Norwegian University
of Life Sciences

Master's Thesis 2021 60 ECTS
Faculty of Chemistry, Biotechnology and Food Science

Does circulation of influenza virus have an impact on birth outcomes? A time series analysis

Elin Wølner Bjørnson
Applied Statistics

Acknowledgements

Working with this thesis has been exciting, interesting, educational, and challenging. I would like to thank my two supervisors Ida Laake and Kathrine Frey Frøslie for their support, good guidance and feedback, and for the opportunity to work on this very interesting project. Thank you Daniel, for all the patience and support, and for keeping my mood up. Six years of studying in Ås is coming to an end, and I am very grateful for all the good memories and friends I have made, and for valuable experiences and education.

Elin Wølner Bjørnson

Summary

Introduction Previous studies on influenza and adverse birth outcomes have produced conflicting results. Individual-level studies can be biased since only the women who seek medical care for their symptoms are registered with influenza, and women with a mild influenza infection are less likely to be classified as influenza cases. I used time series analysis to study the association between influenza circulation and the risk of preterm birth and early pregnancy loss in Norway.

Materials and methods Influenza surveillance data was collected from the Norwegian Institute of Public Health (NIPH), and birth data from the Medical Birth Registry of Norway (MBRN). Poisson regression with cubic splines and regression with ARIMA errors was used to model preterm birth rates and estimate the effects of influenza circulation on these rates from 2006 to 2018. Two approaches were used to model the expected birth rate from 2001 to 2019, and to estimate any excess or deficits in births.

Results Some association was found between influenza circulation and preterm (<37 weeks) and very preterm (<32 weeks) birth, but it was not robust with different models, data sources for influenza circulation, and in sensitivity analysis. A small birth deficit was found in August of 2004, and an excess of births in late 2009.

Discussion The increased risk that was found for preterm birth with increasing influenza circulation in early pregnancy was not confirmed by all the models. But time series analysis may not capture small effects, and perhaps there could be a small effect of mild influenza on the risk of preterm birth in reality. The birth deficit found in 2004 may be a result of early pregnancy loss due to influenza, but was not found in any other influenza seasons. The excess of births in 2009 could correspond to more preterm births during the 2009 pandemic, but it is also likely that other causes are the reason for this excess.

Conclusion Little to no association was found with mild influenza illness and the risk of preterm birth or first-trimester miscarriage.

Sammendrag

Introduksjon Effekten av influensa på fødselsutfall er uklar, og studier av dette har funnet motstridende resultater. Studier på individnivå kan feilklassifisere hvilke kvinner som har hatt influensa, fordi kun de som har oppsøkt lege med influensasymptomer blir registrert med influensa. Dermed vil de som opplever få eller ingen symptomer ha lavere sjanse for å bli klassifisert som influensatilfeller. I denne masteroppgaven brukte jeg tidsrekkeanalyse for å utforske sammenhengen mellom influensasirkulasjon i samfunnet, og risiko for prematur fødsel og tidlig spontanabort i Norge.

Materialer og metoder Data om influensaovervåking ble hentet fra Folkehelseinstituttet (FHI), og fødselsdata fra Medisinsk Fødselsregister (MFR). Poisson-regresjon med tredjegrads splines og regresjon med ARIMA-feilledd ble brukt for å modellere premature fødselsrater og estimere effekten av influensasirkulasjon på disse ratene fra 2006 til 2018. To metoder ble brukt for å modellere forventede fødselsrater, og studere om avvikene samsvarer med tidlige spontanaborter i en influensaperiode fra 2001 til 2019.

Resultater Noe økt risiko ble funnet for prematur (<37 uker) og veldig prematur (<32 uker) fødsel ved økt eksponering av influensa. Men det gjaldt ikke for alle modellene eller med forskjellige datakilder for virusovervåking. I august i 2004 ble det funnet noe færre fødsler enn forventet, og flere enn forventet i slutten av 2009.

Diskusjon Den økte risikoen som ble funnet for premature fødsler ved økende influensasirkulasjon i samfunnet ble ikke bekreftet i sensitivitetsanalyser. Dersom influensa i virkeligheten har en positiv, men svak effekt på risiko for premature fødsler, kan det hende at denne effekten ikke blir fanget opp i en tidsrekkeanalyse. At det ble født færre barn enn forventet i 2004 samsvarer med tidlige spontanaborter i en influensasesong, men dette funnet ble ikke bekreftet i andre sesonger. Det ble funnet flere fødsler enn forventet i 2009, og dette kan samsvare med at det var flere premature fødsler under svineinfluensaen, men andre årsaker til dette fødselsoverskuddet er like trolig.

Konklusjon Lite til ingen sammenheng ble funnet mellom mild influensasykdom og risiko for prematur fødsel eller tidlig spontanabort.

Table of Contents

List of Figures

List of Tables

List of Acronyms

WHO	World Health Organization
ILI	Influenza like illness
NIPH	Norwegian Institute of Public Health
MBRN	Medical Birth Registry of Norway
NorSySS	Norwegian Syndromic Surveillance System
GP	General practitioner
LMP	Last menstrual period
AIC	Akaike's information criterion
PACF	Partial autocorrelation function
RR	Relative risk
ARIMA	Auto-Regressive Integrated Moving Average
CMA	Centered moving average

1. Introduction

1.1 Motivation

Pregnant women may be more susceptible to severe infectious disease [1], and knowledge about how this affects the pregnancy can help prevent adverse birth outcomes. Infectious diseases can have a large impact, both health-wise [2] and economically [3]. Influenza will for many people manifest in a harmless disease, but for certain risk groups, like patients with chronic diseases, elderly, or any person with an increased susceptibility for infections, it can cause severe illness or death [2]. Pregnancy generally makes a woman more vulnerable, which can put her and her fetus at risk of adverse outcomes following an infection [1]. Influenza is a common disease, Hayward et al. [4] found that 18% of the British population was infected on average each winter, suggesting that many pregnant women go through an influenza infection during pregnancy. Knowledge about how influenza infections affect pregnancy and the fetus is important to prevent adverse outcomes.

1.2 Influenza virus

Influenza virus is an RNA virus, typically divided into types A, B, C, and D [2]. Types A and B are most common in humans and cause seasonal epidemics. Influenza type A is classified after the proteins neuraminidase (NA) and hemagglutinin (HA) that lie on the surface of the virus. Type A is the only type that is known to have caused pandemics. The Spanish flu in 1918 and the swine flu in 2009 were both caused by influenza virus A(H1N1). Influenza is characterized by symptoms like fever, cough, sore throat, muscle pains, and headaches. For most people, influenza symptoms pass after a week, but some

experience a more severe infection. The World Health Organization (WHO) estimates that influenza causes between 290 000 and 650 000 deaths annually. Influenza usually occurs every year in the winter season, but this pattern varies between different regions.

In Norway, influenza virus mainly circulates during the winter season. Hauge et al. [5] found that an average of 1.7% of the Norwegian population was diagnosed with influenza-like illness (ILI) in primary care every season between 2008 and 2017. The number was higher (3.9%) in 2009/2010 due to the pandemic. The same study also found that the mean number of hospitalized patients each season was 2470 with a mean age of 56 years, and an average of 3% of these died in hospital. Many influenza-related deaths are not registered with influenza as the cause of death. Gran et al. [6] estimated excess deaths due to influenza, using information on influenza activity and the reproduction number R. They found that 910 deaths per season were due to influenza. The 2009 pandemic had a higher count of younger age groups hospitalized where the average age of influenza-related deaths was 47 years [5]. The medical burden of both seasonal and pandemic influenza can be high, which also leads to large economic costs. An estimation of the direct costs of influenza in Norway was made to 22 million dollars a year, and the number of workdays lost in a year was predicted to 793 000 days [3]. During a pandemic, it would be even more. Another potential cost of influenza is the effects on pregnancy and fetuses.

1.3 Influenza in pregnancy

Several physiological changes occur in the woman during pregnancy, and many infections may have a more severe course of disease in pregnant women [1]. Complications can also affect the fetus, and in the worst case lead to miscarriage or permanent injury of the mother or child. One possible complication is preterm birth, that the child is born too early. It can be a challenging start in the life of a baby, and the survival rate is lower for young children after a preterm delivery [7]. It has also been found that preterm birth can have consequences later in life, with increased risks of death by cardiovascular disease, diabetes, and chronic lung disease [8]. Moster et al. [9] also found that shorter gestational age was associated with medical and social disabilities. Because of these possible consequences, it is important to know what can increase the risk of preterm

birth.

Influenza infection during pregnancy may lead to adverse outcomes. In a review of studies on maternal and fetal outcomes following influenza infection, several studies found an increased risk of being hospitalized among pregnant women relative to the rest of the population [10]. Among hospitalized pregnant women, there was an increased risk of preterm birth. In outpatients, i.e. those who were not admitted to hospitals, influenza was not found to increase the risk of preterm birth. Fell et al. [11]'s review of maternal influenza and birth outcomes found several studies that reported an increased risk of preterm birth following severe maternal illness caused by the 2009 pandemic. However, the evidence on mild to moderate maternal influenza illness and preterm birth was limited. Richards et al. [12] found an increased risk of preterm births among unvaccinated women during the 2009 pandemic, suggesting that pandemic influenza increases the risk of preterm birth. Gunnes et al. [13] and Håberg et al. [14] found an association between maternal influenza during the 2009 pandemic and an increased risk of fetal death. Some increased risk of adverse birth outcomes has been found with severe seasonal influenza or pandemic influenza, but less is known about how mild symptoms affect birth outcomes.

A few studies assessed the association between influenza and adverse pregnancy outcomes using an ecological design, a study at the population level. Fell et al. [15] found no association with influenza and preterm birth, stillbirth, and perinatal death using a time series analysis. Rasmussen et al. [16] also used an ecological design and found no association with spontaneous abortions or stillbirths. Bloom-Feshbach et al. [17] investigated the 1918 pandemic and found a depression of birth rates following the pandemic, followed by a sudden increase. Studies at the population level can be useful at measuring the effects of mild influenza disease on adverse pregnancy outcomes.

1.4 Measuring influenza

Using individual-level data when studying the association of influenza and adverse birth outcomes can lead to an information bias, i.e. errors in collecting information about exposure or outcome variables [18]. In general, only women who were diagnosed with

ILI after seeking medical care are included as influenza cases. Thus, the women who are classified with ILI are more likely to be severe influenza cases, because they have to seek medical care for it, while women with mild influenza illness are less likely to be included among influenza cases. Accurate numbers on influenza cases can be difficult to obtain, because many who have an influenza infection experience mild symptoms or can be asymptomatic [4], showing no symptoms at all. They will then not be officially diagnosed with it, and are not counted as influenza cases even though they might have gone through a mild or asymptomatic infection. And in the case of pregnant women, those who are more likely to have adverse pregnancy outcomes, because of other risk factors, may be more likely to seek medical care. This can lead to a bias in individual-level studies since more women who are at risk of adverse birth outcomes are more likely to be classified as influenza cases than those who are less at risk. This can make it difficult to look at the consequences of mild influenza infections in the whole population of pregnant women, and not just those who are more at risk of severe illness and adverse pregnancy outcomes. An approach that can be used to include the whole population is an ecological design, using time series analysis.

Time series are observations that are collected at equal time intervals. A single time series can be used to study patterns or predict future observations, or multiple time series can be used to study associations. The association between influenza virus and birth outcomes can be studied using time series analysis. The circulation of influenza virus in the general population can be used as a measure of influenza exposure for pregnant women, and the association with incidences of adverse pregnancy outcomes can be studied.

