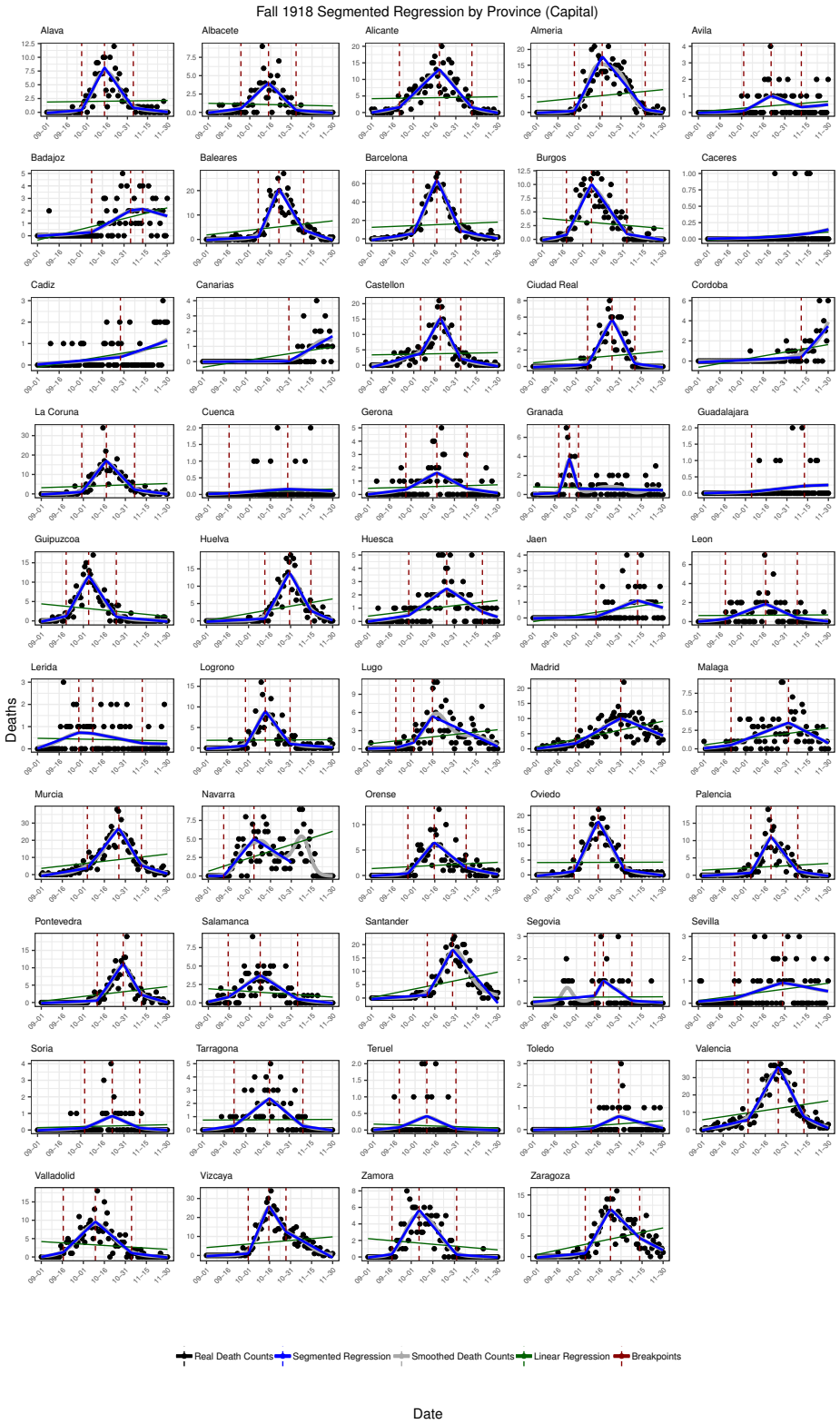

Fig. 2.4 Segmented Regression for capital population of each province in the fall wave 1918 shows estimated breakpoints (red dotted lines) for start, peak, and end dates of the wave in 49 provinces of Madrid. Black dots represent raw mortality rates, gray lines represent smoothed mortality, green line represents linear regression, and blue lines show segmented linear regression prediction.

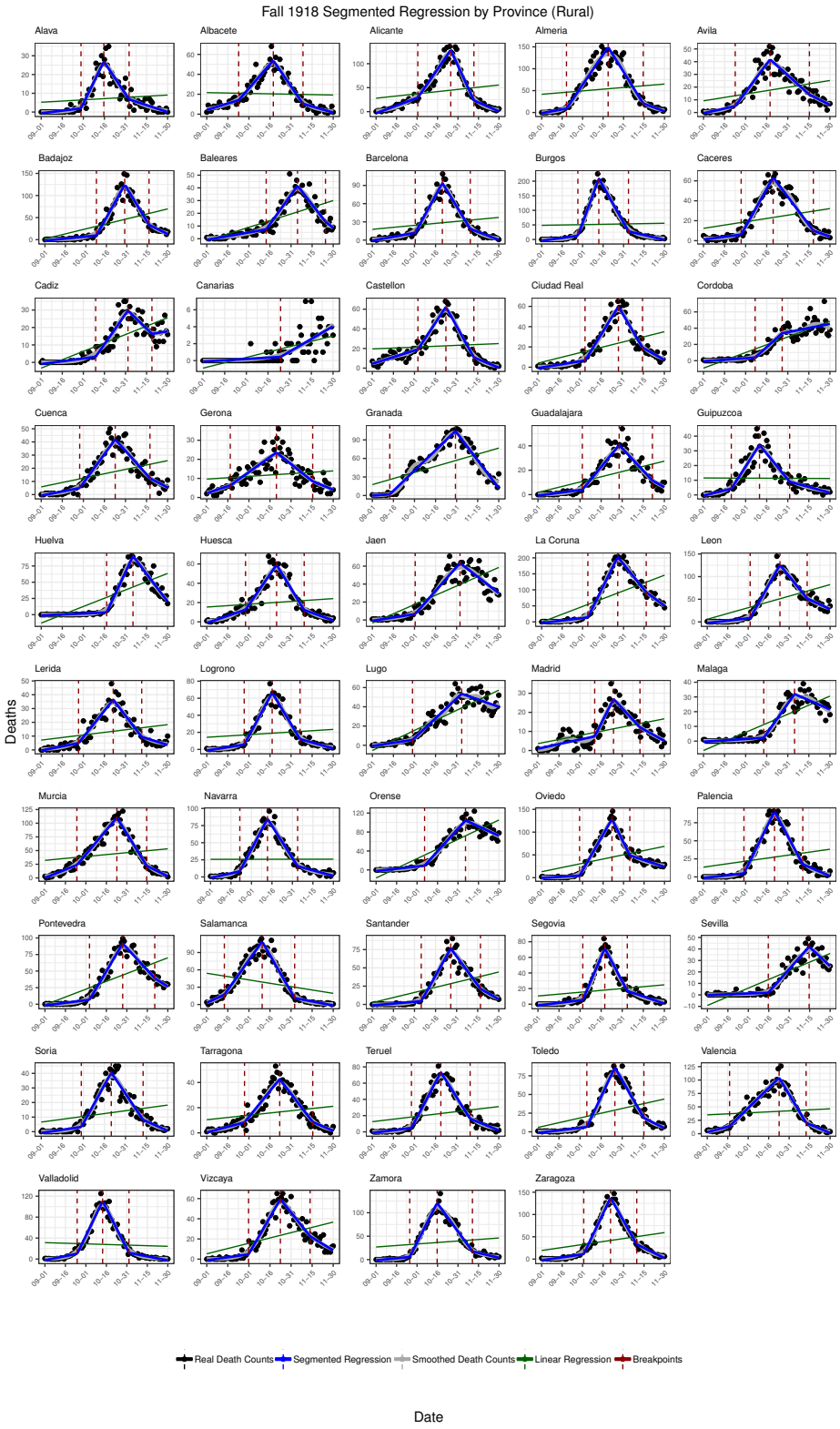

Fig. 2.5 Segmented Regression for rural population of each province in the fall wave 1918 shows estimated breakpoints (red dotted lines) for start, peak, and end dates of the wave in 49 provinces of Madrid. Black dots represent raw mortality rates, gray lines represent smoothed mortality, green line represents linear regression, and blue lines show segmented linear regression prediction.

Total provincial population

Figures 2.6 and 2.7 show the estimated start date and total length of the fall 1918 wave for each of the populations of study. In general, the color patterns of the map reveal a north to south gradient in the progression of the start of the wave, although the provinces of Granada in the south and Salamanca in the west experience the start of the wave earlier than all other provinces. While due to data limitations, the analysis could only calculate waves that lasted up to 89 days between September 1st and November 30th, the breakpoint analysis reveals that in those provinces witnessing the entirety of a wave in this time period, the length lasted between 30 (Segovia) and 70 (Granada) days in all provinces.

Start of Fall 1918 Influenza wave in Spain

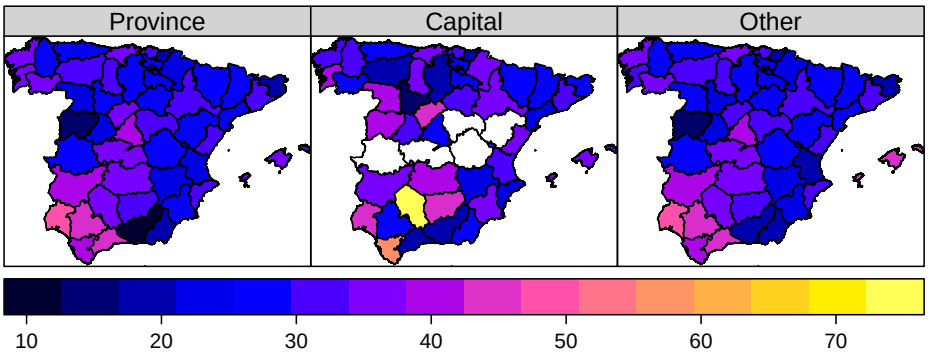

Fig. 2.6 Note: The Canary Islands are not included in this map, but the wave began later here than the other areas. The start is the number of days from September 1, 1918 that the wave began. White-colored areas mean that the start could not be statistically estimated.

Length of Fall 1918 Influenza wave in Spain

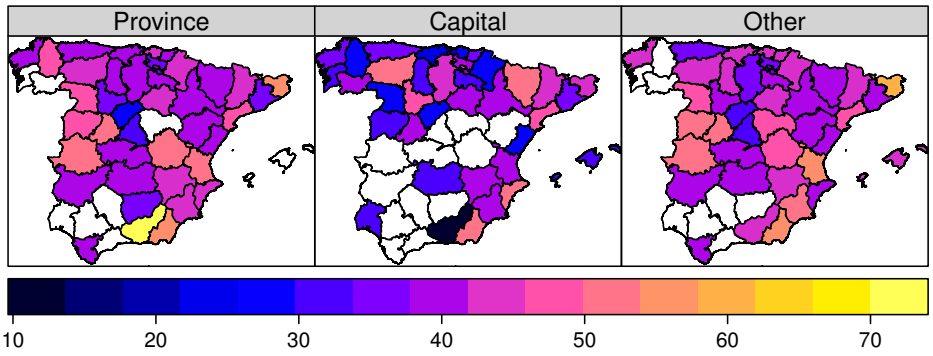

Fig. 2.7 Note: The Canary Islands are not included in this map, but the wave was not completely observed in any population. White-colored areas mean that the length could not be statistically estimated.

Capital and rural populations

When considering only the provincial capitals, which have fewer deaths than the province as a whole, the estimation of breakpoints becomes more difficult. Due to the aforementioned timing limitations of the data, the analysis only revealed start or start and peak points for some capitals. In some capitals, especially in small cities in provinces in the center of Spain, not enough deaths were present to discern a traditional epidemic wave pattern using the applied segmented regression analysis. This rendered some of the breakpoint estimations and further calculations impossible. That said, the same north to south gradient in timing found for the provinces remains when considering only the capitals, although overall, there is more heterogeneity in the capitals. This pattern is most evident when considering that the majority of provincial capitals for which peak and ending dates could not be calculated are situated in the central and southern part of the Iberian peninsula.

Perhaps most interesting, many results between the capitals themselves and their province as a whole are dissimilar. While the province of Granada experienced the longest wave, the wave in the capital lasted the shortest amount of time (13 days). However, while many capitals do experience shorter waves than their province, this is not a characteristic throughout Spain. In terms of all estimated measures calculated from the breakpoint analysis, there is no easily discernible relationship in the manifestation of waves within provinces, neither between the capital and outside population nor between the province as a whole and another population subset.

The answer to this conundrum may likely lie in such factors as capital and rural the population size and density in addition to transportation links connecting areas within and outside of the province. For example, some evidence has been found to suggest that the spring wave spread along transport routes away from Madrid [112], suggesting that connectivity played a large role in transmission. Moreover, total population size and density could also facilitate or hinder transmission mechanisms of the virus within a geographic area. If a capital city experienced a herald wave, the total time and strength of the succeeding fall wave may therefore be longer and weaker than in the surrounding, previously unexposed rural area of the province. Without clear quantitative analyses and evidence to determine the occurrence of a herald wave in the geographies examined here, the true answer remains a puzzle.

Reproduction number

As mentioned in section 2.3.1, the Segmented Regression approach to estimate the timing of the wave also results in information from which the Reproduction Number (R) may be

calculated. Figure 2.8 shows the Reproduction number for the fall wave from left to right by province, their capitals only, and for the population outside of the capital. According to the method used, the Reproduction numbers estimated are substantially lower in the capitals than in the provinces as a whole or their populations outside of the capital. This may be due to a number of reasons. First, the low number of daily deaths in the capitals, even during the wave, makes it difficult to calculate R when considering the change in slope from baseline to peak wave mortality. This type of calculation requires a continued increase in the number of daily deaths from the beginning of the wave until the peak in mortality, but this slope is quite flat (meaning a low R) when capitals have a low number of peak deaths and/or sustain this peak mortality for several days in a row. Inherently, a capital with baseline daily mortality of one and an epidemic peak of three deaths can have a maximum R value of 2, no matter how small the exposure population. Capitals may also have generally lower R values due to acquired immunity within the population. Assuming capitals are better connected to large cities than rural areas of the province, perhaps some of the population became exposed to the virus during a small, even undetectable spring wave, thus providing greater protection during the fall wave from the deadly outbreaks.

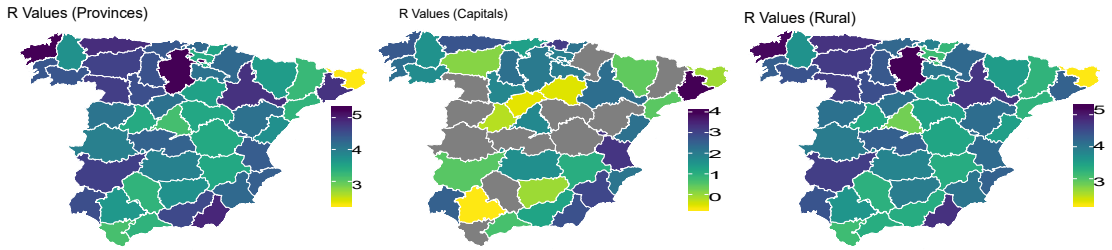

Fig. 2.8 Map showing variation in the Reproduction Number (R) of the fall wave by province, capital, and rural populations. Each scale corresponds to the map on its left, and R could not be calculated for gray-shaded areas.

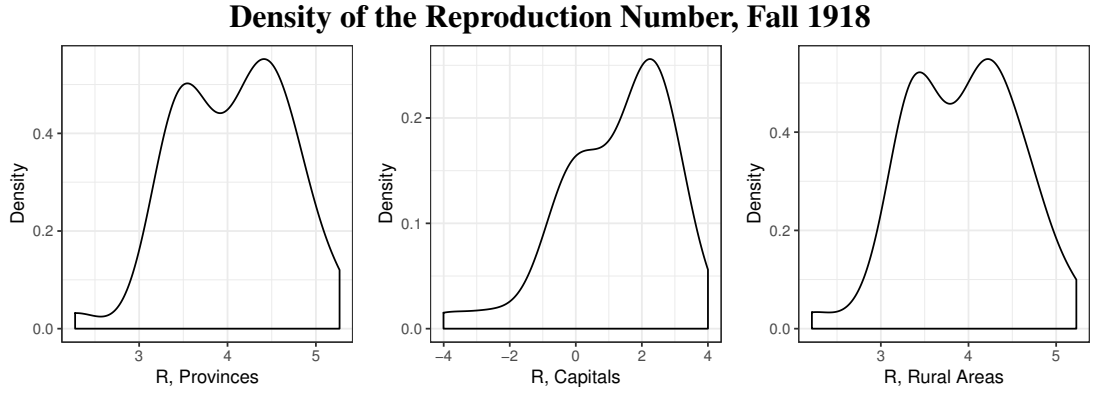

Fig. 2.9 Map showing density of the Reproduction Number (R) for total provinces, capitals, and rural populations.

The Madrid experience

The breakpoint regression results for the city of Madrid for four waves of influenza between 1918 and 1920 can be found in figure 2.10. Results for the fall wave may differ slightly than that portrayed in 2.4, for several reasons. For example, the data itself is from a different source (the monthly bulletins versus the civil register on deaths), and the data used to create figure 2.10 contains deaths to *all* causes rather than only to influenza. Nonetheless, the results show four distinct waves, of which the fall had the highest calculated R due to the pre-outbreak decreasing slope. While much of Spain did not experience the herald wave, its presence in Madrid is clear. Further analyses and discussion on the waves in Madrid are found in chapters 4 and 5 and appendix E.

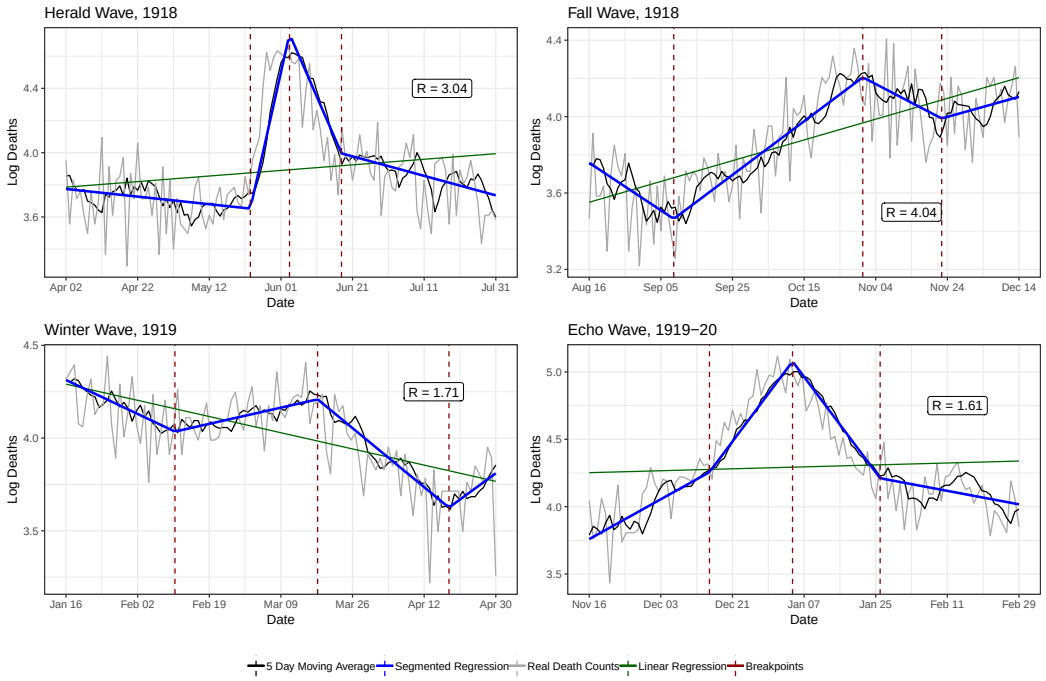

Fig. 2.10 Results of the breakpoint regression for four pandemic waves in the city of Madrid

2.4.2 Sequence Analysis

After defining the state trajectories for each province, capital, and rural area, the sequences were grouped based on two measures of dissimilarity. Each of the two methods used revealed a different optimal number of groups, 5 (Euclidean) and 3 (Chi-2). The results of these two groupings are presented in figures 2.12 and 2.11, and additional figures demonstrating the results of the analysis are found in Appendix B.

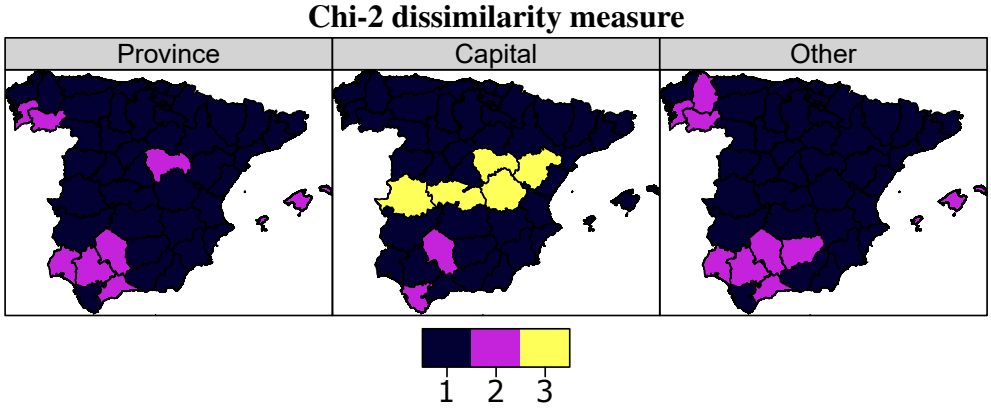

Fig. 2.11 Three optimal groupings based on timing and time spent in each state.

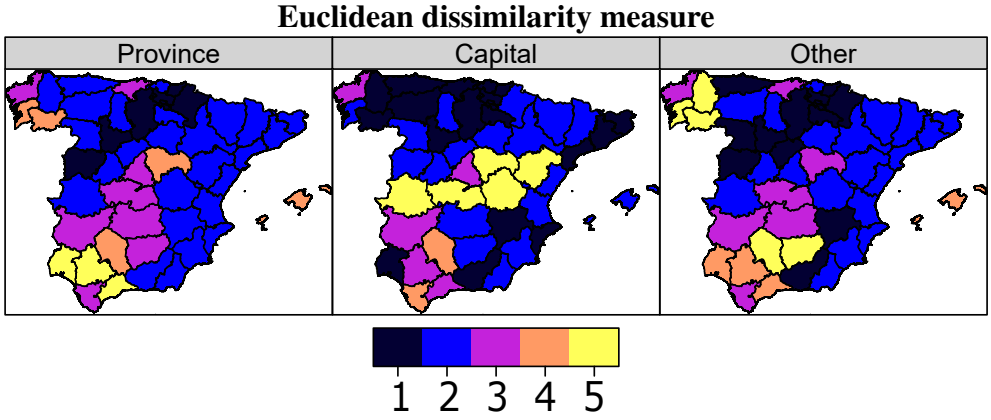

Fig. 2.12 Five optimal groupings based on timing and time spent in each state.

The Chi-2 method, which found three optimal groupings, mainly classified the provinces, capitals, and rural areas according to the extent that the wave passed through. Most generally, those places that experienced all phases of a wave are typified in the first grouping, while those places, particularly the capital cities, in which no wave could be discerned, are classified into the third group. As a whole, this clustering of the wave trajectories does not provide much additional insight into the progression of the wave.

When using the Euclidean distance measure to calculate trajectory dissimilarities, most of the hierarchical clustering algorithms identified five as the ideal number of typologies with which to group the geographic units. These five groupings are generally classified according to the extent that a unit experienced the phases as well as the amount of time spent *in* the phases. Most of these variations are due to differences in the timing of wave onset. For example, considering the entire province, Type 1 and Type 2 differ in the average amount of time spent in the pre- and post- wave states, while Type 3 provinces also experience all parts of the wave, but at a much later time than the first two parts. All geographic levels also

contain at least one grouping that includes provinces that did not experience a complete wave in the period of analysis, and the capital cities include a typology for those (in the middle of the country) for which no wave could be statistically detected.

Especially at the provincial and rural levels, a north-to-south gradient is visually apparent on the map (figure 2.12). However, one can argue that the groups at the capital city levels are also spatially correlated. While it is difficult to calculate measures of spatial autocorrelation for categorical variables (i.e. the typologies), Moran's I calculations for the start dates reveal significant levels of spatial autocorrelation (see table 2.2). Overall, the detected sequence clusters and their visualization on a map reveal statistically what previously had been reported only through anecdotal and qualitative research. These results conclude that the fall wave in Spain began in the north at the western edge of the French border, then spread west and south towards Portugal before diffusing northwest, northeast, and finally towards the southwest corner of the country.

Table 2.2 Moran's I for selected values

Geography	Moran's I (start date)	p value
Total Province	0.3301	3.7e-05
Rural Areas	0.3694	5.0e-06

2.5 Discussion and conclusions

This chapter expands upon a previous work looking at the spatial-temporal excess mortality patterns by examining the spread of influenza through Spain in the fall of 1918 [64]. In this analysis, daily death counts are used to estimate the start and progression of the fall wave of Spanish influenza through provinces, their capitals, and rural populations in Spain. The results point to a clear north to south and slight east to west gradient in the movement of the wave. The province of Granada in the south of Spain, home to a port, also experienced the start of the wave very early. This analysis is important in that it statistically confirms previous qualitative accounts and general theory that in the fall, the flu arrived from migrant workers in southern France returning to their homes in Spain and Portugal [40, 95, 273].

As noted, at the time of the fall 1918 outbreaks, large numbers of migrant workers from southern Spain and Portugal were returning home from temporary work in southern France [40, 236]. With World War One in its final days and the return of soldiers eminent, their efforts in mines were no longer needed. Given the importance of railway connections at the

time, in terms of transport, the spread of the flu in Spain found in these results must be thought of in the context of these arteries of migration and prior anecdotal evidence and accounts [88, 95, 235, 273]. Figure 2.13 shows a network map of the railways in Spain and Portugal in 1921 [290]. Presuming that little change in the structure of these lines changed between 1918 and 1921, it is very easy to see the line that Portuguese return-migrants would have taken home, from San Sebastian and Bilbao down through Burgos, Palencia, and Salamanca. This train path directly correlates to that of the sequence analysis shown in figure 2.12, in which these very same provinces first experienced the fall wave. This very much agrees with what other studies have found regarding the importance of transportation networks and the transmission of disease. Moreover, the timing of the epidemic in Northern Portugal occurred slightly later than in the ports of Lisbon and Porto, and corresponds more to the pattern of timing found in northwest Spain [193].

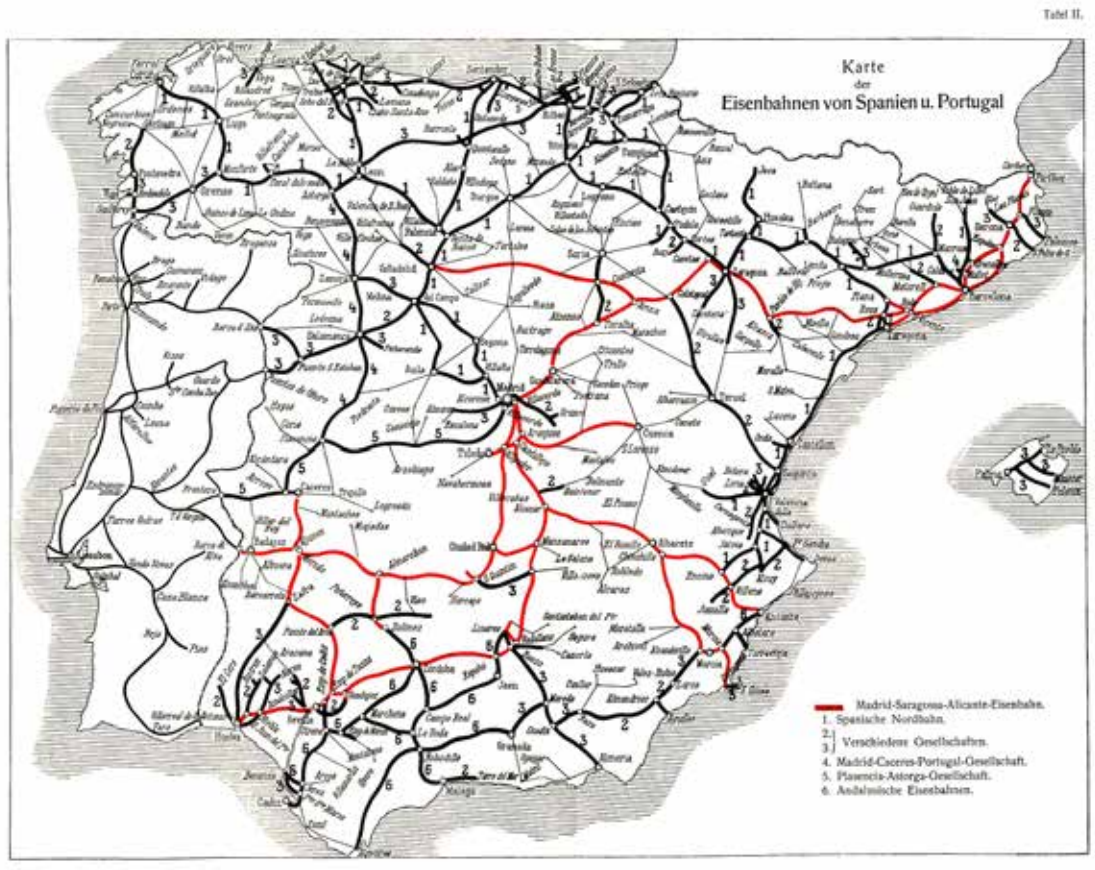

Fig. 2.13 Transport Map of Portuguese and Spanish Railways. Source: Enzyklopädie des Eisenbahnwesens 1921

Clear differences in the manifestation of the wave in provincial capitals and other populations are also apparent. Generally speaking, the Reproduction number (R) is lower in

the capital cities than in the other areas, indicating a slower spread, perhaps to a higher percentage of immune. This would contradict other findings suggesting higher transmission and mortality in urban environments [59, 175, 210]. However, the greater interconnectivity of these urban centers may mean that the herald wave spread to the capitals and their populations became exposed to an earlier strain of the virus [24, 59]. This could theoretically provide some cross protection and immunity in the fall wave, resulting in lower transmission and mortality in the outbreak [24, 37].