In Norway, both influenza circulation and births are well documented. The Norwegian Institute of Public Health (NIPH) monitors influenza using influenza surveillance systems. All births in Norway must be reported to the Medical Birth Registry of Norway (MBRN). MBRN's records can be used by collecting the incidences of preterm birth and study the association with influenza circulation data from NIPH. MBRN does not include pregnancies that last less than 12 weeks, so first trimester miscarriages are not reported [19].

Early miscarriages may not be caught by the system. Miscarriage in the first trimester is

quite common [20] and may not be noticed. But if influenza infection increases the risk of miscarriage, one would expect to see a lower birth rate 6-9 months after a period of influenza. Exploring the birth rates subsequent to influenza pandemics have been done for the 2009 pandemic [15] and the 1918 pandemic [17]. It is possible to do using birth records from MBRN, population data from Statistics Norway, and virus surveillance data from NIPH.

1.5 Aims of the thesis

The aims of this thesis are to use time series analysis to explore:
(i) Whether influenza circulation is associated with an increased risk of preterm birth
(ii) Whether influenza circulation is associated with subsequent lower birth rates

2. Material and Methods

2.1 Collection of data

2.1.1 Influenza virus exposure

NIPH monitors influenza virus circulation in Norway to obtain information about when outbreaks occur, what types of influenza virus is circulating, and the level of immunization in the population [21]. Health care utilization due to influenza, hospitalizations and deaths, and vaccination coverage are also monitored. The influenza monitoring from Norway is also used by WHO to develop influenza vaccines. Two different virus surveillance systems, the Norwegian Syndromic Surveillance System (NorSySS) and laboratory-confirmed influenza surveillance, are used to monitor the influenza circulation in the general population.

Norwegian Syndromic Surveillance System (NorSySS)

NorSySS is monitoring infectious diseases and is run by NIPH. [22]. The purpose of NorSySS is to detect early signs of outbreaks of infectious diseases. The general practitioners (GPs) and primary care facilities report all diagnoses set on patients through the KUHR system which manages reimbursements by the government. 80 different diagnosis codes that are characterized by symptoms of infectious diseases are sent anonymized from KUHR to NorSySS. Type of consultation (physical, over phone, or digital) and information about age, gender, and county of the patient is also included. The diagnosis is most often based on clinical symptoms, and not laboratory-confirmed tests.

I received data from NorSySS with the code R80 in the International Classification of Primary Care-2. R80 consultations are based on clinical symptoms that classify

influenza illness, such as cough, fever, and muscle aches [21]. The weekly number of consultations with an ILI diagnosis and the total number of consultations was received from week 1 in 2006 to week 41 in 2020, on the age groups 20-29 years, 30-39 years, and 40-49 years, the years when women are most likely to have a child.

Influenza virus surveillance

I also received data from NIPH's virus surveillance, which uses laboratory-confirmed influenza cases [21]. They have a network of around 70 GPs who contribute by sending in tests from patients with influenza-like symptoms to NIPH's influenza laboratory. Local laboratories around the country also send weekly reports and influenza specimens to NIPH. Specimens are sent throughout the whole year, but the surveillance is intensified from week 40 to week 20 of the following year. I received data containing the weekly number of tests and number of confirmed influenza cases from week 1 in the year 2000 to week 35 in 2020. For some of the years, information on the type of influenza was also provided.

2.1.2 Preterm births and birth rates

Medical Birth Registry of Norway (MBRN)

MBRN collects information on all births in Norway [19]. MBRN was created after the Thalidomide Catastrophe, where a sleeping drug lead to thousands of children being born with birth defects [23]. The purpose of the registry is to monitor the health of the mother and baby, collect information for research and to advise and inform. Name and social security number of all children and parents must be notified to MBRN, as well as information on maternal health during pregnancy, gestational week of birth, and complications during pregnancy or birth. All live- and stillbirths occurring after week 12 are notifiable.

I received tables from MBRN with information on the number of births and preterm births. Number of births was aggregated by week of birth from week 1 in 2001 to the last week of 2019. Number of births and preterm births were provided, aggregated by week of conception from week 1 in 2001 to the last week in 2018. I also received number of births aggregated by month of birth from 2001 to 2019. Both singleton and multiple

births were included.

Population count

To find birth rates, a population count is needed. This is available at Statistics Norway [24]. Only the population of women of fertile age is required, so the yearly count of women aged 20-44 was collected for the years 2001-2019.

2.1.3 Data processing

Some modifications were needed after receiving the data. The data from MBRN was accumulated with weekly counts. In the raw data, the last week of the year, i.e. the week starting with the last Monday of the year, was split into that week and the first week of the following year. All years where the last week was split into two were handled by summing the two parts and placing them in the last week of the year. Birth rates and preterm rates were required for analysis. Weekly and monthly birth rates were calculated with number of births as the numerator and the yearly population count of women of fertile age as the denominator. Preterm birth was defined as births where the gestational age was less than 37 weeks, and very preterm birth was defined as births where gestational age was less than 32 weeks. Weekly preterm and very preterm conception rates were calculated with the number of weekly conceptions as the denominator. The reason conception date is used is that the population at risk for preterm birth is all the fetuses that are conceived at the same time, thereby being exposed to the same seasonal risk factors [25]. The number of conceptions is also seasonal, changing the number of fetuses-at-risk, which is the denominator. The conception date was defined as 14 days after the last menstrual period (LMP).

The virus surveillance data was received in several excel files, for different time periods. All sheets were imported, variables were given the same names, and they were bound together. NIPH's virus surveillance is mainly active between week 40 and week 20 the following year. Some data was recorded outside of this period, but the number of tests was low, sometimes resulting in a very high percentage of positive tests. For this reason, the number of influenza cases was set to 0 from week 21 to 39. The percentage of positive tests among all tests was calculated in a new variable. The NorSySS data was divided

into three age groups, and three types of consultations. They were summed and the percentage of primary care consultations with an ILI diagnosis was calculated in a new variable.

For all data sets, week numbers were set according to the ISO week standard. I also added a variable for the date, setting the date as the Monday of the week. Having a variable with a date format simplifies the analysis and presentation of time series data in statistical analysis. When all variables were finished, all tables were merged by the date variable.

2.2 Statistical analysis

Time series analysis

When dealing with time series, some function of time must be included in the model. Time series behave in different ways than other types of data and one thing that characterizes time series is the presence of autocorrelation. This means that observations in the time series are correlated with themselves. For example in an influenza season, the number of influenza cases registered one day is correlated with the number of cases the day before. Some time series are seasonal, meaning that they are autocorrelated within a time period. For the influenza example, the number of influenza cases one day is quite similar to the number of cases on the same date one year before, because the influenza season happens around the same time each year. Both autocorrelation and seasonality are examples of non-stationary behavior. For a time series to be stationary, the series' observations must not be dependent on the time it is observed, and there should be no trend or seasonality in the series [26]. Many statistical models are based on the assumption of independent residuals, which is violated by non-stationary time series data. Because of these characteristics, time series analysis is different than other types of statistical analysis.

Association of influenza circulation and preterm births (Aim i)

Poisson regression with cubic splines

One possibility when using time series data is including splines in a regression model. Splines are piecewise polynomials that are fitted to shorter regions on the x-axis. To make a smooth transition between the different polynomials, three constraints are set: the curve and its first and second derivatives must be continuous [27]. The point where they change are called knots, and the knots are often separated by equal intervals of x. Splines are flexible, and may provide a good fit to both short- and long-time patterns in a time series. In other words, they can capture seasonality, and the seasons do not have to occur at exactly the same time each year. They can be used in a regression model by making a set of variables that are functions of the time variable, in this case the weeks, and including them in the model [28].

A Poisson regression model with splines can be used in epidemiological models where one is interested in looking at the effect of an exposure, such as influenza, on an outcome, such as the occurrence of adverse pregnancy outcomes. The adverse pregnancy outcomes that I'm looking into in this analysis are the rates of preterm births and very preterm births. The exposure is the circulation of influenza virus, measured by the percentage of positive influenza tests or percentage of primary care consultations diagnosed with ILI. The exposure window, i.e. what time period of the pregnancy the effect of influenza exposure was studied on, was set to the first month of gestation. The influenza exposure variable was calculated as a 4 week forward moving average [15] as a measure of influenza levels for the first four weeks of pregnancy, from the week of conception:

$$IE_t = (i_t + i_{t+1} + i_{t+2} + i_{t+3})/4 \tag{2.1}$$

where IE_t is the influenza exposure at time t, and i_t is the percentage of positive influenza tests or percentage ILI cases in primary care at time t. i_{t+1} is the measurement of influenza circulation in the subsequent week and so on.

A Poisson regression model was built using a general linear model. The response variable was either number of preterm births (<37 weeks) or very preterm births (<32 weeks)

conceived in each week. To be able to get the rate of preterm births with the denominator as the total number of births conceived, an offset of the log of the denominator was included in the model. This is because a rate is not Poisson distributed, but the count of the occurrences is. The explanatory variables were influenza exposure (equation 2.1), either with data from NorSySS or the laboratory-confirmed virus surveillance, as well as cubic splines (equation 2.2). Cubic splines were built using the bs() function [29], using a time variable from week 1 to week n. The number of knots per year were chosen based on what gave the least autocorrelation in the residuals of the model and using Akaike's information criterion (AIC). Autocorrelation in the residuals was checked using partial autocorrelation function (PACF) plots, which displays the correlation between weeks at different lags, after removing autocorrelation from intervening weeks [29] . Ljung-Box tests were also used on the models to test for residual autocorrelation. The alternative hypothesis for the test is that the residuals are not random, i.e. displaying some remaining patterns of autocorrelation. Different models with different number of knots were also run in the sensitivity analysis.

The analyses were performed using data on virus exposure from both the laboratory-confirmed virus surveillance and NorSySS from 2006 to 2018. Sensitivity analyses were also done on the virus surveillance data from 2009 and after because the surveillance was improved in 2009 when the pandemic hit Norway. Separate models were made using data on preterm conceptions and very preterm conceptions. The models were tested for overdispersion, using a dispersion test from the package AER [30]. Confidence intervals and estimations of relative risks (RR) were calculated using the function ci.lin() from the package Epi ([31]). An RR of 1 can be interpreted as an identical risk of the outcome for the exposed and unexposed groups. An RR of more than 1 means that the exposed group is more at risk of the outcome than the unexposed group. When the exposure is a continuous variable as in this case, an RR of for example 1.1 means that the risk of preterm birth increases by 10% when the influenza exposure increases by 1 unit, i.e. 1 percentage point more positive cases of ILI or influenza.

The Poisson regression model with cubic splines can be written as:

$$ln(E(y_t)) = \beta_0 + \beta_1 IE_t + \sum_{i=2}^{k-1} \beta_i \cdot C_i(t) + ln(BC_t) \tag{2.2}$$

Where y_t is the number of preterm or very preterm births conceived at time t, and BC_t is the offset: total number of births conceived in the week. β*'s* are the regression coefficients, where β_1 is the slope of the influenza exposure variable IE_t (equation 2.1) estimating how much y change when x increases, which the RR estimate is derived from. k is the number of knots, and C_i is the cubic component of the spline that falls in the ith window.

Regression with ARIMA errors

Another tool to model time series data is the ARIMA model. It is built with three different components, the first part being the Auto-Regressive (AR) model, which is a regression model of linear combinations of previous values of the time series, using the fact that the values are autocorrelated [26]. An AR(p)-model of order p can be written as:

$$y_t = c + \phi_1 y_{t-1} + ... + \phi_p y_{t-p} + \epsilon_t \tag{2.3}$$

where y_t is the current value that is being predicted by a constant c, and the previous values of the time series where y_{t-1} is the preceding value and so on. ϕ_i are the coefficients in the regression-like model, and ϵ_t is the error term which should be similar to white noise. The order p is the number of preceding values used in the model.

The second part of the model is I for integrated, which is used if the time series is non-stationary, and needs differencing. Differencing is when previous values of the time series are subtracted from the current value. First order differencing can be considered as the change from each time unit to the next. It can be differenced up to lag d, or with a seasonal lag, for instance if one wants to adjust for seasonality. The moving average (MA) part of the model is a regression model of the linear combination of previous prediction errors. An MA(q)-model can be described as:

$$y_t = c + \epsilon_t + \theta_1 \epsilon_{t-1} + ... + \theta_q \epsilon_{t-q} \tag{2.4}$$

Where y_t is the value at time t that is being predicted by a constant c and the preceding white noise error terms ϵ_t up to order q. θ_i are the coefficients in the model.

An ARIMA model is often used in forecasting, but can also be used in a combination with a regression model [32]. I used the function auto.arima() from the forecast package [33], which fits a regression model with ARIMA errors. The function determines the best model for the errors, and the order of AR, I, and MA according to AIC, AICc, or BIC. Influenza exposure was defined with the same exposure window as in the Poisson regression model and added to the model as a regressor (equation 2.5). This model was also run with the two data sources of virus surveillance like the Poisson regression model, and in sensitivity analysis using laboratory-confirmed virus surveillance after 2009. The rate of preterm or very preterm conceptions were the response variables for the models. RR estimates and confidence intervals were also calculated for this regression model in the same manner as the Poisson regression model. The models were also tested for autocorrelation with the Ljung-Box test, and with PACF plots.

The model can be written as:

$$y_t = \beta_0 + \beta_1 IE_t + \eta_t \tag{2.5}$$

Where y_t is the preterm rate, β*'s* are the regression coefficients, IE_t is the influenza exposure variable (equation 2.1) and η_t is the errors modeled using ARIMA(p, d, q):

$$\eta_t = \phi_1 \eta_{t-1} + ... + \phi_p \eta_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + ... + \theta_q \epsilon_{t-q} \tag{2.6}$$

and

$$\epsilon_t \sim \text{iid } \mathrm{N}(0, \sigma^2)$$

This is the ARMA part of the model, where the coefficients are the same as in equations 2.3 and 2.4. If the data is non-stationary, the differenced series can be modeled as an ARMA model to get an ARIMA model.

Association of influenza circulation and birth rates (Aim ii)

To investigate whether influenza exposure increases the risk of pregnancy loss in the first trimester, potentially affecting the birth rates, two methods were used. Both methods include making a baseline birth rate, or expected birth rate, and then looking into the residuals to find when there was an excess or deficit of births.

Linear model

Two components were built from the observed monthly birth rates: a centered moving average (CMA) captures the long-term trend and a seasonal component (S) that captures seasonal variation. Two windows of moving average were used, 6 years (73 months) and 2 years (25 months). I performed additional analyses where data from a period after the pandemic was excluded from the calculation of CMA. The period that was excluded was the year 2010, 6-9 months after the pandemic peak plus a few extra surrounding months. The seasonal component was calculated by taking the average difference between each month and the moving average [17].

The centered moving average was calculated with the following equation, shown with a 6-year window:

$$MA_t = \frac{1}{73} \sum_{i=t-36}^{t+36} BR_i \tag{2.7}$$

Where MA_t is the moving average at time t, and BR_i are the observed birth rates in the surrounding months.

The seasonal component was calculated according to the following equation, shown with 19 years of data:

$$S_m = \frac{1}{19} \sum_{y=0}^{18} BR_{m+12 \cdot y} - MA_{m+12 \cdot y} \tag{2.8}$$

where S_m is the seasonal component in month m, BR is the observed monthly birth rate, and MA is the moving average. A linear model is then derived from these two variables [15]:

$$B_t = \beta_0 + \beta_1 MA_t + \beta_2 S_m + \epsilon_t \tag{2.9}$$

where B_t is the estimated birth rate at time t, β's are the coefficients, MA_t is the centered moving average at time t, S_m is the seasonal component of month m and ϵ_t is the error term. A 95% confidence interval for the residuals was calculated with the following equation: $CI = 0 \pm z_\alpha \cdot \frac{\sigma}{\sqrt{n}}$, assuming a normal distribution. The variance is calculated as the sum of the squared residuals, excluding residuals from the pandemic

period for the models where the pandemic period is not used in the moving average calculations [17]. If there are at least two consecutive months with observed values above or below the confidence interval, these are to be counted as excess or deficit births, respectively.

FluMOMO-model

A collaboration between 30 partners in 27 European countries, called the EuroMOMO network, has made an algorithm to monitor mortality. They have developed a model to monitor excess mortality due to influenza, called FluMOMO. It calculates a baseline, or expected, mortality and excess mortality due to influenza virus. Part of this model was used with some modifications, to fit the data of expected birth rates [34].

The FluMOMO model consists of a baseline mortality rate, and then uses the baseline in a model including influenza activity and extreme temperatures to predict the number of deaths attributable to influenza and extreme temperatures. I only used the baseline model, following the same line of thought as the linear model where deviations from the expected number of births are investigated [17]. The baseline model for birth rates is composed using a general linear model with a quasipoisson distribution. Two pairs of sine/cosine functions are used to detect the seasonality, where one is yearly, equation 2.10, and one is half-yearly, equation 2.11:

$$sin52(t) = sin\left(\frac{2\pi \cdot t}{365.25/7}\right) + cos\left(\frac{2\pi \cdot t}{365.25/7}\right) \tag{2.10}$$

$$sin26(t) = sin\left(\frac{4\pi \cdot t}{365.25/7}\right) + cos\left(\frac{4\pi \cdot t}{365.25/7}\right) \tag{2.11}$$

The FluMOMO model was advised to only be used for a period of maximum 5 years, as it is not designed to model long-term trends. To avoid this problem, a second-degree global polynomial was added to the model to capture the long-term trend. The baseline model can be shown as:

$$BR_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 sin52 + \beta_4 sin26 + \epsilon_t \tag{2.12}$$

Where BR_t is the birth rate, β's are the coefficients, and t is the time variable going from week 1 to week n. Sine/cosine pairs are shown in equations 2.10 and 2.11.

The FluMOMO algorithm calculates the residual variance and uses a 2/3-power correction for skewness [35]. A detailed description can be found in the appendix of the article by Nielsen et al. [36]. Using the normalized residual variance, a pointwise confidence interval is calculated. R-code can be found at the EuroMOMO website (www.euromomo.eu). Significant deviations from the baseline model are investigated and compared to the timing of the 2009 pandemic and seasonal influenza.

All analyses were done using version 4.0.2 of R [37].

3. Results

3.1 Descriptive analysis

The influenza surveillance showed a strong seasonal pattern, where the main waves are in the winter season as seen in figure 3.1. The NorSySS surveillance data had a similar pattern to the laboratory-confirmed virus surveillance. The percentages of ILI patients were lower in the NorSySS data, where the denominator was the total number of consultations in primary health care. The percentages were higher for the laboratory-confirmed surveillance, where the denominator was the number of ILI cases sent to laboratory testing. Both data sources showed a peak in 2009. Summary statistics for the time series are shown in table 3.1, where the highest number of positive influenza tests from the influenza surveillance was 43.9% in 2009.

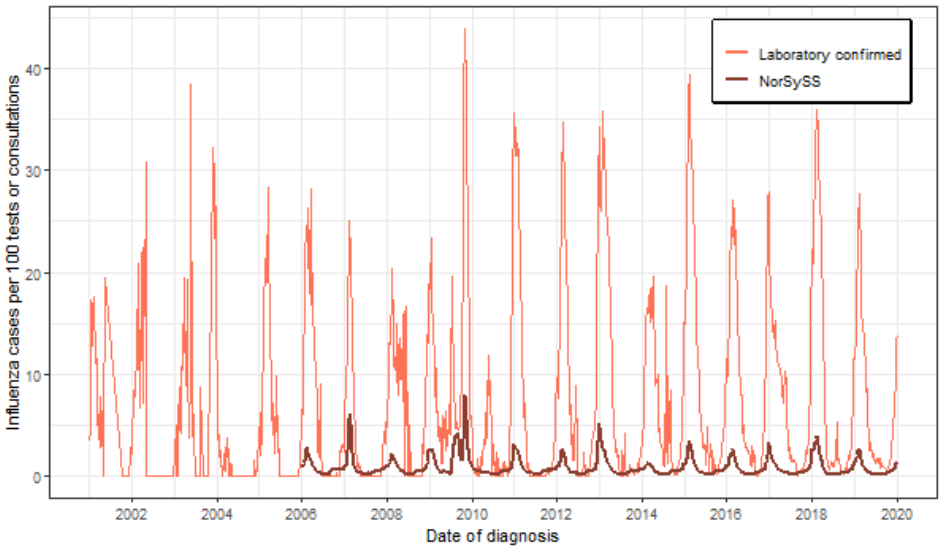


Figure 3.1: Weekly data of percentages of influenza cases from influenza surveillance with both laboratory-confirmed and primary health care consultations(NorSySS) from year 2001 to 2019.

	Min	Max	Mean	N weeks	Years
Positive influenza per 100 ILI tests	0.00	43.87	6.20	990	2001-2019
ILI per 100 consultations	0.13	7.98	0.90	729	2006-2019
Birth rate per 100 women	0.08	0.16	0.13	990	2001-2019
Preterm births per 100 conceptions	3.47	9.33	5.79	939	2001-2018
Very preterm births per 100 conceptions	0.17	1.79	0.86	939	2001-2018

Table 3.1: Summary statistics of the weekly time series used in the analyses.

Descriptive characteristics of births and conceptions are shown in table 3.2 for the years 2001 to 2019. The total number of births for the 19 years was $1,099,411$. The number of births increased from 2001 to 2009 and then decreased, while the number of preterm births showed a somewhat decreasing trend. Both the number of births and preterm births exhibited seasonality. Births peaked around July and conceptions peaked 8-9 months earlier in October, while rates of preterm conceptions peaked around February-April. Summary statistics of the weekly time series are shown in table 3.1. The percentage of conceptions being born preterm ranged from 3.47% to 9.33%, and very preterm ranged from 0.17% to 1.79% for the years 2001 to 2018.