Due to the small amount of deaths and the incomplete coverage of the herald wave in the spring and summer of 1918, this study does not quantitatively consider the role spring wave mortality may have played in the larger fall wave. However, purely visual evidence in figure 2.1 shows that those highlighted cities with a clear mortality peak during the spring wave appear to have a muted fall wave (i.e. Barcelona, Madrid, Cordoba, Jaen) compared to those cities without an obvious peak in the spring who then experienced a larger wave in the fall (i.e. Burgos, La Coroña, Vizcaya). Contemporary studies find that urban areas and transportation hubs play a particularly large role in the transmission of influenza throughout the world [75]. The provinces with a clear spring peak tend to contain much larger cities at the time and served as transportation hubs; their exposure to the initial wave may have provided cross protection in the fall, contributing to lower mortality and influenza transmission [37].

Furthermore, while this chapter has identified the timing and strength of each wave, as well as the total progression through the country, more work can be done to ascertain some of the inter-group differences in the sequence analysis. As Spain was, and remains today, a diverse place of different cultures, industries, and climates, provincial, and urban and rural differences should, in the future, be examined from a greater lens. Did the demographic, social, and economic make up of an area, in conjunction with climactic and geographic differences, play a role in the strength of each wave. Moreover, given the emphasis of consecutive waves in this work, more information must be found detailing the possibility of a spring outbreak in each province to determine the role it may have played in the fall.

This study clearly identifies the geographical spread of influenza through Spain and finds differences in the force of transmission between sub-geographies through the calculation of the Reproduction number, suggesting a relationship between the presence of a herald wave and the strength of the fall 1918 wave. Studies focusing on modern transmission of influenza continue to focus on the spread of the virus both throughout the world and within smaller environments, often identifying human transportation flows and their density as a main method of transmission [75, 283]. Other findings advocate that population heterogeneity may play a role in the spread of influenza [55, 176]. While determinants of influenza spread are important, epidemics often consist of more than a single wave, thereby changing the initial

mechanisms of transmission. These results are a reminder that those modeling and seeking to understand epidemic spread and strength must also consider the role successive waves may play in the strength and mortality impact of an outbreak.

Chapter 3

Estimating a seasonal mortality baseline from limited data

Abstract

Quantifying the strength and timing of epidemics requires a reasonable expectation of seasonal baseline mortality. However, in historical and some subgroups of contemporary populations, it is difficult to find this information at weekly or daily intervals. Using several data sources of varying temporal aggregations (individual death records, weekly, and monthly aggregated death counts primarily related to the Spanish flu), this chapter explores traditional methods of baseline and excess mortality calculations as well as some adaptations. Then additional ways to calculate and refine the seasonality in a yearly mortality baseline using both Metropolis-hastings MCMC and interpolation are presented that provide a probability based approach to estimation. Baseline and excess mortality estimates are shown from all models, followed by a discussion of the merits and practicality of each method.

Resumen

El capítulo tres, trata de cuantificar la fuerza y el momento de la epidemia, estimando la mortalidad estacional de referencia. Sin embargo, históricamente y en algunos subgrupos de población contemporánea, es difícil encontrar esta información a intervalos semanales o diarios. Utilizando varias fuentes de datos con agregaciones temporales variables (registros de defunción individuales, recuentos de defunción agregados semanales y mensuales relacionados principalmente con la gripe española), este capítulo explora los métodos tradicionales de cálculo de la mortalidad base de referencia y con ello el exceso de mortalidad, así como algunas adaptaciones. Posteriormente se presentan formas adicionales de calcular y refinar los valores estacionales de referencia utilizando tanto el método Metropolis-hastings MCMC como la interpolación que proporciona un enfoque basado en una estimación probabilistica. Se muestran las estimaciones de la mortalidad estacional de referencia y el exceso de mortalidad de todos los modelos, seguidas de una explicación de dichos modelos y la utilidad de cada uno de ellos.

3.1 Background: influenza epidemics and the seasonal mortality baseline

Contemporary estimations claim the influenza pandemic events between 1918 and 1921, the so-called “Spanish" flu, account for the deaths of more than 50 million people throughout the world [141]. The series of successive influenza virus outbreaks gripped the world beginning in early 1918, however results of various phylogenetic and molecular-clock analyses indicate the initial circulation of the virus from avian or swine and other mammal species to humans may have occurred as early as 1911 [258, 298]. The difference in influenza-related mortality by year and flu sub-type is often examined by using age specific mortality rates as both an absolute value and as a ratio of excess mortality rates between vulnerable (young and old) and lesser affected populations [256]. Ergo, a reliable estimation of baseline overall and baseline mortality is essential.

The first major world-wide influenza pandemic during the modern age of transportation was that of the “Russian" flu in 1889-1890, spreading to every corner of Europe in only 6 weeks and throughout the world as the winter progressed [101, 218]. Due to this shift in the speed of which influenza pandemics spread and their increase in scope, the study of pandemic timing has also become a topic of interest. However, within the context of the 1918 Spanish flu, quantitative research about the specific timing of each wave is difficult, as at smaller geographic levels, most longitudinal mortality data is aggregated. Thus, while some areas collected daily information about the flu as it progressed, it is difficult to ascertain how its specific timing may have differed from baseline flares of seasonal influenza. Smaller intervals of baseline mortality data in these areas may provide a better indication of the specific time-frame from which the virus spread not only into Spain, but through the rest the world. While the location of the first human infection remains unclear, the virus likely moved to Spain via Spanish and Portuguese labor migrants returning to the Iberia peninsula from Southern France during the World War I [273].

In fact, the actual timeline and progression leading to the virus’s emergence is debated, though the H1N1 strains responsible for the Spanish flu may have been present in both swine and humans more than 5 years before the first waves in 1918 [258]. While strains of the H1N1 virus continue to circulate in the form of seasonal influenza viruses, biological remnants of the particularly deadly 1918 strains are still found in avian species via the presence of specific encoded proteins [294]. In this manner, continued research into the timing of the Spanish flu, as well as its health and mortality impacts on different populations is essential to understanding the potential effects that a virulent influenza strain could have on the global population today. To quantify these impacts at a refined level, a reasonable and reliable

mortality baseline (at small time intervals of time) in non-epidemic years from which excess mortality may be determined is vital.

3.1.1 Seasonality in mortality

Distinct seasonal fluctuations in mortality vary greatly according to some such factors as geographic differences and specific cause of death. These intra-year variations have been noted and studied in different contexts, and these peaks change over time and may be linked to the effects of modernization and population aging [153, 222]. Traditionally, countries with different climates experienced mortality peaks at different times and by different magnitudes; before technology allowed for better inside climate control, summer peaks could be found in warmer environments, while winter seasonal peak of mortality were generally higher in colder locales [153]. It should be noted that the degree of seasonal difference varies from place to place and is not necessarily dependent on climactic environment or level of development. As the age structure of the population in more advanced countries changes and ages, the general pattern of mortality seasonality again changes, and this may be driven by causes of death more likely to affect those at advanced ages [100, 222]. For example, in contemporary developed countries, respiratory-related causes of death tend to affect those at older ages to a greater degree than those in young- and mid-adulthood, and these diseases face a higher degree of seasonality [100]. While historical populations faced different population dynamics than this modern example, the importance of understanding and modeling this seasonality when quantifying an epidemic impact cannot be understated.

Figure 3.1 shows the raw average weekly death rate in Madrid from 1917-1922, years before, during, and after the waves of the Spanish flu epidemic (designated by gray blocks). The red line is the average level of mortality throughout the period; this line is clearly different from the seasonal pattern of mortality visible during the time period.

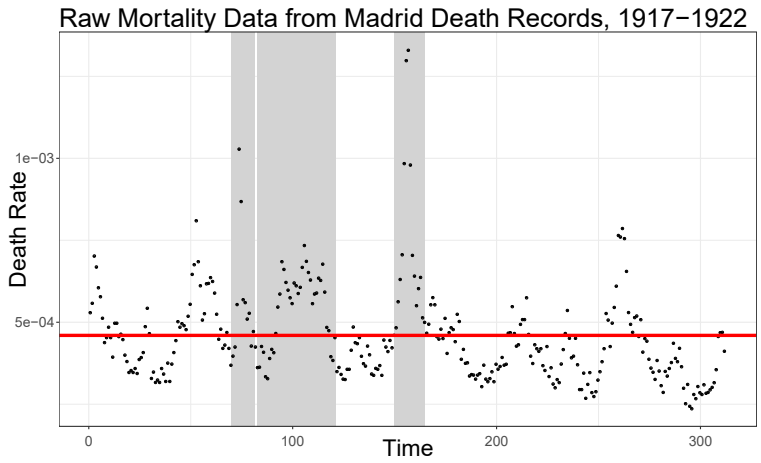

Fig. 3.1 All-age all-cause raw mortality in Madrid, 1917-1922

When further disaggregating yearly death rates by demographic variables, it becomes obvious that the effects seasonal timing plays in mortality vary greatly across age. Figure 3.2 demonstrates this; clearly the youngest and oldest ages (less than 5 and greater than 70) are much more sensitive to seasonal patterns. These groups also have substantially higher mortality than their young-adult and middle-age counterparts, who on average have lower weekly morality rates than the yearly average.

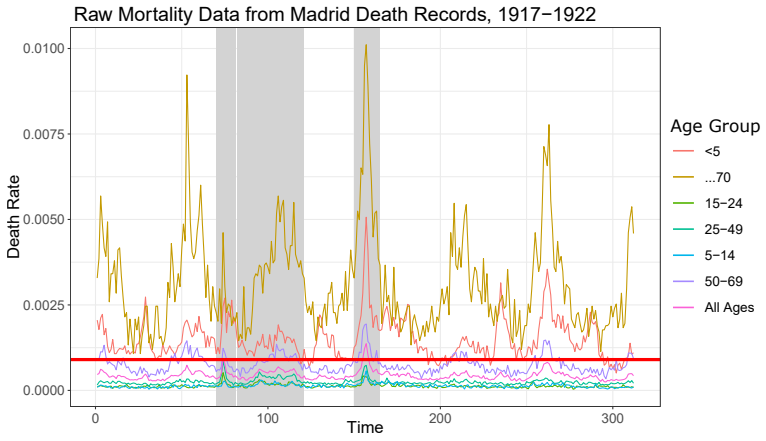

Fig. 3.2 Age-specific all-cause raw mortality in Madrid, 1917-1922

Moreover, because epidemic waves during the 1918 pandemic often struck more than once in a location and at different times of the year, calculations of excess mortality depend heavily on the estimated underlying level of mortality. For example, in the 1918 herald wave in Madrid, some weeks had excess levels well above the seasonal level, but others were barely–or not at all–above what would have been a normal seasonal peak in the fall or winter [72].

To contextualize the seasonal baseline within the study of epidemics and excess mortality, the effect of a reliable mortality baseline can shape the debate surrounding the specific topics of interest in the flu. In the wealth of information and research on Spanish and other influenza outbreaks, much examination has focused on how age-specific mortality differs from seasonal outbreaks. Often, analyses featuring little-to-no baseline mortality information show high rates of excess mortality for young adults, which in the past led to some discussion of a "w"-shaped curve of excess mortality (for example, [110, 158, 284]). However, other research calls into question this large peak; while the standard mortality ratio (SMR) continues to follow the shape of an inverted "V"–that is, the probability of dying in the epidemic period relative to a non-epidemic period is higher in the young adult ages–, the actual amount of excess mortality peaks only at the lowest and highest ages. While this is not to say only one shape of age-specific mortality curve is possible, the extent to which a baseline is calculated and implemented can have a large impact on results. Given the ongoing debate about total mortality related to the Spanish flu and its associated pattern, I found it fitting, one hundred (and one) years following the pandemic, to reexamine traditional baseline estimation methods and the application of an interpolation technique to refine aggregate data for timing analyses on a smaller-scale.

As such, the rest of this chapter explores various baseline estimations using data from the city of Madrid before, during, and after the 1918 influenza pandemic and is structured as follows. First, the data used in these calculations is briefly outlined, whose peculiarities inspired this examination. Next, the "standard" Serfling Regression model used to estimate seasonal mortality patterns is quickly introduced, then potential applications of this method are provided that may ensure a better fit and quantify uncertainty when faced with with limited pre-epidemic mortality data. The use of a Metropolis-hastings MCMC approach to estimate the distribution of deaths throughout the year by optimizing the Serfling parameters is also considered. Finally, the essay switches gears to explore how monthly-aggregated data may be interpolated in order to provide a "best-guess" of weekly mortality patterns throughout the year. After providing the results of these methods when applied to the Madrid data, excess mortality is calculated according to each method and the similarities and differences of each completed baseline as well as the applications of these methods in the future discussed.

3.2 Data used for baseline estimation

The Madrid Civil Register of Deaths provides excellent, detailed information from death records, including age, sex, marital status, location, data, and cause of death [224].[1] However, currently, the typed, analysis-ready version of these death records is only available for the years 1917-1922. This is enough to cover the full period of Spanish-flu related epidemic waves in Madrid, but only one year of data from before the first outbreak is available from which to calculate a baseline.

Other mortality data sources during the period of study exist, including the Boletín Mensual Estadístico Demográfico-Sanitaria, which provides monthly mortality information by selected causes for the entire Madrid province (both the city proper and surrounding rural area in the administrative region) [299]. While the data does not provide individual level information, these monthly death counts published by the Ministry of Government in Spain for the years 1915-1919 encapsulate the first three waves of influenza in Spain. In addition to total deaths, this data also provides counts for several causes; thus, baseline estimations for both overall and influenza-specific mortality in the city and province of Madrid could be created and compared, though they are not done so in this chapter.

Finally, information about the total city population was obtained from the Yearly Statistical Books of Madrid. At the time of the influenza outbreaks, the evolution of the city's population was recorded via a quasi-register based system, of which a census-equivalent (padrón) was taken every five years. Published on an annual basis, the volumes provide population by district both for the city and the region of Madrid given reported and estimated changes. Because the mortality data describes two different geographic areas and thus, different population numbers, mortality estimates are calibrated according to the at risk population of both the city and the region it encompassed.

3.3 Methods: Creating a mortality baseline

Several methods exist to infer a seasonal mortality baseline using the single year of 1917 data, but without additional years of death time-series information, one cannot be sure that 1917 is representative of general mortality patterns during this time. Monthly counts from 1915-1917 also allow a way to see if mortality in 1917 is significantly different from prior years, but do not allow for the same level of refinement a weekly series may provide.

[1] The Madrid Civil Register of Deaths is covered in more detail in the following chapter (4) and its appendix (C).

The Serfling regression model which incorporates parameters for time and seasonal trends in mortality is briefly reviewed. Then, some additional modifications to the traditional model are outlined that may better fit Madrid's non-traditional mortality patterns and overcome data limitations before explaining how the application of a Metropolis-Hastings Markov Chain Monte Carlo model can also estimate the mortality baseline and its upper bound. Lastly, an analytical exploration into how interpolating monthly-aggregated data can provide additional insight into mortality patterns at the weekly level is preformed.

While for epidemiologists, the "gold standard" of mortality baseline estimation remains the noted Serfling regression model and "current model" for count data, additional methods exists to calculate changing mortality patterns across the year. Several studies have begun to estimate seasonal mortality through the use of other methods, such as with Poisson counts or cubic splines (i.e. [118, 268, 269, 293]), which also allow for baseline changes during the observation period, such for the introduction of a vaccine. These methods are not specifically reviewed in this chapter. The chapter does however expand to discuss the use of MCMCs to model seasonal variations in mortality and determine the presence of epidemic waves through examining the likelihood that an observed point of weekly mortality occurs based on the parameterized baseline distribution [129].

3.3.1 Serfling regression

Often, in an effort to quantify seasonal mortality and smooth a baseline across several years of data, researchers employ a Serfling cyclical regression model [250], which provides an average mortality level incorporating time and seasonal peaks through cosine and sine parameters.

$$\frac{Deaths_{x_t}}{Population_{x_t}} = u + \alpha * (t) + \beta * sin(\tfrac{2\pi}{52.17} * t) + \beta * cos(\tfrac{2\pi}{52.17} * t)$$

After calculation of the baseline, observed deaths during the period of analysis above the upper 95% confidence interval bound of the expected values are defined as an epidemic period, from which excess mortality is calculated. However, the traditional Serfling baseline approach generally relies on the presence of several years of pre-epidemic mortality data–normally at least three years–to account for the fluidity in seasonal mortality. Additionally, shorter time periods could result in a loss of continuity between the beginning and end of the year. In the case of the 1917 data, this is displayed in Figure 3.3, where due to the change in the timing of the seasonal peak between the winter of 1916-1917 and 1917-1918, the baseline suggests there is significantly higher mortality in week fifty-two of the year than in week one.

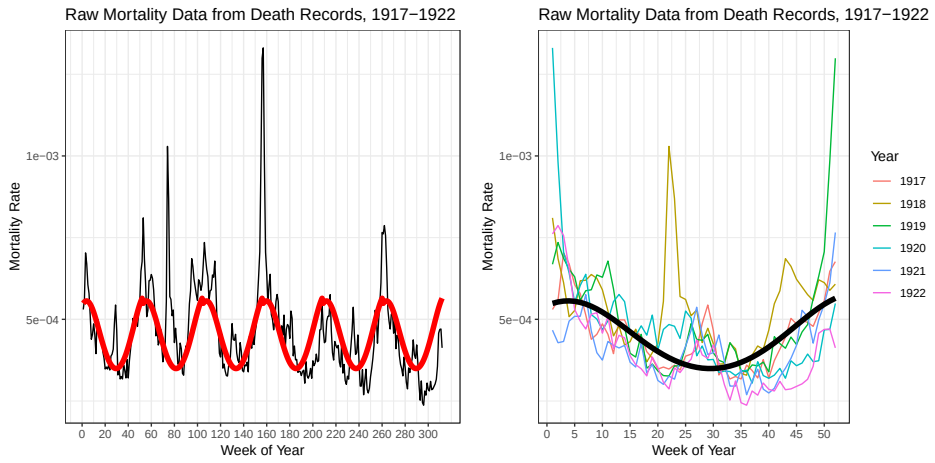

Fig. 3.3 The black line shows the real weekly mortality rates from 1917-1922, while the solid red line displays the predicted mortality values based on a simple Serfling regression (left). Lines show the real weekly mortality rates from 1917-1922, while the solid black line displays the predicted mortality value based on a simple Serfling regression (right).

Other issues can arise when mortality does not follow the parameter patterns of a traditional Serfling model. For example, Madrid experiences a small but noticeable summer mortality peak which the basic Serfling model does not account for (see Figure 3.3), common in “undeveloped" or pre-industrialized populations [153]. The simple Serfling model does not account for this bump, but this may be rectified through the addition of more time and seasonality parameters to better fit the mortality pattern [74]. While statistical research tends to avoid over-fitting the data, in the creation of a mortality baseline, the true representation of seasonal differences in mortality is paramount. If the mortality data used to create the baseline is trusted to be a correct representation of actual mortality, and that actual mortality is assumed to follow a normal pattern in the area of study, than the regression model with additional parameters used to create the baseline can be better in providing a realistic expectation of normal mortality than the simple approach. In the case of the 1917 mortality records, the visible mortality peak in the summer can be represented through the addition of parameters such that the baseline equation is written as

$$\begin{aligned}\frac{Deaths_{x_t}}{Population_{x_t}} = \quad & u+\alpha*(t)+\alpha*(\tfrac{100}{t})^2+ \\ & \beta*sin(\tfrac{2\pi}{52.17}*t)+\beta*sin(\tfrac{4\pi}{52.17}*t)+ \\ & \beta*sin(\tfrac{8\pi}{52.17}*t)+\gamma*cos(\tfrac{2\pi}{52.17}*t)+ \\ & \gamma*cos(\tfrac{4\pi}{52.17}*t)+\gamma*cos(\tfrac{8\pi}{52.17}*t)\end{aligned}$$

The added coefficients in the model account for both linear and non-linear time (α) and seasonal (β & γ) variations in normal mortality activity that create the oscillations present in the data.

Serfling Regression with parametric bootstrapping

While visual analysis reveals similarity in the yearly pre- and post-Spanish flu epidemic mortality data, only mortality information from *before* the epidemic is used to construct the baseline, as post-outbreak mortality may be influenced by the increased number of deaths during the flu onslaught [242, 256]. In fact, figure 3.3 does appear to show some decrease in overall mortality in the years following the epidemic outbreaks. Thus, it could be that using only the one year of available death records (1917) may ultimately provide an incorrect estimation of the baseline, as it forced the assumption that mortality in 1917 followed a normal pattern at all ages. By applying parametric bootstrapping to the 1917 Serfling model with added parameters, some uncertainty is generated to account for the potential variability of the 1917 data from typical mortality patterns [96].

To do so, data is first simulated before fitting the above regression model, accounting for the possibility of aforementioned fluctuation in the annual timing of winter and summer mortality peaks. A single set of mortality points from which the bootstrapped points were estimated consisted of six consecutive iterations of total weekly deaths in 1917, to mimic the six years of mortality data used in this analysis.[2] For each of these 312 week sets of weekly death counts, the number of expected deaths is simulated assuming a Poisson count distribution. The Poisson estimations assumed the mean and variance of a week were equal the observed total number of deaths in that week of 1917.

From each of simulated six-year datasets, α, β, and γ parameters are estimated according to the modified seasonal regression model above. The mean values of the coefficients from the models are used to compute the upper baseline from the upper quartile value of the 95% confidence interval of coefficients. As in previous literature, weeks with mortality above the upper baseline are deemed “epidemic" [69, 74, 160, 257, 269].

3.3.2 Metropolis-Hastings Markov Chain Monte Carlo

In the example above, numbers of weekly deaths are simulated according to parameters of their observed values in 1917 to account for the lack of multi-year baseline data. From these simulated vectors of weekly mortality data,a baseline and upper 95% certainty epidemic

[2] Each year is considered to have 52 weeks, and calculate the “total" deaths in the 52nd week of the year as the $\frac{7}{8}$ or $\frac{7}{9}$ of deaths in the final week and associated excess day(s) of the year (1920 was a leap year).

threshold are created according to Serfling regression. However, other ways exist to estimate these baseline parameters that may better take into account how each parameter affects mortality at different times. Here, an application of the Metropolis-Hastings algorithm via Monte Carlo Markov Chains recreates the distribution of deaths throughout the year [177].

The application of these methods are particularly appropriate when considering the inter-year seasonality in mortality. As modeled in Serfling regression, mortality throughout the year depends heavily on parameters that define seasonal peaks according to time points within the year. Rather than maximizing the likelihood of the regression function, the Metropolis-Hastings algorithm approaches the vector of weekly deaths as a probability distribution of deaths throughout the year. That is to say, the distribution can be expressed mathematically as the integral of the parameters from time 0 to the end of the year, or, in the case of Madrid, as:

$$\int_0^{52} t + (\frac{100}{t})^2 + sin(\frac{2\pi t}{52.14}) + sin(\frac{4\pi t}{52.14}) + sin(\frac{8\pi t}{52.14}) + cos(\frac{2\pi t}{52.14}) + cos(\frac{4\pi t}{52.14}) + cos(\frac{8\pi t}{52.14}) + \varepsilon$$

By sampling the distribution parameters within the state space, defined as the space of one year, the algorithm converges to a specific distribution by comparing the log-likelihoods of the functions for current and randomly selected proposed values [239]. After the distribution and its parameters have been optimized, the 95% probability of the distribution can be calculated from the parameter quantiles of accepted values during the random walk. As in the other methods, weeks in which total deaths are higher than this 95% threshold are assumed to be “epidemic."

Thinking of the yearly baseline mortality pattern as a distribution is quite useful. For example, one can consider that the mortality rate in a year is relatively constant–i.e., from year to year, the total number of deaths in a stable population (akin to the total density of the distribution) will not change. However, the distribution of these deaths changes throughout the year, resulting in a non-normal distribution that generally has two modes at the beginning and end of the year. In the case of Madrid, the deaths appear to have a distinct tri-modal distribution with peaks at the beginning (winter), middle (small summer peak), and end of the year. This general distribution, for which the Metropolis-hastings algorithm seeks to define parameters, follows the same general pattern from year to year. While the peaks are defined mostly according to the observed points in 1917, the uncertainty allows for the possibility that the exact timing of these peaks may vary from year to year, as does the onset of seasonal diseases and weather patterns.

Implementation

To implement the Metropolis-Hastings procedure, the target distribution is defined according to the adapted Serfling regression parameters explored in section 3.3.1, such that 9 parameters are proposed and tested with each iteration. This includes a value for the intercept, two for the time variables, and three-each for sine and cosine values. The initial proposed distribution gives a value of average mortality to the intercept and 9 to the other seasonal and time-varying parameters such that the mortality distribution is represented as a straight line, equal throughout the state space (year). A Markov Chain samples new parameters selected from a random uniform distribution, then accepts or rejects these new parameters based on the change in log-likelihood of the function [239]. The log-likelihood value is preferred over the normal likelihood as it is more numerically stable. If the log-likelihood of the proposed parameters is better than that of the previous ones by a randomly specified amount, the proposed parameters become the new values from which a new set of randomly selected proposed values are generated. Over time, the parameters convene to an optimal distribution based on the 1917 weekly mortality. This method can be generalized for all seasonal mortality distributions by optimized the log-likelihood of an appropriate set of parameters.

3.3.3 Interpolation of monthly data

The above methods explored involve mathematically determining a baseline from data at smaller intervals in order to quantify seasonal patterns in *weekly* mortality. Especially in the context of both historical and population- and geographic-specific subgroups, mortality data is not always available in such refined time intervals. Thus, here an idea to interpolating data such that weekly death counts can be inferred from monthly aggregated mortality information is presented [293]. Several methods exist to interpolate values form smaller intervals of time from aggregated values. Within the realm of fertility research, this often involves determining single-year age specific fertility rates from grouped five-year intervals. Here, similar methods to those outlined in the Human Fertility Database [137] are used to interpolate yearly age-specific rates from aggregated data.

Monotonicity is a requirement of most interpolation techniques, but both yearly mortality and the age-specific fertility curve do not follow a pattern of strictly increasing or decreasing values through time (see Figure 3.4). However, by using the aggregate amount of expected births (or deaths, in this case) across the time period of study, a strictly increasing number of total deaths can be observed; that is, there will never be *fewer* total deaths during a year on one day than on the day before.