Year	Births	Conceptions	Preterm	Very preterm
2001	55786	54204	3574	573
2002	54788	55163	3403	542
2003	55959	55873	3541	537
2004	57173	57010	3558	506
2005	56003	57158	3387	481
2006	57858	57097	3621	526
2007	57786	58949	3553	543
2008	59826	60725	3515	526
2009	62562	61842	3528	537
2010	61177	60063	3211	462
2011	59968	59499	3350	478
2012	59857	58662	3165	481
2013	58633	58474	3179	482
2014	58682	58702	3264	486
2015	59535	59384	3248	472
2016	58616	56516	3028	442
2017	56281	55070	2963	413
2018	54787	54227	2924	414
2019	54134			

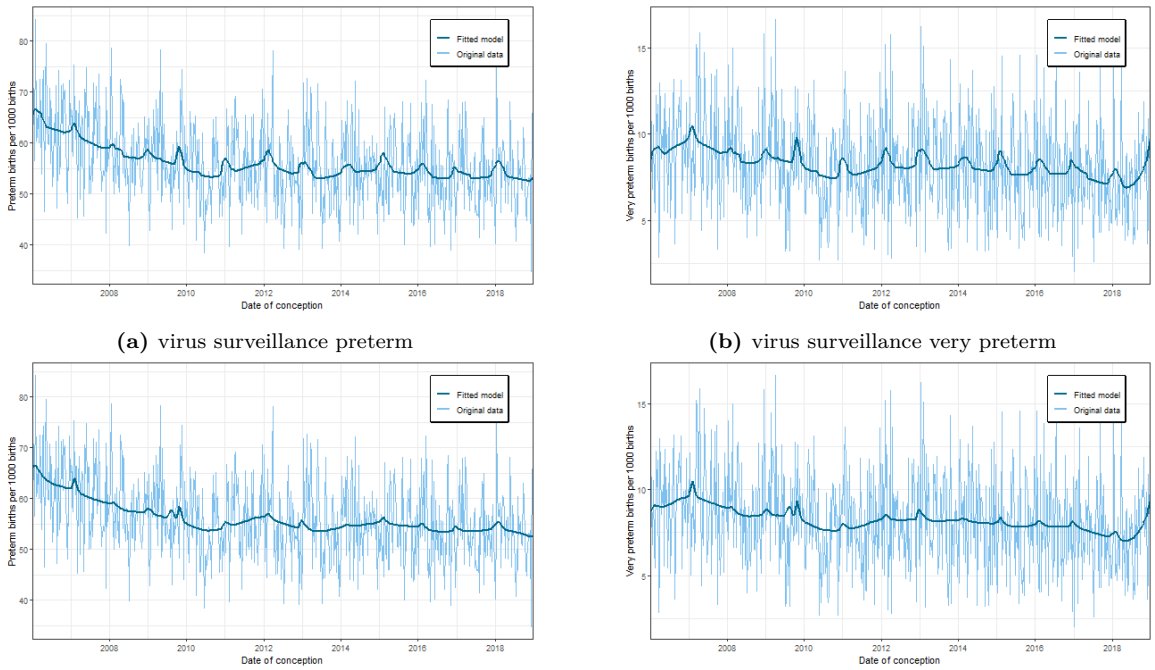
Table 3.2: Summary table of number of births, conceptions and preterm deliveries in Norway in 2001-2019.

3.2 Models of preterm births (Aim i)

Poisson regression model with cubic splines

Models were fitted using Poisson regression with cubic splines, where preterm or very preterm birth rates were the outcomes. The exposure variable of interest was influenza exposure in the first month of pregnancy. The models were fitted with 1, 2, and 3 knots per year. The models with 1 knot per year are shown in figure 3.2 together with the observed data. The models used data from 2006 to 2018, with influenza exposure data

from both primary care ILI patients (NorSySS) and laboratory-confirmed tests. Models with different numbers of knots per year were tested, and the models with the least number of knots per year gave the lowest AIC and least autocorrelation in the residuals. AIC was lowest in the models with only 1 knot, and the models with NorSySS data had a lower AIC for 3 knots than 2 knots per year, while the laboratory-confirmed data had a lower AIC for 2 knots than for 3 knots per year. Autocorrelation plots can be found in appendix A, and residual plots in appendix B. The models with 1 and 2 knots per year showed little autocorrelation outside the confidence interval, while the model with 3 knots show some more, especially for the laboratory-confirmed virus data. The models were tested for autocorrelation using the Ljung-Box test, and all of the Poisson models displayed some remaining autocorrelation in the residuals ($p < 0.05$). The residual plots show no specific patterns. The models did not display overdispersion ($p < 0.05$).



(a) virus surveillance preterm

(b) virus surveillance very preterm

(c) NorSySS preterm

(d) NorSySS very preterm

Figure 3.2: Fitted Poisson regression models with cubic splines and original data with 1 knot per year. The number of preterm or very preterm births per 1000 conceptions are shown on the y axis, and date of conception on the x axis from 2006 to 2018. **(a)** Virus surveillance data and preterm birth as the outcome. **(b)** Virus surveillance data and very preterm birth as the outcome. **(c)** NorSySS data and preterm births as the outcome. **(d)** NorSySS data and very preterm births as the outcome.

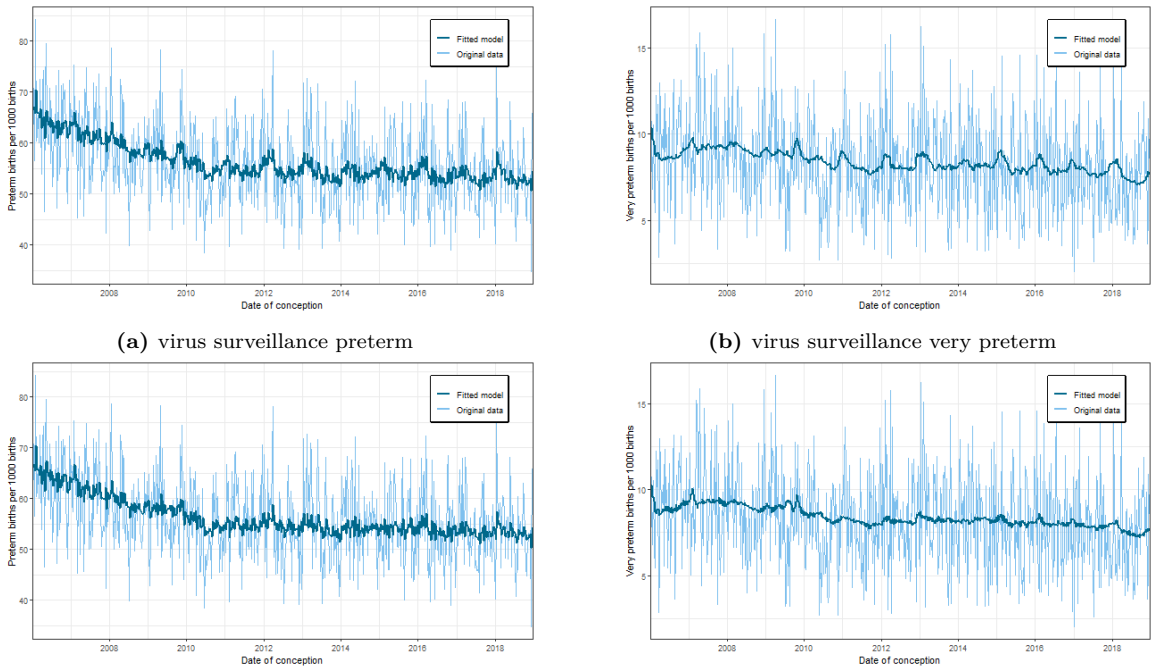
Knots	Outcome	Data Source	Estimated RR	Lower Bound	Upper bound
1 knot per year	Preterm	Virus surveillance	1.001719	1.000656	1.002783
		NorSySS	1.008736	0.998004	1.019583
	Very preterm	Virus surveillance	1.004318	1.001578	1.007066
		NorSySS	1.019863	0.992209	1.048288
2 knots per year	Preterm	Virus surveillance	1.001539	1.000297	1.002784
		NorSySS	1.004523	0.991756	1.017454
	Very preterm	Virus surveillance	1.004576	1.001351	1.007812
		NorSySS	1.017539	0.984572	1.051610
3 knots per year	Preterm	Virus surveillance	1.000196	0.998290	1.002106
		NorSySS	0.996549	0.980466	1.012896
	Very preterm	Virus surveillance	1.001545	0.996617	1.006498
		NorSySS	1.001650	0.960517	1.044546

Table 3.3: Poisson models RR estimates and 95% confidence intervals with 1, 2 and 3 knots per year.

Table 3.3 shows relative risk (RR) estimates and confidence intervals for the Poisson regression models with cubic splines. RRs showed an increased risk of preterm birth and very preterm births for influenza exposed pregnancies (RR $>$ 1), though only in the models that used laboratory-confirmed influenza data, and only with 1 or 2 knots per year. When only virus surveillance data after 2009 was used, there was also a weak association (data not shown, RR $>$ 1). This indicates a small increase in the risk of preterm birth or very preterm birth when influenza circulation increases. All models with the NorSySS data and the models with 3 knots per year had a confidence interval that included 1, i.e. identical risk for influenza exposed and unexposed groups.

Regression with ARIMA errors

Similar models were fitted using a different approach, regression with ARIMA errors. The fitted models and observed data of weekly preterm and very preterm conceptions are shown in figure 3.3. The two data sources of influenza surveillance were used, and the years studied were 2006 to 2018. The models that were fit for the virus surveillance data had ARIMA(1, 1, 1) and ARIMA(2, 1, 1) errors, for preterm and very preterm births, respectively, as the outcome. The models where NorSySS data was used had ARIMA(1, 1, 1) and ARIMA(4, 1, 2) errors for preterm and very preterm births, respectively, as the outcome. PACF plots showed little autocorrelation in the residuals(appendix A) and the residual plots showed no clear patterns(appendix B). None of the regression models with ARIMA errors showed autocorrelation with the Ljung-Box test ($p > 0.05$)



(a) virus surveillance preterm

(b) virus surveillance very preterm

(c) NorSySS preterm

(d) NorSySS very preterm

Figure 3.3: Fitted regression models with ARIMA errors and original data. Rate of preterm or very preterm births per 1000 births on y axis and date of conception on x axis from 2006 to 2018. **(a)** Virus surveillance data and preterm birth as the outcome. **(b)** Virus surveillance data and very preterm birth as the outcome. **(c)** NorSySS data and preterm births as the outcome. **(d)** NorSySS data and very preterm births as the outcome.

Risk ratios and their confidence intervals from the ARIMA models are shown in table 3.4. No increased risk of preterm or very preterm birth was found for pregnant women exposed to influenza in the first month of gestation. All models gave estimated RRs of

about 1, and all confidence intervals included the number 1. Sensitivity analysis using virus surveillance data only from 2009 gave similar results (data not shown).

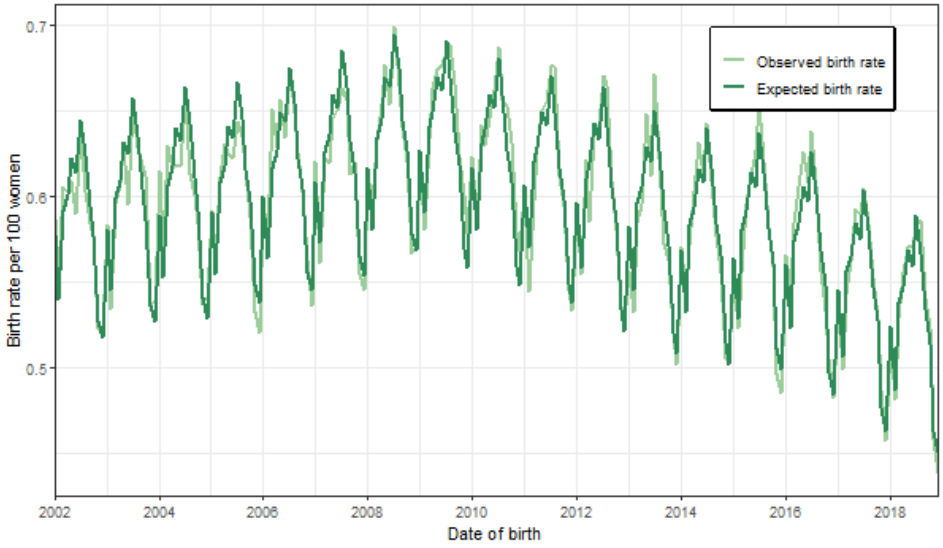
Outcome	Data Source	Estimated RR	Lower Bound	Upper bound
Preterm	Virus surveillance	1.000101	0.999999	1.0002043
	NorSySS	1.000529	0.999822	1.001237
Very preterm	Virus surveillance	1.000030	0.999952	1.000108
	NorSySS	1.000158	0.999935	1.000381

Table 3.4: ARIMA models RR estimates and 95% confidence intervals.

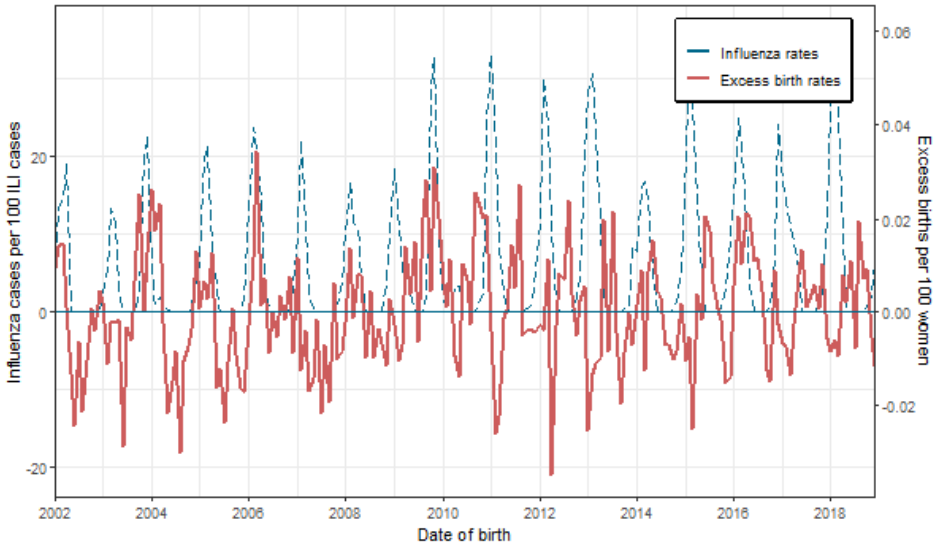
3.3 Models of birth rates (Aim ii)

Linear model

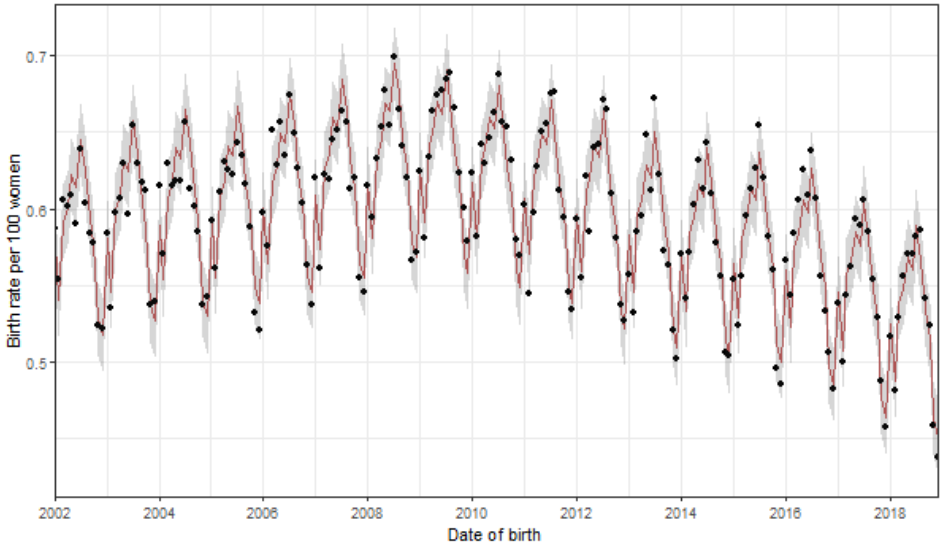
A baseline of monthly birth rates was modeled using a moving average component and a seasonal component. The models were made using two different windows of moving average, two years and six years, and with or without birth data in a period following the 2009 pandemic. Figure 3.4a shows the linear model with a 2-year moving average, where the pandemic period is excluded from the moving average. The data showed a strong seasonality with peaks in the middle of the year. The residuals from the model are shown in figure 3.4b, plotted together with the influenza activity. The residuals are to be considered as excess or deficit births in the model and are shown as the number of monthly excess/deficit births per 100 women, and they vary from around -0.03 to 0.04 births per 100 women. There were not many consecutive births with excess births outside of the 95% confidence interval(3.4c), but some continuous deficit of birth was found in early 2011. More than one model found birth deficits in August 2004, April 2012, January 2013, and March 2015. Birth excess was found in all models in the autumn of 2009. More than one model found positive excess in October 2003, January 2004, March 2006, September 2010, and August 2012.



(a) Linear model of monthly birth rates



(b) Excess births linear model

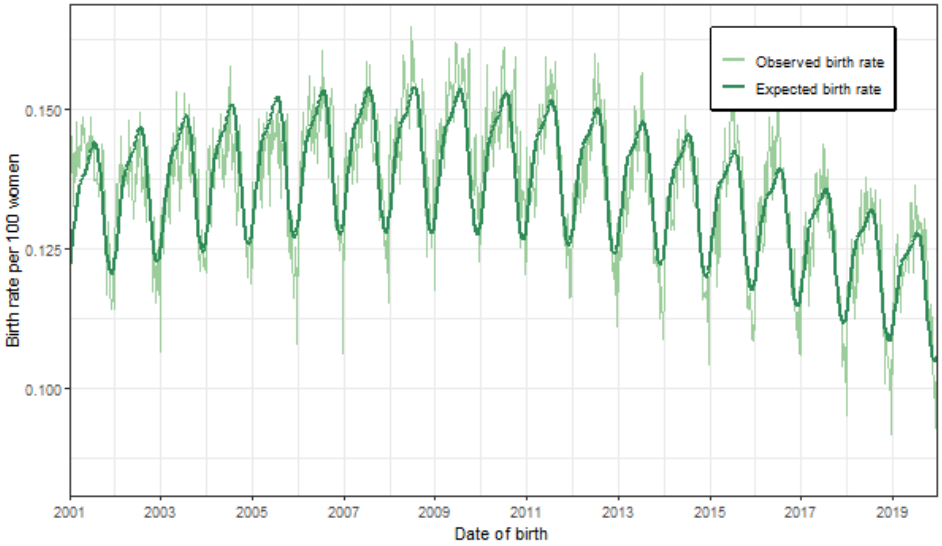


(c) Linear model with confidence band

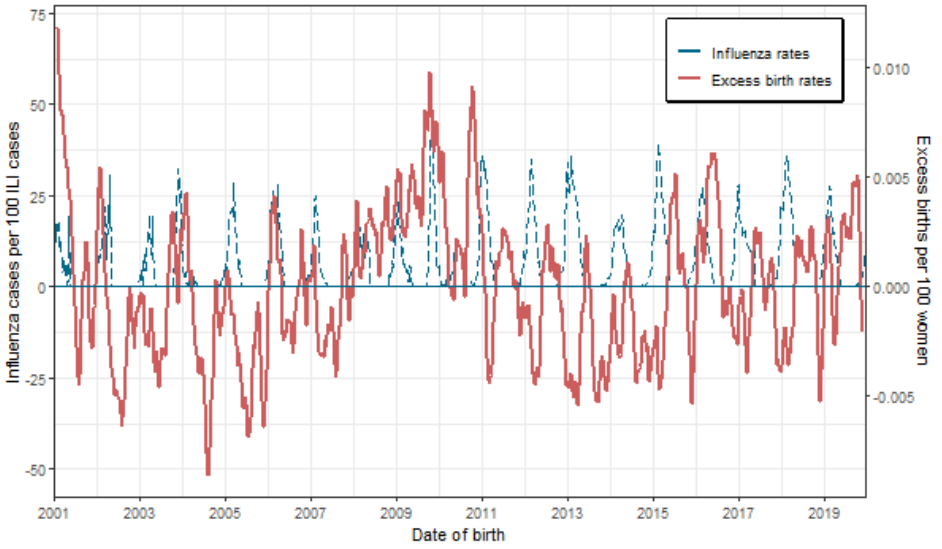
Figure 3.4: Linear model with a 2 year moving average, and excluded pandemic period . **(a)** Model and observed data. **(b)** Excess births. **(c)** Linear model with confidence band.

FluMOMO model

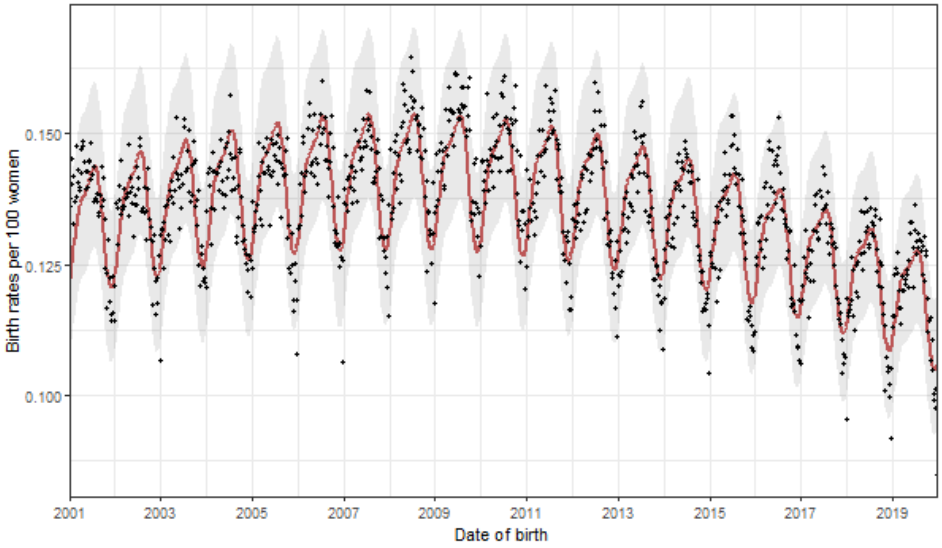
A baseline of weekly birthrates was modeled using the part of the FluMOMO algorithm that was designed to model a baseline for mortality [34]. Harmonics were used in a Poisson model, to model the expected birth rates in the years 2001 to 2019. A second-degree polynomial was also added to take into account the long-term trend of the birth rates. The fitted model is shown in figure 3.5a together with the observed weekly birth rates. Following the same line of thought as the linear model, the residuals of the models are considered excess births(3.5b), and a 95% confidence interval was calculated to see which observed values deviate from the baseline of birth rates(3.5c). The FluMOMO model found continuous birth deficits in two consecutive weeks in august of 2004. Birth excess was found in the 4 first weeks of 2001, and in the autumn of 2009 and 2010. The deficit in 2004 and excess in the autumn of 2009 and 2010 correspond to the results of the linear model.



(a) FluMOMO model of weekly birth rates



(b) Excess births FluMOMO model



(c) FluMOMO model with confidence band

Figure 3.5: FluMOMO model. **(a)** Model and observed data. **(b)** Excess births. **(c)** FluMOMO model with confidence band.

4. Discussion

In this thesis I wanted to explore whether influenza circulation was associated with an increased risk of preterm birth, or with first trimester miscarriages, resulting in lower birth rates. Using an ecological time series approach, little to no association was found between circulating influenza and the adverse birth outcomes investigated.