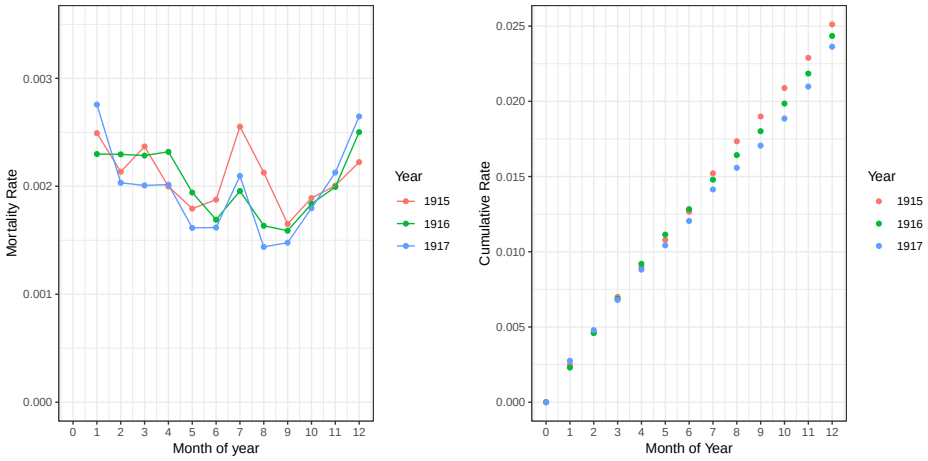

Fig. 3.4 Aggregate mortality data by month (left) and cumulative (right)

To interpolate, the logit of the cumulative (year-to-date) monthly mortality rates is calculated as $Y(t) = log[\frac{M_t}{M_{tmax}} - M_t]$, where Y is equal to the logit of the cumulative mortality rate (M) at time t. T ranges from 0, interpreted as the beginning of the year (before the nonoccurence of any deaths), to 12 (the end of December), when all deaths have occurred, equal to M_t. At these extreme t points, the logit value is replaced with reasonable values. As according to the method in the Human Fertility Database, I then perform cubic spline interpolation according to the *"Interp1"* function of the *"signal"* package in *R*.[3] The interpolation is preformed for 52 points along the logit function of cumulative mortality in order to represent the weeks of the year.

Inverse logit transformation is used to return the interpolated values to cumulative mortality rates, using the formula $M(x)_{hat} = [\frac{e^{Y_{hat}(x)}}{1+e^{Y_{hat}(x)}}] * M_{x_{max}}$. From the difference of these cumulative points, determine weekly mortality estimates are determined for three years of the monthly aggregated mortality data. The baseline and its 95% estimate are created by taking the upper bound of a Poisson distribution containing each the three interpolated points.

[3] HFD uses Hermite spline interpolation, but cubic splines are used, as they are a better fit and mathematical estimation

3.4 Results and discussion

3.4.1 Baseline results from Serfling regression with parametric bootstrapping

The estimation technique for the mortality baseline created a general linear pattern mimicking mortality patterns for the year 1917. In the years following the main epidemic waves, the all-cause baseline continues to mimic the general shape and pattern of the baseline, but the overall mortality rate falls relative to 1917. Given the additional variables present in the individual data of individual deaths, this method can also be used to create refined mortality baselines for additional population sub-groups and estimate baseline mortality accordingly.

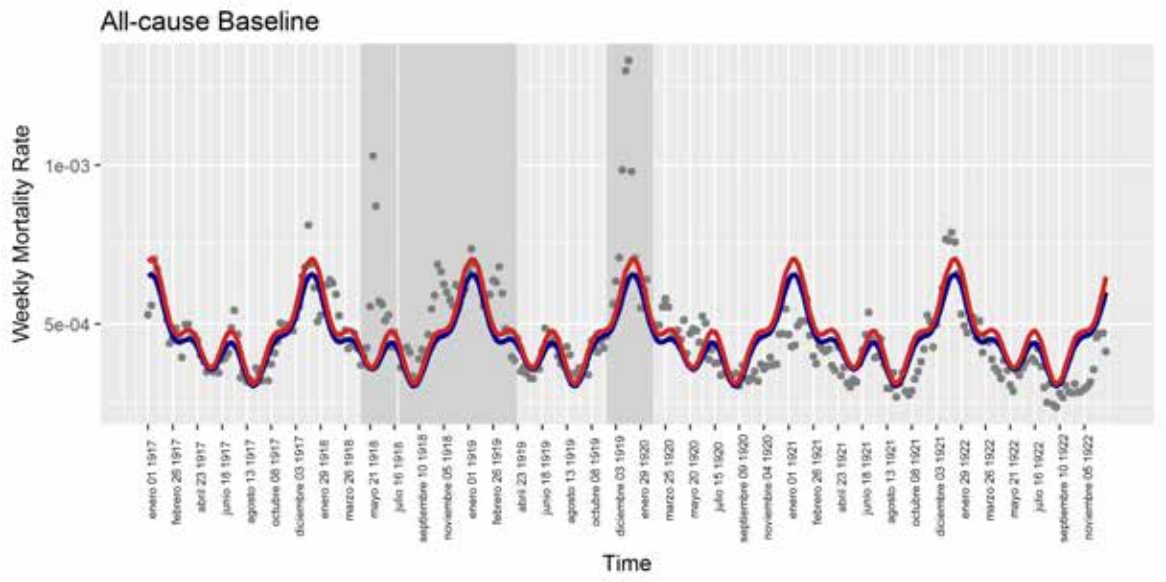

Fig. 3.5 Grey points show the real weekly mortality rates from 1917-1922, while the blue and red lines display mean and upper 95% bound baseline from simulated 1917 deaths data. Shaded gray blocks represent the three epidemic wave periods.

3.4.2 Metropolis Hastings Markov Chain Monte Carlo optimization

Figure 3.6 shows graphical results of the Metropolis-Hastings MCMC for two potential distributions. The gray-scale shows, after the a series of “warm up" steps, the accepted distribution parameters during the optimization process from start (gray) to end (black). The white line shows the mean distribution, and the red lines represent the upper and lower 95% confidence intervals. The first figure demonstrates the random walk of the optimization

process using the Madrid 1917 data. The second figure shows the results of this algorithm to a different distribution of deaths, similar to the standard Serfling regression parameters, such that:

$$\int_0^{52} t + (\frac{100}{t})^2 + sin(\frac{2\pi t}{52.14}) + cos(\frac{2\pi t}{52.14}) + \varepsilon$$

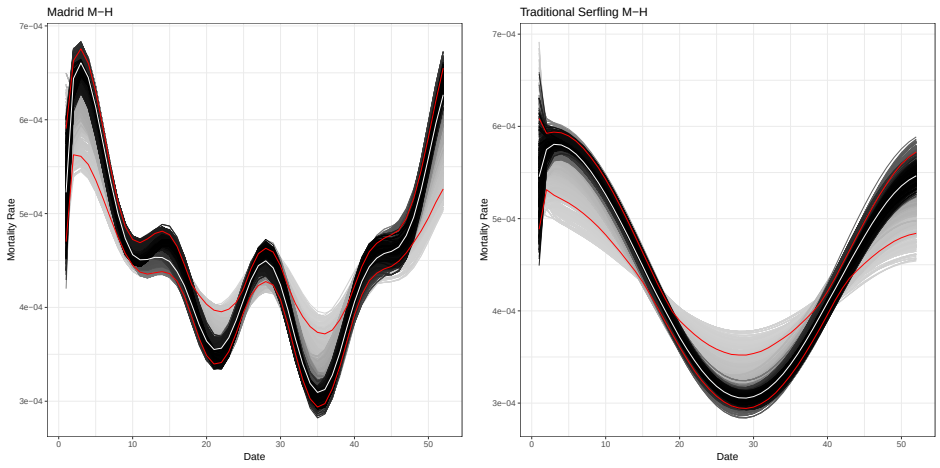

Fig. 3.6 MH MCMC results for 1917 Madrid data (left) and other distribution (right)

In both cases, the algorithm slowly optimizes the distribution by sampling and accepting or rejecting new parameter values. In outbreak periods, observed weekly rates higher than the red 95% confidence interval of the accepted parameters will be considered epidemic weeks, and the level of excess is the difference between the observed value and mean expected value (white line).

3.4.3 Baseline according to interpolation method

Initial results obtained by strictly adhering to the HFD methods revealed realistic estimations of the baseline for all months, with the exception of the first and last several month of the year. In the case of time $t = 0$ to time $t = 4$, the cumulative mortality rate is equal to zero at the beginning of the year and some value of January mortality at $t = 4$, but the interpolation function interprets this as zero deaths occurred at time 0, rather than simply that zero deaths during the year had occurred. Likewise, the end of the year shows a flattening of mortality rates towards zero as the values of the logits move towards an asymptote of the function where cumulative mortality remains the same. This is not realistic, as cumulative mortality will continue to grow even after the end of the early observation period.

To rectify this,“phantom deaths" are added to the beginning and end of the year before calculating cumulative mortality, the logits, and performing interpolation. This can be interpreted as providing information about mortality patterns before and after the end of the year (i.e. December mortality in the year prior to the baseline and January of the following year). These these six (two each for 1915, 1916, and 1917) values are randomly generated according to a Poisson distribution fit from the number of monthly deaths during the months of January and December present in the data. Thus, the initial problems of the interpolated values remain at the extremes of the interpolated function, but all values of cumulative mortality during the observation year fall within a part of the increasing logit function such that the resulting values are realistic.

Both cases of interpolation results are visually depicted in Figure 3.7, where the quick rise of mortality in the beginning of the year and plateau at the end are easily visible. The right side of the figure also depicts the change in the expected number of deaths at the beginning and end of the year without and with the addition of the phantom deaths.

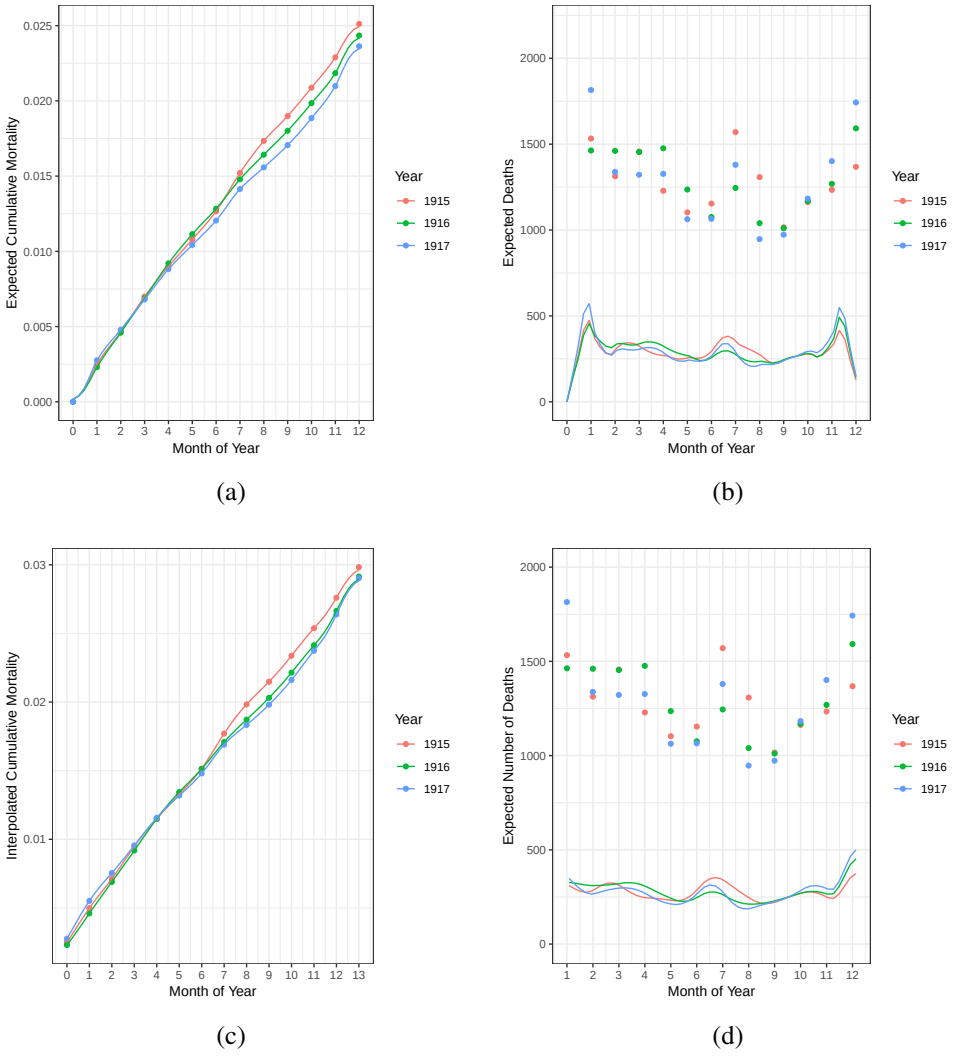

Fig. 3.7 **Interpolation issues and adaptation:** Initial technique–(a) estimated weekly values of cumulative yearly mortality and (b) mis-estimation of deaths at the beginning and end of the year. Phantom deaths–(c) interpolated cumulative mortality including additional December-January deaths and (d) realistic estimation of deaths from January to December of observed year

3.4.4 Comparison of baselines

Figure 3.8 shows both baselines overlain on all mortality data as well as for the year 1917 according to both monthly and death record data.[4] There are many similarities between the three; all share the same basic shape and take into account the summer mortality peak, and

[4]The presented results tables (e.g. table 3.1) show excess mortality estimates according to the interpolation methods for both the entire province of Madrid, as well as only the city. This is to facilitate comparison and discussion with another paper on excess mortality in Spain that used monthly estimates. The figures depicting the interpolated baselines do so according to the monthly mortality data at the city-only level.

the baselines follow nearly identical paths and values in the second half of the year. Overall, the baselines reinforce each other's legitimacy, as the three years of interpolated data validate the relative shape of the curve determined by only 1917 data. Moreover, the thickness of the confidence intervals show that neither the simulated data and the number of simulations involved in the bootstrapping nor the iterations of the MH MCMC optimization over-fit the variability in the baseline estimation.

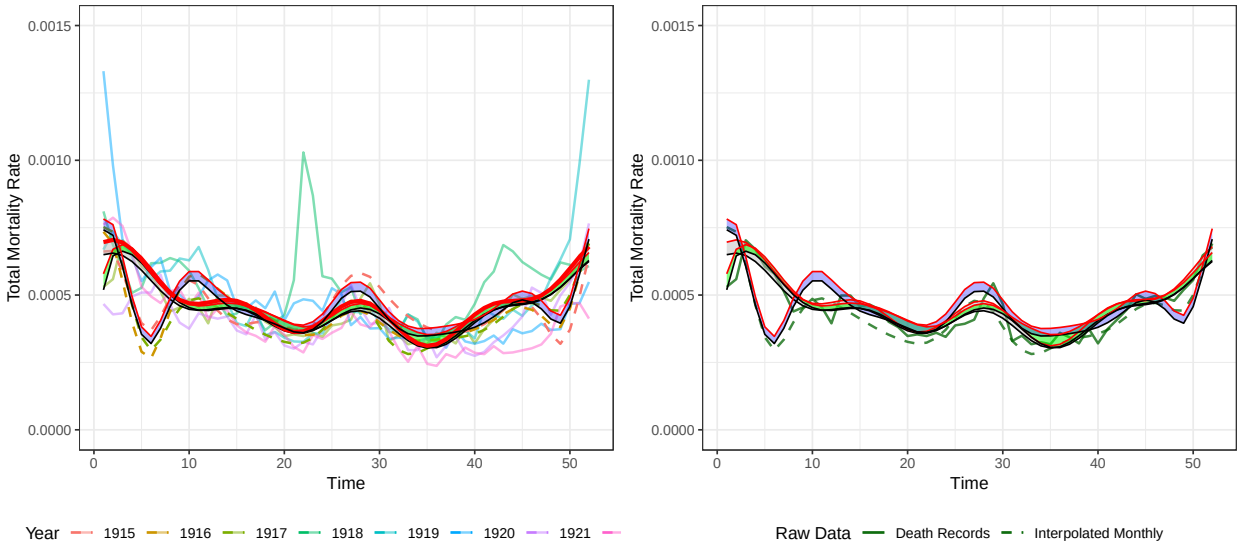

Fig. 3.8 Blue shaded region shows baseline and upper 95% of interpolated monthly data, gray shaded region shows the same for parametric bootstrapping of death records, green shaded region shows same for the M-H MCMC estimated baseline; mortality rates for all years of data plotted on left, two years of 1917 (death records and interpolated provincial counts) on right

Outside of its calculation method, the MH MCMC baseline largely differs from the parametric bootstrapping technique in that it predicts the mean and 95% uncertainty interval of the summer and winter peaks with greater certainty and the spring and fall mortality troughs with less uncertainty. This is due to the selection mechanism for the proposed parameters; it should be noted that a greater number of simulations and increase in the specified burnout period would further decrease the level of uncertainty and lower the confidence intervals. Overall, the two share very similar mean values of predicted mortality levels, though this is expected given they are derived from the dame data and are calculated using the same parameter values, even if via different methods.

Nonetheless, there are also noticeable differences in the baselines. Because the data used to construct the single-year baseline follows the distribution of deaths in 1917, the resulting shape of the mortality curve, despite the introduced uncertainty, adheres to the timing and

strength of the peaks in 1917. That is to say, the winter seasonal peak in 1917 occurred mid-way through January, and this is reflected through the early peak in the 1917-only baselines, while the interpolated monthly data peaks at the start of the year. Likewise, the summer peak was lower in 1917 than what was observed in the aggregated monthly data. However, this could also possibly be attributed to the timing of the summer peak. Whereas the individual death dates can pinpoint an exact peak, the monthly data may show a muted rise if the peak occurred during the end of one month and the beginning of another.

In general, the methods using only the 1917 data highlight the importance of using more than one year of baseline mortality data; both place winter mortality peaks in the middle of January *and* at the end of December. In a five-year continuous time-series of these rates, this would indicate a small winter “trough" in the beginning of January. However, as a general pattern of seasonal mortality, this valley is unrealistic. However, it also cannot be assured with certainty that the monthly data is interpolated correctly such that the location of the peaks within a month is where it actually occurred. The interpolation infers a smoothed rate of increase towards and away from the peak according to the monthly points, and without more precise information, researchers must have reason to believe this assumption is true.

Several additional interpolation methods exist that should be considered in order to best quantify a weekly baseline from aggregated data. Other types of methods, such as polynomial regression, were briefly considered [148], but due to the summer mortality peak, the high-order function produced results that were more complex than the seasonal parameterized baseline. Another option to determine the weekly baseline involves using a calibrated splines method to estimate four segments of the mortality curve across the year [248]. Though this technique would require estimation of several segments of the baseline, its applications to fertility data (if the age-specific fertility curve is thought of in a similar manner to the time-of-year-specific mortality curve) appear to produce at least comparable, if not better results than the current Human Fertility Database technique [125] and thus might be considered as a possible additional procedure to interpolate mortality data.

3.4.5 Excess mortality

Table 3.1 shows five estimates of excess mortality in Madrid for the influenza waves in the spring of 1918 and the combined fall/winter waves in 1918-1919. The first estimates are derived from the technique of parametric bootstrapping the single year of 1917 mortality data. The second column is calculated from the 95% probability interval of the MH MCMC derived distribution of deaths, and the third and fourth sets are estimated from the interpolated data, using the upper 95% interval of an assumed Poisson distribution, where the third corresponds to the city only, and the fourth corresponds to the entire month. The final set of estimates

come from a previously published paper on excess mortality in Spain during the pandemic [64]. These excess mortality rates are based on a simple three year Serfling model using the monthly data in its aggregated form.

Table 3.1 Estimates of Excess Mortality with Different Baseline Estimations (absolute excess rates per 10,000)

Wave	Serfling with Param Bootstrapping	MH MCMC	Interp Data & (City)	Interp Data (Province)	Monthly Data
Spring 1918	19.42	17.70	16.11	15.20	11.7
Fall/Winter 1918-1919	33.50	44.28	63.01	52.23	55
Overall	52.92	61.97	79.13	67.42	66.7

Despite the differences in the overall baseline discussed above, most of the baseline estimations presented here reveal very similar numerical results of excess mortality during the two waves in Madrid. With little exception, the estimates of excess mortality calculated from the baselines are much more similar to each other than to those of the previous study of excess monthly mortality in all provinces of Spain [64]. For example, the excess mortality from both of the baselines calculated here is much lower in the fall/winter wave, and higher in the spring. When the interpolated data is used to calculate a baseline with the MH MCMC method, higher results of excess mortality than in the other methods is also found, most notably in the combined fall/winter wave. Part of this could come from the lower expected mortality of interpolated data at the beginning of the year, but it is difficult to ascertain what specifically contributes to these differences. Also, it should be mentioned the previous estimates use the same monthly data as in the three-year baseline (but in 1918 and 1919) to estimate the excess, while here, the death records are used exclusively to estimate the excess from calibrated baselines in the first four methods of table 3.1.

The mortality data sources used in this analysis also cover slightly different geographic areas, which could account for some of the discrepancies in excess mortality. The overall timing of the baseline mortality rate for the city of Madrid may be different from the province due to varying levels of influenza immunity. Previous exposure to influenza virus and a person's triggered immune defense changes the probability of (a) contracting a specific strain of influenza from exposure and (b) the body's reaction to fighting such a virus [109]. The city itself experienced a strong spring wave and this *is* present in the provincial data due to its inclusion of the city. However, the population in the countryside of the surrounding region likely did not face the same exposure to the spring wave, meaning the monthly provincial data may not perfectly reflect the mortality conditions in the city during the outbreaks.

Additionally, by far the largest estimate of excess mortality comes from the monthly interpolated counts that were recorded within the city during the combination of the fall

and winter waves in 1918-19. This excess estimate of 63.01 is almost larger than all of the other *total* estimates from the calculations. Because the actual mathematics have been checked, then checked again, this number, should the data have been recorded correctly, stands. However, the nature of the data sources do challenge the results and may provide at least a partial explanation for this high number. The Madrid Civil Register, which provided the information leading to these high interpolated counts, accounts for *all* deaths in the city, irregardless of the place of origin or residence of the deceased. The Boletín Mensual Estadístico Demográfico-Sanitaria was compiled from self reported statistics by physicians and often lacked complete coverage of the province for which it was reported [219]. The reporting percentage and amount of coverage also varied from year to year. Moreover, reporting delays may have furthered skewed these self-reported statistics, but the extent to which that may have happened is not examined here. Thus, it is possible the number of deaths to flu were misreported either in the baseline period, during the waves of flu themselves, or at both times, but by different amounts.

In many urban centers that experienced a stronger herald wave in the spring or summer of 1918, the severity of the succeeding fall wave is less pronounced relative to the rest of the world. Often, the total effect of the epidemic was lower in cities with herald waves. While the spring wave lacked the virulence of the successive fall outbreaks, its overall transmission rates were quite high, and where present, the virus tended to spread through much of the population [24, 198]. However, excess mortality rates specific to the spring waves are generally lower than in the fall. Thus, the differences in the province-wide mortality from the city estimates may be due to a lack of exposure to the spring wave, followed by a strong fall wave to which the provincial population had no mortality. While further sensitivity analysis will be done to look at the differences in the city and surrounding populations, the current results reveal a stark contrast of more than 50% higher mortality in the province itself.

3.4.6 Final notes: seasonal mortality baseline calibration

While here, this chapter explores novel ways to calibrate a mortality baseline, caution should be taken when deciding the best approach to take. When producing a mortality baseline, it is most essential that the seasonal rates are a reasonable expectation of “normal" mortality patterns for the population of analysis. As mentioned, the monthly data used in the comparison paper covers the entire province of Madrid, whereas the individual-level death records provide information from the city only. While the provincial-level data takes into account the population of the city as well, the differences in mortality during the baseline period of 1917 are stark, specifically during the winter peaks. Both baselines accurately reflect their input data–that is, according to available data, they both paint a believable

picture of mortality–, but there are large, significant differences in the epidemic threshold, particularly in the first two thirds of the year.

The content of this chapter focuses on the city of Madrid during several waves of Spanish influenza, but as noted, these estimation techniques are applicable to all types of aggregated mortality data. Or, it can be used with other types of aggregate data that need to be examined at a finer level, provided there is no reason to believe a large spike may occur within an aggregated period. That is to say, while the method is applicable in a wide range of cases to several types of aggregated data, caution should be exercised, and it should be noted that the interpolation method will always produce a smoothed result. To further understand the benefits and limitations of the interpolation and baseline re-calculation process, future analyses should incorporate data from other locations and time periods.

Yet, in this analysis, the results show that reasonable mortality time-series at finer time intervals can be created through interpolating higher-level aggregated data. For example, the severity of the 1918 Influenza Pandemic throughout the world led many cities, counties, and other administrative areas to collect daily and/or weekly surveillance and mortality data during the epidemic periods, especially in the fall of 1918. However, the continued usefulness of this data in contemporary studies relies heavily on the estimation of baseline mortality during the time and the discernment of how much larger an outbreak was compared to seasonal flu. In these same areas with detailed 1918 information, often, previous years of mortality data are available only in a larger aggregate scale. Thus, interpolating this information, as shown through these estimations, provides a viable method through which a finer baseline may be calculated. From this, researchers can better understand some of the intricacies of the influenza pandemic and other events still debated today.

Chapter 4

Age-specific excess mortality patterns during the Spanish influenza pandemic in Madrid, Spain

Abstract

Although much progress has been made to uncover age-specific mortality patterns of the 1918 influenza pandemic in populations around the world, more studies in different populations are needed to make sense of the heterogeneous mortality impact of this deadly pandemic. Here the absolute and relative magnitude of 3 pandemic waves in city of Madrid between 1918-1920 are assessed based on age-specific all-cause and respiratory excess death rates. Excess death rates are estimated using a Serfling model with a parametric bootstrapping approach to calibrate baseline mortality levels with quantified uncertainty. Then excess all-cause and pneumonia and influenza mortality rates for different pandemic waves and age groups are calculated and presented. Age-specific analyses reveal the youngest and oldest experienced the highest excess mortality rates, and young adults faced the highest standardized mortality risk. Waves differed in strength; the peak standardized mortality risk occurred during the herald wave in spring 1918, but the highest excess rates occurred during the fall and winter 1918-1919. The analysis finds little evidence to support a 'w'-shaped age-specific curve. The results indicate acquired immunity may have tempered a protracted fall wave, but recrudescent waves following the initial two outbreaks heightened the total mortality impact of the pandemic.

Resumen

El capítulo cuatro, se centra en el análisis de la mortalidad por edad durante la pandemia. Aunque se han hecho muchos progresos para mostrar los patrones de mortalidad por edad de la pandemia de gripe de 1918 en poblaciones de todo el mundo, son necessarios más estudios en diferentes poblaciones para dar sentido a la hetereogenidad de esa letal pandemia en cuanto al patrón por edad de la mortalidad. Aquí se evalúa la magnitud absoluta y relativa de tres de las olas pandémicas de gripe en la ciudad de Madrid entre 1918-1920 en base a las tasas de mortalidad por todas las causas y causas de tipo respiratorio específicas de cada edad. Las tasas de mortalidad extraordinaria se estiman utilizando un modelo estadístico de Serfling con un enfoque de bootstrap paramétrico para calibrar los niveles de mortalidad de referencia con una incertidumbre cuantificada, es decir, permitiendo establecer unos márgenes de confianza. Luego se calculan y presentan las tasas de mortalidad por todas las

causas, neumonía y gripe para las diferentes olas pandémicas y por grupos de edad. Los análisis específicos por edad revelan que los más jóvenes y los mayores experimentaron las mayores tasas de exceso de mortalidad, y que los adultos jóvenes afrontaron el mayor riesgo de mortalidad estandarizada. Cada ola difería en fuerza; el pico de riesgo de mortalidad estandarizada ocurrió durante la ola de primavera de 1918, pero las tasas de exceso más altas ocurrieron durante el otoño y el invierno de 1918-1919. El análisis encuentra poca evidencia que apoye una curva específica por edad en forma de "w". Los resultados indican que la inmunidad adquirida puede haber atenuado unas olas prolongadas, pero las olas posteriores que siguieron a los dos brotes iniciales aumentaron el impacto total de la pandemia en la mortalidad.

4.1 Background: differences in epidemic waves and age-specific mortality patterns

The 1918-1920 influenza pandemic or the so-called “Spanish" flu, is responsible for millions of deaths worldwide [141, 186]. In Europe, the excess mortality rate associated with the 1918-19 influenza pandemic has been estimated at 1.1% or approximately an 86% elevation in all-cause mortality [26]. This pandemic rapidly spread around the world in a series of pandemic waves that gripped the world beginning in early 1918 [255]. However, results of various phylogenetic and molecular-clock analyses suggest that the initial circulation of the virus from avian or swine and other mammal species to humans may have occurred a few years earlier [206, 258, 298]. Moreover, the symptoms and age-specific mortality patterns associated with this particular pandemic are particularly unique. For example, the most severe patients were often young adults presenting with heliotrope cyanosis and acute respiratory distress. In fact, young adults consistently exhibited the highest excess mortality rates from a number of detailed historical investigations. Contrast this to seasonal influenza epidemics which primarily affect the very young and elderly [178, 256].