4.1 Preterm births (Aim i)

In the first outcome of preterm and very preterm birth, a weak association was found with influenza exposure in some of the Poisson regression models. However, it was not robust in sensitivity analysis, with different numbers of knots per year, and with different sources of data for virus exposure. The associations that were found were weak, with less than a 1% increased risk of preterm (<37 weeks) or very preterm (<32 weeks) birth when influenza circulation increased by 1 percentage point in the first month of gestation. Partial autocorrelation function (PACF) plots showed that the Poisson models with fewer knots had less autocorrelation (appendix A), but none of the models passed the Ljung-Box test, indicating that some autocorrelation is left in the models. In the regression model with ARIMA errors, no significant association between influenza exposure and preterm birth was found. This was also the case in sensitivity analysis with the laboratory-confirmed virus surveillance after 2009. Autocorrelation in the residuals of these models was not an issue (appendix A), confirmed by the Ljung-Box tests. Using two different models, Poisson regression with cubic splines and regression with ARIMA errors, can increase the robustness of the analysis. But the two models gave slightly different results, and so it cannot be inferred that the weak association found in the Poisson models is real. The regression models with ARIMA errors might be a better fit,

as it does not display signs of remaining autocorrelation.

Some issues can arise with the time series design. Confounding by other time-varying factors is a possibility if they co-vary with exposure and outcome. Conceptions exhibit seasonal variation, and this could be a potential confounder. But because conceptions have been used as the denominator, and not the number of births, the rate of preterm births are from the pool of fetuses-at-risk. There could also be individual factors, leading to differences in seasonality of conceptions among different sociodemographic groups. If one of these groups has a different risk of preterm births, and the seasonality of conceptions in the group co-vary with the influenza season, other factors could influence the results. Darrow et al. [25] found different seasonal patterns in conceptions and preterm rates among different education levels, ethnic groups, and marital status of the mother. Similar patterns may exist in Norway, potentially leading to confounding in a time series analysis.

A limitation to the study is that only one type of adverse outcome is investigated. Only livebirths that are preterm are included, and other outcomes could be studied, such as stillbirth or perinatal death. A weakness of the study design is that influenza circulation is not the same as influenza exposure. Influenza circulation in the population is just a measure of how likely the woman is to be infected, but it is not an accurate measure of actual infection. Influenza exposure was only studied in the first month of gestation. Because the data was aggregated on conception date to get fetuses-at-risk as the denominator, it was not possible to look at an exposure window at the end of gestation. Other exposure windows might give different results.

Other studies have also looked at the association between maternal influenza and the risk of preterm birth. Few of them have used time series analysis, but Fell et al. [15] also used a Poisson regression with cubic splines. They did not find an increased risk of preterm birth. Meijer et al. [10] found in their review article that the risk of preterm births among hospitalized women was 8-55% higher, but among outpatients the risk was not increased. Many of the hospitalized women had other underlying conditions. This supports the hypothesis that individual-level studies find higher associations because of information bias. The hospitalized women are more likely to have other underlying conditions, increasing both the probability of severe influenza illness and of adverse

pregnancy outcomes. Different studies found an increased risk of preterm birth following 2009 pandemic maternal illness [11, 12]. This suggests that there could be an increased risk of preterm birth following 2009 pandemic maternal illness. In an ecological study over many years, the possible effect of the 2009 pandemic on preterm birth could be too small to be detected. Perhaps the results from the Poisson regression model with virus surveillance could indicate a small association.

4.2 Birth rates (Aim ii)

For the second outcome of birth rates, little evidence was found of continuous deficits of births that could correspond to early miscarriages during an influenza season. The only birth deficits that were found following an influenza peak were in August of 2004, which fits the time slot of 6-9 months after the seasonal influenza wave in 2003 with a peak in December. Two different models were used to model a baseline birth rate. For the linear model, two different windows of moving average were used, with and without data around the 2009 pandemic. All models gave the same results; that there were no persistent deficits of births following the pandemic influenza and only one occurrence of birth deficit following a seasonal influenza wave.

The deficit found in 2004 could be caused by influenza infections leading to early miscarriages. The data from the virus surveillance shows a peak of influenza levels around December of 2003, with 32% positive tests at most. Influenza type A was dominating in this period. More information about the 2003 seasonal influenza period is difficult to obtain, and the virus surveillance was less extensive before 2009, but perhaps this was a type of influenza that affected pregnancy differently than in other seasons, or that more pregnant women were infected. But because similar effects on the birth rate were not found for any of the other influenza periods, it may be more likely that this finding is random or is caused by other exposures or trends.

One possible reason that no birth deficits were found after the 2009 pandemic, could be health awareness. There was a substantial amount of news articles about the 2009 pandemic, possibly leading to increased health awareness. Maybe it made women plan their pregnancies differently, or seek more health care when feeling ill. One consequence

of health awareness was vaccination, and many pregnant women were vaccinated during the pandemic. 54% of women who were pregnant during the 2009 pandemic were vaccinated, but the vaccine was recommended for pregnant women in the second or third trimester, and most vaccinations were set in this period of the pregnancy [14]. Many women were not protected in the first trimester, but some were vaccinated before pregnancy, and this could have balanced out a possible negative effect that pandemic influenza had on the birth rates.

An excess of births was found in the last months of 2009. One possible explanation for this could be that the 2009 pandemic, which was at its peak around this time, lead to more preterm births, giving a higher birth rate than normal. But there is also a rising trend in birth rates in general towards 2009-2010, and then a decreasing trend after that. The excess of births observed in 2009 could be a result of a long term rising trend, peaking because of a decreasing trend that could be explained by other reasons, such as economic insecurity following the financial crisis between 2007 and 2009 and fewer families choosing to have more than two children [38]. Thus, it is highly likely that the excess births found in the analysis are due to other trends in birth rates than an increased rate of preterm births.

Method selection for this analysis has been challenging. The model was supposed to capture the long-term trend, but not fit the data too well, as the chosen method was based on studying the deviations from the trend. Similar analyses done by others [17, 15, 34] were not always described in adequate detail, but I attempted to fit similar models to the available data. These types of models give no information about the effect of influenza on pregnancy loss, or about the causality for deviating birth rates, but are merely an indication of association. Perhaps a similar approach as aim (i) could be more reliable, with a regression model with the number of births as the outcome, and influenza circulation as an exposure variable. The influenza circulation variable would need to have a lag of 6-9 months in order to study the effect on the pregnancy loss in the first trimester. Possible confounders could also have been included in the model, like data on temperature which can affect the number of conceptions [25].

Bloom-Feshbach et al. [17] found a deficit of births after the 1918 pandemic. Because of this, it was of interest to see if this could apply to the 2009 pandemic or seasonal

influenza. No such association was found. Fell et al. [15] also looked at the birth rates after the 2009 pandemic and found no decrease corresponding to early pregnancy losses due to influenza illness. The difference in results here could indicate that the influenza virus in the 2009 pandemic had a different pathology, with different severity of disease or spreading abilities, or that the health care was more advanced in 2009. Differences in models could also be the case, but the linear model was inspired by the methods of Bloom-Feshbach et al. [17].

Rasmussen et al. [16] and Fell et al. [15] studied the association between influenza and fetal death, also using a time series approach. They did not find an association with influenza circulation. Even though they did not include first trimester miscarriages, this could also support the result that mild influenza disease is not associated with fetal death. Gunnes et al. [13] and Håberg et al. [14] found an increased risk of fetal death in the second and third trimester following ILI during the 2009 pandemic. Conflicting results might be because of different study designs, as they used cohort studies which could be a better measure of severe pandemic disease. Also, if the effect of influenza on fetal death was only with the 2009 pandemic, the potential effect might not be visible in the other studies with time series over many years.

4.3 Implications

Information about what causes pregnancy complications and adverse birth outcomes is important to have in order to prevent adverse outcomes. Because most people experience mild or asymptomatic disease [4], an ecological study is mainly a measure of mild influenza illness. My results indicate that there is no clear association between mild influenza illness and preterm birth or early pregnancy loss. These results are obtained using methods that avoid a bias where women with a higher risk of pregnancy complications might be more likely to be classified as an influenza case. An individual study might be more suited to study risks of adverse pregnancy outcomes following severe influenza disease and can give different results than an ecological design. The effects of influenza infection on adverse pregnancy outcomes could also be too weak to be detected at a population level. This thesis contributes to the knowledge on how mild influenza infection can affect adverse birth outcomes, and how time series analysis can be used to

study associations between seasonal exposures and health outcomes in epidemiology.

5. Conclusions

The analyses in this thesis did not support the hypothesis that influenza circulation is associated with preterm birth or early pregnancy loss. Seasonality was found in preterm birth rates, but in this thesis, it was not explained further by influenza. A small increased risk of preterm birth was found in some of the analyses, but the results were not robust with different models and data sources of influenza circulation. One instance of a deficit in birth rates that corresponded with the timing of early pregnancy loss during an influenza period (in 2003) was found. No other occurrences of this were found, leading to the conclusion that other causes of the birth deficit are more likely. Using an ecological design with influenza circulation in the general population as an indicator of influenza exposure is mainly a measure of mild influenza disease. It can not be excluded that severe influenza illness increases the risk of adverse birth outcomes.

References

[1] A. P. Kourtis, J. S. Read, and D. J. Jamieson. Pregnancy and Infection. *New England Journal of Medicine* 370 (23 June 2014): 2211–2218. DOI: 10.1056/NEJMra1213566.

[2] WHO. *Influenza (Seasonal)*. Accessed: 2021-03-16. Nov. 2018. URL: https://www.who.int/en/news-room/fact-sheets/detail/influenza-(seasonal).

[3] Y. Xue, I. S. Kristiansen, and B. F. D. Blasio. Modeling the cost of influenza: The impact of missing costs of unreported complications and sick leave. *BMC Public Health* 10 (2010). DOI: 10.1186/1471-2458-10-724.

[4] A. C. Hayward, E. B. Fragaszy, A. Bermingham, L. Wang, A. Copas, W. J. Edmunds, N. Ferguson, N. Goonetilleke, G. Harvey, J. Kovar, M. S. Lim, A. McMichael, E. R. Millett, J. S. Nguyen-Van-Tam, I. Nazareth, R. Pebody, F. Tabassum, J. M. Watson, F. B. Wurie, A. M. Johnson, and M. Zambon. Comparative community burden and severity of seasonal and pandemic influenza: Results of the Flu Watch cohort study. *The Lancet Respiratory Medicine* 2 (6 June 2014): 445–454. DOI: 10.1016/S2213-2600(14)70034-7.

[5] S. H. Hauge, I. J. Bakken, B. F. de Blasio, and S. E. Håberg. Burden of medically attended influenza in Norway 2008-2017. *Influenza and other Respiratory Viruses* 13 (3 May 2018): 240–247. DOI: 10.1111/irv.12627.

[6] J. M. Gran, B. Iversen, O. Hungnes, and O. O. Aalen. Estimating influenza-related excess mortality and reproduction numbers for seasonal influenza in Norway, 1975-2004. *Epidemiology and Infection* 138 (11 Nov. 2010): 1559–1568. DOI: 10.1017/S0950268810000671.

[7] J. Bryce, C. Boschi-Pinto, K. Shibuya, and R. E. Black. WHO estimates of the causes of death in children. *Lancet* 365 (9465 Mar. 2005): 1147–1152. DOI: 10.1016/S0140-6736(05)71877-8.

[8] K. Risnes, J. F. Bilsteen, P. Brown, A. Pulakka, A.-M. N. Andersen, S. Opdahl, E. Kajantie, and S. Sandin. Mortality Among Young Adults Born Preterm and Early Term in 4 Nordic Nations. *JAMA Network Open* 4 (1 Jan. 2021): e2032779. DOI: 10.1001/jamanetworkopen.2020.32779.

[9] D. Moster, R. T. Lie, and T. Markestad. Long-Term Medical and Social Consequences of Preterm Birth. *New England Journal of Medicine* 359 (3 July 2008): 262–273. DOI: 10.1056/nejmoa0706475.

[10] W. J. Meijer, A. G. V. Noortwijk, H. W. Bruinse, and A. M. Wensing. *Influenza virus infection in pregnancy: A review*. Aug. 2015. DOI: 10.1111/aogs.12680.

[11] D. B. Fell, D. A. Savitz, M. S. Kramer, B. D. Gessner, M. A. Katz, M. Knight, J. M. Luteijn, H. Marshall, N. Bhat, M. G. Gravett, B. Skidmore, and J. R. Ortiz. *Maternal influenza and birth outcomes: systematic review of comparative studies*. Jan. 2017. DOI: 10.1111/1471-0528.14143.

[12] J. L. Richards, C. Hansen, C. Bredfeldt, R. A. Bednarczyk, M. C. Steinhoff, D. Adjaye-Gbewonyo, K. Ault, M. Gallagher, W. Orenstein, R. L. Davis, and S. B. Omer. Neonatal outcomes after antenatal influenza immunization during the 2009 H1N1 influenza pandemic: Impact on preterm birth, birth weight, and small for gestational age birth. *Clinical Infectious Diseases* 56 (9 May 2013): 1216–1222. DOI: 10.1093/cid/cit045.

[13] N. Gunnes, H. K. Gjessing, I. J. Bakken, S. Ghaderi, J. M. Gran, O. Hungnes, P. Magnus, S. O. Samuelsen, A. Skrondal, C. Stoltenberg, L. Trogstad, A. J. Wilcox, and S. E. Håberg. Seasonal and pandemic influenza during pregnancy and risk of fetal death: A Norwegian registry-based cohort study. *European Journal of Epidemiology* 35 (4 Apr. 2020): 371–379. DOI: 10.1007/s10654-020-00600-z.

[14] S. E. Håberg, L. Trogstad, N. Gunnes, A. J. Wilcox, H. K. Gjessing, S. O. Samuelsen, A. Skrondal, I. Cappelen, A. Engeland, P. Aavitsland, S. Madsen, I. Buajordet, K. Furu, P. Nafstad, S. E. Vollset, B. Feiring, H. Nøkleby, P. Magnus, and C. Stoltenberg. Risk of Fetal Death after Pandemic Influenza Virus Infection or Vaccination. *New England Journal of Medicine* 368 (4 Jan. 2013): 333–340. DOI: 10.1056/nejmoa1207210.

[15] D. B. Fell, D. L. Buckeridge, R. W. Platt, J. S. Kaufman, O. Basso, and K. Wilson. Circulating Influenza Virus and Adverse Pregnancy Outcomes: A Time-Series Study. *American Journal of Epidemiology* 184 (3 Aug. 2016): 163–175. DOI: 10.1093/aje/kww044.

[16] I. S. Rasmussen, L. H. Mortensen, T. G. Krause, and A. M. N. Andersen. The association between seasonal influenza-like illness cases and foetal death: A time series analysis. *Epidemiology and Infection* 147 (2019). DOI: 10.1017/S0950268818003254.

[17] K. Bloom-Feshbach, L. Simonsen, C. Viboud, K. Mølbak, M. A. Miller, M. Gottfredsson, and V. Andreasen. Natality decline and miscarriages associated with the 1918 influenza pandemic: The scandinavian and United States experiences. *Journal of Infectious Diseases* 204 (8 Oct. 2011): 1157–1164. DOI: 10.1093/infdis/jir510.

[18] M. Porta. *A dictionary of epidemiology*. Oxford university press, 2014.

[19] NIPH. *Medical Birth Registry of Norway*. Accessed: 2021-03-09. Dec. 2019. URL: https://www.fhi.no/en/hn/health-registries/medical-birth-registry-of-norway/medical-birth-registry-of-norway/.

[20] A. J. Wilcox, C. R. Weinberg, J. F. O'Connor, D. D. Baird, J. P. Schlatterer, R. E. Canfield, E. G. Armstrong, and B. C. Nisula. Incidence of Early Loss of Pregnancy. *New England Journal of Medicine* 319 (4 July 1988): 189–194. DOI: 10.1056/nejm198807283190401.

[21] NIPH. *Overvåkingssystemet for influensa - FHI*. Accessed 2021-03-19. In Norwegian only. Nov. 2019. URL: https://www.fhi.no/sv/influensa/influensaovervaking/overvakingssystemet-for-influensa-/.

[22] NIPH. *NorSySS*. Accessed: 2021-03-09. Feb. 2019. URL: https://www.fhi.no/en/hn/statistics/NorSySS/about-the-norwegian-syndromic-surveillance-system/.

[23] NIPH. *Medisinsk fødselsregister – historiske røtter - FHI*. Accessed 2021-03-19. In Norwegian only. Mar. 2013. URL: https://www.fhi.no/hn/helseregistre-og-registre/mfr/medisinsk-fodselsregister--historis/.

[24] Statistics Norway. *07459: Population, by sex and one-year age groups (M) 1986 - 2021. Statbank Norway*. Accessed 2020-11-12. URL: https://www.ssb.no/en/statbank/table/07459/.

[25] L. A. Darrow, M. J. Strickland, M. Klein, L. A. Waller, W. D. Flanders, A. Correa, M. Marcus, and P. E. Tolbert. Seasonality of Birth and Implications for Temporal Stud-

ies of Preterm Birth. *Epidemiology* 20 (5 Sept. 2009): 699–706. DOI: 10.1097/EDE.0b013e3181a66e96.

[26] R. J. Hyndman and G. Athanasopoulos. *Forecasting: Principles and Practice (2nd ed)*. 2018. URL: https://otexts.com/fpp2/.

[27] G. James, D. Witten, T. Hastie, and R. Tibshirani. *An introduction to statistical learning : with applications in R*. 2013.

[28] K. Bhaskaran, A. Gasparrini, S. Hajat, L. Smeeth, and B. Armstrong. Time series regression studies in environmental epidemiology. *International Journal of Epidemiology* 42 (4 Aug. 2013): 1187–1195. DOI: 10.1093/ije/dyt092.

[29] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2020. URL: https://www.R-project.org/.

[30] C. Kleiber and A. Zeileis. *Applied Econometrics with R*. ISBN 978-0-387-77316-2. New York: Springer-Verlag, 2008. URL: https://CRAN.R-project.org/package=AER.

[31] B. Carstensen, M. Plummer, E. Laara, and M. Hills. *Epi: A Package for Statistical Analysis in Epidemiology*. R package version 2.42. 2020. URL: https://CRAN.R-project.org/package=Epi.

[32] F. Carrat and A.-J. Valleron. Influenza mortality among the elderly in France, 1980-90: how many deaths may have been avoided through vaccination? *J7ournal of Epidemiology and Community Health* 49 (1995): 419–425. DOI: 10.1136/jech.49.4.419.

[33] R. J. Hyndman and Y. Khandakar. Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software* 26 (3) (2008): 1–22. URL: https://www.jstatsoft.org/article/view/v027i03.

[34] EuroMOMO. *FluMOMO*. The FluMOMO model. Accessed: 2021-02-26. Oct. 2018. URL: https://www.euromomo.eu/how-it-works/flumomo/.

[35] C. Farrington, N. J. Andrews, A. Beale, and M. Catchpole. A statistical algorithm for the early detection of outbreaks of infectious disease. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 159 (3) (1996): 547–563.

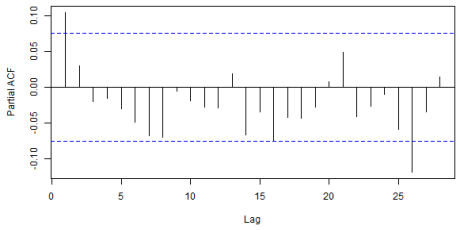
[36] J. Nielsen, T. G. Krause, and K. Mølbak. Influenza-associated mortality determined from all-cause mortality, Denmark 2010/11-2016/17: The FluMOMO model. *Influenza and other Respiratory Viruses* 12 (5 Sept. 2018): 591–604. DOI: 10.1111/irv.12564.

[37] RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, PBC. Boston, MA, 2020. URL: http://www.rstudio.com/.

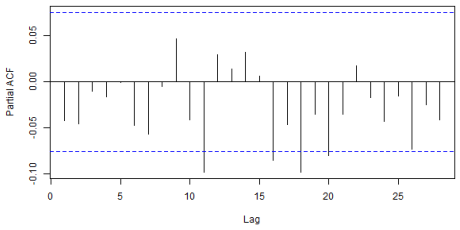
[38] L. Dommermuth and T. Lappegård. *Nedgangen i fruktbarheten fra 2010. Betydningen av utdanning, økonomisk aktivitet og økonomiske ressurser for førstefødsler og tredjefødsler*. In Norwegian only. Accessed 2021-05-09. Statistics Norway, May 2017. URL: https://www.ssb.no/befolkning/artikler-og-publikasjoner/nedgangen-i-fruktbarheten-fra-2010.

Appendix A. PACF plots preterm models

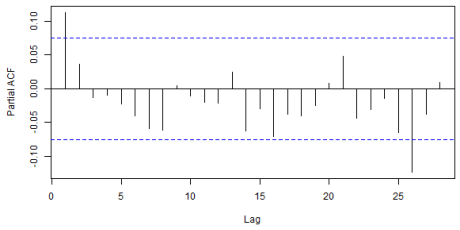
Poisson model with 1 knot per year



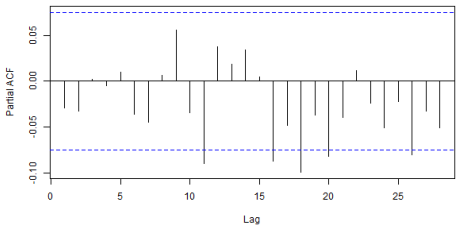
(a) virus surveillance preterm



(b) virus surveillance very preterm



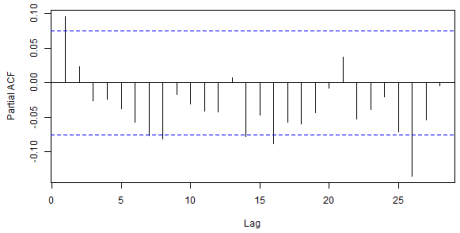
(c) NorSySS preterm



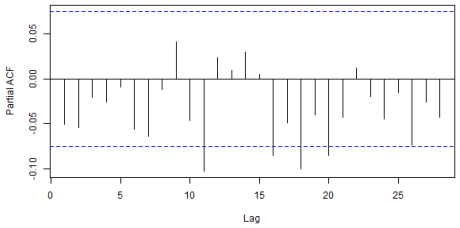
(d) NorSySS very preterm

Figure A.1: Partial Autocorrelation Function plots for Poisson regression models with 1 knot per year.