The name “Spanish" flu comes from the first news reports of influenza-like-illness in Madrid in the late spring of 1918. However, this pandemic gained its moniker because the first mentions of the virus were published in Spain, where the press faced no censorship during World War I due to the country’s neutrality [95]. San Isidro, the festivals surrounding Madrid’s patron saint, lasted several weeks in early and mid May, and the close contact of Madrileños at the festivities probably contributed to its transmission [112, 273]. Many fell ill with respiratory symptoms in May 1918, including King Alfonso XIII, the prime minister, and other top government officials, which was well documented in the press [95, 273]. The outbreaks were strong, rapid, and affected many individuals, though overall, mortality was reported to be relatively low. During this time, many public and private offices closed [40, 95].

Yet when the flu returned to the city in fall of 1918, public officials in Madrid seemed reluctant to acknowledge its presence and argued over its prevalence. The provincial health inspector noted the existence of a second flu outbreak in September [4]. However, even into mid and late October, officials in positions of varying importance argued publicly, through several days of newspaper editorials and official announcements over the extent to which an influenza epidemic was present in the city [112, 114]. A consensus could not be reached as to whether the outbreak was limited to a few cases in public quarters such as prisons, hospitals, and orphanages, or if it truly was widespread across the city. As the outbreaks continued, their pervasiveness could not be denied.

Because respiratory disease outbreaks occurred in neighboring France as early as April 1918, it is likely that the virus was introduced into Spain via Spanish and Portuguese labor migrants in Southern France [273]. Prior research provides abundant information regarding the timing, severity, and excess mortality of the 1918 influenza pandemic in Spain [64, 95, 273] as well as some estimates of transmission potential of the virus within the city of Madrid [64, 99, 196]. This previous work on the effect of the Spanish Influenza in Madrid and in Spain was strongly based on the historiographic analysis of the pandemic and used aggregated mortality data to analyze the mortality, social, and administrative changes in the city during and due to the pandemic [84, 95, 111, 112, 113]. The work here and in the following chapter does not pretend to replicate the important and impressive work already performed by previous scholars but instead aims to expand the analysis by looking more in detail at mortality by age and place, complementing and adding more evidence to what the pioneering work found.

Nevertheless, these analyses provide a primarily descriptive picture of the pandemic in Spain through the lens of period press reports and mid-century publications including on the evolution of sanitation and health in Spain [94, 95, 120, 273], though newly digitized data sources provide increased opportunities to quantify the impact of the pandemic on the Spanish population [64]. For instance, estimates of pandemic excess respiratory death rates have ranged from 6.1 per $10,000$ for the Canary Islands to 169.7 per $10,000$ for Burgos [64]. Moreover, spatial factors such as latitude, population density and the proportion of children have explained about 40% of between-province variation in cumulative excess death rates in Spain during 1918-1919 [64]. Few of these analyses take into account a recrudescent wave in Spain, which peaked in Madrid in late December 1919 and in later months in the rest of Spain [26], resulting in an additional 17,841 deaths specifically from influenza and primarily affected the young [94, 120].

Although much progress has been made in uncovering the age-specific mortality patterns of this pandemic in a number of populations in Latin America [65, 67, 68, 69] as well as in US and European settings [24, 82, 198, 233, 278, 284], more studies are needed to make sense of the heterogeneous mortality impact of this deadly pandemic across different populations around the world. For instance, characterizing and comparing the age-specific excess death rates across pandemic waves during 1918-20 in different populations could suggest alternative hypotheses on the drivers of pandemic mortality risk at the time and place more emphasis on lesser studied phenomena associated with the pandemic.

Despite previous efforts to characterize the impact of the 1918 influenza pandemic in Spain, prior studies have not systematically investigated differences in mortality impact between age groups and pandemic waves. This chapter takes a step forward and analyze

detailed series of deaths after retrieving over 70,000 individual death certificates representing all cause deaths during 1917-1920. Here, the analysis aims to assess the timing of pandemic waves and their magnitude in absolute and relative terms based on all-cause and respiratory excess death rates across across age groups and 3 pandemic waves in the city of Madrid during 1918-1920, including a recrudescent wave in winter 1919-1920.

4.2 Methods

4.2.1 Madrid mortality data

All death certificates from the Madrid Civil Registry, the same source used for some of the baseline estimation in chapter 3, between 1917 and 1920 were retrieved to construct time series of mortality during the 1918-20 influenza pandemic (Figure 4.1). Each record provides specific details of the deceased, including the date of death, age, and causes of death. For years 1917-1920, the registry holds a total of 70,061 death records (an average 17,650 deaths per year). Cause of death information for each death record allowed us to extract influenza and respiratory deaths.

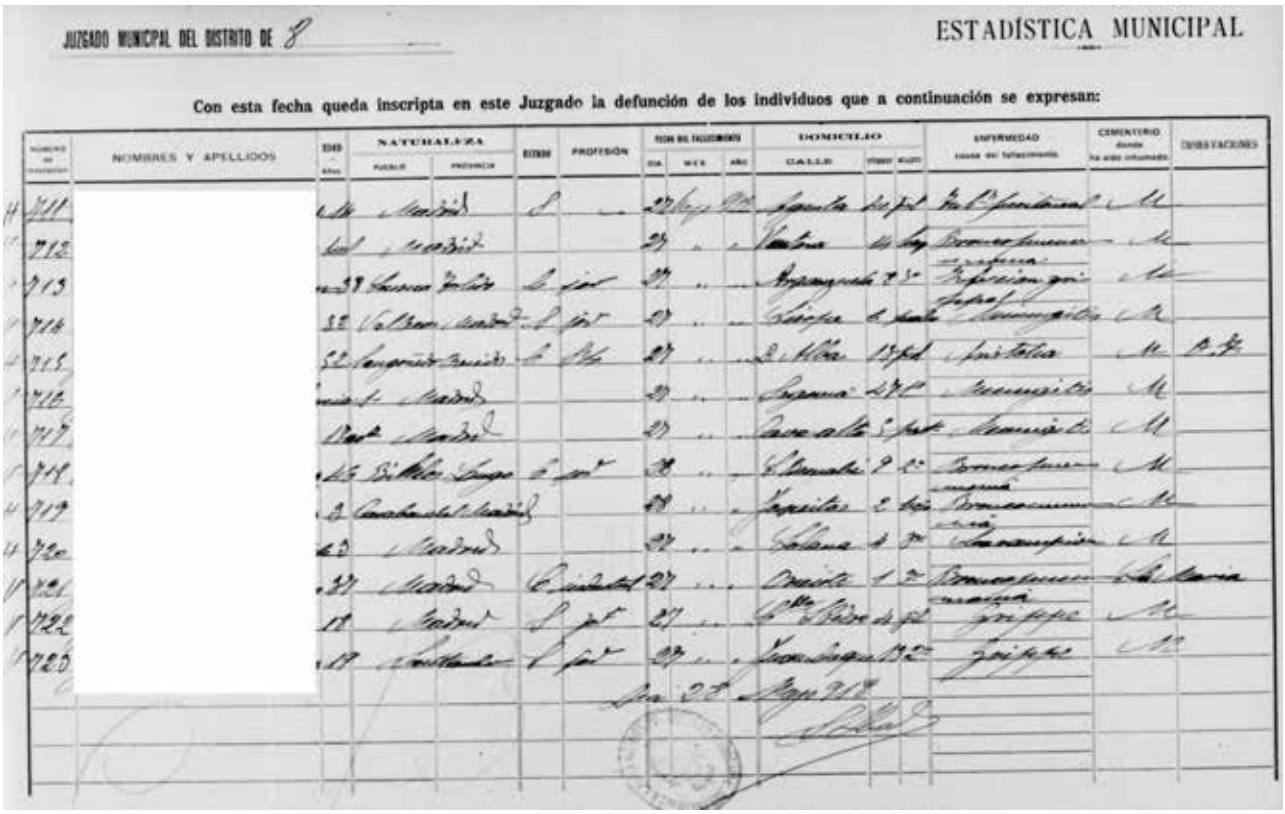

Fig. 4.1 Sample of death records from the 27-May-1918 from the Civil Register of Madrid. Source: Archivo Villa de Madrid

Causes of death

It is now well recognized that a significant fraction of the pandemic deaths resulted from secondary respiratory ailments (e.g., most commonly bacterial pneumonia) following influenza infection (rather from influenza infection alone) [181, 182]. Additional influenza-related deaths have been attributed to other types of bacterial infections and severe-acute respiratory distress, often evidenced by the appearance of blueish-gray skin shortly before death [181]. As such, estimates of respiratory related mortality also provide key information regarding the effects of influenza-specific mortality. As in prior studies (e.g., [68]), excess death rates are estimated for all-cause deaths and for pneumonia and influenza related mortality, a death category that comprises all death records containing influenza, pneumonia, bronchopneumonia, or bronchitis as a cause of death after removing death certificates that contained tuberculosis as a cause of death.

However, additional causes of death, seemingly unrelated to influenza mortality, have been found to rise and peak at dates exactly coinciding with those of influenza waves. These patterns are especially prevalent in Madrid, and they have been dissected and analyzed by

two cause of death coding systems. The information regarding the magnitude and timing of changes in causes of death during the time of the pandemic can be found in Appendix C.1.

Additional information

Further, information regarding the population composition of Madrid was obtained from the city's yearly population books to estimate death rates [214]. With this information, the results could be standardized according to the age structure of the population of Madrid at the time. Spain experienced one of the highest excess mortality rates during the 1918 influenza pandemic in Europe [26], although this country did not take part in World War I. Perhaps this pandemic outcome is associated with the fact that Spain was going through a demographic transition experiencing elevated mortality rates that were only comparable to those of Eastern Europe. Of note, the life expectancy in Spain was 41 years in 1910 and 40 in 1920 [92].

4.2.2 Estimating mortality baselines with quantified uncertainty

A more detailed explanation of estimating mortality baselines can be found in chapter 3, but the parametric bootstrapping technique applied to the data is briefly retold here. Initial estimations of a baseline using a simple cyclical Serfling linear regression model failed to capture a small but noticeable summer mortality peak [250]. To account for this variation, modifications to the initial Serfling parameters, as in another study of the 1957 influenza pandemic in Maricopa County, Arizona [74]. accounted for both time (α) and seasonal (β & γ) variations in normal flu activity. The new model identifies the oscillations (at time t) according to:

$$\begin{aligned}\frac{Deaths_{x_t}}{Population_{x_t}} = \quad & u + \alpha * (t) + \alpha * (\tfrac{100}{t})^2 + \\ & \beta * sin(2 * \tfrac{\pi}{52.17} * t) + \beta * sin(4 * \tfrac{\pi}{52.17} * t) + \beta * sin(8 * \tfrac{\pi}{52.17} * t) + \\ & \gamma * cos(2 * \tfrac{\pi}{52.17} * t) + \gamma * cos(4 * \tfrac{\pi}{52.17} * t) + \gamma * cos(8 * \tfrac{\pi}{52.17} * t)\end{aligned}$$

Parametric bootstrapping was used to further account for uncertainty in the 1917 baseline mortality level [96]. The complete five year baseline was calculated from the mean values of the coefficients from simulated models, and the upper baseline from the upper quartile value of the 95% Confidence Interval of coefficients. To calculate excess mortality, weeks with mortality above the upper baseline are classified as "pandemic weeks" [69, 74, 160]. As the difference between the fall and winter wave was difficult to discern, three distinct wave periods are defined as May-July 1918, August 1918- April 1919, and November 1919-February 1920. That is to say, while there is evidence to suggest the city of Madrid experienced one 1918 fall wave and a 1918-19 winter wave, these become unclear when disaggregating the data into smaller categories such as age groups. For this reason and to

facilitate comparisons with prior studies [64, 101], the successive fall and winter increases in excess mortality are calculated as if they were one pandemic wave.

Excess mortality for each wave is calculated by summing the total deaths above the baseline during the epidemic periods. To aid in the comparison of these results with other research, the provided relative estimates for each wave and age group allow relative comparisons across age groups [64, 251]. For each wave, the relative risk is defined as the ratio of total excess mortality observed to expected baseline mortality during pandemic weeks, when total mortality exceeded the 95% confidence interval of the baseline. This aids in the direct comparison of the total mortality impact of the flu between study groups, as baseline mortality varies substantially by age group [69].

4.3 Results: quantified age-specific excess

The analyses of weekly mortality rates from January 1917 to December 1921 revealed three distinct periods of pandemic-related mortality: a brief but well-defined spring wave (May-July 1918), an intense fall-winter wave during August 1918- April 1919 and a recrudescent winter wave during November 1919-February 1920 (Figures 4.2 & 4.3). Overall, peaks in respiratory and all-cause death rates were well synchronized. All- cause and respiratory related excess deaths for all age groups generally followed the same pattern of excess mortality by wave; the fall/winter wave has the highest excess rates, followed by the third recrudescent wave, then the herald wave in spring 1918. Additionally, the shape of the age-specific standardized mortality risk (SMR) remains the same, but the total elevated risk in all waves is much more pronounced when considering only respiratory mortality.[1] The cumulative estimates of excess mortality for these three pandemic waves were 86.8 per 10,000 from all-cause mortality and 44.6 per 10,000 from respiratory mortality, or approximately 6,500 total excess deaths, of which 3,300 were respiratory related.

[1] Given the recognized high peak of mortality among those aged 25-30, this group was also individually analyzed. The results are in C.2.

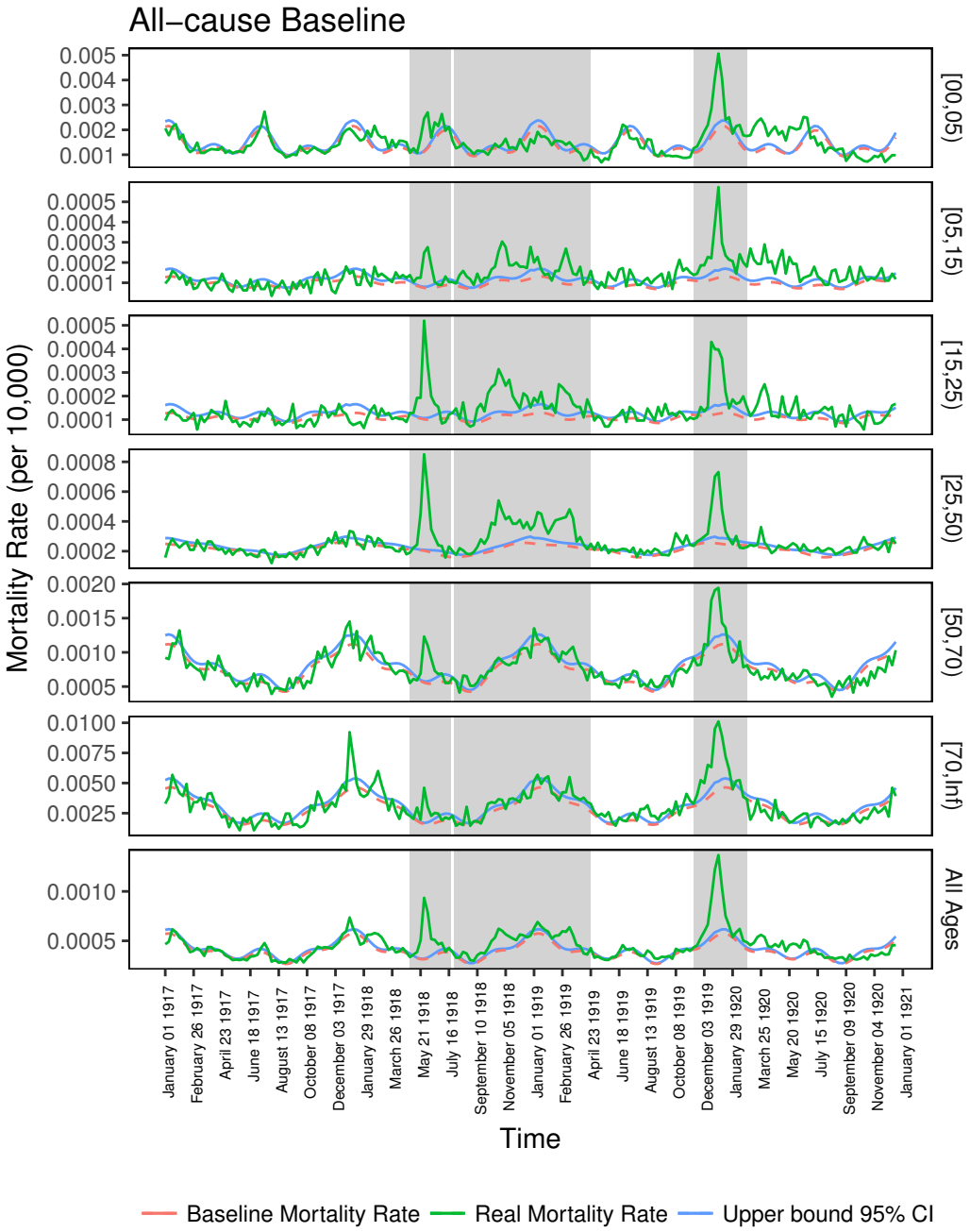

Fig. 4.2 Lines show the real weekly mortality rates from 1917-1922 (green), and mean (red) and upper 95% bound (blue) baseline from simulated 1917 deaths data. Shaded gray blocks represent the three epidemic wave periods.

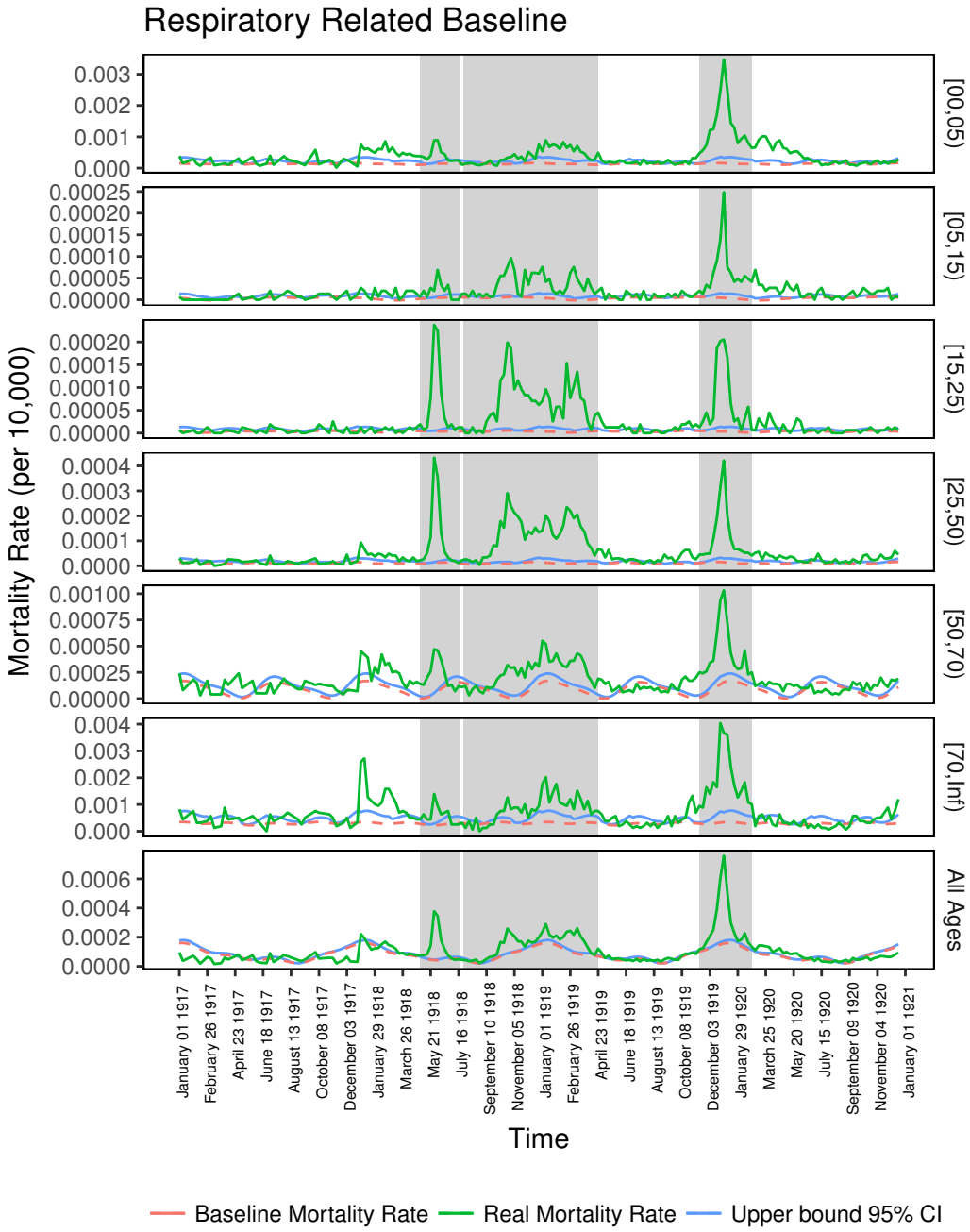

Fig. 4.3 Lines show the real weekly mortality rates from 1917-1922 (green), and mean (red) and upper 95% bound (blue) baseline from simulated 1917 deaths data. Shaded gray blocks represent the three epidemic wave periods.

Total excess mortality for epidemic weeks during the observed period is highest during the second fall-winter wave in 1918-19. A total excess rate of about 33.5 deaths per 10,000 is calculated based on all-cause deaths and 22.3 per 10,000 based on respiratory deaths. In

contrast, the spring-summer wave was associated with an excess death rate at 8.2 per 10,000 based on respiratory mortality and 19 per 10,000 for all-cause mortality. It is interesting that the third wave in winter of 1919-20 generated a substantial death rate at 34 deaths per 10,000 based on all-cause deaths, which is comparable to that of the intense fall-winter 1918-19 wave. However, it is worth noting that the first and third waves were relatively brief and pointy while some age groups exhibited two well-defined peaks during the protracted second wave in fall-winter 1918-19 (Figure 4.3).

Table 4.1 Age-Specific Excess Mortality by Wave

Age Group	Total Excess Deaths	Weekly Excess Mortality Rate (Per 10,000)	Standardized Mortality Risk
Spring Wave, 1918			
Overall	1456	19.42	1.57
$[00,05)$	375	57.57	1.40
$[05,15)$	95	6.58	2.03
$[15,25)$	165	10.59	2.28
$[25,50)$	486	18.13	1.95
$[50,70)$	213	21.32	1.55
$[70,Inf)$	127	80.39	1.68
Fall and Winter, 1918 -1919			
Overall	2511	33.50	1.27
$[00,05)$	293	44.90	1.22
$[05,15)$	364	25.11	1.82
$[15,25)$	401	25.73	1.83
$[25,50)$	1250	46.63	1.58
$[50,70)$	262	26.22	1.24
$[70,Inf)$	275	173.58	1.24
Winter Wave 1919-1920			
Overall	2538	33.86	1.52
$[00,05)$	823	126.34	1.59
$[05,15)$	261	17.98	2.18
$[15,25)$	235	15.08	2.14
$[25,50)$	467	17.41	1.63
$[50,70)$	344	34.45	1.41
$[70,Inf)$	485	306.51	1.63

In general, age-specific excess mortality rates were lowest during the Spring wave and highest during the protracted second wave, as shown in Figure 4.4. Compared to the first two pandemic waves, the youngest and oldest groups were particularly affected during the recrudescent wave in the winter of 1919-1920. In fact, during the third wave, those above 70 faced excess all-cause and respiratory mortality rates that were more than three times

higher than in the first wave. Further, during the last wave, infants and children aged up to 15 experienced more than double the all-cause and respiratory excess mortality rates estimated for the first two waves. The 5-15 and 15-24 year age groups maintained similar patterns in each of the waves, facing the lowest excess rates in the spring herald wave and highest in the combined fall and winter waves of 1918-1919. The 25-50 year old group faced their highest excess mortality rates in the second wave in fall-winter 1918-19.

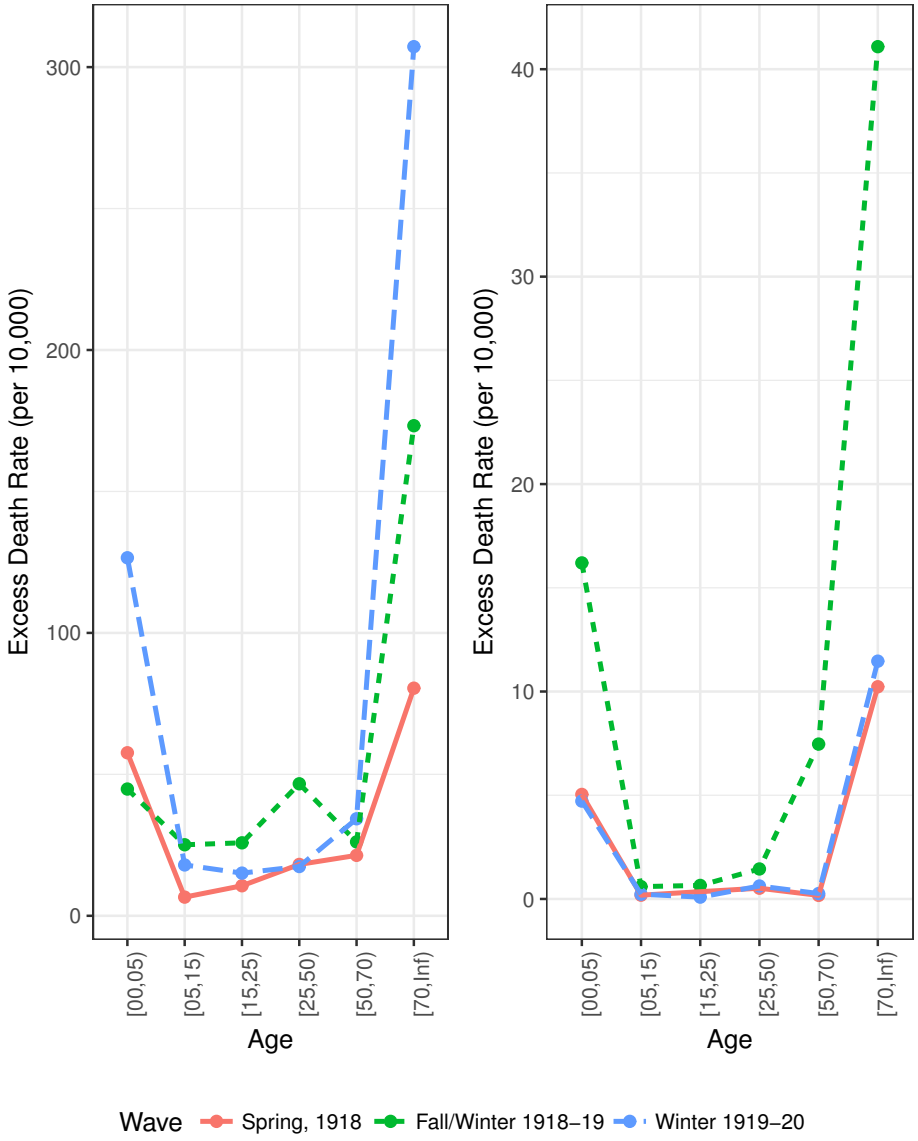

Fig. 4.4 Total excess mortality rates per 10,000 for all-cause (left) and respiratory related (right) deaths plotted by age group for each wave.

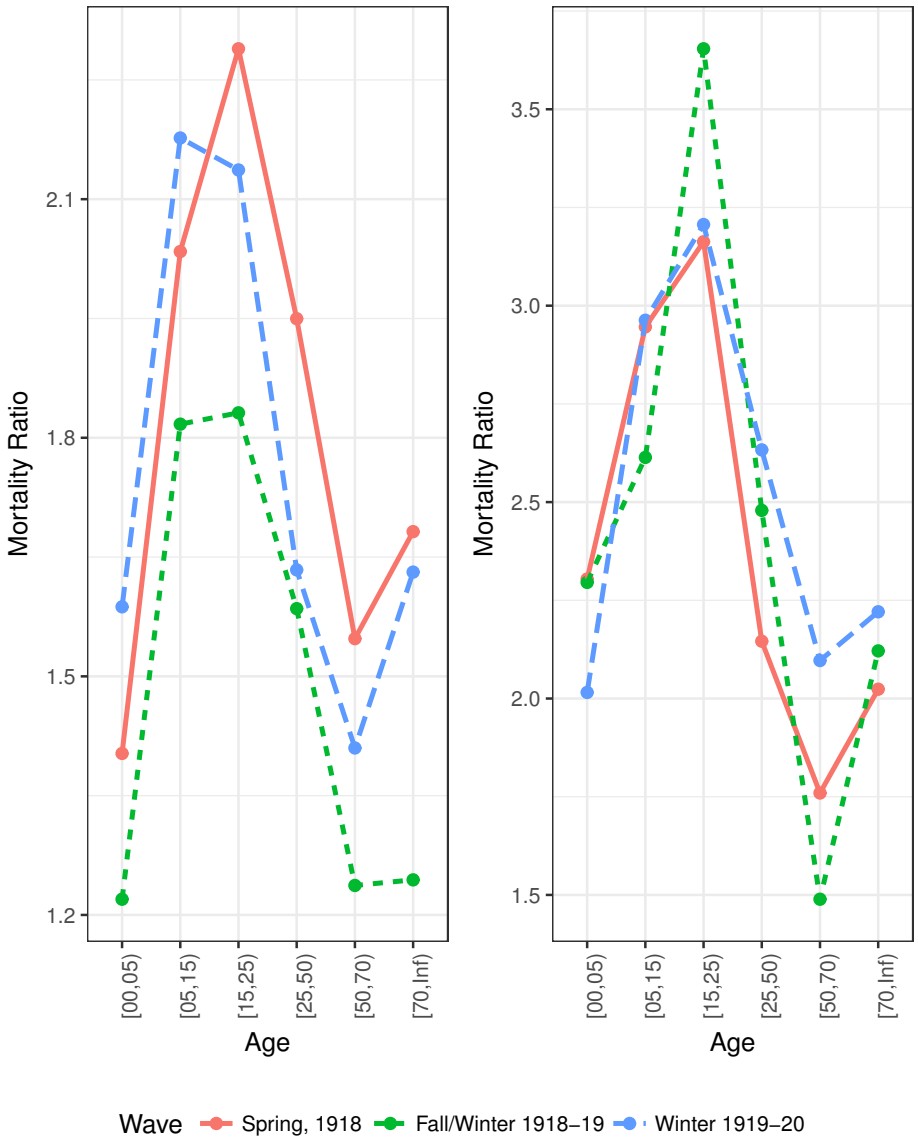

Fig. 4.5 Standardized Mortality Ratio for all-cause (left) and respiratory related (right) deaths plotted by age group for each wave.

While the herald spring wave accounts for slightly more than 20% of all total excess deaths, the SMR during this period was higher than in the succeeding waves, due to lower baseline mortality during spring and summer (see figure 4.5). Remarkably, although individuals 15-25 yrs. exhibited low excess mortality rates relative to other age groups, this age group exhibited the highest SMR across all pandemic waves. Generally, the age-specific pattern of the SMR is that of an upside-down 'v,' with the exception of the oldest age group. During the first and third waves, those above 70 yrs. experienced a higher SMR than individuals 50-70 yrs. Most generally, the highest SMRs occur in the first and last waves, though the highest

calculated SMRs for respiratory-related mortality (4.4 and 4.2) are found for individuals 15-25 yrs. in the first and second waves.

4.4 Discussion: implications of Madrid excess mortality in the context of other research

While estimates of excess mortality reveal variability in age-specific patterns throughout the world, the results are unique in that the highest absolute excess rates occurred among older populations (*age*70) compared to previous reports from Europe and the US [24, 198, 256]. Specifically, the Madrid age-specific excess dominant pattern resembles that of seasonal influenza epidemics where the highest excess rates occur in the youngest and oldest groups [68, 69, 162]. However, much of the elderly population of Madrid would have been exposed to other viruses; for example, in the decades preceding the Spanish flu, the "Russian" Influenza pandemic that struck Madrid in the winter of 1889-90 produced overall all-cause excess mortality rates of 58.3 per 10,000 and produced an age-specific excess mortality pattern that is similar to each of the three pandemic waves in Madrid [218].

The results also confirm earlier analyses of a particularly lethal spring wave in Madrid relative to smaller mortality peaks, but high incidence rates in some locations such as Norway and Denmark [24, 198, 255]. In Madrid, weekly excess death rates during the spring wave nearly rivaled that of the protracted fall/winter 1918-1919 wave.

This analysis allows the timing and excess mortality of this first wave in Madrid to be contextualized relative to herald pandemic waves in North America and civilian outbreaks in Europe. Many of the first purported spring outbreaks occurred in U.S. military camps, spreading to larger cities in April and May, before the herald wave in Madrid [255]. However, the mid-late May outbreak was the first reported in civilian Europe. In the following months, reported influenza outbreaks in Europe occurred east- and north-ward to other parts of Spain and Italy, then England, Sweden and Norway, and Switzerland and Poland [22, 24, 64, 82, 133, 181, 255, 284]. However, it remains difficult to distinguish to what extent the virus spread through military rather than general population movement [255].

Analyses of hospitalization, mortality, and other surveillance sources in both military and civilian settings have found evidence of cross-protection between spring and fall influenza outbreaks during waves of the 1918-1919 epidemic [37, 89, 169, 237]. The high pandemic mortality found in the analysis, together with evidence of high incidence rates during spring-summer waves [85, 94, 95] could have provided some immunity and cross-protection to the strain of virus in the succeeding fall wave. Conversely, in New York City, a noticeable

age-shift in influenza mortality patterns occurred in early 1918, perhaps suggesting the presence of the new virus strain, yet there is little total excess mortality until the strong fall wave, which killed more than 9 times as many people [198]. This may partially explain the slower growth and protracted wave in Madrid that begins in September 1918 and continues through the winter and early spring of 1919.

Following the enduring second wave, the results provide clear evidence of a powerful recrudescent wave that peaked at the very end of 1919 and appears in locations throughout the world in the spring of 1920 [26, 65, 67, 68, 82, 133, 141, 198, 230, 256]. In Madrid, the all cause excess rates of this echo wave are on par with that of the elongated second wave, and all-cause and respiratory excess mortality rates higher than in the Spring 1918 wave. In other countries and cities where this wave has been documented, a slight shift in the age-specific mortality often occurs with a return to high excess mortality among those above 65 [67, 68, 82, 198, 256]. As in the case of these estimates, the mortality of young adults often drops slightly but remains persistently high and well above the pre-pandemic level.

In line with previous reviews and mortality estimates focusing on the impact within Spain, during this fourth wave, infants and young children suffered particularly high rates of mortality [94, 95, 120]. Because high rates of excess mortality exist in all age groups, lack of acquired immunity from earlier waves may only explain the excess mortality among infants and young children. However, this theory can also be debated. In the United States, for example, births during the outbreak in fall 1918 dropped considerably, but returned to normal levels several months after the outbreak before again dropping 9-10 months after [54, 83]. Those children born in the 5-6 months following an outbreak would have been in utero, and therefore possibly exposed to the virus via their mother, during the earlier epidemic wave. Antigenic shift or mutation in the virus could also describe why the mortality remains elevated across all ages, but currently, it remains difficult to ascertain exactly why these mortality patterns changed during this strong wave.

Though not specifically analyzed here, one more all-cause and respiratory-related mortality peak occurs in late December 1921, present in all age groups but predominately in those older than 50 and younger than 5 [61]. Recrudescent waves can still occur years after the initial and main pandemic waves, echoing the initial impact of an outbreak, such as in the 2011 A/H1N1 epidemic recurrence in Mexico following the 2009 A/H1N1 pandemic [62]. Their presence and impact should continue to be studied and quantified, as they may significantly change the overall mortality impact of the influenza pandemic.

These estimates of pandemic mortality impact in Madrid can be compared to those derived from a previous study that analyzed excess monthly all-cause and respiratory mortality in all provinces of Spain during the herald spring wave and second fall/winter wave [64]. Here, the

analysis finds higher overall excess rates in the Spring wave (19.4 v. 11.7 per 10,000) but lower in the second wave (33.5 v. 55 per 10,000). This study also calculates lower excess respiratory mortality in both the herald and protracted second waves. These differences may be attributed to the fact that both studies employed different sources of mortality data and the previous study analyzed pandemic impact in the entire province of Madrid while the current study focuses on the capital city alone. Moreover, the spring wave may have largely affected the city itself (which was analyzed) and the surrounding province (the subject of the prior study) to a lesser degree, resulting in the differences in excess mortality estimates. This could extend to the second wave; perhaps, those living in the city gained some immunity from exposure to the first wave, while those without this exposure had no cross-protection. Future work could aim to disentangle additional factors driving these differences.

Considering the pandemic events collectively known as the Spanish influenza, the case of Madrid provides additional insights into how, in a large urban environment, individual waves and their progression contributed to the overall mortality impact on the city. While other analyses look at herald waves and question the impacts of acquired immunity from spring to fall [24, 198], the force of the spring wave in Madrid, relative to the successive fall and winter outbreaks, appears to indicate some type of protective effect, perhaps due to a small amount of antigenic shift in the virus between the two waves. Currently, only strains from the spring and fall waves of 1918 have been studied, meaning that the extent earlier and later strains differed cannot be confirmed [134, 253, 258]. Yet, continued analyses of successive waves using new data sources and innovative approaches should be undertaken to better understand acquired immunity and the protection it may provide against successive outbreaks. Using contemporary and historic demographic mortality and surveillance data of recent and past epidemics, further insights into the ways early outbreaks affect the immunity and transmission can affect the way public health officials respond to contain future outbreaks.

Table 4.2 Age-Specific Excess Respiratory Mortality by Wave

Age Group	Total Excess Deaths	Weekly Excess Mortality Rate (Per 10,000)	Standardized Mortality Risk
Spring Wave, 1918			
Overall	613	8.17	2.59
$[00,05)$	253	38.81	2.62
$[05,15)$	19	1.31	3.11
$[15,25)$	49	3.13	4.43
$[25,50)$	114	4.25	2.89
$[50,70)$	100	9.98	2.27
$[70,Inf)$	77	48.72	3.15
Fall and Winter, 1918 -1919			
Overall	1670	22.28	1.82
$[00,05)$	308	47.25	1.69
$[05,15)$	82	5.64	2.27
$[15,25)$	185	11.87	4.20
$[25,50)$	524	19.56	2.77
$[50,70)$	346	34.66	1.65
$[70,Inf)$	250	157.79	1.88
Winter Wave 1919-1920			
Overall	1061	14.15	1.86
$[00,05)$	397	61.01	2.04
$[05,15)$	58	3.97	2.35
$[15,25)$	83	5.33	3.56
$[25,50)$	180	6.71	2.18
$[50,70)$	168	16.85	1.67
$[70,Inf)$	193	121.84	1.89

Chapter 5

Neighborhood variation in excess mortality across three waves of influenza in Madrid

Abstract

Responsible for 50-100 million deaths worldwide, the influenza pandemic events of 1918-1920 continue to be studied. While recent studies examine the contribution of socioeconomic status to influenza mortality, this analysis looks at three consecutive epidemic waves and adds a geographic element to look at the intersection of structural and social characteristics within neighborhoods. Two sources of data (the Madrid padrón from 1915 and death records from 1917-1922) are employed to help disentangle the roles neighborhood socioeconomic and structural factors played in mortality during three epidemic waves of Spanish flu from May 1918 to February of 1919. Focusing on 91 neighborhoods in the city, this chapter explores the role that the composition of each neighborhood (such as literacy level, average rental prices, career make-up of inhabitants, density measures, etc.) may have played in the variation of excess mortality (aggregate and wave-specific). Findings show significant differences in model results by wave, underlining the potential importance of acquired immunity within a population and the effect of greater medical resources in sustained epidemic waves.

Resumen

El capítulo cinco se centra en el análisis de la Madrid dentro de la propia ciudad de Madrid. Aunque estudios recientes examinan la contribución de las características socioeconómicas a la mortalidad por gripe, este trabajo analiza tres olas epidémicas consecutivas y agrega un elemento geográfico para observar la intersección de las características estructurales y sociales dentro de los barrios. Se emplean dos fuentes de datos (el padrón de habitantes de Madrid de 1915 y los registros de defunciones de 1917-1922) para ayudar a descifrar el papel que los factores socioeconómicos y estructurales del barrio desempeñaron en la mortalidad durante las tres olas epidémicas de la gripe española acontecidas entre mayo de 1918 y febrero de 1919. Centrándose en 91 barrios de la ciudad, este capítulo explora el papel que la composición de cada barrio (es decir el nivel de alfabetización, los precios medios de alquiler, la composición profesional de los habitantes, las medidas de densidad, etc.) puede haber desempeñado en la variación del exceso de mortalidad (agregada y específica de las olas). Los resultados muestran diferencias significativas en los resultados de los modelos

por ola, lo que subraya la importancia potencial de la inmunidad adquirida dentro de una población y el efecto de mayores recursos médicos en olas epidémicas sostenidas.

5.1 Background: Social variation in excess mortality during influenza epidemics

Though it occurred more than 100 years ago, researchers continue to study the strong mortality impact of the Spanish Influenza pandemic, particularly the variation in its manifestation and severity throughout the world. Though studied extensively, questions remain about the ways in which spatial, temporal, and social differences affect influenza mortality, particularly during pandemics.

While climatic differences between tropical and northern countries play a role in seasonal influenza activity [265], the high world-wide mortality of the 1918 Spanish flu pandemic [141] indicates these climactic patterns did not play a major role in regional death outcome. Typically characterized into three waves (Spring, Fall, and Winter), the extent to which each wave existed and its severity differed by location, creating debate regarding transmission mechanisms and the role acquired immunity in consecutive breakouts may have played in the tempering of each successive wave. While a virus will spread more slowly in a population with some immunity (i.e. the reproduction number $R_{effective}$ in a partially-immune population will be lower than R_0), the mortality of those exposed to the flu may differ according to their exposure and environmental surroundings [168]. Furthermore, mutation and evolution of the virus may further play a role in the extent to which a person, once immune to an earlier strain, is again susceptible to a circulating virus [168, 255].

Some prior studies focusing on entire cities (i.e. New York, Copenhagen, Mexico City) examines transmission mechanisms and the role acquired immunity in consecutive breakouts may have played in the tempering of each successive wave of Spanish Influenza [24, 68, 134, 181, 198]. These mechanisms may also have been different between economic classes [163]. However, without biological evidence, this link is difficult to prove. Other research clearly points toward a social gradient in the transmission and strength of seasonal influenza outbreaks and preventative vaccination campaigns in the US and across the world, especially in elderly and minority populations [77, 187]. Yet the extent that these patterns remain the same or differ during pandemic events is debated. For example, many of the earlier analyses of the 1918 influenza pandemic led researchers to postulate that the viral strain present in the 1918-1920 pandemic events was so virulent that aside from affecting all age groups, the airborne nature of the disease outweighed the potential of any other social variables to create class mortality differentials [78, 271, 277]. Several examples of these studies, such as in a Great Britain Ministry of Health survey about fatality and social status in 1920, can be found in a 2006 review [160, 271].

In the age of technology, historical data sources can be digitized so that researchers can analyze mortality dynamics in past pandemics to be better prepared for potential outbreaks in the future One study using individual level data did find a class gradient to a fall wave of influenza in Sweden, finding that farmers were least affected and low skilled laborers were most susceptible to death in 1918 [38]. However, these results were less clear among females and did not offer a strict gradient. Other recent research explores the correlation between district level demographic and social levels and flu mortality during the 1918 pandemic, yet these analyses generally focus on single waves. For example, a census tract analysis found that during the strong fall wave in Chicago (the end of September to November 1918), influenza and related mortality was higher in places with greater illiteracy [123]. Neighborhood of residence and apartment size, as a proxy for household wealth, were found to have had effects on influenza survival in Kristiana during 1918 [160]. Even outside the realm of the 1918 pandemic, Reproduction Number has been found to be correlated with the average number of people in a residence [275].

These small-area studies within cities often focus on one or two waves of pandemic outbreak, but this can ignore the changing transmission mechanisms and role acquired immunity plays in consecutive breakouts [134, 181, 198]. Few papers combine the strength and timing of each wave with the impact of the socioeconomic status, as this requires information about the individual's or group's income and/or occupation across more than one wave in a location. Moreover, the recorded presence of herald waves and their measured strength is limited to some large cities, and social data can be difficult to discern in an effected place. A study of influenza prevalence in Bergen, Norway suggests a change in the effect of social status (measured as apartment size), where those with smaller apartments experienced higher morbidity in a herald wave and less in the fall wave [163]. The role compositional factors played in mortality during each of the three influenza waves in Madrid may have changed according to changing immunity in the population and virus makeup.

5.1.1 Influenza in Madrid

The progression and mortality impact of the Spanish influenza in Madrid has been detailed in previous chapters. Four waves hit the city between May 1918 and February 1920, causing thousands of excess deaths in the population. Thus far in this thesis, the age-specific mortality pattern for each wave has been detailed, as well as the role consecutive waves may have played in tempering mortality in later waves. For example, the strength of the herald wave is postulated to have minimized the mortality of the succeeding fall wave, which, though protracted with three distinct peaks in October, December, and February, was relatively mild compared to the outbreaks during this time in Europe and North America [95, 218]. The less

studied fourth wave of the pandemic (sometimes referred to an an “echo" wave) produced the largest mortality peak of the period in Madrid from December 1919 to January 1920 [94].

During these outbreaks, age-specific excess mortality also follows a different pattern; with the exception of the fall wave (which experienced a small mortality bump), young adult absolute excess remains low. Nonetheless, due to low baseline rates among these adults, the highest *relative* excess rates are observed in those between 5 and 70. Prior research has likewise found that those districts with the highest mortality rates in the years before the influenza outbreaks similarly experienced the highest levels of excess mortality during epidemic [112, 113]. Was it strictly the age-structure in these areas that contributed to heightened excess? What other factors may have played a role in mortality risk?

Madrid, in the middle of a large period of growth during a stage of urbanization in Spain, also had a large degree of social and economic district variation (for a summary, see [23]). The city was officially organized into 100 neighborhoods, contained in 10 districts, in various stages of development and urbanization, with high levels of inequality and distinct population compositions. All this may have played a role in the spread of influenza and spatial mortality differences [160, 186]. Given the role of environment, exposure, and social factors associated with influenza mortality, as well as the wave and age specific patterns of influenza in Madrid, this chapter seeks to understand and disentangle these relationships at the neighborhood level. Here, mortality during the outbreaks is examined to determine how the type of buildings, apartments, and population in each neighborhood related to varied mortality levels, and if these factors changed by wave.

5.2 Data and methods

5.2.1 Madrid civil register of deaths

The “domicilio" (address) variable from the Civil Register records on deaths in Madrid is used to match each record to the city geography and determine the excess mortality by district and neighborhood (“barrio"). In addition to address, each record contains a death date, demographic and some contain limited socio-economic information, such as occupation. This unique, individual-level data allows for analyses to demonstrate the flow and characteristics of each wave across districts, as well as the changes in death by age group and specific cause within and between each location.

During the influenza epidemic, it is likely many people died at another location than where they were living. Unfortunately, many death certificates are coded with the address of death rather than the deceased’s address. Thus, it is also likely that due to the epidemic, many

neighborhoods with large hospitals or first aid centers experienced inflated excess mortality, due to higher numbers of deaths in these medical centers from those actually living outside the neighborhood. In order to eliminate this confounding effect in the results, this analysis has been conducted with two sets of data; one included all deaths, as coded on the death certificate, and the other included only deaths that are believed to have occurred outside of a medical institution. To create the non-first aid subset, address information on all institutions in Madrid, not limited to hospitals, first aid clinics, other medical centers was collected. From this list of 324 locations, additional descriptive data provided information that led to the decision that 140 of these institutions that may have experienced a disproportionate increase in mortality during the epidemic waves, leading to falsely inflated excess mortality in the neighborhood. It was important to understand which institutions may have led to artificially increased neighborhood mortality, as this may have confounded the results concerning the relationship between neighborhood mortality and the structure and population that lived there. All deaths whose addresses matched one of these 140 institutions were excluded from the analysis.

Initially, the data consisted of 103,323 typed death records between 1917-1922, but when ascribing a neighborhood to each record, several additional problems became apparent, both due to missing values on the records and current data limitations. Some death records do not have an address written from which to determine a neighborhood, and a few have no death date written. Using the official 1900 cadastral map of Madrid [126] to code addresses into neighborhoods created additional problems due to the large growth of the city between the year of the digitized cadastre, 1900, and the period of observation. A number of death records have a written address also not listed in an additional 1915 database. This also hindered the ability to include several deaths on the outskirts in all years of the analysis.

In total, due to the limitations of the data (roughly equally distributed), about 2500 deaths are not coded to a neighborhood each year. Other issues with some missing data reveal inaccuracies in some neighborhoods during the time period of study. Thus, the following analysis uses 91 of the 100 total neighborhoods in Madrid at the time, which are equally distributed throughout the geographic landscape of the city and maintain a high accuracy of death registration throughout the baseline period of 1917 and during each outbreak. However, the accuracy level in coding deaths to a district is quite high, due to the structure of the registration system at the time. Before the results of the primary neighborhood analysis are presented below, a series of analyses at the district level were preformed in order to provide some larger-scale background to the neighborhood results. These are available in appendix D.

As mentioned, in order to verify the role hospital deaths may have inflated excess mortality in some neighborhoods, the analysis is limited to only deaths for which the address is *not* at one of 140 identified medical centers. From these records, neighborhood absolute and relative excess mortality rates are calculated for all epidemic periods combined, as well as individual waves from May to August 1918, September 1918 to April 1919, and November 1919 to February 1920. The level of excess mortality is equal to the difference between observed deaths during the periods of each wave in 1917 and the number of deaths during each waves. Following the elimination of these medical center deaths and those in neighborhoods with incompletely coded deaths during the baseline or outbreak period, 86,704 deaths between 1917 and 1922 remain from which the following analysis is completed. [1] Thus, it is likely that in general, the estimates under-report excess mortality in neighborhoods, because deaths removed from the analysis due to there occurrence at a hospital are not included as a death in another neighborhood, likely that of where the deceased individual lived. The hospital deaths that were removed from the analysis were not reallocated to other neighborhoods (that is, put back into the analysis via a probabilistic distribution), as there was difficulty in determining exactly how these deaths could be divided into neighborhoods to an acceptable degree of precision. Thus, the decision was made to use the estimates of calculated excess mortality without hospital death reallocation under the assumption that the results would overall be more reliable, if conservative.

5.2.2 1915 Padrón

The 1915 padrón of Madrid provides information about the constructed composition of the city, such as the number and type of buildings, rental prices, and measures of density. The padrón also provides limited data about the social composition of the districts and neighborhoods, including the proportion of literate population and tabulations of occupations. While the level of detail in the padrón provides a fascinating glimpse into the city composition during this time, certain elements must be combined and altered in order to create variables that bring meaning to the analysis. Thus, for the neighborhood-level analysis, several changes were implemented.

The padrón provides neighborhood-level information on building use by total number of units, including types of commercial space and size and building position of apartments in the city. Generally speaking, houses located in the basement and the highest floors were smaller

[1] As visible in the Figure 5.1, which shows raw weekly mortality rates from 1917-1922 by neighborhood in the city, some neighborhoods with hospitals continue to experience extreme levels of mortality even in the absence of these deaths, such as the neighborhood of Doctor Forquet, slightly south of the city center.[2] This particular neighborhood is also home to an asylum and several orphanages, which had underlying rates of high mortality.

and of lower quality, while those apartments on the first and second floors were occupied by those who could afford higher rental prices. By aggregating housing generally occupied by those with the lowest income and social position as a percentage of total housing units in the neighborhood, a single variable contains information on the amount of housing generally occupied by lower-income renters. In conjunction with this supplied housing data, a single variable calculating the weighted average of all neighborhood housing unit prices from a 14-category table with counts of apartment rental prices in each neighborhood creates a single variable to describe rental prices from the raw data. In addition, all types of commercial units are grouped into a “commercial" variable that denotes the total amount of private businesses in the neighborhood as a percentage of all types of building units.

Further manipulation the raw padrón data led to the creation of variables describing the social interaction and stratification in each neighborhood with a HISCAM score [150]. After matching a listing of occupations in the neighborhoods to HISCO codes, they were then linked to HISCAM codes using both the male-only universal scale U2 (using only the listed male occupations) and the later period scale of both men and women L (using all occupations in the neighborhood). After creating an average HISCAM score for each neighborhood through a weighted average of the count of each code in the neighborhood, the final neighborhood-level variable equals the average of the aggregate HISCAM scores. Two HISCAM coding schemes were used for several reasons, not limited to: a) the HISCAM numbers were used from data not found in southern European countries, b) while most married women did not work, many single females were employed, c) while the results showed that the two schemes were highly correlated, there were some differences in neighborhoods between the two schemes’ scores [150].

Because age-specific mortality follows a specific pattern in the city of Madrid [218], it is necessary to control for differences in population structure within neighborhoods. Thus, the proportion of the population in each neighborhood that falls within the age range of 5 to 75 is also calculated from a larger group of age-specific numbers. The broader 5 to 75 group was chosen as during each of the waves, this group had the lowest rates of absolute and highest rates of relative excess mortality.

Figure 5.2 shows the total and wave-specific standardized mortality ratio for each of the 91 included neighborhoods in the analysis, while summary statistics of other included variables are found in table 5.1. Further descriptive visuals in the Appendix D highlight the relationship between the explanatory variables and excess mortality (absolute and relative).

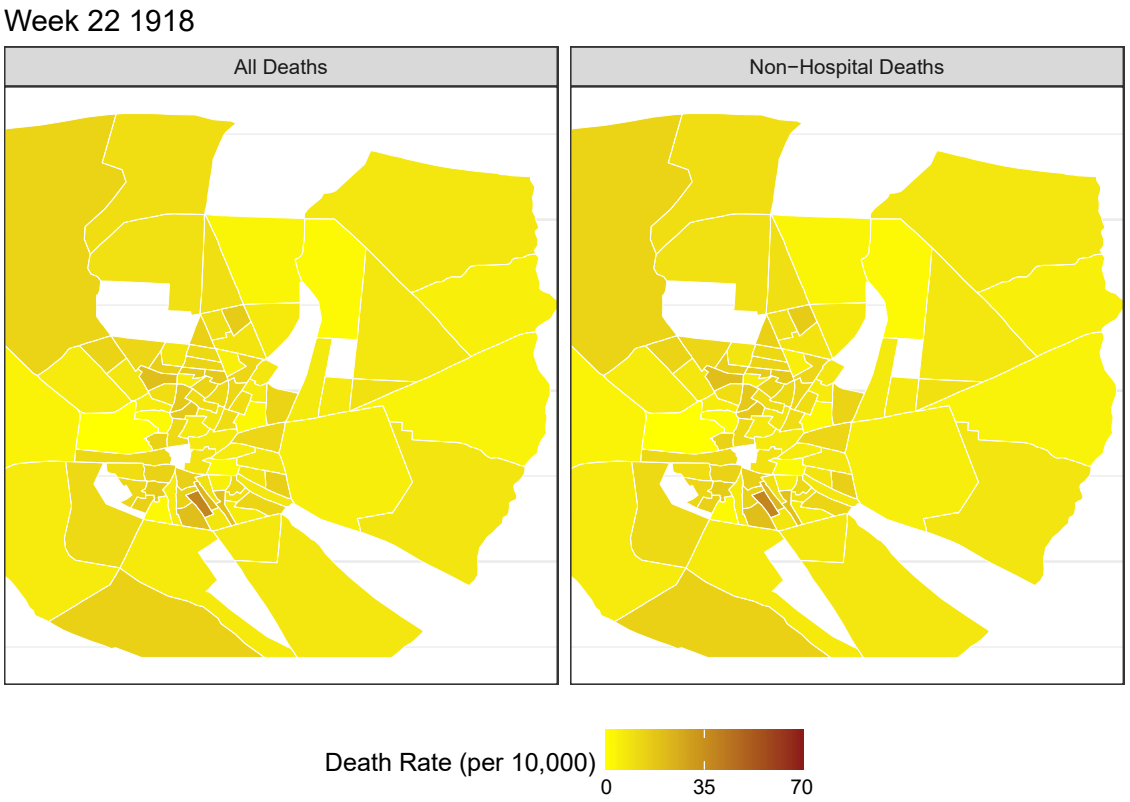

Fig. 5.1 Weekly mortality rates with (right) and without (left) hospital and medical center deaths during a peak-outbreak week during the Herald wave in 1918.

Fig. 5.2 Standardized Mortality Ratio by neighborhood, all-cause deaths

Table 5.1 Selected Descriptive Statistics

	Mean	Std. Dev.	Min	Max
Total Excess	0.30	0.26	-0.20	1.12
Spring Excess	0.73	0.67	-0.40	3.00
Fall Excess	0.14	0.24	-0.34	0.93
Winter Excess	0.52	0.43	-0.16	2.37
Literacy	0.82	0.07	0.65	0.95
HISCAM	51.68	3.67	44.68	58.32
Poor Housing	0.18	0.14	0.06	0.70
People per Building	49.84	17.26	11.36	103.64
Average Rent	52.98	37.97	9.68	206.13
Commercial	0.10	0.04	0.02	0.21
Population 5 to 75	0.91	0.02	0.86	0.95

Checking for spatial autocorrelation within the city

The administrative neighborhoods of the city are clearly defined, but the extent to which each neighborhood can be considered an independent observation must be determined. That is to say, the proximity of neighborhoods to each other may play a role in the variables observed in each neighborhood, leading to spatial autocorrelation in the data and model issues [97, 292]. It should also be noted that many of the explanatory variables used in the the neighborhood analysis *are* highly spatially correlated. Thus, after computing the standard mortality ratio for total and each wave in all neighborhoods, Moran's I for the absolute and relative mortality in each neighborhood was computed, finding no evidence of spatial autocorrelation. Further analysis of the residuals and other measures found also found a lack of spatial dependence.

5.2.3 Modeling relationship between neighborhood composition and excess mortality

While the strength of an outbreak my be quantified in terms of relative and absolute excess mortality rates, modeling the impact of factors on disease-related mortality requires a slightly modified technique. The focus of this analysis is the examination of the probability of experiencing the observed number of cases in a neighborhood based on the expected (baseline) amount of cases. More than absolute and relative excess mortality values, this takes into account the underlying mortality distribution of disease and population size [70, 121]. For example, variations in population size across neighborhoods mean that a single death in a lower populated neighborhood raises the mortality rate more than a single extra death in a larger neighborhood. In this case, the number of observed cases (dependent variable) is offset by the log of the total number of expected deaths in a count-data regression model

[122]. The coefficients can be roughly interpreted as the factor of change to the incidence rate of the disease.

An assumption of Poisson models is that the mean and variance within the distribution are equal to each other. Following tests, this condition of the Poisson distribution is clearly violated in the data, as it is very over-dispersed. Ignoring this violation will result in small standard errors and overestimated significance of the model. A negative binomial regression model can account for this violation, producing more precise results due to the inclusion of a theta term θ that accounts for the unobserved heterogeneity in the data and over-dispersion [45]. More specifically, the negative binomial and Poisson regressions differ in that the mean and variance estimation in a Poisson regression are:

$$E[Y_i|\mathbf{x}_i] = \lambda_i \text{ and } Var[Y_i|\mathbf{x}_i] = \lambda_i$$

However, the negative binomial model allows for over dispersion in the data by:

$$E[Y_i|\mathbf{x}_i, \varepsilon_i] = e^{(\alpha+\beta\mathbf{x}_i'+\varepsilon_i)} = h_i\lambda_i$$

The h_i parameter has an assumed gamma distribution with a mean equal to 1 and a variance equal to $\frac{1}{\theta}$ [124]. Estimation of the parameters is completed using the MASS package of R, which fits the additional h_i parameter through maximum likelihood estimation and provides the θ value of h_i's distribution [282]. As they are nested models, the comparison of the Poisson and negative binomial models is done through a likelihood ratio test. Poisson and negative binomial regressions were compared for all models.

5.3 Results: excess mortality variation by neighborhood composition

Results of negative-binomial regressions for total and wave-specific mortality with all variables are presented in table 5.2. The coefficients are presented as exponents for ease in interpretation; see figure 5.3 for a plot of the standardized coefficients of the presented models. In all models, the data fit the negative-binomial distribution with parameter θ significantly better than the Poisson distribution. The models presented in table 5.2 each contain all of the descriptive variables outlined above; addition and deletion of variables in other models does not change the significance of any of the others.

Considering total mortality across all epidemic periods, a single variable can be said to have a strong relationship with excess mortality: the amount of “lower status" housing. As the percentage of small basement and upper level/attic apartments in a neighborhood

increases, the rate of excess mortality across all waves also increases. This relationship holds *and* strengthens in the fall wave, where the magnitude of this variable doubles from overall mortality. Not including mortality from previous waves, this housing variable has a weak relationship in the winter wave.

In each model for the fall and winter wave, the amount of excess mortality in the immediately preceding wave is highly correlated with the level of excess experienced in the same neighborhood during the studied wave. In the fall, the coefficient (exponentiated, almost 1) is nearly proportional to the amount of spring mortality, meaning the level of excess is similar from one way to the next. However, in the winter wave, the expected excess mortality is more than doubled from the fall, likely demonstrative of the magnitude of the winter wave. This indicates that despite the other variables in the model, neighborhoods with high mortality in one wave generally also had high levels of mortality (relative to other neighborhoods) in the other waves.

While the model focusing only on the herald wave mortality finds no relationship between any of the variables and excess mortality, some significant relationships do exist in the subsequent waves, most notably in the fall. In addition to the quality of housing, the proportional amount of those older than 5 and less than 75, as well as the amount of commercial property in the neighborhood increase the amount of overall wave mortality in the fall. This relationship strengthens with the addition of the amount of excess mortality experienced in the spring herald wave. Conversely, the significance of the proportion of “poor" housing and literacy during the initial winter wave model is tempered by the addition of both spring and fall excess. In the winter model which only includes excess mortality for the immediately preceding fall wave, the positive relationship between the percentage businesses and excess increases in strength (but with smaller magnitude) from the initial model.

However, there are some additional inferences that can be made from the results of model fit to encourage continued work on the subject. For example, while no model appears to aptly describe the relationship between the descriptive variables and the excess mortality, one can hypothesize that characteristics of the neighborhood population played different roles in the individual waves. Both the first and last model, analyzing the SMR of the total and last winter 1919-1920 wave, fits best when looking contextually at the built environment and total population. Both find that the lower density of built area per person in neighborhoods plays a large role in higher influenza-related relative mortality. While this variable remains slightly significant in models for the protracted fall-winter 1918-1919 waves (as well as the total population size), the model fits more poorly. In both this and the spring wave, results find that demographic indicators of population size and age composition better fit a model

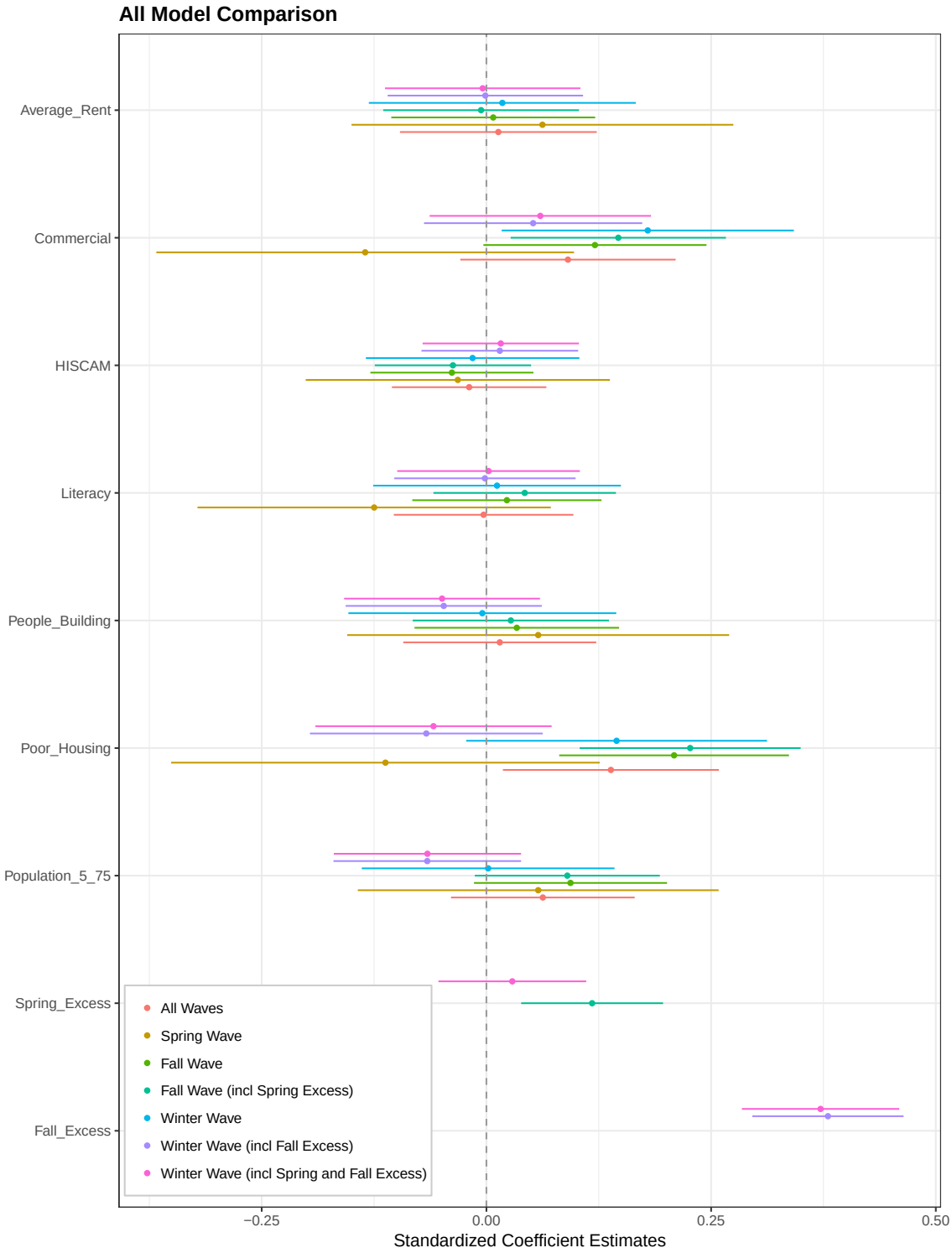

Fig. 5.3 Plot of standardized regression coefficients for presented models

Table 5.2 Regression Results for All Models

Models with all variables (exponentiated coefficients, original standard errors)

	Dependent variable:						
	Total Observed Deaths	Spring Observed Deaths	Fall Observed Deaths		Winter Observed Deaths		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Poor_Housing	1.658*	0.664	2.143**	2.288***	1.697′	0.783	0.806
	(0.224)	(0.444)	(0.238)	(0.229)	(0.312)	(0.241)	(0.245)
Literacy	0.977	0.396	1.184	1.372	1.091**	0.988	1.017
	(0.378)	(0.744)	(0.399)	(0.384)	(0.522)	(0.382)	(0.385)
HISCAM	0.997	0.996	0.995	0.995	0.998	1.002	1.002
	(0.006)	(0.012)	(0.006)	(0.006)	(0.008)	(0.006)	(0.006)
People_Building	1.000	1.002	1.001	1.001	1.000	0.999	0.999
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Average_Rent	1.000	1.001	1.000	1.000	1.000	1.000	1.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Commercial	3.473	0.157	5.243′	7.489*	11.746*	2.041**	2.277
	(0.839)	(1.627)	(0.869)	(0.838)	(1.138)	(0.850)	(0.862)
Population_5_75	5.684	4.931	13.324′	12.073′	1.055	0.162	0.162
	(1.443)	(2.835)	(1.517)	(1.453)	(1.986)	(1.474)	(1.471)
Spring_Excess				1.091**			1.022
				(0.030)			(0.031)
Fall_Excess						2.174***	2.138***
						(0.088)	(0.091)
Constant	0.249	1.244	0.086	0.075	1.044	6.584	6.189
	(1.212)	(2.375)	(1.270)	(1.216)	(1.663)	(1.240)	(1.241)
Observations	91	91	91	91	91	91	91
Log Likelihood	−463.460	−553.871	−783.259	−779.254	−702.659	−673.766	−673.521
θ	34.886 (5.962)	7.909 (1.191)	26.916 (3.982)	29.376 (4.351)	15.757 (2.331)	29.659 (4.441)	29.819 (4.465)
Akaike Inf. Crit.	942.919	1,123.742	1,582.518	1,576.509	1,421.319	1,365.531	1,367.042

*p<0.1; **p<0.05; ***p<0.01

describing the relative excess. Despite this interesting finding regarding model fit, one must remember that despite this, there is no real relationship between any of the variables and the mortality in the spring herald wave.

Despite other recent findings in Chicago [123], literacy is not present in any best-fit model as an indicator (significantly or not) of mortality differences between baseline and epidemic mortality.

5.4 Discussion: understanding social dimension of neighborhood and individual level mortality risk during epidemics

The relationship between the dependent variables and regression model differ in each wave. While certainly, much variation in excess mortality is not explained via these models, the results may be indicative of varying levels of immunity throughout the population acquired from previous waves, especially provided that the level of mortality in each preceding wave has the strongest relationship to the observed excess mortality in the wave of interest. However, the strength of the preceding wave as an indication of current wave excess mortality may also simply indicate only that those neighborhoods with high excess mortality during the outbreaks continued to have high excess mortality when the epidemic returned. To this effect, the relationship between excess mortality during the spring wave and the protracted fall wave is strong, but nearly proportional. Also, this relationship is always positive, meaning that the the higher the excess mortality in the preceding wave, the studied wave's mortality is predicted to also be higher.

In the Madrid-specific case, the idea that higher neighborhood mortality in one wave is associated with higher mortality in succeeding waves is not unexpected. The nature of the model used in this analysis takes into account underlying mortality in each neighborhood through the inclusion of an offset variable. Thus, it is difficult to directly associate high mortality neighborhoods before the epidemic with those neighborhoods with the highest excess. Nonetheless, this previously found association (see [112, 113]) provides support to these findings regarding successive waves of high excess in neighborhoods with characteristically high mortality.

In the spring, there is no significant relationship between the variables and the amount of neighborhood excess mortality. While other cities have reported strong, but not particularly lethal herald waves, the initial outbreak in May 1918 produced high levels of excess mortality, the wave's peak being much higher than any observed week in the subsequent wave from fall

1918 through winter 1919. Perhaps the virulence and transmission speed of this particular wave was such that it did affect all parts of the city with equal ferocity. This would support initial post-outbreak research that argued the flu killed indiscriminatingly [271, 277].

In terms of mortality impact, the protracted fall wave, compared to other parts of Spain and the world, was muted in Madrid, continuing with three small peaks dispersed between September 1918 and April 1919. It is also the wave were a slight absolute excess mortality bump does occur in young adults, as has been found in many populations throughout the world. Here, the strongest relationships can be found between the composition of a neighborhood and its mortality experience in preceding waves. As has been indicated in Chicago and Kristiana, areas and individuals with lower social status may have experienced a heightened mortality risk [123, 160]. Here, an indication of this effect may be found through the amount and type of housing found in a neighborhood, but this is not observed in other variables that may have described the social makeup of the neighborhood, such as the mean HISCAM score, the average rental price, and the percentage of population that could both read and write.

Yet, we should note that the results presented here, as in much of the other literature, focus on a single urban area, in which a large population of all classes lived within several kilometers of each other. However, other analyses considering larger geographical units (even adjusted for differing geographies within these units) have some times found, both at the level of the individual or larger administrative unit, differences between social status or baseline mortality and excess mortality harder to disentangle[38, 64]. The extent to which, in these larger areas versus single cities, transmission dynamics and previous exposure may have affected mortality during the outbreaks may also play a role in the observed results.

The last wave of pandemic influenza struck with a vengeance between December 1919 and January 1920. In this wave, there is also little found social impact on mortality, beyond that of the increased mortality in neighborhoods with higher numbers of stores and businesses as a percentage of total building units. The interpretation of this may be vague, but it could be a result of transmission dynamics in neighborhoods where people traveled to for work or to run errands.

While the results do not directly contradict recent findings, they also are not able to provide strong further evidence of a social dimension to influenza mortality during the Spanish flu outbreaks. Certainly, the historical nature of and available data for the analysis limit the extent to which one can examine the relationship. Yet as the overall mortality experience of the flu differed throughout the world, the extent to which social status within a single urban environment influenced mortality also likely varied. This analysis provides another example from which researchers can advance research to disentangle this relationship

between environment, social and material resources, and health. The continued effort to understand this relationship can lead to contemporary solutions in future influenza epidemics.

Chapter 6

Scotland during the Spanish flu: new perspectives on age-specific mortality changes

Abstract

Newly digitized sources with detailed death information provide researchers the opportunity to examine conundrums about past pandemics, including the 1918 Spanish influenza outbreaks, such as the unique age-specific mortality curve in which young adults died at abnormally high rates, particularly to respiratory causes. Here, mortality records for the entire country of Scotland are used to look at all cause age-specific excess during four distinct epidemic waves between 1918 and 1922. After introducing the concept of and creating “seasonal death tables," changes in wave-specific measures of life expectancy (e_0) and lifespan inequality ($e^\dagger$) are decomposed by individual wave groups. The results show large changes in these aggregate measures between pre- and during- wave periods driven by increased young adult mortality and highlight how underlying frailty in the population, due to such diseases as tuberculosis, led to a harvesting effect in the strong fall wave. The discussion focuses on the necessity of targeting susceptible groups in vaccination and other prevention programs before and during future pandemics.

Resumen

El capítulo seis trata de mostrar una comparativa internacional. Las fuentes recientemente digitalizadas con información detallada sobre muertes proporcionan a los investigadores la oportunidad de examinar los enigmas sobre pandemias pasadas. Aquí, los registros de mortalidad de toda Escocia se utilizan para examinar el exceso de mortalidad por edad durante cuatro olas epidémicas distintas entre 1918 y 1922, examinando el alcance de la curva de mortalidad por edad en la que los adultos jóvenes murieron a tasas anormalmente altas Después de introducir el concepto y crear "tablas de mortalidad estacional", los cambios en las medidas específicas de la esperanza de vida (e_0) y la desigualdad en la esperanza de vida (e^{dagger}) se descomponen por grupos de olas individuales. Los resultados muestran grandes cambios en estas medidas agregadas entre los períodos previos y durante la ola, impulsados por el aumento de la mortalidad de los adultos jóvenes, y ponen de relieve cómo la fragilidad subyacente de la población, debida a enfermedades como la tuberculosis, condujo a un exceso de mortalidad en la fuerte ola de otoño. El debate se centra en la

necesidad de dirigirse a los grupos susceptibles en los programas de vacunación y otros programas de prevención, antes y durante pandemias futuras.

6.1 Background: Spanish flu in Scotland

6.1.1 Timing in the country

Former studies on the global path of the outbreaks put the arrival of a herald wave flu in Britain in June of 1918 [140, 209]. Though in this herald wave, the virus in Britain is thought to have been brought via an incoming ship in a port city, its exact origin is unknown. In Scotland, this small wave peaked in July 1918, but did not reach the entire country. In fact, these initial outbreaks went largely unnoticed in much of the country and have not been studied statistically. Monthly data indicates that as a whole, the fall wave occurred later that year between October and January and was responsible for the highest death counts during the outbreaks. Scotland also experienced a lethal winter wave stretching from late January through April 1919 [140]. An echo wave gripped the country in December 1921 and January 1922, but it seems little or no research has been dedicated to this outbreak, nor has its impact been quantified.

Scotland's fall and winter experience differed from that of England and Wales. These two countries combined had a much higher relative fall peak than in the subsequent winter wave. Nearly three times as many deaths occurred in the hardest hit fall month, November 1918, than in March, the peak month of the 1919 winter wave [140].

Within the "principal towns" of Scotland, the peak week of each wave also varied, with the largest and most well connected towns experiencing the highest absolute weekly rates, nearly all of which occurred in the third winter wave. In prior literature, no focus is given to the geographic regions primarily affected by the spring wave, which creates the question, if these transport hubs had higher mortality in the third wave, did the spring wave hit only these cities, and did the impact of this exposure create some immunity in the population during the fall wave?

6.1.2 Age-specific mortality patterns in Scotland and beyond

When comparing seasonal and pandemic influenza, young adults often face increased mortality during epidemics relative to seasonal flu, when infants, young children, and elderly are disproportionately affected. Many studies have shown high young adult age peaks during the Spanish Influenza pandemic, though this pattern is not universal (see section 1.4.3 and chapter 4). Given the circulation of both the H1N1 strain associated with the pandemic and other strains of seasonal influenza viruses circulating before, concurrently, and after the outbreaks, additional research has looked at the change in age specific mortality patterns in an attempt to identify the appearance of the virus in a population [198, 255]. Echo waves are

known to have occurred in many locations throughout the world. Generally, excess mortality during these re-occurrences returns to the traditional seasonal pattern, particularly affecting 0-1 year olds, who would not have been alive during the initial waves.

Using a newly available source of data, individual records of all registered civilian deaths in Scotland from 1916 to 1923, this paper re-examines the mortality experience of Scotland during four distinct waves of epidemic flu between 1918 and 1922. The paper first quantifies excess mortality in several age groups for each wave to add to the research on age-specific mortality excess. Then, traditional demographic methods are used to create and decompose “seasonal" death tables to further understand the ways in which changes in age-specific mortality during the individual waves contributed to population-level changes in life expectancy and lifespan disparity.

6.2 Excess mortality calculation

6.2.1 Data: Digitising Scotland project

The Digitising Scotland project, many years in the making, is a large-scale collaborative effort to digitize all birth, marriage, and death records in Scotland between 1856 and 1973 for analysis [91]. As the data keying and verification process is still on-going, the project presented in part here is the first to be able to take advantage of these records for research. Individual registered records of death between 1916 and 1923 include information such as when and where the person died (and in which district the death was registered), sex, age at death, self and parents occupation information, and cause of death.[1] A sample of a death record, according to the form used between 1861 and 1965 is located in figure 6.1. Of the total 564,230 records in the dataset, 564,008 occur between 1916 and 1923, and a total of 1821 of these (0.29%) do not contain information about the age at death. Thus, for all-cause age specific analyses, a total of 562,356 records are used to make the estimations. Additionally, total population numbers, used to determine death rates, come from annual statistics published by the National Records of Scotland.

Age at death distribution (by some causes)

Before estimating excess deaths, simple count frequencies by wave provide an idea of the magnitude of each wave without reference to baseline mortality (see Table 6.1).

[1] Civilian deaths; those who died away from Scotland in WWI are not present in these records. Nearly always, the district where the individual dies is the same as that in which the death is registered.

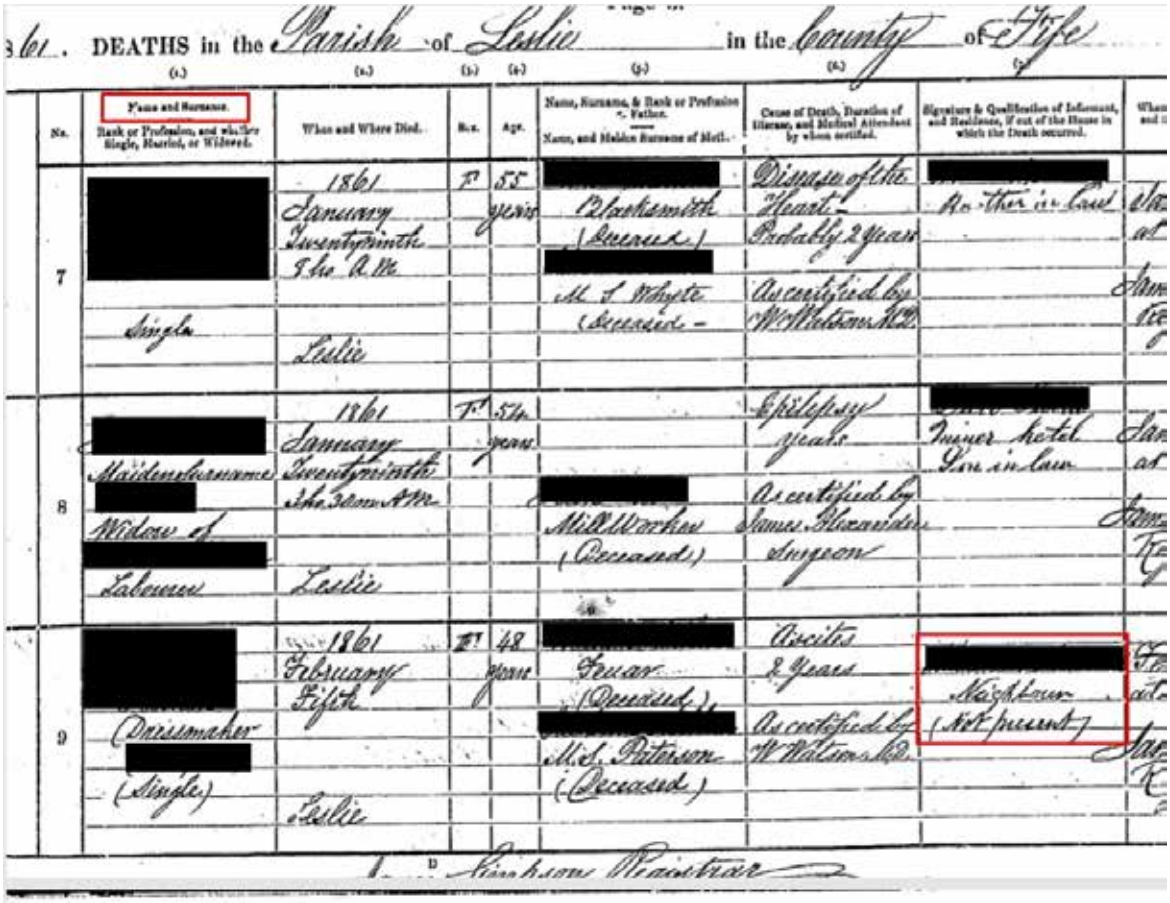

Fig. 6.1 Sample Scottish death record from 1861 with names redacted. Source: National Records of Scotland [197].

Table 6.1 Total Deaths During Each Wave

Sex	Wave	Deaths	Mortality Rate per 10,000
Male	Herald	30406	131
Female	Herald	29614	118
Male	Fall	52109	225
Female	Fall	54526	218
Male	Winter	54631	236
Female	Winter	56333	224

Of all waves, the winter outbreak has the highest number of deaths, especially considering the youngest and older ages. However, the fall outbreak contributes the most total deaths to each age between 5 and 40. Unsurprisingly, the spring herald wave and 1922-23 account for the fewest deaths of the epidemic waves. However, considering that these crude death rates are not indicative of the baseline mortality at the time, the total excess due to each wave must by calculated in consideration of expected mortality absent the influenza outbreaks.

6.2.2 Methods: baseline and excess mortality estimates

In order to understand the mortality impact of Spanish flu on the total population and several subgroups, the underlying mortality level must first be quantified. Unlike the unique mortality patterns in Madrid (see chapter 3), the seasonal baseline in Scotland could be easily parameterized with a simple Serfling regression. As noted in chapter 3, the equation takes into account seasonal (α) and time (β) trends in mortality, where M_t is the mortality rate at time t: $M_t = \alpha * sin(\frac{2\pi}{52.17*t}) + \alpha * cos(\frac{2\pi}{52.17*t}) + \beta * t$ Separate models were estimated for each age group and by sex for all causes of death. From each model, the 95% confidence interval of the baseline is calculated, from which waves are determined as periods with sustained mortality above the upper bound. For each wave, absolute excess mortality is quantified as the difference between the observed rate and the baseline mortality estimate. Relative excess mortality equals the observed rate divided by the expected rate of the baseline.

6.2.3 Results: excess mortality calculations

Sex differences in mortality

281,124 deaths are women, and 283,024 are to men. As noted, mortality has been found to peak at around age 28, specifically among men. This general finding is also found for men in the data, but more interestingly, a slightly *younger* mortality peak in women occurs at age 26. At ages less than 35, death counts for women are much higher, however this might be easily

explained by the large numbers of men fighting in World War I at the time. Their physical absence in Scotland and the fact that they are not included in the death records is the likely reason so many more deaths occur among women during the epidemic waves. While the number of female deaths continues to remain higher than men in the years following the conclusion of the war; this could also remain as a relic of the smaller population of men (due to large casualties of the cohorts that fought in the war).

The largest differences between male and female excess occur in the echo wave, when men, especially those in late youth and middle age, have much lower excess rates than women. This may be an artifact of the data, which may reflect the return of soldiers and therefore increase the population denominator, *or* perhaps, the results show an effect of immunity. It is plausible to imagine that these soldiers were exposed to different strains of the virus while away, making them less susceptible to death in this fall wave.

Age-specific mortality time series

Figure 6.2 reveals three distinct waves of Spanish flu between 1918 and 1919 as well as one echo wave between 1921 and 22. The herald wave in July 1918 is mild, quite small, and not universally experienced; the mortality associated with this first wave is primarily observed between the ages of 10 and 50. Despite the low overall excess in this wave, the age-specific structure of mortality does strongly suggest that the mortality spike is a result of a small outbreak of influenza associated with the fall pandemics. Other research has shown that the first small outbreaks in the spring contain mortality peaks at younger ages, more consistent with the characteristic fall age-specific mortality pattern than in the previous year's seasonal influenza mortality [198].

Additionally, two clearly defined waves in the fall of 1918 and winter 1919 account for the majority of mortality, with two large, distinct peaks in all ages except for those above 70. These waves also exhibit differing mortality patterns by age; up to age 40, the peak week of absolute mortality occurs in the fall wave, but beginning with the 40 to 50 age group, the apex of weekly excess mortality occurs in the winter wave. An echo wave in 1921-22 produces varying amounts of mortality excess according to age group, but primarily, the youngest (those under 5) and older adults are affected.

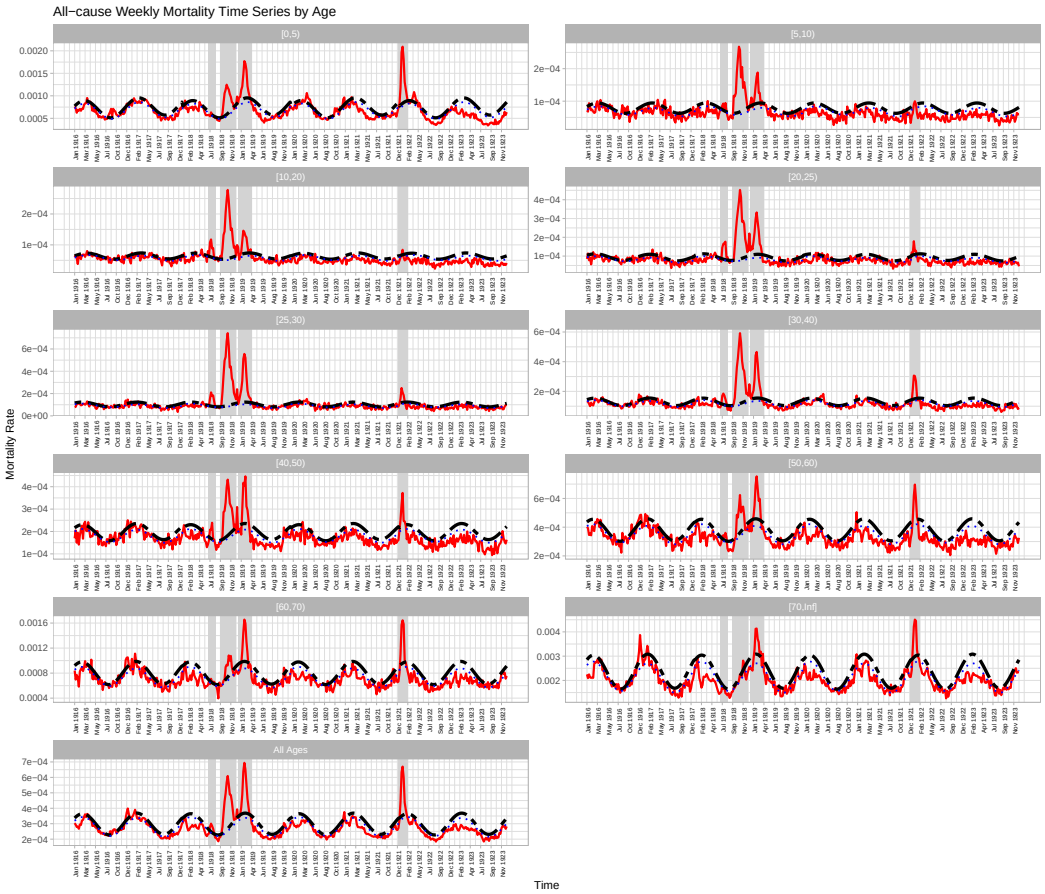

Fig. 6.2 The red line shows the real weekly mortality rates from 1916-1923, while the black line displays the predicted mortality values based on a simple Serfling regression.

Age-specific excess mortality

Table 6.2 shows excess mortality estimates by age-group and wave, including the total excess amount, for both sexes. Figure 6.3 graphically depicts these numerical estimates. Sex-specific estimates can be found in Appendix E. Total excess mortality of all four waves in Scotland is quantified as 26,685 excess all-cause deaths (excess rate of 55 per 10,000).

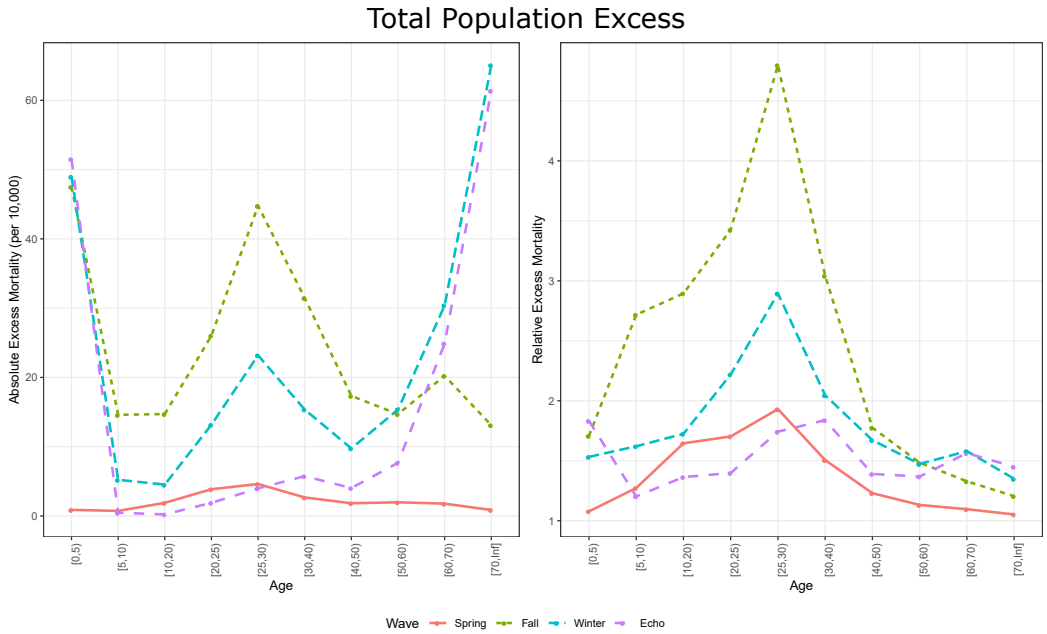

Fig. 6.3 Age-specific Absolute and Relative Excess Mortality by Wave, Scotland 1918-1922

The virulent fall wave witnessed across the world also produced the most excess deaths in the whole of Scotland, followed by the subsequent winter wave. The spring herald wave produced the smallest amount of total excess mortality, despite occurring during the period with the lowest baseline rates, and the echo wave produced slightly less than half the excess deaths as the strongest (fall) wave.

Throughout Scotland, wave and age specific mortality reveals the familiar "w" shaped excess mortality pattern in the two largest waves of influenza, fall 1918 and winter 1919. The highest rates of excess are present in the fall wave, when those 25 to 30 had more than 4.5 times the expected levels of mortality. As seen in the weekly time series of mortality rates, the herald wave in July of 1918 produced small levels of excess mortality, especially at the youngest and oldest ages, but the excess mortality hump in young adults is consistent with the subsequent fall and winter waves. Notably, in the echo wave, the pattern of absolute excess mortality returns to a pattern relatively consistent with that of seasonal flu; the youngest and oldest age group experience the highest rates of excess mortality, while this level remains low for children and young adults.

Sex- and age-specific excess mortality

Figure 6.4 shows absolute and relative excess mortality by wave for each sex for ten age groups. Given that the youngest children and elderly have the highest baseline mortality and the patterns of absolute excess, it is no surprise that they exhibit the lowest levels of relative

mortality in all waves. Moreover, they have the highest Standardized Mortality Ratio (SMR) in the echo wave, more than that in any of the preceding waves of Spanish flu. With one exception, the highest levels of relative excess are found among 25 to 30 year olds, similar to other places finding a mortality peak around age 28 [110, 284]. This relative rate implies that during the fall wave, nearly 5 times as many people died than in the same period given the calculated baseline levels of mortality. These patterns are generally consistent across sexes.

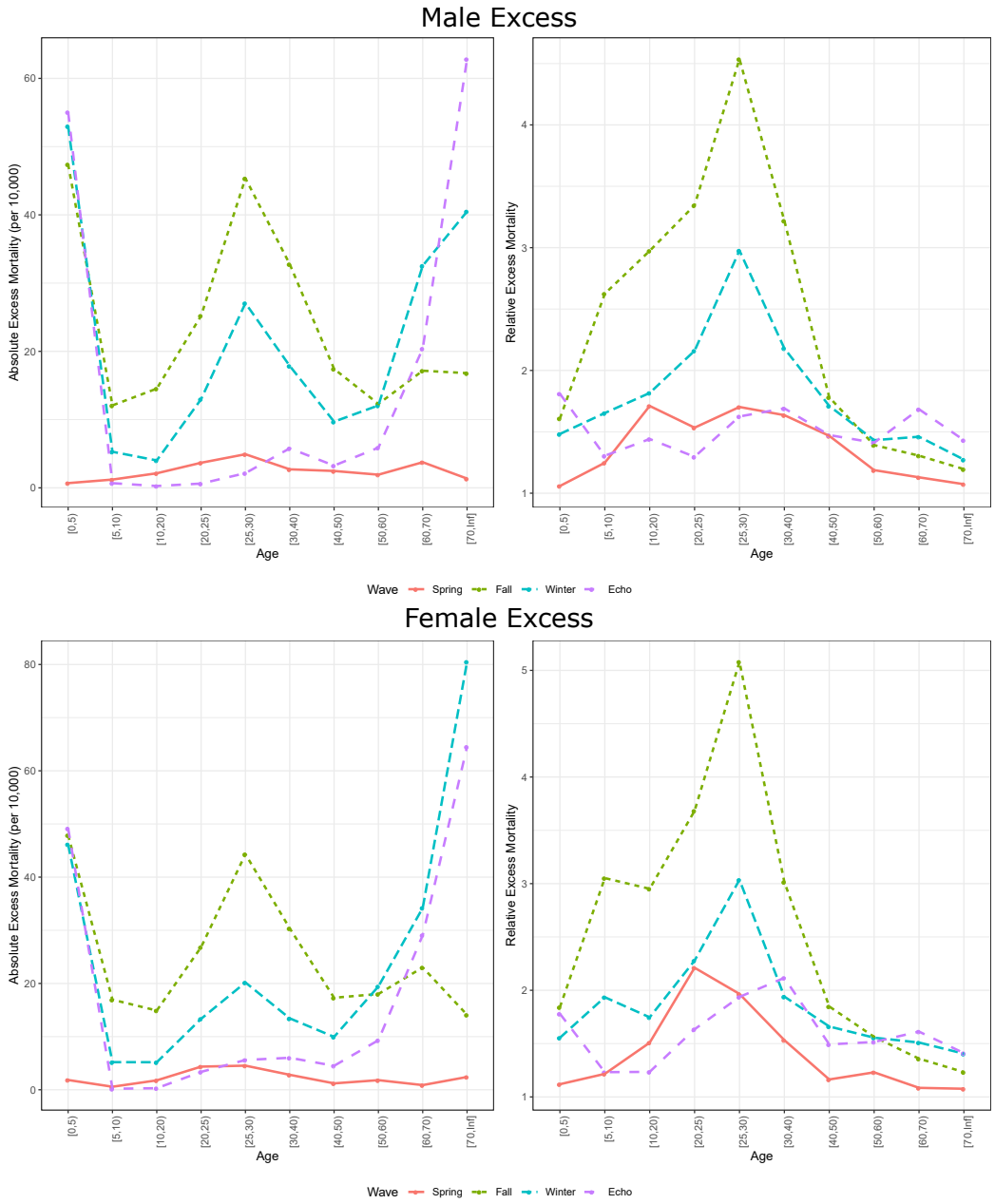

Fig. 6.4 Age-specific Absolute and Relative Excess Mortality by Wave and Sex, Scotland 1918-1922.

It is worth noting that while the winter wave in 1919 and echo wave in 1921-22 produced higher gross weekly rates of mortality, due to higher baseline rates during this time of year (due to seasonality in mortality) and smaller time-frame of outbreak (particularly in the echo wave), the total levels of excess (both absolute and relative) are smaller than those observed in the fall.

6.3 Seasonal Death Tables[2]

As is well documented, mortality patterns generally change throughout the year, and in the context of the Spanish flu, the age-specific mortality patterns shifted before, during, and after the pandemic. A simple visualization of the density of age at death in those over 5 (Figure 6.5) reveals large changes in the age structure of death in Scotland between 1916 and 1923, particularly during 1918 and 1919, the two years in which the largest epidemic waves hit. Moreover, the density of deaths by age in each wave differs, demonstrating changing age-specific mortality impact during the first, second, third, and echo waves (Figure 6.6) and the strain of virus mutates and much of the population acquires some form of immunity.

[2]A conscious decision has been made to use the term “death table" rather than “life table." This is largely due to two reasons. First, the seasonal nature of the table, in that it does not accurately capture the entire yearly mortality experience of the population. Secondly, their creation and use is to examine large increases in deaths during the pandemic rather than focusing on years of life *lived*.

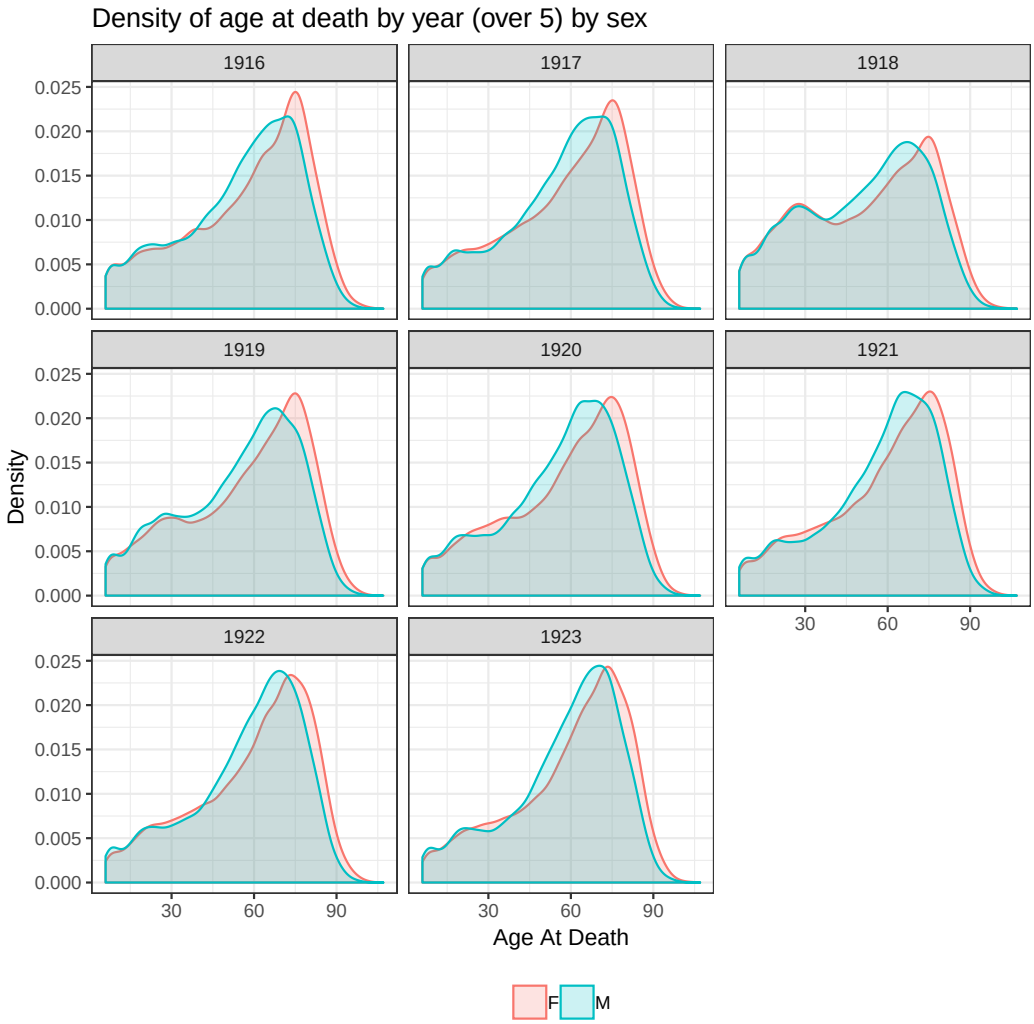

Fig. 6.5 Density of Deaths by Age Over 5 Years Old, 1916-1923

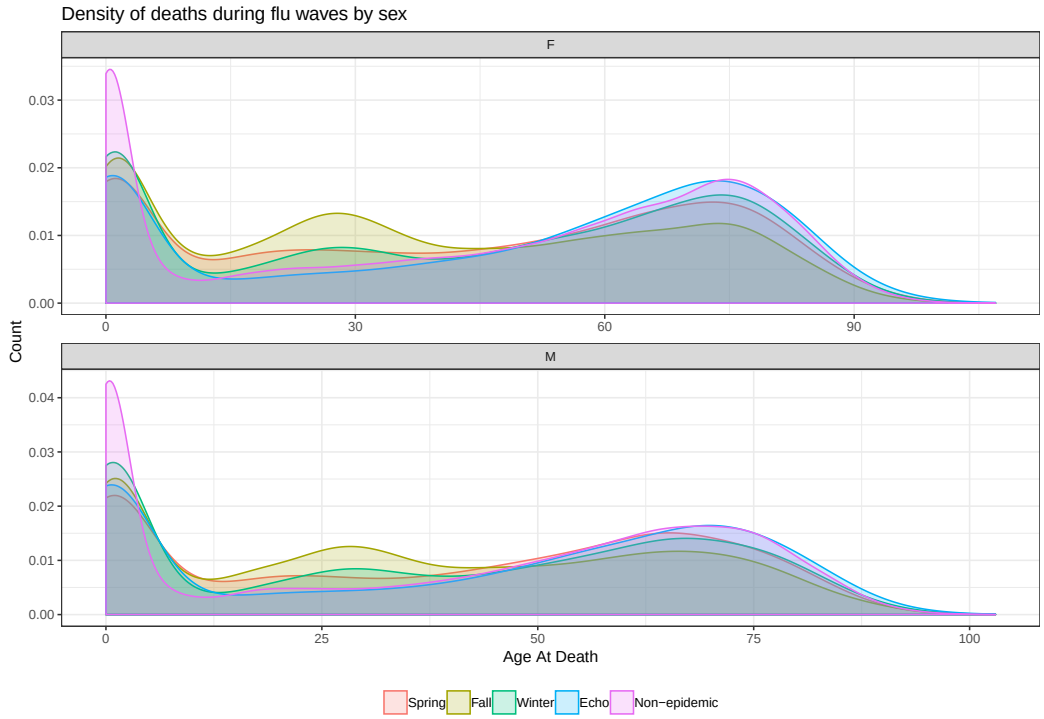

Fig. 6.6 Density of Deaths by Age and Wave

Thus, in order to understand and quantify the mortality impact of a single wave of influenza, the underlying expected mortality during the time of the year at which the outbreak occurs must be considered in conjunction with the wave.

By tabulating the number of deaths and population exposure by age during the period of each wave during the years before, during, and after the outbreaks (see figure 6.7), so-called "seasonal death tables" can summarize the mortality schedule of a particular time of year and provide detailed information about its changes during the period of a specific epidemic wave (for example, see figure 6.8). Once in the lifetable framework, subsequent calculations and decomposition techniques can reveal at which specific ages and by how much did the mortality structure change and contribute to the aggregate change in mortality schedule. These seasonal death tables differ from traditional mortality schedules in that they focus on the mortality at a particular time of year; while many analyses have found the change in life expectancy during 1918 as a whole [127], the change may be even more dramatic during a particular wave.

It should be stressed that because these death tables do not represent the total mortality pattern within a year, they must be used through an abstract lens of mortality characterized within a specific period. While generally, period lifetables reflect the mortality patterns at a given time, they can also be used to approximate the average expected life expectancy of an

individual born in the time, given unchanging mortality rates over time. Because these death tables only represent mortality patterns at a specific point of the year, their use should always be explained and justified in the context of the proposed research. In this context, without an influenza outbreak, pre-epidemic period mortality rates would be expected to remain reasonably similar to those only one year later. By comparing the age-specific mortality rates during the same period of the year before and during each outbreak, the mortality impact of individual waves of influenza can be understood in a different way.

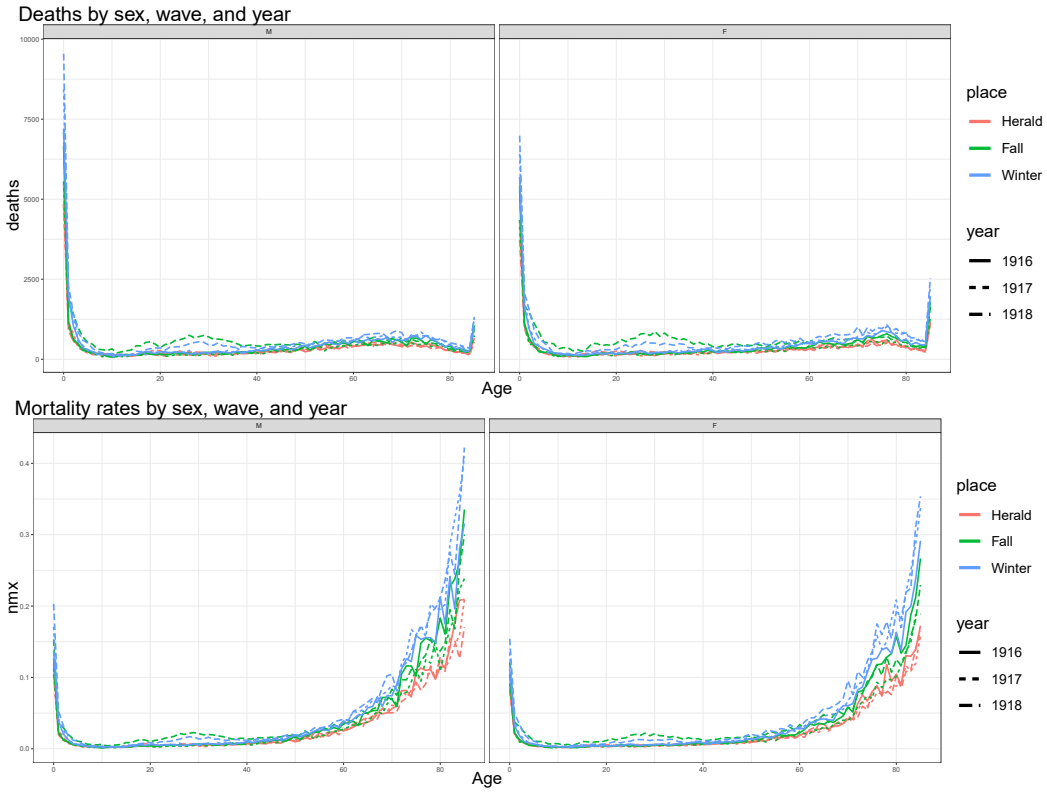

Fig. 6.7 Deaths and Mortality Rate by Age, Wave Period, and Year

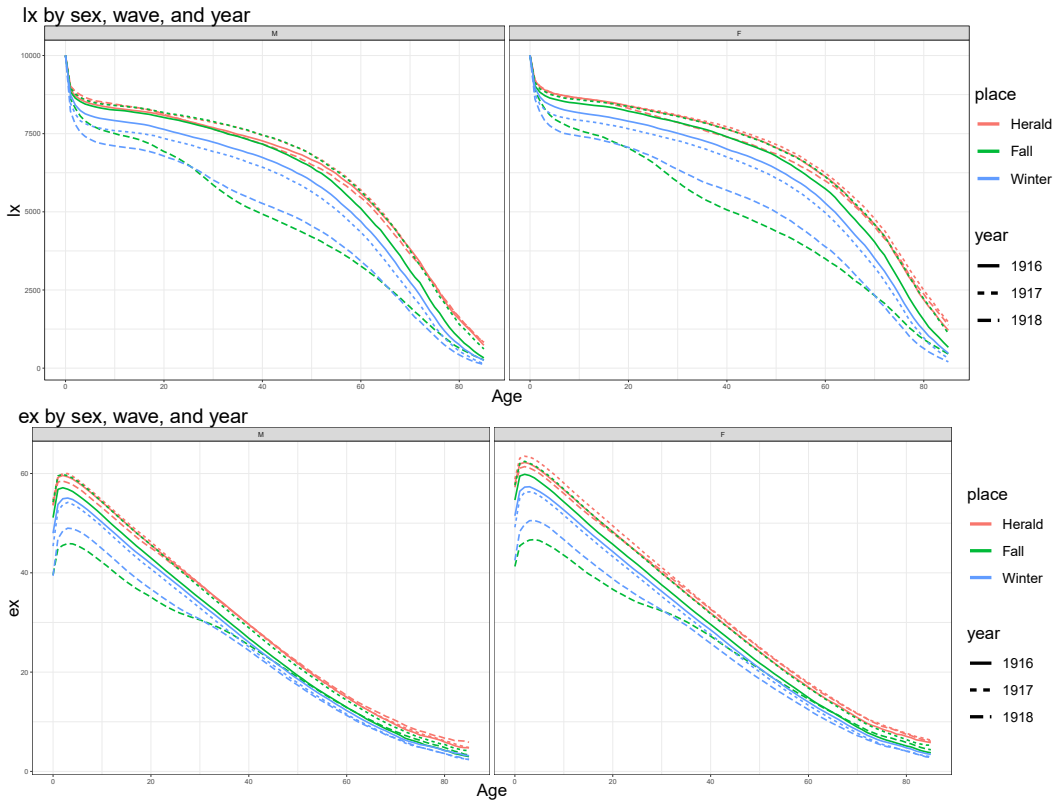

Fig. 6.8 l_x and e_x by Wave Period and Year

Sex-specific period death tables were created for the period of each wave, as well as the same period in the years preceding the outbreaks, as presented in Appendix E and visualized in figure 6.8. Between 1916 and 1917, life expectancy improves, but mortality is much higher in the periods in 1918 and 1919 (winter wave) during each outbreak, leading to lower life expectancy. Across all years and ages, the life expectancy, based on the crude death rates, is highest during the period of time corresponding to the years before the herald outbreak, while, generally, the winter period has the lowest. This is expected, as it corresponds to the seasonality of mortality, whereby more people die in the winter months. The exception occurs during the fall epidemic wave, life expectancy falls to its lowest in all periods–31 years in both males and females.

While still lower than the pre-epidemic period, overall, the mortality during the herald wave is the closest to its “normal" expected mortality (the years preceding the outbreak). Moreover, the shape of the mortality and life expectancy curves are similar in all years during the herald period. This is unsurprising; the age-specific curves of absolute and relative excess rates showed little notable differences during the spring wave, especially relative to the fall.

However, large differences can be observed in the death tables that reflect the mortality patterns during the fall and winter waves. Generally, as populations shift to today’s current

mortality structure in which more people die at older ages, a “rectangularization" of the survivor or l_x curve occurs, meaning that more people survive to older ages, at which a sharp drop off occurs as the modal peak grows and more individuals die around this age. However, the Kaplan-Meier estimates of the seasonal death tables, shown in figure 6.9, show a “de-rectangularization" effect of the survivor cure during the periods associated with influenza waves (particularly the fall and winter), due to the characteristic deaths to those at younger ages. The effect is muted in the small Herald wave and even less present in the last echo wave, in which as noted in section 6.2.3, the age-pattern in excess mortality returned to a pattern more consistent with seasonal influenza.

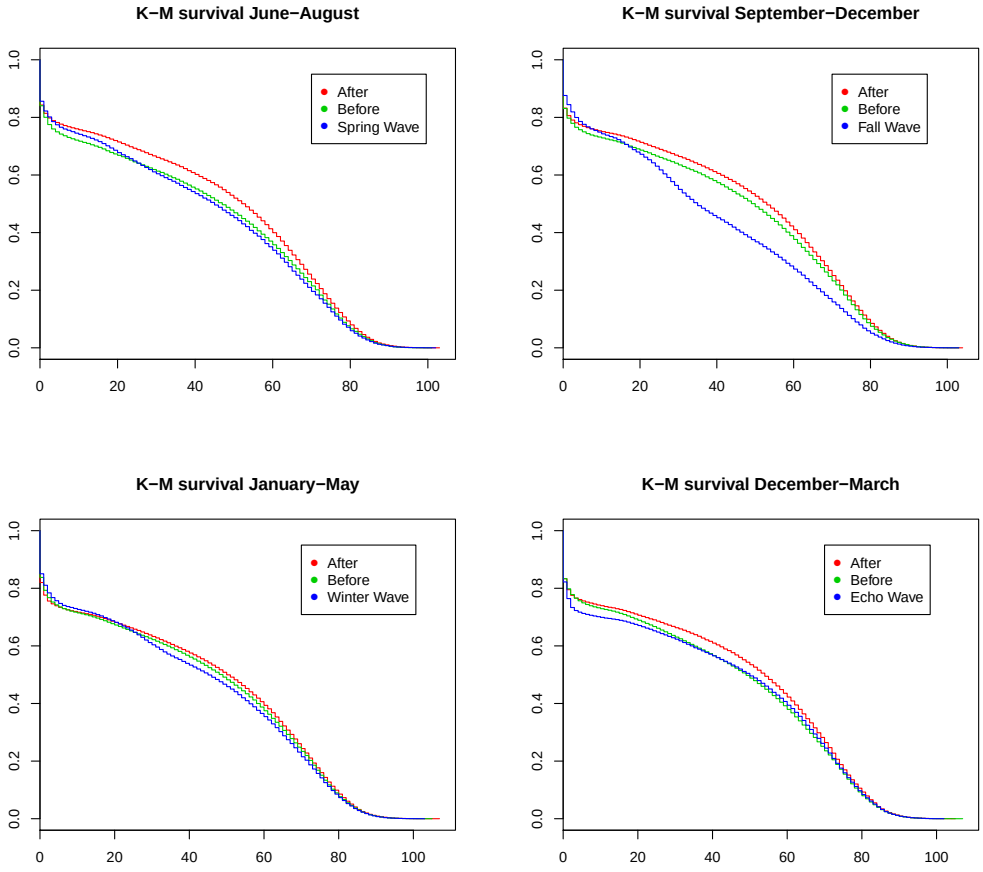

Fig. 6.9 Kaplan-Meier estimates, corresponding to periods before, during, and after each wave

6.3.1 Age-specific contribution to change in life expectancy: Arriaga decomposition

The seasonal death tables allow mortality changes to be compared and visualized, but the numeric extent to which each age contributed to aggregate measures describing mortality within the population can better be understood using decomposition techniques [49]. Here, Arriaga decomposition is used to demonstrate how the changes in life expectancy and life disparity in each individual wave can be understood by changes in age-specific mortality patterns. Arriaga's formula, published in 1984 was one of a number of similar discrete methods for lifetable decomposition published in the 1980s, along with Andreev, and Pressat [281]. The formula decomposes the total age-specific contributions to the change in life expectancy into *Direct* and *Indirect* contributions [27]. That is to say, because e_0 is actually a function of the total number of life years lived, T_0, or the sum of ${}_nL_x$ values across all age intervals, the number of survivors that changes in a given age interval between death tables affects the total calculation of e_0 in two ways. The *direct* effect of age-specific mortality differences to the change in life expectancy at birth measures the effect that the change of mortality rates within a single age group had on the change in the number of years lived in that group, ${}_nL_x$. The *indirect* effect can be described as the change in T_x in each subsequent age group due to the change in survivorship in previous age groups. In the case of the Spanish flu, younger age groups die at higher rates, thereby indirectly decreasing the total number of people years lived in subsequent age groups. Together, the direct and indirect contributions to the change in e_0 total the overall change in average life expectancy from one period ot the next.

It has been documented that many died not simply from "influenza" but from a variety of respiratory-related diseases, such as pneumonia and bronchitis. Furthermore, much research has been completed that demonstrates a particular relationship between mortality during the Spanish flu pandemic and tuberculosis [14, 195, 305], although the sex-specific effects have been debated [43, 191, 192]. Not only did those with tuberculosis tend to die at higher rates [164], this overall effect contributed to a harvesting of deaths that would have been to tuberculosis, leading to much lower rates of tuberculosis mortality in the years following the influenza outbreaks in many parts of the world. It is possible to divide each age-specific contribution to show how each specific cause of death contributed to differences in life expectancy. In order to do so, one must simply multiply the age-specific contribution by the proportion of deaths to each cause by age. The results are divided into deaths to (a) respiratory

causes, including influenza, (b) tuberculosis,[3] and (c) other causes and are included in figure 6.10.

In the case of each individual wave,the decomposition highlights the role that the increased mortality in young adult age groups contributed greatly to the overall mortality schedule of the population. Figure 6.10 shows the age-specific contributions to the overall change in life expectancy according to period rates during each wave and by sex. Table 6.3 shows the total contribution for each sex to the decrease in life expectancy in each wave by cause. Life expectancy falls in each wave, but especially so in the fall and winter wave of 1918 and 1919. Female life expectancy drops by nearly 15 years in the fall for women, and between 7 and 8 years for both sexes in the winter waves. While the lower mortality of the spring wave also leads to a smaller change in life expectancy, the role of young-adult age groups in decreasing the life expectancy during this time, consistent with the age-specific mortality patterns of the flu, is quite visible.

With regards to causes of death, the results show a distinct pattern. Whereby the “other" deaths make up the majority of the contribution to changing life expectancy, those ascribed to respiratory illnesses also provide a large amount of the difference. Given the similarity in age-specific contributions of the two, this visualization is a good demonstration of the importance of considering both respiratory and all-cause causes of death. However, most striking is the contribution of tuberculosis as a cause of death contributing to change in e_0. The majority of the contribution occurs during the fall wave. In the subsequent winter wave, there is little contribution of tuberculosis deaths (as there were few) to reduced life expectancy.

Notably, the differences in the amount of direct and indirect age-specific contributions is large. Generally speaking, the direct effect within each age group towards the total change in life expectancy is much smaller than the indirect change. This pattern is quite easily understood; because mortality is different (higher) at child and young adult ages, this *indirectly* affects the amount of total person years that can be lived in the population at later ages, effectively lowering the total number of person years lived in the life table. At the oldest ages, the changes in mortality rates at each age directly effect the change in life expectancy more so than the measured indirect contributions. Because there are fewer in each age group–ergo fewer total life years left to live–left at the end of the death table, each change in mortality rates has a higher total impact on the remaining life expectancy and thereby, within the age group, the contribution due to the direct change in rates is higher than the indirect contribution.

[3]Tuberculosis deaths include: Sturma, Scrofula, Phthisis, consumption, Lupus Vulgaris, Pott’s disease, Meningitis, and Tuberculosis

Table 6.2 Excess mortality in Scotland by age and wave

Age Group	Wave	Excess Mortality Rate	Excess Deaths	Standardized Mortality Ratio
[0,5)	Total	148.87	6882	1.64
[5,10)	Total	21.03	1043	1.96
[10,20)	Total	21.36	2056	2.22
[20,25)	Total	44.97	1869	2.41
[25,30)	Total	76.56	2844	3.23
[30,40)	Total	55.26	3723	2.31
[40,50)	Total	32.96	1942	1.60
[50,60)	Total	39.70	1703	1.40
[60,70)	Total	77.29	2124	1.44
70+	Total	140.44	2311	1.35
[0,5)	Spring	0.88	41	1.08
[5,10)	Spring	0.73	36	1.27
[10,20)	Spring	1.90	182	1.65
[20,25)	Spring	3.86	160	1.70
[25,30)	Spring	4.62	171	1.93
[30,40)	Spring	2.68	181	1.51
[40,50)	Spring	1.84	1085	1.23
[50,60)	Spring	1.98	83	1.13
[60,70)	Spring	1.78	48	1.10
70+	Spring	0.87	14	1.05
[0,5)	Fall	47.50	2224	1.71
[5,10)	Fall	14.60	726	2.72
[10,20)	Fall	14.71	1414	2.90
[20,25)	Fall	26.03	1076	3.43
[25,30)	Fall	44.77	1659	4.80
[30,40)	Fall	31.47	2126	3.04
[40,50)	Fall	17.34	1016	1.78
[50,60)	Fall	14.70	620	1.48
[60,70)	Fall	20.23	542	1.33
70+	Fall	13.10	210	1.20
[0,5)	Winter	48.98	2174	1.53
[5,10)	Winter	5.22	258	1.62
[10,20)	Winter	4.53	438	1.72
[20,25)	Winter	13.17	551	2.22
[25,30)	Winter	23.19	864	2.89
[30,40)	Winter	15.36	1036	2.05
[40,50)	Winter	9.75	577	1.67
[50,60)	Winter	15.36	658	1.47
[60,70)	Winter	30.39	826	1.58
70+	Winter	65.08	1051	1.35
[0,5)	Echo	51.51	2443	1.83
[5,10)	Echo	0.47	23	1.20
[10,20)	Echo	0.22	22	1.36
[20,25)	Echo	1.90	82	1.40
[25,30)	Echo	3.98	150	1.74
[30,40)	Echo	5.75	381	1.84
[40,50)	Echo	4.03	241	1.39
[50,60)	Echo	7.67	342	1.37
[60,70)	Echo	24.89	708	1.56
70+	Echo	61.39	1036	1.45

Table 6.3 Total contributions towards change in e_0 by wave, sex, and cause

Sex	Wave	Cause	Total contribution to decrease in e_0
Female	Herald	All Causes	-1.12
Female	Herald	Other	-0.57
Female	Herald	Respiratory	-0.52
Female	Herald	Tuberculosis	-0.03
Female	Fall	All Causes	-14.59
Female	Fall	Other	-6.85
Female	Fall	Respiratory	-6.33
Female	Fall	Tuberculosis	-1.41
Female	Winter	All Causes	-8.16
Female	Winter	Other	-5.55
Female	Winter	Respiratory	-2.65
Female	Winter	Tuberculosis	0.04
Male	Herald	All Causes	-3.67
Male	Herald	Other	-3.08
Male	Herald	Respiratory	-0.57
Male	Herald	Tuberculosis	-0.02
Male	Fall	All Causes	-14.81
Male	Fall	Other	-7.08
Male	Fall	Respiratory	-6.34
Male	Fall	Tuberculosis	-1.39
Male	Winter	All Causes	-9.08
Male	Winter	Other	-6.39
Male	Winter	Respiratory	-2.77
Male	Winter	Tuberculosis	0.08

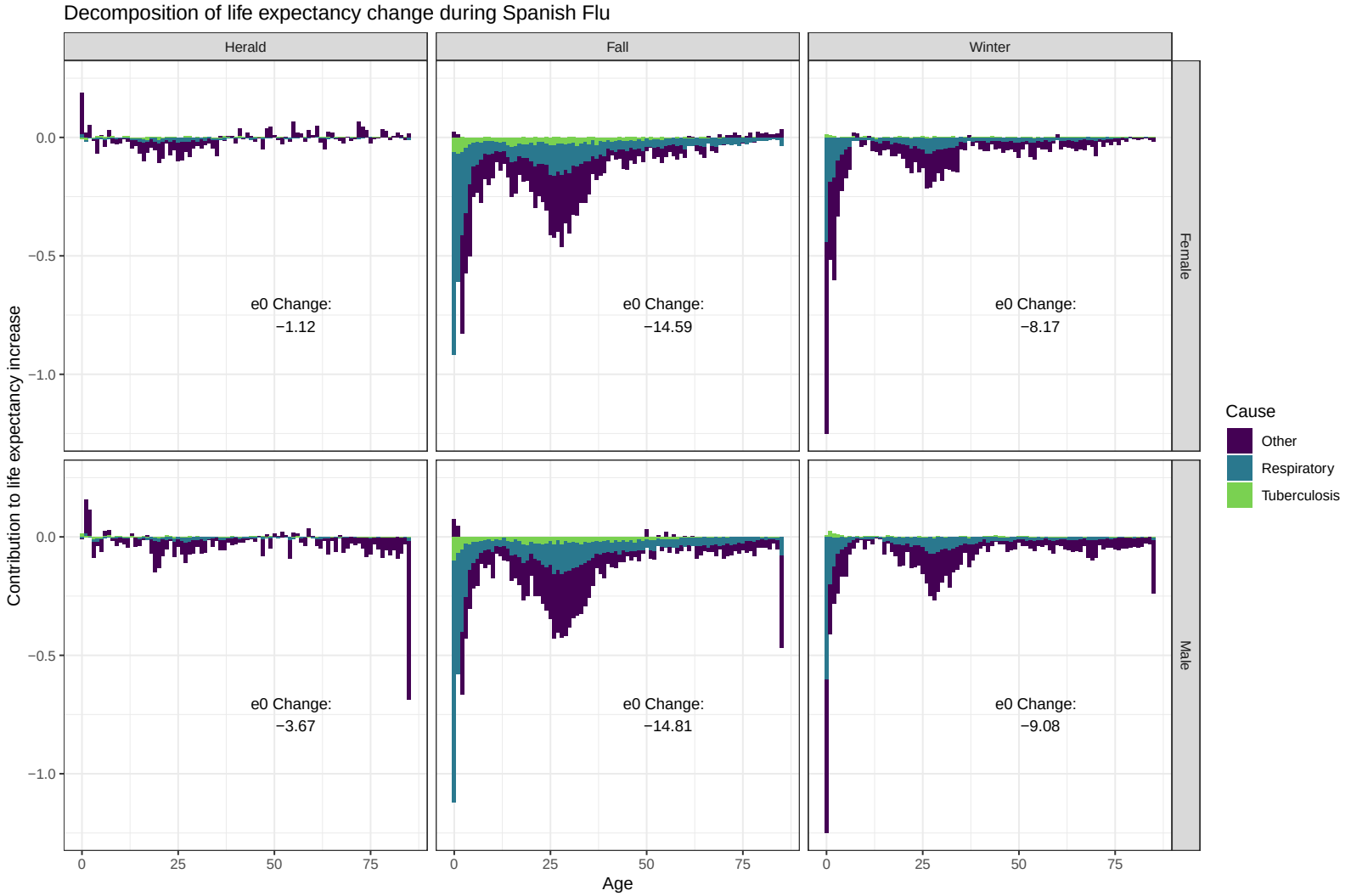

Fig. 6.10 Decomposition of Life Expectancy Changes by Sex

6.3.2 Age-specific contribution to change in E-dagger ($e^{\dagger}$): step-wise decomposition

One of the main principles in Arriaga's formula is that changes in each individual age-specific rate each play a different role in the overall difference of life expectancy between two populations. Step-wise decomposition takes a similar approach to understanding these population-level differences, but it also allows the decomposition for more complicated lifetable functions, beyond e_x. Basically, by changing the age-specific mortality rates of a population (a) to another (b) one at a time and then recalculating the aggregate measure of interest, the contribution of the change in each age-group towards the life table function is decomposed [25].

Here, the measure of life span disparity, also known as e-dagger, or $e^{\dagger}$, is decomposed by the step-wise technique to display the effect of each wave on the "equalness" of the lengths of life lived [49]. In contemporary mortality schedules, demographers generally see improvements in $e^{\dagger}$ as mortality becomes concentrated around the modal age of death, and then as the standard deviation around this mode decreases [106]. However, in historical populations, e-dagger was still heavily influenced by the level of infant and child mortality. During the Spanish flu, heightened in mortality in middle ages also contributed to higher levels of lifespan disparity. That is to say, while before the flu, the age-specific mortality schedule featured a bi-modal hump, during the 1918-19 outbreaks, particularly in the fall and winter, the addition of higher rates in middle ages, especially around the additional third mode (as seen in figure 6.6), increased $e^{\dagger}$. During the waves, the total person years of life to live at birth (T_0) decreased, thereby increasing the lifespan disparity. By step-wise decomposing this measure by age, it is possible to see exactly to what extent the additional mortality at each age played a role in added lifespan inequalities during the waves and led to the decrease in total person years lived. Figure 6.11 shows these results by the selected causes of death, and table 6.4 shows the total contribution to the change in each wave by cause for each sex.

As expected the total lifespan disparity according to $e^{\dagger}$ increases for both sexes across all three measured waves. The age-specific contributions follow the same patterns across the fall and winter waves, generally representing an expansion of mortality across all ages rather than a shift in mortality rates to an older age (mode) or compression around that mean (standard deviation). The youngest ages contribute the most to each change in disparity, though their role differs by wave. In the spring wave, because the mortality at the youngest ages is relatively unaffected by the outbreak, the youngest ages actually contribute to a decrease in lifespan inequality. This makes sense, as outside of the Spanish flu, infant and

Table 6.4 Total contributions towards change in $e^{\dagger}$ by wave, sex, and cause

Sex	Wave	Cause	Total contribution to decrease in $e_{\dagger}$
Female	Herald	All Causes	-1.12
Female	Herald	Other	-0.57
Female	Herald	Respiratory	-0.52
Female	Herald	Tuberculosis	-0.03
Female	Fall	All Causes	-14.59
Female	Fall	Other	-6.85
Female	Fall	Respiratory	-6.33
Female	Fall	Tuberculosis	-1.41
Female	Winter	All Causes	-8.16
Female	Winter	Other	-5.55
Female	Winter	Respiratory	-2.65
Female	Winter	Tuberculosis	0.04
Male	Herald	All Causes	-3.67
Male	Herald	Other	-3.08
Male	Herald	Respiratory	-0.57
Male	Herald	Tuberculosis	-0.02
Male	Fall	All Causes	-14.81
Male	Fall	Other	-7.08
Male	Fall	Respiratory	-6.34
Male	Fall	Tuberculosis	-1.39
Male	Winter	All Causes	-9.08
Male	Winter	Other	-6.39
Male	Winter	Respiratory	-2.77
Male	Winter	Tuberculosis	0.08

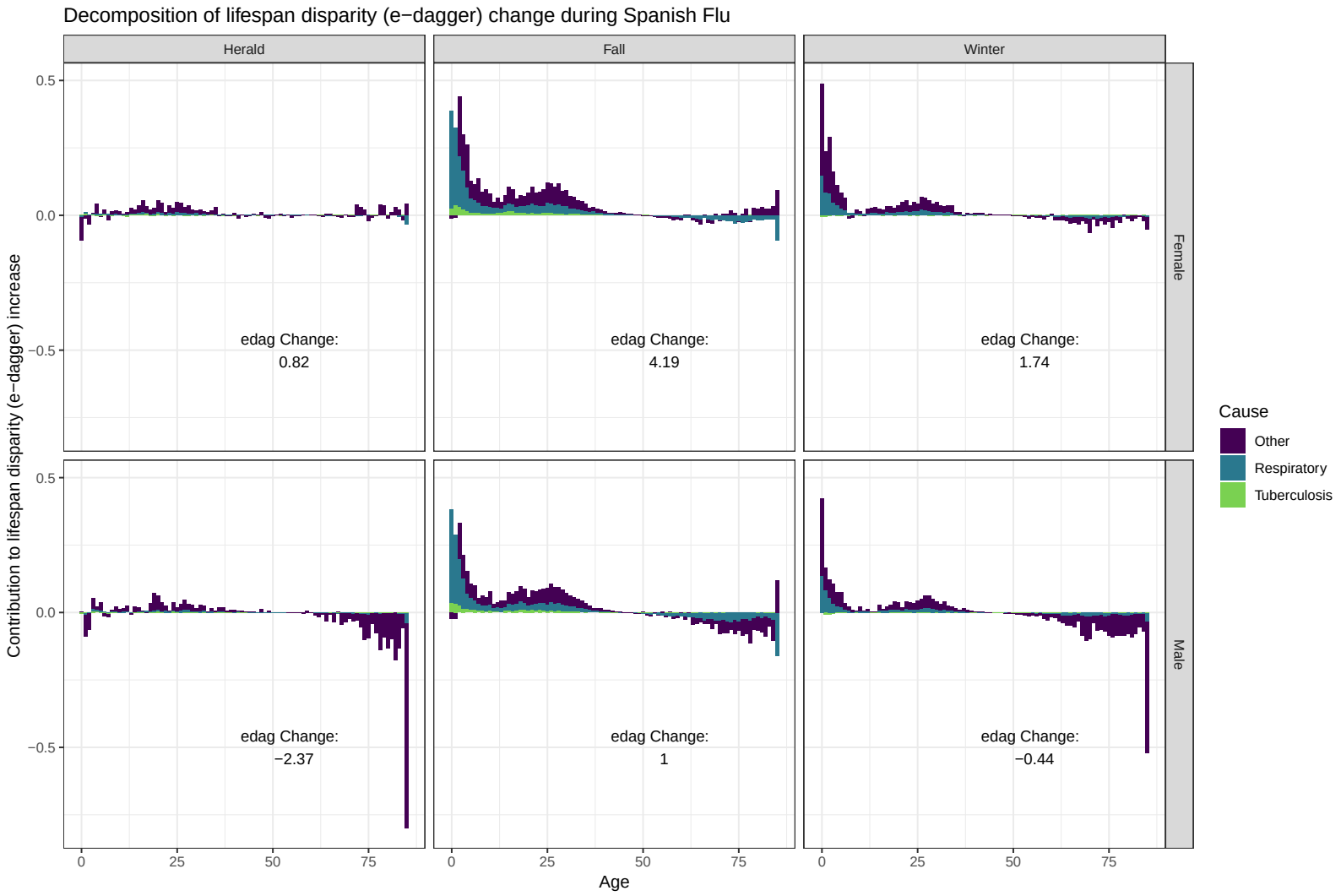

Fig. 6.11 Decomposition of Life Span Equality Changes by Sex

child mortality rates were on the decline. However, in the fall and winter waves, in which infant and children also died at high rates, their increased mortality contributes to a larger amount of lifespan inequality. Though the analysis is not completed here, it is likely the age-specific contributions to infant and child deaths to e dagger were even greater in the echo wave, as they died in excess at the highest rates and therefore contributed most to the decrease in total and potential number of person years lived.

Similar to the decomposition of life expectancy, a peak in contribution to changes in lifespan disparity can be found around young adult ages, mimicking the local mode of mortality during the waves. In the fall and (to a smaller extent) winter waves, the contributions to increased lifespan disparity follows a positive slope up to the local mode, followed by a steeper descent of smaller age-specific contributions towards the life expectancy at birth. The age-specific contributions on the left hand side of the mode are prevalent over more age groups, as the value of l_x and T_x at these age groups is higher, and therefore more person years are lived, despite increased mortality. Following the peak, the quicker decrease in contributions can be explained due to the lower value of l_x, meaning that fewer individuals in the death table contribute to the measure of life span inequality.

With regard to cause of death, a similar pattern to that of the change in e_0 exists. A visible amount of the contribution to lifespan equality changes due to tuberculosis deaths in the second fall wave, but this disappears in the winter wave. While levels of excess mortality still exist in older age groups, the contribution of these ages to lifespan equality actually decreases during the wave. This is a mechanism of the modal age at death during the period before the outbreaks. The trough represents the old age mortality hump prior to each influenza wave. Because the heightened mortality rates imply that more deaths occur around the ages of highest adult mortality pre-outbreak, these ages actually lower the level of lifespan disparity, meaning that in the death table, a higher density of deaths happen in the same age groups. Especially noticeable in the herald, but also slightly visible in the fall waves, an oft observed pattern in lifespan disparity at the oldest ages can be found. Those living to the highest age contribute to an *increase* in lifespan inequality because they live many more years than the highest age at death.

6.4 Concluding remarks

With regard to age-specific mortality, the results presented here mirror a “textbook" example of Spanish flu. The outbreaks began with a small, mild herald wave that did not reach the entire country. Then a large fall wave affected the entire population, but a mortality hump existed among those between 25 and 30. This pattern continued in a winter wave,

beginning in January 1919 and lasting into the Spring of that same year. As seen in the weekly time series of mortality rates, the herald wave in July of 1918 produced small levels of excess mortality, but the excess mortality hump in young adults is largely consistent with the subsequent fall and winter waves, if a bit smaller, suggesting this was in fact a similar strain of virus. The echo wave in 1921 was large, but primarily reflected a return to seasonal age-specific patterns.

With regards to the standardized mortality ratio, the results show much higher relative deaths in the fall wave than in others. However, these large ratios are largely due to a lower seasonal baseline in the fall. In the echo wave, there is a local peak among the 0-5 year olds, although there remains slightly elevated relative excess amount the middle ages, suggesting both that the lethal virus was still circulating in the population, and that among this population, there were still many at risk.

Perhaps the most interesting results lie in the creation and decomposition of death tables according to the mortality rates during the waves. Understanding that the herald wave was both small in size and did not reach all parts of the country, it makes sense that the overall drop in life expectancy was lower that the other waves. Similarly, given the lower fall mortality (relative to winter), it also makes sense that this wave shows the largest drop in period life expectancy. However, the cause-specific decomposition provides interesting results that highlight the frailty of those in the population suffering and dying from tuberculosis.

Those with already weakened immune systems, especially those with tuberculosis, were likely to die from the virus at higher rates, and additional evidence finds that the influenza epidemic likely contributed to a large decrease in tuberculosis rates in the years succeeding the epidemic, likely due to the harvesting effect of the flu [14, 195, 305]. Lab research has since confirmed that tuberculosis and influenza virus infection hinders survival in mice [223]. The results of the decomposition presented in this chapter provide further support to this theory (see table 6.4). Here, the decomposition by cause of death shows that even between the second and third waves of flu, the contribution of tuberculosis deaths towards the change in life expectancy disappears.

In the herald wave, the total mortality impact was minimal, and in fact, was not universally present in the country. Even still, there is evidence that deaths to tuberculosis contributed to the small decrease in life expectancy and increase in lifespan disparity. However, the second two powerful and wide-reaching waves spread throughout the country and provided similar patterns of age-specific excess mortality. Given their comparable features, the differences by cause of death are stark–there is a clear harvesting effect in the fall wave of those with tuberculosis. In the fall wave, life expectancy is reduced by roughly 1.4 years for both men and women due only to an increase in tuberculosis deaths.

However, in the winter wave, tuberculosis has no negative contribution to the change in life expectancy between the pre-influenza period of the wave and the wave itself. Moreover, the net contribution of tuberculosis towards the change in lifetable measures between the same period of the year before and during the winter wave is *positive*, adding 0.04 and 0.08 years of life expectancy to women and men respectively. This implies the presence of a harvesting effect, wherein due to the increase in mortality of those infected with tuberculosis in the fall wave, the sum of the age-specific mortality rates for tuberculosis during the winter wave was higher than in the same period of the year before the winter wave.

One thing should be noted with respect to the findings. During the pre-influenza period in the winter wave, there *were* higher rates of tuberculosis deaths than in the pre-influenza period associated with the fall wave. Nonetheless, the disappearance of the tuberculosis contribution can also be compared relative to the contribution to the change in life expectancy and lifespan equality for non-tuberculosis respiratory-related causes. In the pre-influenza period of the fall, the overall death rates of respiratory causes were *also* lower than in the pre-wave winter period. However, they continue to contribute a high percentage to the decrease in life expectancy in both waves, thereby highlighting the stark drop in tuberculosis deaths between the fall and winter waves. This reinforces the idea that those with tuberculosis were not only more susceptible to death from the contraction of influenza, but those that contracted and succumbed to the virus did so overwhelmingly and solely in the large second fall wave.

Today, though incidence continues to decline, tuberculosis remains one of the highest causes of death in the world, and the World Heath Organization estimates more than 10 million individuals contracted and 1.6 million people died from some form of tuberculosis in 2017 [10]. Moreover, while the disease is curable, several strains of the bacterium are drug resistant, making their treatment more difficult [202]. Some regions of the world with high tuberculosis rates are also underdeveloped areas where a large influenza outbreak would likely have a higher mortality impact due to a lack of infrastructure to treat and limit transmission of the virus [139]. Thus, even 100 years after the Spanish influenza outbreaks, the prevalence of tuberculosis, particularly in those with depleted immune systems and areas with fewer resources to combat a potential pandemic is high. Officials involved in pandemic preparedness plans must consider the heightened risk of those with tuberculosis and other diseases that limit respiratory and immune capacity in the context of influenza mortality during future outbreaks.

Finally, further work can also be completed within the context of the Digitising Scotland data to understand differences in mortality risk during the Spanish flu by pre-existing health conditions. Tuberculosis is certainly not the only disease, diagnosed or not, that limits

respiratory function. Coal mines were of large economic importance in the central belt of Scotland in the early 20^{th} century, providing many jobs to men who worked within mines and coal pits [48]. Coal miners are known to suffer from several other chronic lung diseases due to extended exposure to coal dust [28, 151]. Future work can examine and highlight differences in mortality based on occupations, especially in the context of compromised lung capacity. As in the case of tuberculosis-specific mortality, findings can alert public health, medical, and policy officials to additional sub-populations that may have heightened mortality risk in future pandemic outbreaks, when strains are more virulent. Continued research into past influenza pandemics, especially the hard-hitting Spanish flu, will help researchers understand how to prepare for and minimize excess mortality in future outbreaks.