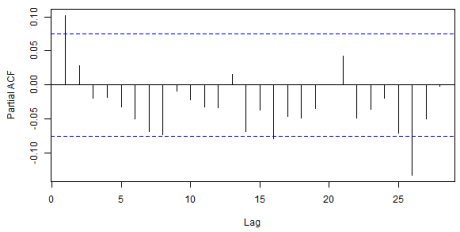
Poisson model with 2 knots per year



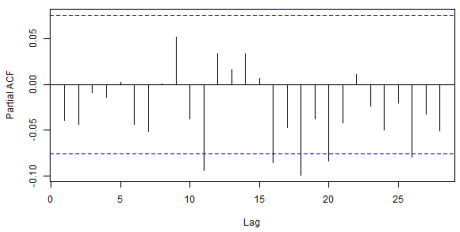
(a) virus surveillance preterm



(b) virus surveillance very preterm



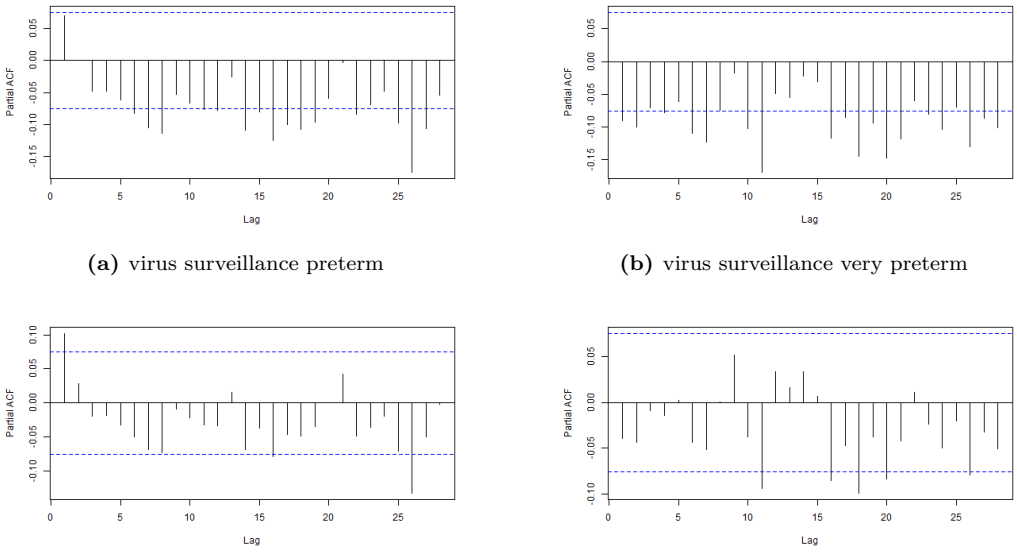
(c) NorSySS preterm



(d) NorSySS very preterm

Figure A.2: Partial Autocorrelation Function plots for Poisson regression models with 2 knots per year.

Poisson model with 3 knots per year



(a) virus surveillance preterm

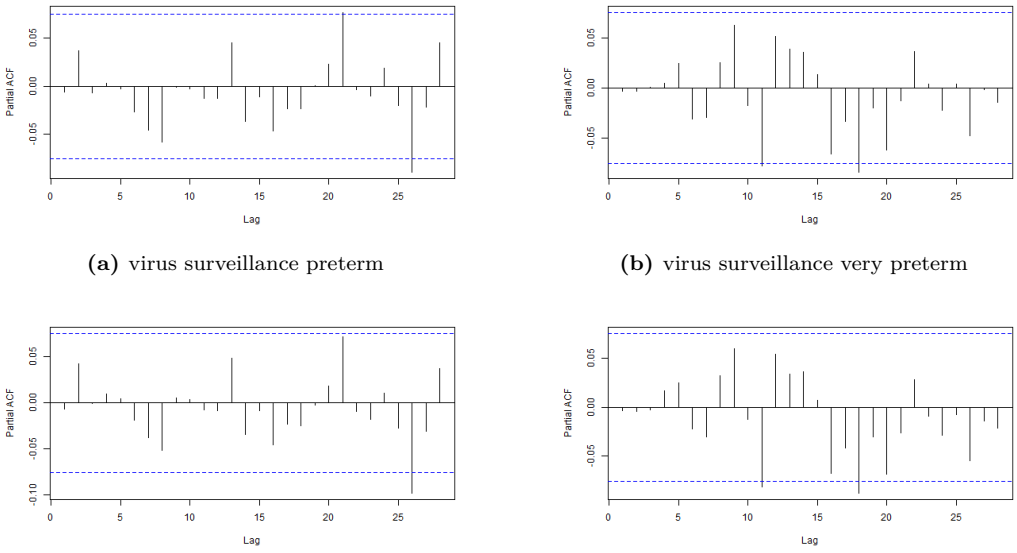
(b) virus surveillance very preterm

(c) NorSySS preterm

(d) NorSySS very preterm

Figure A.3: Partial Autocorrelation Function plots for Poisson regression models with 3 knots per year.

Regression with ARIMA errors



(a) virus surveillance preterm

(b) virus surveillance very preterm

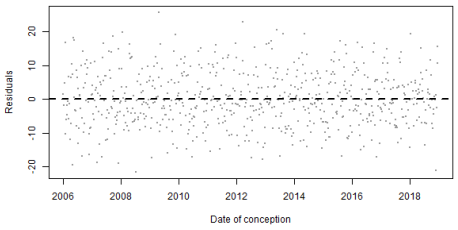
(c) NorSySS preterm

(d) NorSySS very preterm

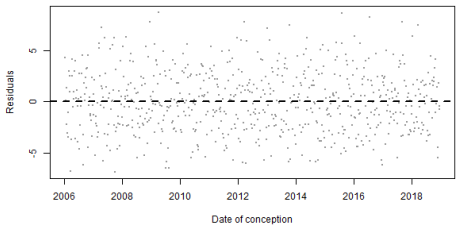
Figure A.4: Partial Autocorrelation Function plots for regression models with ARIMA errors.

Appendix B. Residual plots preterm models

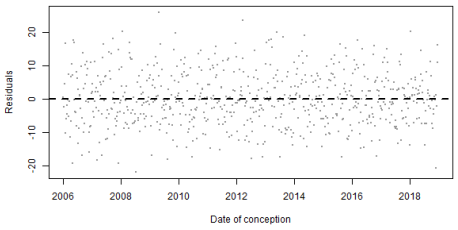
Poisson model with 1 knot per year



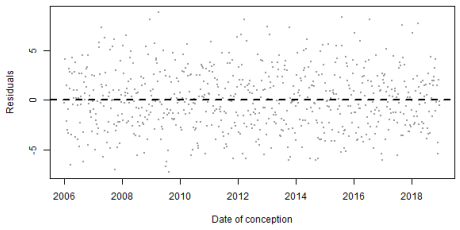
(a) virus surveillance preterm



(b) virus surveillance very preterm



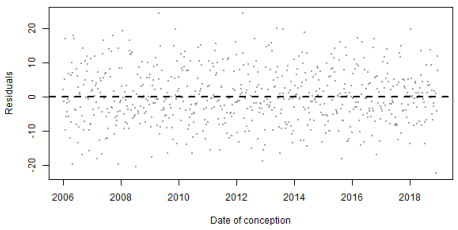
(c) NorSySS preterm



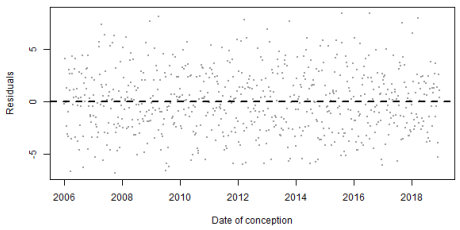
(d) NorSySS very preterm

Figure B.1: Residual plots for Poisson models with 1 knot per year.

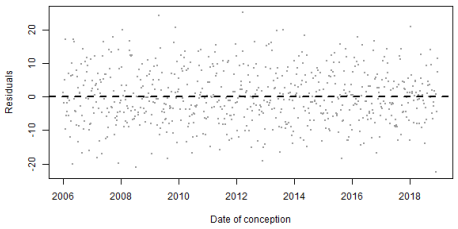
Poisson model with 2 knots per year



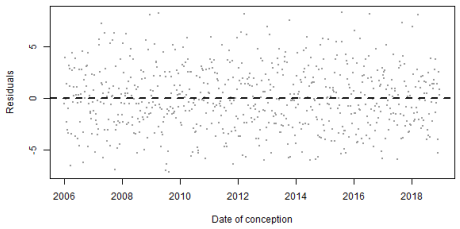
(a) virus surveillance preterm



(b) virus surveillance very preterm



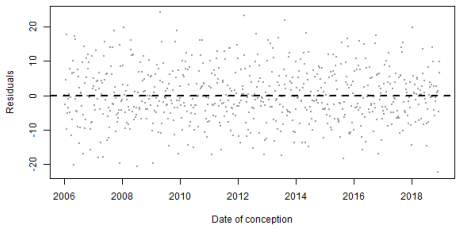
(c) NorSySS preterm



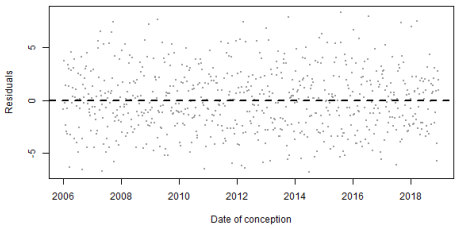
(d) NorSySS very preterm

Figure B.2: Residual plots for Poisson models with 2 knots per year.

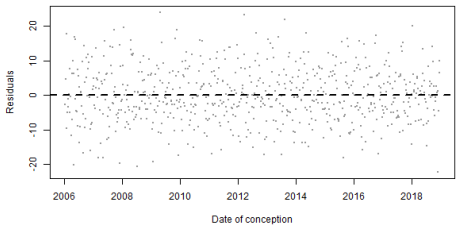
Poisson model with 3 knots per year



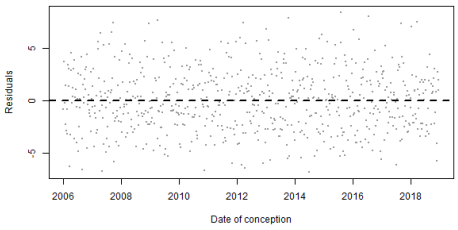
(a) virus surveillance preterm



(b) virus surveillance very preterm



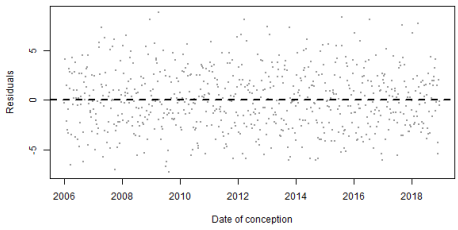
(c) NorSySS preterm



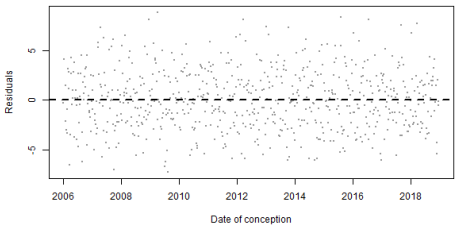
(d) NorSySS very preterm

Figure B.3: Residual plots for Poisson models with 3 knots per year.

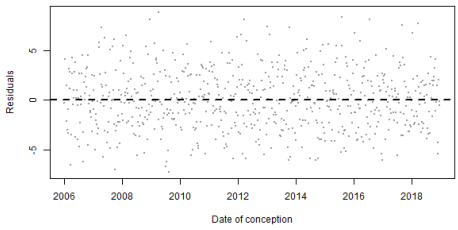
Regression with ARIMA errors



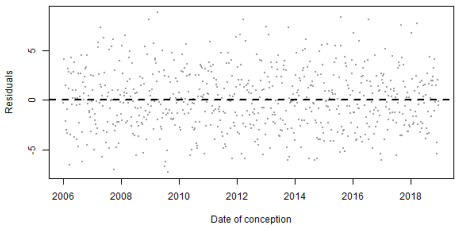
(a) virus surveillance preterm



(b) virus surveillance very preterm

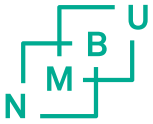


(c) NorSySS preterm



(d) NorSySS very preterm

Figure B.4: Residual plots for regression models with ARIMA errors.



Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway