Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

Simen Nordeide

Analysing multimorbidity using registry data

Master's thesis in Industrial Mathematics
Supervisor: Andreas Asheim
June 2023

NTNU
Norwegian University of
Science and Technology

Simen Nordeide

Analysing multimorbidity using registry data

Master's thesis in Industrial Mathematics
Supervisor: Andreas Asheim
June 2023

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

NTNU
Norwegian University of
Science and Technology

Abstract

In the next decades, the western world is anticipating an aging population, giving rise to a higher prevalence of complex multimorbidity. Since hospitals are primarily organized around specialized disciplines, this will have a significant effect on the complexity of care of patients, and the overall capacity of the system. This master thesis aims to investigate the risk of being diagnosed with a new disease associated with having specific underlying diseases, as well as how various combinations of chronic diagnoses are associated with increased health service utilization. The data material used in this project contains pseudoanonymized information of all initial hospital admissions in Norway associated with specific chronic diagnoses during the period 2008-2021 from the Norwegian Patient Registry (NPR), in addition to date of death of each patient from the Norwegian Cause of Death Registry (DÅR), the number of days of utilization of general practitioner and out-of-hour general practitioner from Control and Payment of Health Reimbursement register (KUHR), and days of acute visits to specialized health services and days of hospital stay from NPR. We implemented log-linear Poisson regression models, with offset defined as number of days of exposure, including covariates sex of patient, and age using cubic-splines. Our findings coincides with existing knowledge in the field, which gives validation to our model. Based on our analysis, we have identified several patterns which may indicate significant potential for further research.

Sammendrag

I løpet av de neste tiårne er det forventet at den vestlige verden får en aldrende befolkning, som vil føre til flere tilfeller av kompleks multimorbiditet. Siden sykehusene hovedsaklig er organisert rundt spesialiteter, vil dette ha en signifikant effekt på kompleksiteten av behandlingen av pasienter, samt den totale kapasiteten av systemet. Denne masteroppgaven undersøker risikoen for å bli diagnostisert med en ny sykdom assosiert med å ha spesifikke underliggende sykdommer, samt hvordan forskjellige kombinasjoner av kroniske sykdommer er assosiert med økt helsetjenestebruk. Datamateriale som brukes i dette prosjektet består av pseudoanonymisert informasjon om alle første sykehusinnleggelser i Norge assosiert med spesifikke kroniske diagnoser i løpet av perioden 2008-2021 fra Norges Pasientregister (NPR), i tillegg til dødsdato for hver pasient fra Dødsårsaksregisteret (DÅR), antall dager med besøk hos fastlege og legevakt fra Kontroll og utbetaling av helserefusjoner (KUHR), og antall dager med akuttbesøk og liggedager på sykehus fra NPR. Vi implementerte log-lineære Poisson-regresjonsmodeller, med offset definert som antall dager eksponert, inkludert kovariater kjønn, og alder ved bruk av cubic-splines. Funnene våre stemmer overens med eksisterende kunnskap om feltet, noe som styrker vår tro på modellen. Basert på analysene våre så har vi identifisert flere mønstre som kan indikere signifikant potensiale for videre forskning.

Acknowledgement

I am delighted to present this master thesis, and would like to take this opportunity to express my gratitude towards the people who have supported me throughout this journey.

I would first like to extend my heartfelt appreciation to my supervisor Andreas Asheim for his unwavering support and great collaboration. Your expertise, patience and steady flow of ideas has helped my immensely in navigating and shaping this project.

I also want to extend my gratitude to Regforsk by Johan H. Bjørngaard, a project financed by NFR, for making this project possible. Additionally, I would like to thank Kristin H. Vinjerui for providing me with insight into the field of multimorbidity.

I would also like to thank my partner Marlén Mannsåker for her invaluable mental support during these years. I have been both tired and burnt out, but I have always been able to count on you for your encouragement and belief in my abilities.

Furthermore, I am immensely grateful for the master's group at SiT for their camaraderie, sharing of frustrations and continuous motivation. I am lucky to have been able to share in the journey towards our final theses with you.

Finally, I would like to thank both my family, and the family of my partner, for their continuous support. Your love, encouragement and faith in my abilities has been great sources of strength through this endeavor.

This master thesis would not have been possible without the collaborative effort and support of these incredible people. I am forever grateful for their presence in my life.

Thank you all!

Simen Nordeide

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Problem description and motivation

One of the great challenges going forward is the aging population in the world [1]. This has been caused by several factors, including improved living conditions, advances in health care and declining birth rates. Along with age, multiple chronic diseases may start to accumulate. This is referred to as multimorbidity [2], and leads to higher mortality [3] and more complex patients.

Norwegian hospitals are organised around specialties, focusing on single diseases. When a patient is admitted to a hospital, they are usually transferred to the appropriate ward. For instance is a patient who is admitted for cancer transferred to an oncologist. If the same patient has a history of other diseases, e.g. heart failure, they may require special treatment. Consider a patient with chronic obstructive pulmonary disease (COPD). Treatment often include oral corticosteroids, which increases blood-sugar. If we consider that the same patient also has diabetes mellitus, this increase in blood-sugar might be harmful.

The interaction between several morbidities is handled by the general practitioner (GP). In Norway there is an increased load on the GP systems, and around 150 000 citizens did not have access to a regular general practitioner in February 2022 [4]. When the primary care services are inadequate, a consequence may be that patients resort to the use of acute specialist health services for non-emergency situations. This places increased stress on these services.

Additionally, having a single disease may increase the risk of developing other diseases. Some diseases may for example cause chronic inflammation, including irritable bowel syndrome. This may increase the risk of certain diseases, like cardiovascular disease and diabetes. Other diseases may weaken the immune system, for example HIV/AIDS, which may lead to a higher risk of opportunistic infections. Yet other diseases may act as risk factor for various other diseases. Obesity and high blood pressure are examples, which may increase the risk of heart disease, diabetes and specific types of cancer.

In addition to physical diseases, multimorbidity also includes mental disorders, like depression and schizophrenia [2]. There are significant interactions between physical and mental disorders [5]. Patients with mental health disorders can e.g. have decreased physical health due to medication, and be less motivated or able to care for their physical health by following complex treatment regiments and medication schedules. Patients with physical chronic diseases can have psychiatric side effects of medication, and experience negative effects of hormone imbalances and living with a chronic disease.

1.2 Background

There exists a wide range of literature on multimorbidity. Some studies focus on a wide range of diseases [6, 7, 8], while others are limited to a few specific diseases often considered central in multimorbidity [9]. Others again are mainly focusing on the prevalence of multimorbidity, or the number of chronic diseases, rather than specific diseases [10, 11]. A range of different methods has been used, e.g. factor analysis [12, 6, 8], network analysis [9, 13, 14], clustering [15, 6, 8] and logistic regression [16, 10].

Numerous publications focus on finding specific clusters of diseases often occurring together, using a range of clustering techniques. The clusters and patterns varies depending on the technique used [6]. A systematic review [17] identified 97 patterns of 2 or more diseases. From these, three groups of patterns were demonstrated to be relevant. "The first one comprised a combination of cardiovascular and metabolic diseases, the second one was related with mental health problems, and the third one with musculoskeletal disorders" [17].

Studies have found that multimorbidity leads to a substantial increase in health service utilization and cost [18, 19, 20, 21]. One study found that patient having two or more diseases may lead to both increase and decrease in cost [22], depending on the specific diseases.

Furthermore, multimorbidity has been found to be associated with increased primary health care use, e.g. consultations with a general practitioner [23, 24, 11], as well as secondary health care [25, 26]. A systematic literature review found that almost all studies in their review indicated a positive relation between multiple chronic conditions and health care use [27].

Finally, several studies have questioned the effectiveness of the single-disease management practiced in hospitals for patients with multiple chronic conditions [28, 29]. Some patients may benefit from disease-specific interventions, while patients with multimorbidity would benefit from a multi-disease program [30].

1.3 Analysis design

Our goal was to utilize registry data, with the limitations they have, to investigate the association between several chronic diseases, and how specific combinations of these relate to the amount of utilization of health services. We wanted to look for associations that are already established within the medical field, and potentially seek to discover novel connections.

We aimed to estimate the increase or decrease in risk of being diagnosed with a specific chronic disease, associated with having history of other diagnoses. Consider a patient with a history of diabetes. The risk of being diagnosed with e.g. obesity may be higher for them than for patients with no history of diabetes.

Figure 1 shows the timeline of the first model. We registered all main and secondary diagnoses per patient associated with admissions to hospitals during the period 2008-2011. During the follow-up period 2012 - 2021, we modelled the risk of a new diagnosis using a Poisson model with binary data [31, 32]. In other words, for each pair of diagnoses (A, B) we aimed to estimate the change in risk of being diagnosed with B, given a prior history with A.

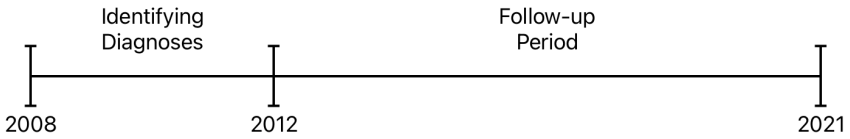


Figure 1: Model 1 - We followed each person in Norway above the age of 60 in 2013. We identified all diagnoses associated with admissions to a hospital between 2008 and 2011, which we recorded as underlying. We then modelled the risk of being admitted with a new diagnosis from 2012 until 2021, given the underlying diagnoses.

We also aimed to estimate the effect that specific combinations of diseases has on health service use, in particular the following outcomes, measured in days; visit to general practitioner, visits to out-of-hour general practitioner, number of acute contacts with specialist health services, and days hospitalized. Consider for example patients with diabetes and anxiety. We aimed to estimate the effect that the specific combination of those has on health service utilization.

Consider the timeline in Figure 2. Similar to model 1, we identified diagnoses from 2008 - 2011, followed by a 2 year quarantine period for potential new diagnoses to stabilize in health care use. During the follow-up period 2013 - 2021 we recorded the number of days of health care use of each patient.

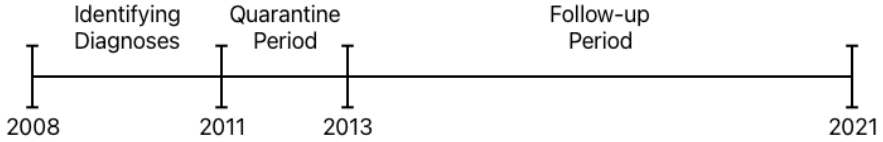


Figure 2: Model 2 - We followed each person in Norway above the age of 60 in 2013. We identified all diagnoses associated with admissions to a hospital between 2008 and 2011. After a 2 year quarantine period we modelled the health service use from 2013 until 2021.

The diagnoses included in the analysis were taken from the list of diagnoses included in the list of complex multimorbidities by The Norwegian Directorate of Health [33], in addition to a range of mental health disorders. All diagnoses included are presented in Table 1. We used ICD-10 diagnosis codes, i.e. the International Classification of Disease, 10th Revision, which is a standardized way of classifying diseases using alphanumerical codes [34]. It is maintained by the World Health Organization (WHO) and is widely used for epidemiological, clinical, and research purposes.

Short name	**Diagnoses**	**ICD-10 codes**
neoplasms	Neoplasms	C00-C97,D00-D09
blood-forming	Diseases in blood and blood-forming organs	D50-D77
immune	Immune diseases	D80-D89,B20-B24
thyroid	Diseases in thyroid gland	E00-E07
diabetes	Diabetes	E10-E14
endocrine	Endocrine and metabolic diseases	E15-E16,E20-E35,E70-E90
obesity	Obesity	E66
dementia	Dementia	F00-F03,G30-G32
neuromusc	Neuromuscular diseases	G10-G14,G20-G26,G35,G40-G41,G70-G73
paralytic	Paralytic syndroms	G80-G83
eye	Diseases in the eye	H10-H59
hypertensive	Hypertensive diseases	I10-I15
ischaemic	Ischaemic heart diseases	I20-I25
arrhythmia	Arrhythmia	I44-I49
chf	Heart failure	I50
cerebrovasc	Cerebrovascular diseases	I60-I64,G45
lower-respiratory	Chronic lower respiratory diseases	J40-J47
enteritis-colitis	Noninfective enteritis and colitis	K50-K52
ibs	Irritable bowel syndrome	K58
liver-gall	Diseases in liver and gallbladder	K70-K77,K80-K87
joint	Joint disorders	M05-M14,M30-M36
osteopathies	Osteopathies	M80-M85
kidney	Chronic kidney disease	N18
prostate	Hyperplasia of prostate	N40
openwound	Open wounds	T01,S01,S11,S21,S31,S41,S51,S61,S71,S81,S91
affective	Mood affective disorders	F30-F39
neurotic	Neurotic and somatoform disorders	F40-F48
schizo	Schizophrenia	F20-F29
intox	Substance-induced mental disorders	F10-F19

Table 1: Diagnoses included in the list of complex multimorbidities by The Norwegian Directorate of Health [33], in addition to a range of mental health disorders, with corresponding ICD-10 codes.

2 Setting and data

Norway has a universal public health care system, which includes both primary and secondary care services. General practitioners and out-of-hour general practitioners acts as a first point of contact for all patients. They provide primary health services, including diagnosis and treatment, and make referral to hospitals or specialists when necessary. Access to specialist care is mainly possible only after being referred by a GP or during emergencies.

The data available in this study were acquired from several Norwegian national registers. A personal identification number unique to all Norwegian inhabitants was used for linking information from different registers.

2.1 Norwegian health registry data

All contacts with public systems in Norway are registered in the Norwegian Patient Registry (NPR), including a list of primary and secondary discharge diagnoses associated with each admission (episode of care), encoded using International Classification of Disease (ICD- 10) diagnosis codes [34]. The NPR data contains 1 line for each ward episode, with a separate list of all codes associated with each episode, connected with one key for each episode. For this project, the information available included limited, pseudoanonymized information of all first admissions to hospitals in Norway with primary and secondary diagnoses corresponding to the list in Table 1, during the years 2008 - 2021 for patients above the age of 60 in 2013. Specifically, for each patient we had access to the first year each code appeared in the data set as well as the age and gender of the patient. Thus our data contained the first year each diagnosis was observed among all admissions to a hospital in Norway.

Information of all-cause mortality for each patient was included from Norwegian Cause of Death Registry. From Control and Payment of Health Reimbursement register (KUHR) we had access to the number of days each patient visited a general practitioner (GP), as well as out-of-hour general practitioner, each year from 2012 - 2021. Similarly, from the Norwegian Patient Registry (NPR) we had access to the number of days of acute contacts with specialized health services, as well as the number of days each patient was hospitalized.

2.2 Ethical approval and privacy

Use of the data in this project was approved by the Regional Committee of Ethics in Medical Research (reference number 2016/2159). Participant consent was not required. Only relevant variables were available.

The data we analysed was only accessible through a virtual computer within the Hunt-Cloud system. Hunt-Cloud is a cloud service provided by NTNU (Norwegian University of Science and Technology), which provides secure access to sensitive health service data for research and development purposes. Each researcher has the ability to set up their own workspace within the virtual lab, where they can access the data. Researchers are not permitted to take access out of the lab, and are only given access to project relevant data. This ensures a safe and secure environment for accessing the sensitive data.

3 Theory

3.1 Poisson distribution

One of the most common probability distributions for modelling count data is the Poisson distribution [35], which describes the probability of observing an event or several events within a given time frame. There are a wide range of use cases for Poisson distribution, like the number of hospital admissions, product defects or traffic accidents within a given time interval. There are three assumptions which need to be satisfied for a process to be deemed a Poisson process. The number of events happening within a time frame must be independent to the number of events in any other disjoint time frame. The probability of an event happening within a time interval also needs to be proportionate to the length of that interval. Finally, the probability of two or more events occurring within a small time interval is negligible.

The Poisson distribution has a single parameter λ, which is sometimes called the intensity parameter, and explains the average rate of events occurring within an interval. Suppose that $Y \sim \text{Pois}(\lambda)$ is a Poisson distributed stochastic variable with expectation λ. Then its probability density function is

$$f(y;\lambda) = \text{P}(Y = y|\lambda) = \frac{e^{-\lambda}\lambda^y}{y!}, \quad y = 0, 1, \dots \tag{1}$$

3.2 Generalized Linear Model

Poisson regression is a statistical technique that is often used for modelling count data, when the outcome variable follows a Poisson distribution. Using Poisson regression it is possible to analyse the association between one or more covariates and the outcome variable, i.e. the expected count of events. The Poisson regression model is a generalized linear model (GLM) [36].

Rewriting equation 1, it can be shown that the Poisson distribution is part of the Exponential Family of Distributions, an can thus be written on the following form:

$$f(x|\theta) = \exp\left(\frac{\theta y - b(\theta)}{\phi} w + c(y, \phi, w)\right). \tag{2}$$

where

- θ is called the natural or canonical parameter, the parameter of interest.
- ϕ is often considered the nuisance parameter, and is equal to 1 for the Poisson distribution.
- w is a weight function, and is usually 1.
- b and c are known functions

To show that Poisson distribution can be written in the specified form, we first take the logarithm, rewrite the equation, and then exponentiate the equation.

$$\ln f(y) = y\ln\lambda - \lambda - \ln(y!)$$

$$f(y) = \exp\Big(y\ln\lambda - \lambda - \ln(y!)\Big)$$

We set the following parameters; $\theta = \ln\lambda$, $b(\theta) = \lambda$, $\phi = 1$ and $c(y, \phi, w) = -\ln(y!)$, which brings us to the form of the exponential family.

In particular, Poisson models use a log-link function to connect the mean of the response variable to the predictor variables. This transformation ensures that the predicted response variables are always positive, which is important in count data. Thus the log of the mean of the response variable is used to model a linear combination of the predictor variables, $\boldsymbol{x_i^T\beta}$, indicated in the following formula:

$$E(Y_i) = \lambda_i = \exp(\boldsymbol{x_i^T\beta}) = \exp(\beta_0) \cdot \exp(\beta_1 x_{i1}) \cdot ... \cdot \exp(\beta_k x_{ik})$$

or

$$\ln(\lambda_i) = \boldsymbol{x_i^T\beta} = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik}$$

3.3 Estimation & Testing

3.3.1 Maximum likelihood estimation

In Poisson regression, the parameters are often estimated using maximum likelihood estimation (MLE). The discrete density, or the likelihood $L_i(\beta)$, of a Poisson distributed response variable is given by the following [36]

$$f(y_i|\beta) = \frac{\lambda_i^{y_i}\exp(-\lambda_i)}{y_i!}, \quad \mathrm{E}(y_i) = \lambda_i$$

where $\lambda_i = \exp(\mathbf{x}_i'\beta)$.

In maximum likelihood estimation, the parameters are estimated by maximizing the likelihood of observing the given data. The log-likelihood is used to simplify calculations, given by

$$l(\beta) = \sum_{i=1}^{n} l_i(\beta) = \sum_{i=1}^{n} y_i(\mathbf{x}_i'\beta) - \exp(\mathbf{x}_i'\beta)$$

The next step is to calculate the score function, which is found by partially differentiating the log-likelihood with respect to β, which results in

$$\mathbf{s}(\beta) = \sum_{i=1}^{n} \mathbf{x}_i(y_i - \lambda_i).$$

To estimate the maximum likelihood estimate $\hat{\beta}$, we need to solve the set of non-linear equations:

$$s(\hat{\beta}) = 0$$

This can be done using the Fisher scoring algorithm, which is a numerical optimization algorithm commonly used for estimating MLE in statistical models. The last step before being able to estimate the ML estimates $\hat{\beta}$ is calculating the Fisher information matrix $\mathrm{F}(\beta)$, which equals the covariance matrix of the score function. Using the relation $\mathrm{E}(y_i - \lambda_i)^2 = \mathrm{Var}(y_i) = \lambda_i$, the Fisher information is obtained

$$\mathbf{F}(\beta) = E(\mathbf{s}(\beta)\mathbf{s}'(\beta)) = \sum_{i=1}^{n} \mathbf{x}_i\mathbf{x}_i'\lambda_i$$

Finally the maximum likelihood estimators $\hat{\beta}$ can be found by using the Fisher scoring algorithm:

$$\hat{\beta}^{(t+1)} = \hat{\beta}^{(t)} + F^{-1}(\hat{\beta}^{(t)})s(\hat{\beta}^{(t)}), \quad t = 0, 1, 2, ...$$

3.3.2 Hypothesis test

After estimating the parameters in the model, p-values should be calculated to assess the statistical significance of each parameter. The first step in this is to set up a null hypothesis, which assume that the specific parameter is zero, i.e. there is no effect from the predictor variable on the expected count. Hypothesis testing for a generalized linear model is usually performed using one of three statistics; likelihood ratio, Wald or score statistic. In both likelihood ratio tests and Wald tests the p-values under the null-hypotheses are generally calculated in the upper tail of the χ^2 distribution.

P-values can be interpreted as the probability of observing the result, given that the null hypothesis is true. If a p-value is small, usually below a significance level of 0.05, then the null hypothesis is rejected. The parameter is said to be statistically significant, meaning that the result, or a more extreme outcome, is unlikely to happen by chance.

A common challenge of significance tests is that as the number of simultaneous tests increases, the likelihood of Type I error also increases. In other words, there is an increased probability of at least one significant result (or a more extreme outcome) occurring by chance. The Bonferonni correction adjusts the p values by dividing the desired significance level α by the number of comparisons being made. Suppose we desire $\alpha = 0.05$, with a model including 900 covariates. After applying Bonferonni correction, the covariates needs to satisfy the new p-value of $0.05/900 = 0.0000556$. Note that using this method may lead to a higher risk of Type II errors, i.e. potential significant effects may be missed [37].

"In equation form, if

$$P(T_i \text{ passes } | H_0) \leq \frac{\alpha}{n}$$

for $1 \leq i \leq n$, then

$$P(\text{some } T_i \text{ passes } | H_0) \leq \alpha$$

which follows from the Bonferonni inequalities."

3.4 Overdispersion

The Poisson model assumes that the mean and variance of the response variable are equal, referred to as equidispersion:

$$\lambda_i = \mathrm{E}(y_i) = \mathrm{Var}(y_i)$$

Often when Poisson models are used to fit real data, the observed variation in the data is larger than the conditional expectation. In other words, there is more variability in the data than can be explained by the model, violating the assumption of equidispersion for Poisson models. This may be due to several reasons, for example that the data is correlated, i.e. the observations are not independent, or that there are extra sources of variation that is not accounted for in the model.

The overdispersion present in a Poisson model can be accounted for in various ways. One often used method is to switch the distribution to instead using a negative binomial regression model, which is also used for count data, but is especially suitable for overdispersed data. This is done by including an extra parameter to account for the variance, called the dispersion parameter.

3.5 Poisson regression with binary outcome

Usually for developing models with binary outcomes, logistic regression or Cox regression is used. However, Guangyong Zou has provided a modification to Poisson regression for handling binary outcomes [31, 32]. In logistic regression log-odds ratios, and thus odds ratios, are computed, while the modified Poisson regression computes relative risks, which are usually more intuitive and straight forward to interpret. When using Poisson models for binary data, the estimated relative risks are often overestimated. Modified Poisson regression handles this with a robust error variance procedure known as sandwich estimation.

3.6 Offset

In addition to using Poisson regression for modelling count data, it can also easily be applied for modelling rates. This is done by adding a term to adjust for the difference in exposure or observation time. In our case we are interested in modelling the number of hospital visits, based on patients with different exposure time. By dividing the response variable with the exposure time t_i, and rearranging the terms, the model is describing rate data instead. The term $ln(t_i)$ is referred to as an offset.

$$ln(\frac{y}{t_i}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p$$

$$ln(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + ln(t_i)$$

which, after exponentiation results in

$$y = t_i e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p}$$

$$y = e^{\beta_0} \cdot e^{\beta_1 \cdot x_1} \cdot e^{\beta_2 \cdot x_2} \cdot \ldots \cdot e^{\beta_p \cdot x_p} \cdot t_i$$

The term e^{β_i} is referred to as the incidence rate ratio (IRR). This can be interpreted as the relative change in the count or rate, associated with one unit change in x_i. A IRR greater than 1 suggests an increase in the count or rate, and a IRR less than 1 suggests a decrease. Consider an IRR of 1.2, which suggests a 20% increase in the count/rate associated with a unit increase in x_i. Note that for binary outcomes, e^{β_i} is interpreted as RR - relative risk, being $\frac{P_a}{P_b}$, where a is the exposed group, and b is the unexposed group.

3.7 Splines

In some cases, there may be a need to capture non-linear relationships between the predictor variable and the outcome variable. There are various ways of extending linear models, including polynomial regression, where a polynomial is used to extend the predictor variable. An example of polynomial regression is cubic regression, where the original predictor is extended to include three variables; X, X^2 and X^3. Another method is called step functions, or piecewise-constant functions, and consists of cutting the range of a variable into K distinct segments. These two methods are special cases of what is called basis functions, which refers to the approach of having a family of functions or transformations that can be applied to X, to extend it from a linear model [38].

A more sophisticated and flexible approach is based on polynomial and piece-wise functions, called regression splines. This includes splitting the range of the variable into K distinct segments, and fitting a polynomial to each region. At the points that each pair of polynomials join together, in

the knots, constraints are imposed so that they join smoothly. Consider for example piecewise cubic polynomial, where a cubic polynomial is fitted to each segments of the form

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \epsilon.$$

which means a piecewise cubic spline with a single knot can be represented by [38]

$$y = \begin{cases} \beta_{01} + \beta_{11} x + \beta_{21} x^2 + \beta_{31} x^3 + \epsilon, & \text{if } x < c; \\ \beta_{02} + \beta_{12} x + \beta_{22} x^2 + \beta_{32} x^3 + \epsilon, & \text{if } x \geq c. \end{cases}$$

The following figure shows an example of a piecewise cubic spline, fit with a single knot at $x = 50$ [38]. For cubic splines, the polynomial are contrained to be continuous at the knots, in addition to the first and second derivatives also being continuous at the knots. This adds up to 5 degrees of freedom with a single knot, or in general, $4 + K$ degrees of freedom for cubic splines with K knots.

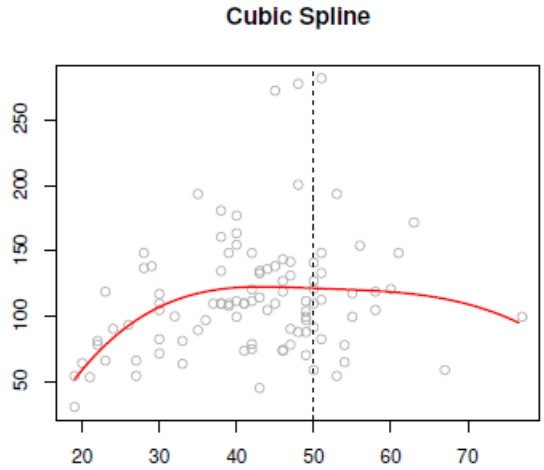


Figure 3: Example of a piecewise cubic spline fit to data with a single knot at $x = 50$. [38]

4 Methods

We started the analysis by exploring the data. We were interested in the number of patients having a history of each diagnosis, as well as the number of patients having 2, 3, ... different diagnoses. Furthermore, we examined which diagnoses often occurred together in the same patient. Which pairs of diagnoses were most common, and which triples were most common.

Next, we were interested in examining information about the patients included in the data, mainly the age and sex of the patients with specific diagnoses. If there were significant differences in the age and sex of patients with different diagnoses, these covariates needed to be included in the models. We used box plots for these, grouping by diagnoses. The age of each patient were based on the year each diagnosis was first observed in the patient. Thus patients with more than 1 diagnosis was registered with different ages for each diagnosis in the box plot. We suspected a non-linear relationship to age, which figure 7 indicates that it was. Therefore we approximated the relationship using natural cubic splines.

Before implementing the model, we examined the distribution of the outcome variables for model 2, the number of days of the following: acute contacts with specialized health services, hospitalization, general practitioner and out-of-hour general practitioner. We did that by creating histograms of the respective distributions.

4.1 Poisson regression

We implemented 2 Poisson models. The first model was made to investigate the increase or decrease in risk of a new diagnosis, for patients with different underlying diagnoses. With the second model we aimed to examine the change in health service utilization associated with specific combinations of chronic diagnoses, i.e. multimorbidity. The prevalence of each diagnoses in the list of diagnoses in Table 1 was included in both models. Additionally model 2 included all pairwise interactions between the diagnoses.

We also included the sex of the patient, as well as age. To account for the non-linear relationship of age shown in Figure 7, we implemented natural splines of the age, using 10 splines. Furthermore, to account for patients dying within the follow-up period, we added an exposure variable. This measured the number of days from the beginning of follow-up until either the end of follow-up, or the death of the patient in the case of the patient dying within the follow-up. This exposure was included as offset in our models.

4.1.1 Model 1 - Risk of new diagnosis

The first model we implemented aimed to investigate the relationship between having a diagnosis, and the risk of being diagnosed with a new one. We used a log-linear Poisson model for finding the increase/decrease in risk of having a new diagnosis, associated with having a history of a different diagnosis. We used the event of an admission after 2011 associated with a new diagnosis as the outcome. For each new diagnosis we fitted a model with the prevalence of each underlying diagnosis. As mentioned earlier, the underlying diagnoses were registered during the year 2008 - 2011. We included an offset, with the exposure being the time until a patient received the new diagnosis or died.

The model can be explained by the following formula:

$$
\begin{aligned}
\ln(\mathrm{E}[y]) \sim \beta_0 &+ \beta_{\text{diabetes}} \cdot x_{\text{diabetes}} \\
&+ \beta_{\text{dementia}} \cdot x_{\text{dementia}} \\
&\vdots \\
&+ \beta_{\text{neoplasm}} \cdot x_{\text{neoplasm}} \\
&+ \beta_{\text{sex}} \cdot x_{\text{sex}} + \beta_{\text{age}} \cdot x_{\text{age}} + t_i
\end{aligned}
\tag{3}
$$

where t_i is the exposure time.

Consider x_{neoplasm}, x_{diabetes}, ... to be the prevalence of the diagnoses. Consider then $e^{\beta_{\text{diabetes}}}$. This is the relative risk of the event y, an admission with a new diagnosis, for patients with diabetes. In other words this is the relative risk of the new diagnosis for the group of patients with a history of diabetes, compared to the group of patients with no history of diabetes, adjusted for all other diagnoses. Each diagnosis in the follow-up have an association with every potential underlying diagnosis.

4.1.2 Model 2 - Health service utilization

Next we examined the association between multimorbidity and health service utilization. We implemented a log-linear Poisson model for the incidence-rate ratio of various measures for health service use. Specifically, the following measures was used separately as the outcome variable, measured in the number of days of contact: acute contacts with specialist health services, days hospitalized, visits to general practitioner and visits to out-of-hour general practitioner. We included the age and sex of patients similarly to model 1, as well as the offset.

Recall that the diagnoses for each patient were registered during the period 2008 - 2011. After a 2 year quarantine period, we measured the number of days through the follow-up period. The model can be described as follows:

$$
\begin{aligned}
\ln(\mathrm{E}[y]) \sim \beta_0 &+ \beta_{\text{neoplasms}} \cdot x_{\text{neoplasms}} \\
&+ \beta_{\text{arrhythmia}} \cdot x_{\text{arrhythmia}} \\
&+ \beta_{\text{diabetes}} \cdot x_{\text{diabetes}} \\
&\vdots \\
&+ \beta_{\text{neoplasms \& arrhythmia}} \cdot x_{\text{neoplasms \& arrhythmia}} \\
&+ \beta_{\text{neoplasms \& diabetes}} \cdot x_{\text{neoplasms \& diabetes}} \\
&\vdots \\
&+ \beta_{\text{sex}} \cdot x_{\text{sex}} + \beta_{\text{age}} \cdot x_{\text{age}} + t_i
\end{aligned}
\tag{4}
$$

Consider x_{neoplasm}, x_{diabetes}, ... to be the prevalence of the diagnoses. $e^{\beta_{\text{neoplasm}}}$, $e^{\beta_{\text{diabetes}}}$ are then the relative increase in number of days of health services use associated with having a history of neoplasm compared to not having a history of neoplasm. These were included to adjust for the association of the single diagnoses, such that the interaction terms, e.g. $e^{\beta_{\text{neoplasm \& diabetes}}}$ explains the association of specifically having both diagnoses, without the association of having them separately.

4.2 Visualization

The results were visualized using graph drawing by force-directed placement [39], which is a technique for creating graphs by simulating physical forces between nodes. This was done by considering the diagnoses as particles, and the relative risks/incidence rate-ratios as springs. By iterating

until equilibrium, diagnoses with stronger relations was visualized closer together. In model 1 we drew arrows from underlying diagnoses to the diagnoses with increased risk, with thickness of the arrow corresponding to the magnitude of the increase in risk, removing non-significant p-values using Bonferonni correction. In model 2 we drew lines between diagnoses with strong relations and significant p-values, also using Bonferonni correction. The relative risks/incidence rate-ratios from the models were also visualized with heatmaps, including the values of the RR/IRR.

5 Results

5.1 Descriptive analysis

The transformed data set consisted of 1 147 426 patients. The mean age of all patients in 2013 was 72, while the mean age of patients with each specific diagnosis are shown in Figure 2, with age based on first admission associated with the diagnosis. 271 453 patients died before the start of the follow-up period, and were excluded from the models.

Diagnosis	**Patients**	**Mean age**	**Prop. women**
Diseases in the eye	239340	73	0.6
Hypertensive diseases	184692	71	0.53
Ischaemic heart diseases	118301	72	0.39
Neoplasms	117479	71	0.49
Arrhythmia	103475	73	0.45
Diabetes	83290	69	0.46
Endocrine and metabolic diseases	75306	71	0.57
Chronic lower respiratory	68830	70	0.53
Joint disorders	55207	70	0.65
Hyperplasia of prostate	45901	71	0
Cerebrovascular diseases	43473	73	0.49
Diseases in blood and blood-forming organs	40895	74	0.58
Osteopathies	35854	72	0.87
Heart failure	33537	76	0.45
Diseases in thyroid gland	32686	71	0.82
Mood affective disorders	28324	67	0.65
Neuromuscular diseases	27265	69	0.52
Diseases in liver and gallbladder	25112	70	0.59
Open wounds	24372	70	0.44
Neurotic and somatoform disorders	22220	66	0.68
Dementia	20855	79	0.65
Chronic kidney disease	18127	75	0.38
Noninfective enteritis and colitis	13069	68	0.55
Substance-induced mental disorders	11879	64	0.35
Obesity	8990	64	0.58
Irritable bowel syndrome	6650	66	0.68
Schizophrenia	4702	67	0.61
Immune diseases	3266	66	0.49
Paralytic syndroms	2752	68	0.45

Table 2: The number of patients above the age of 60 in 2013, who were admitted to Norwegian hospitals during 2008-2011, grouped by diagnosis. Mean age at the time of the first admission during the period with each diagnosis. Proportion of women for each diagnosis is also included. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

The plot in Figure 4 shows the number of patients at the end of the data period having each diagnosis. We can see that diseases in the eye have the highest number of patients, with around 240 000, which is 20% of all patients. The second most prevalent diagnosis is hypertensive diseases, with almost 190 000 patients. Other diagnoses with a lot of patients are ischaemic heart diseases, neoplasms, arrhythmia and diabetes. On the other hand, the least prevelant diagnoses in the data set contains paralytic syndroms, immune diseases, schizophrenia, irritable bowel syndrome and obesity.

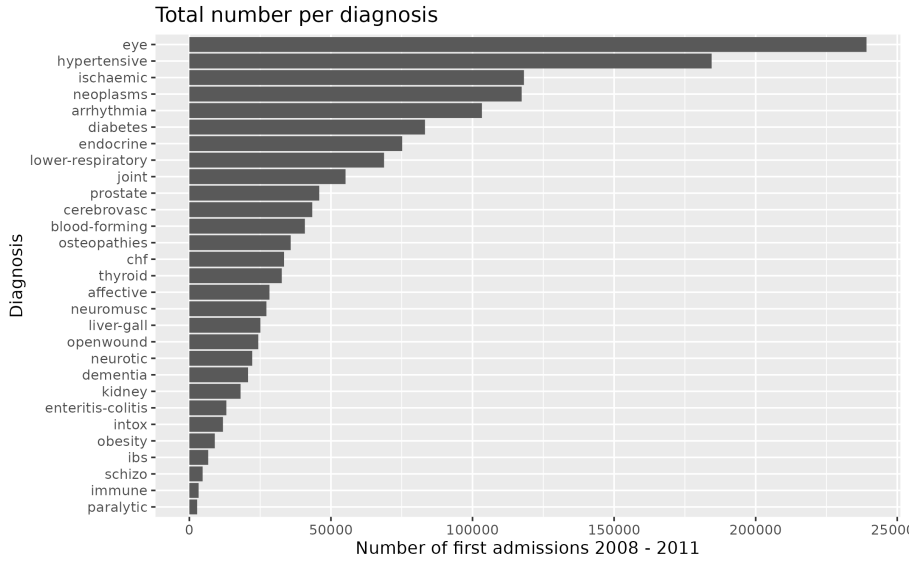


Figure 4: The number of patients above the age of 60 in 2013, who were admitted to Norwegian hospitals during 2008-2021, grouped by diagnosis. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

5.1.1 Most common pairs and triplets

Diag1	**Diag2**	**Prevalence**
hypertensive	eye	0.34
arrhythmia	hypertensive	0.32
endocrine	hypertensive	0.31
hypertensive	ischaemic	0.3
arrhythmia	ischaemic	0.29
neoplasms	eye	0.28
arrhythmia	chf	0.28
arrhythmia	eye	0.28
chf	ischaemic	0.26
hypertensive	neoplasms	0.25
blood-forming	endocrine	0.25
arrhythmia	endocrine	0.25
endocrine	eye	0.24
endocrine	ischaemic	0.24
ischaemic	eye	0.23
endocrine	neoplasms	0.22
arrhythmia	neoplasms	0.21
blood-forming	hypertensive	0.21
affective	neurotic	0.21
blood-forming	neoplasms	0.21

Table 3: Most prevalent combinations of 2 chronic diagnoses in patients admitted to Norwegian hospitals during the period 2008-2021. Prevalence indicates the proportion of individuals with 2 diagnoses among individuals with at least one of them. Data from the Norwegian Patient Registry (NPR). The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

Diag1	Diag2	Diag3	Prevalence
arrhythmia	hypertensive	eye	0.11
arrhythmia	hypertensive	ischaemic	0.11
endocrine	hypertensive	eye	0.1
arrhythmia	chf	ischaemic	0.1
hypertensive	ischaemic	eye	0.1
arrhythmia	endocrine	hypertensive	0.1
endocrine	hypertensive	ischaemic	0.1
hypertensive	neoplasms	eye	0.1
arrhythmia	chf	hypertensive	0.09
arrhythmia	ischaemic	eye	0.09
arrhythmia	endocrine	ischaemic	0.08
chf	hypertensive	ischaemic	0.08
blood-forming	endocrine	hypertensive	0.08
arrhythmia	endocrine	eye	0.08
arrhythmia	chf	endocrine	0.08
endocrine	hypertensive	neoplasms	0.08
arrhythmia	hypertensive	neoplasms	0.08
arrhythmia	neoplasms	eye	0.08
arrhythmia	blood-forming	hypertensive	0.08
arrhythmia	blood-forming	chf	0.08

Table 4: Most prevalent combinations of 3 chronic diagnoses in patients admitted to Norwegian hospitals during the period 2008-2021. Prevalence indicates the proportion of individuals with 3 diagnoses among individuals with at least one of them. Data is from the Norwegian Patient Registry (NPR). The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

5.1.2 Distribution

The plot in Figure 5 shows how each diagnosis is distributed between sex. Hyperplasia of prostate is of course exclusive to men. Some other diagnoses more common in men are ischaemic heart diseases, chronic kidney disease, paralytic syndroms and diabetes. The psychiatric diagnoses seems to be split between the sexes. Mental disorders due to psychoactive substance use is most prevalent in men, while neurotic and somatoform disorders, mood affective disorders and schizophrenia are more prevalent in women. Other diagnosis with a significant majority of women are osteopathies, thyroid gland diseases and irritable bowel syndrome.

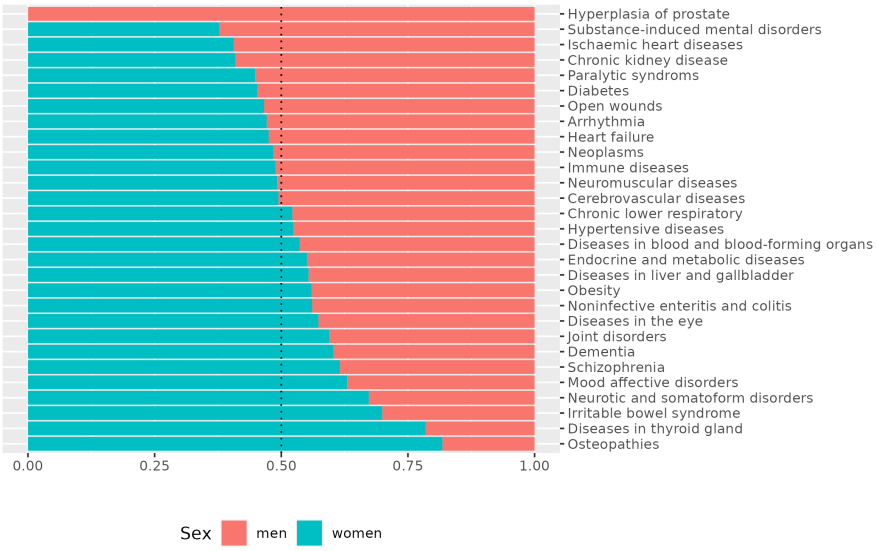


Figure 5: The sex of patients above the age of 60 in 2013, who were admitted to hospitals in Norway from 2008-2021 with at least one chronic diagnosis, plotted for each diagnosis group. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included. Data is from the Norwegian Patient Registry (NPR).

Similarly, the box plot in Figure 6 shows the age distribution of patients for each diagnosis. Note that the age of dementia patients are the oldest group. Other diagnoses with a high median age are heart failure and chronic kidney disease. The diagnoses with the youngest patients at the time of first admission include mental disorders due to psychoactive substance use, obesity, immune diseases and irritable bowel disease.

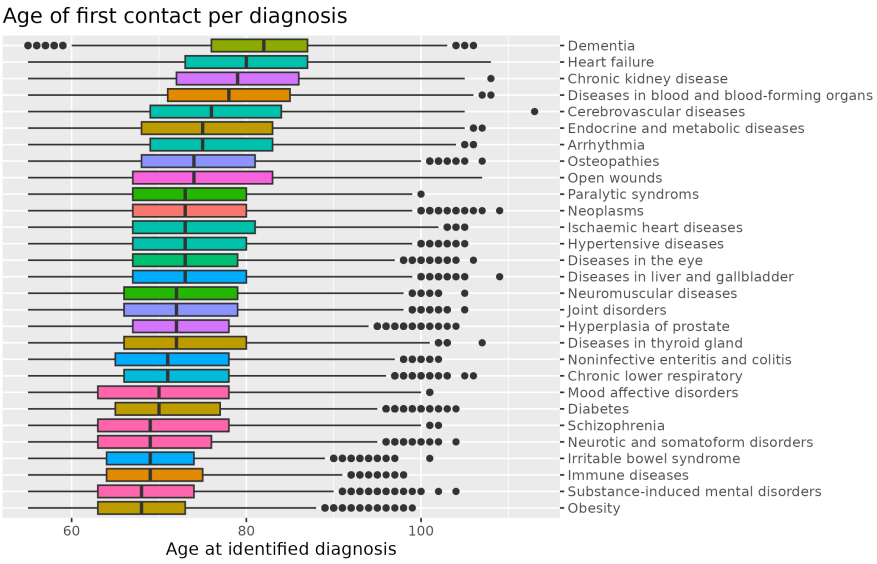


Figure 6: The age of patients above the age of 60 in 2013, who were admitted to hospitals in Norway from 2008-2021 with at least one chronic diagnosis, plotted for each diagnosis group. The age is based on the age of the patient at the time of first admission assosiated with the diagnosis. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included. Data is from the Norwegian Patient Registry (NPR).

5.1.3 Non-linear age with splines

Consider Figure 7, containing the association of age on the predicted number of GP days in our model for patients with arrhythmia. We also observed similar associations for the other diagnoses

and outcomes. Note the clear non-linear relationship of age, validating the need for the use of cubic splines in adjusting for age.

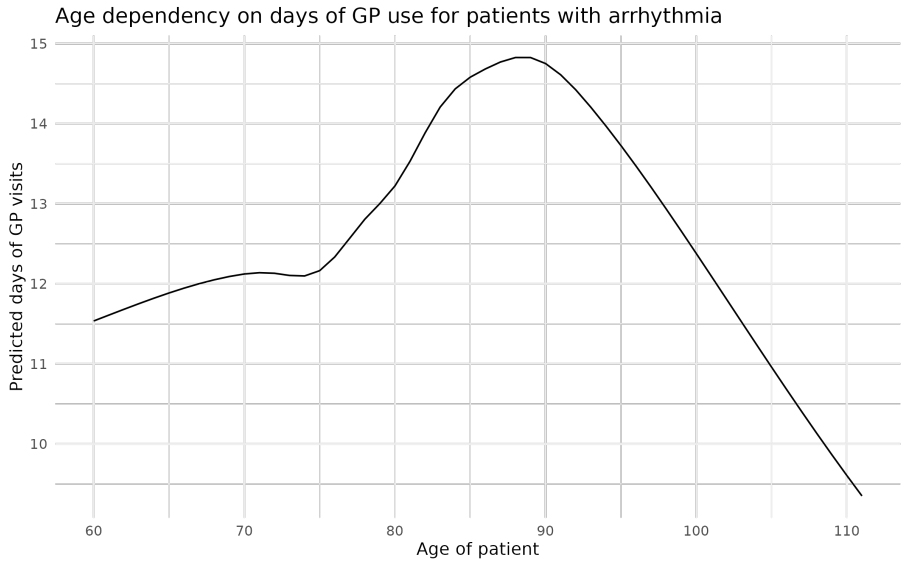


Figure 7: The predicted number of days of GP visits during 1 year for patients in Norway with arrhythmia, as a function of the age of the patient. Predictions is calculated using Poisson regression model, accounting for exposure time. Data is from the Norwegian Patient Registry (NPR) and Control and Payment Health Reimbursement register (KUHR).

5.1.4 Forecast of population

The following plot clearly illustrates the aging population. Note that the number of people in Norway above the age of 65 is predicted to double during the next 60 years, while the number of people of working age is predicted to decrease.

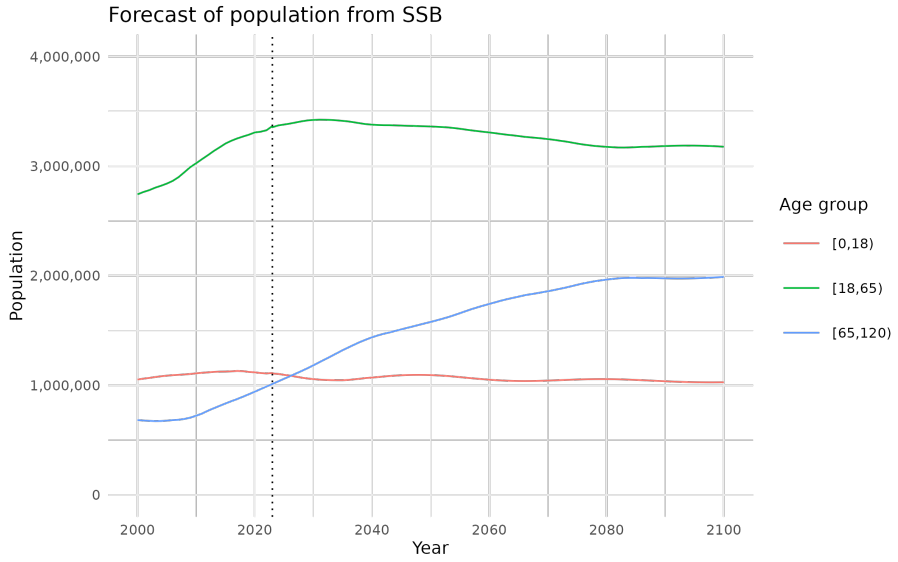


Figure 8: History and forecast of the population size for the age groups [0,18), [18,65), [65,120) from the year 2000 until 2100, based on data from National population projections by SSB [40].

5.1.5 Number of diagnoses per age

Consider now Figure 9, showing how the number of chronic diseases are distributed among four different age groups. Note that about 60 % of patients between 60 - 70 had no diagnoses identified

during the entire period 2008-2021. Along with the age groups, the distribution changed to more and more diagnoses. For patients older than 90 years, almost half of the patients had 2 or more diagnoses. This is consistent with the notion that chronic diseases start to accumulate along with age.

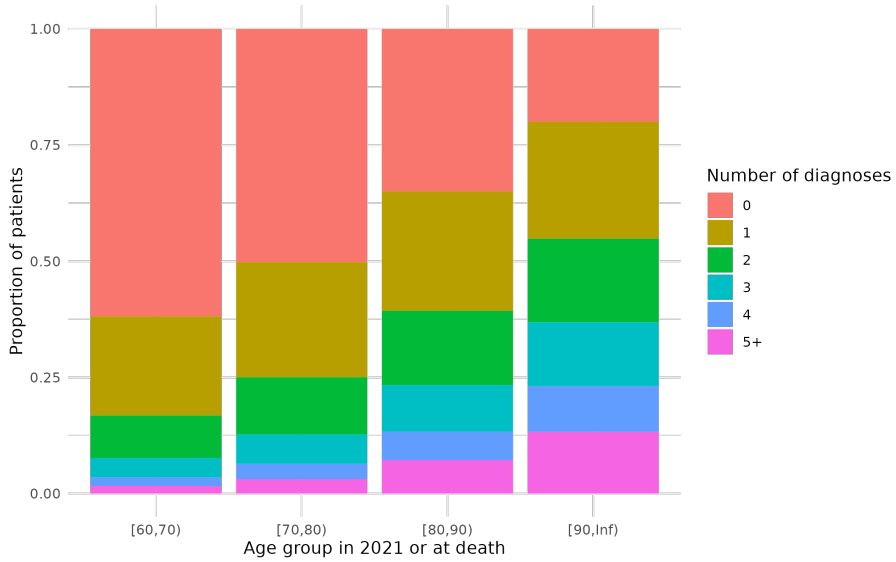
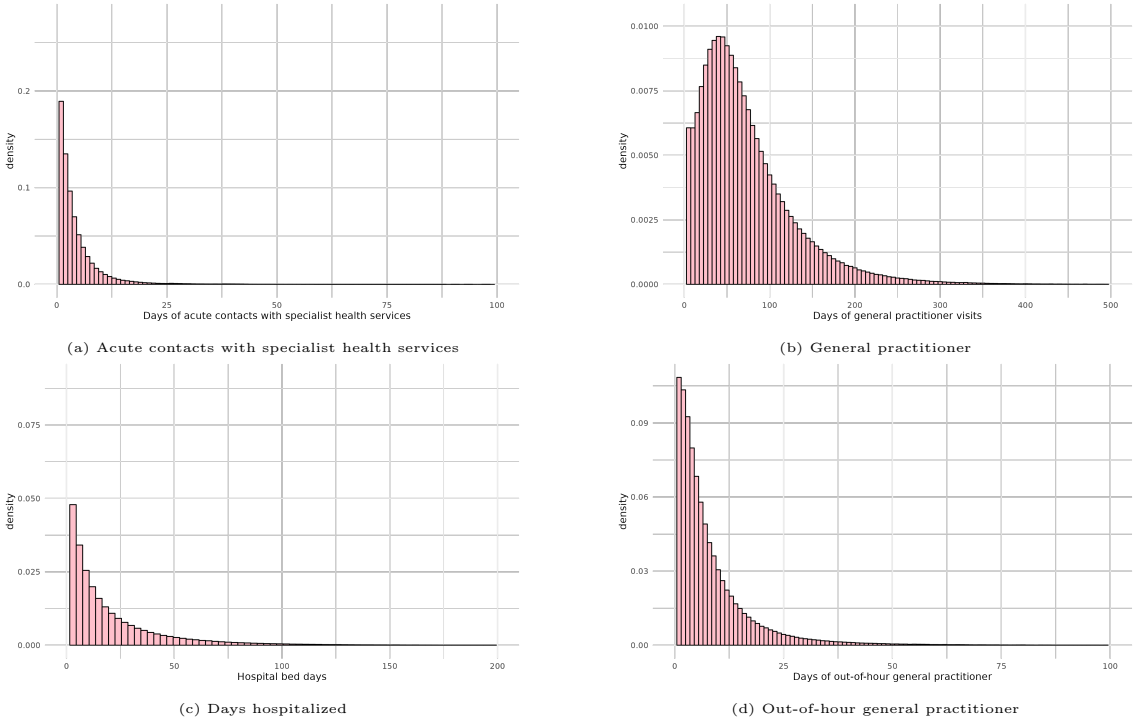


Figure 9: The number of diagnoses registered per patient as main or secondary diagnosis for admissions to Norwegian hospitals during the period 2008-2021, plotted against the age groups of the patients. The data contains all admissions to hospitals in Norway, for patients above 60 years in 2013. The age groups are based on the age of the patients in either 2021 or at time of death. Data is from the Norwegian Patient Registry (NPR)

5.1.6 Distribution of health service use

The plots in Figure 10 shows the distributions of the outcome variables. It should be noted that they appear to be distributed according to a Poisson distribution.



(a) Acute contacts with specialist health services

(b) General practitioner

(c) Days hospitalized

(d) Out-of-hour general practitioner

Figure 10: Distribution of number of days of health service use during the period 2008-2021 for Norwegian patients above the age of 60 in 2013. Data is from Control and Payment of Health Reimbursement register (KUHR) and ...

5.2 Model 1 - Risk of diagnosis

The results from model 1 can be seen in Figure 11. The first results to note are the arrows from cerebrovascular and neuromuscular pointing to paralytic, with 166% and 141% respectively, meaning a higher risk of being diagnosed with paralysis for patients with neuromuscular and cerebrovascular diseases. Another strong relation is the two way link between diabetes and obesity, with 135% for underlying diabetes and 161% for underlying obesity, which can be interpreted as a higher risk of obtaining the other for patients with either diabetes or obesity. Furthermore, the model shows that patients with diabetes have a 104% increased risk of being diagnosed with kidney diseases, and patients with kidney and joint diseases have 227% and 122% increased risk of diagnosis with immune diseases. Another significant cluster consists of several psyciatric diseases, including schitzophrenia, affective, neurotic and intox. It is also notable that neoplasms are further away from all diseases, with a single link pointing to immune diseases. Other diseases with few to no strong relations are dementia, endocrine and lower-respiratory. Prostate also has a single two way link with irritable bowel syndrome.

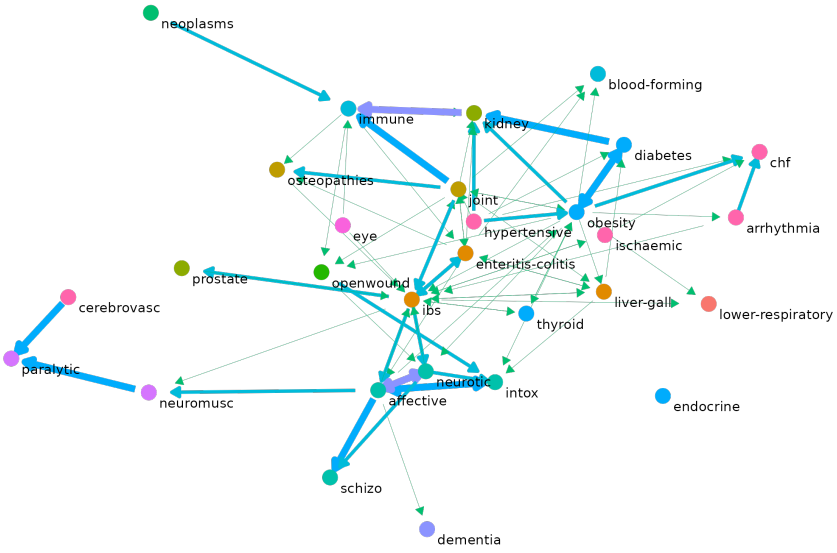


Figure 11: Network map showing associations with increased risk of new diagnosis for patients above the age of 60 in 2013, with specific underlying diagnoses. Arrows points to the diagnosis with increased risk given a history of the diagnosis the arrow points from. The interactions are incidence-rate ratios based on log-linear Poisson model, with covariates being prevalence of underlying diagnoses, sex and age of patients, including offset of exposure time from 2013 until either death or 2021. Outcome variable is the event of the patient having an admission to a hospital associated with a new diagnosis. Data set based on all admissions to Norwegian hospitals between 2008-2021 for patients older than 60 years in 2013. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

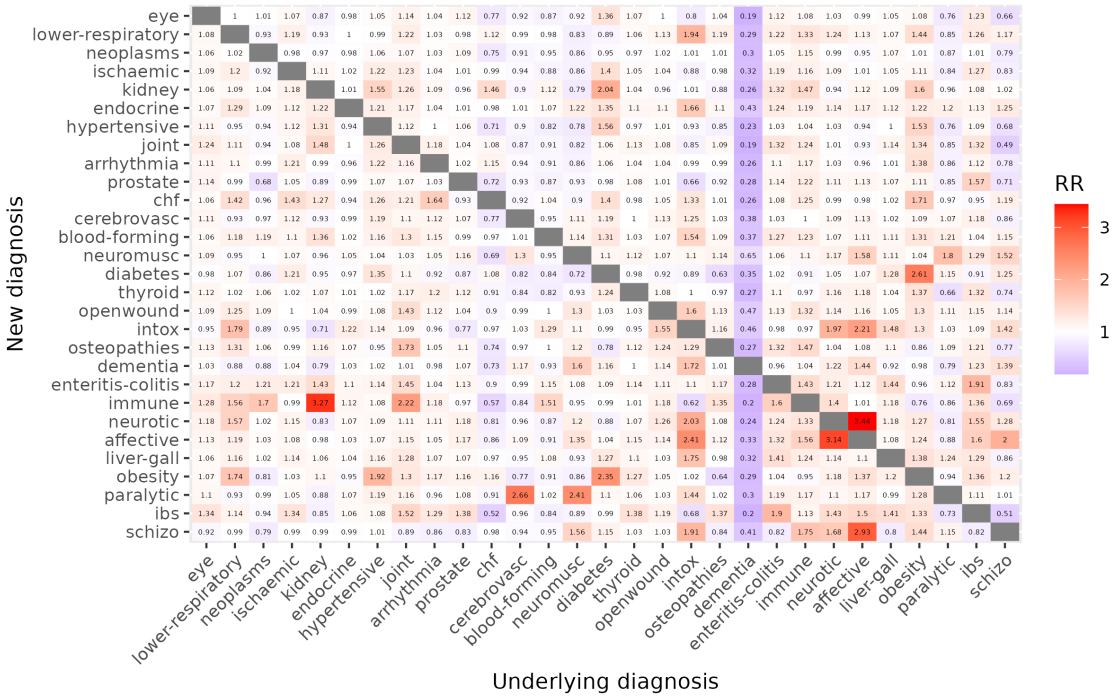


Figure 12: Heatmap showing associations with relative risk of new diagnosis for patients above the age of 60 in 2013, with specific underlying diagnoses. The interactions are incidence-rate ratios based on log-linear Poisson model, with covariates being prevalence of underlying diagnoses, sex and age of patients, including offset of exposure time from 2013 until either death or 2021. Outcome variable is the event of the patient having an admission to a hospital associated with a new diagnosis. Data set based on all admissions to Norwegian hospitals between 2008-2021 for patients older than 60 years in 2013. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

5.3 Model 2 - Health Service Utilization

5.3.1 Acute

We begin with Figure 13. The lines between diseases can be interpreted as pairs of diseases that when being present in the same patient can cause increased use of acute specialist health services. Note that irritable bowel syndrome has 6 lines connecting it to other diagnoses. Thus patients with irritable bowel syndrome who also have those specific diseases are prone to visit acute health services at least 7.5% more. Other diagnoses with more interactions are paralysis, immune diseases, obesity and prostate. The strongest increase is in patients with immune diseases, who also have neurotic/somatoform disorders (49%) or substance-induced mental disorders (42%). Other strong interactions are present in patients with obesity, who also are schizophrenic (33%) or have neuromuscular diseases (35%). Some diagnoses, like diabetes, neoplasms and arrhythmia have no links in the network map, indicating that the model found no significant increase in use of acute specialist health services for patients associated with multimorbidity due to those particular diagnoses.

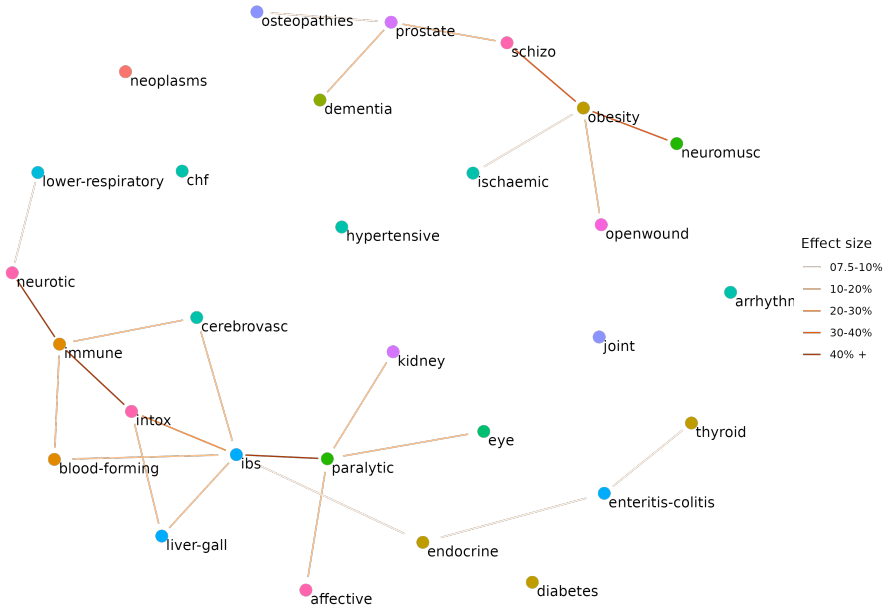


Figure 13: Network map showing pairs of diagnoses associated with increase in number of days with acute health service use per patient above the age of 60. Lines between two diagnoses corresponds to at least 7.5% increase for patients with history of both diagnoses between 2008 - 2011. The interactions are incidence-rate ratios based on log-linear Poisson model, with covariates being prevalence of underlying diagnoses, pairwise interactions between the diagnoses, sex of patients, non-linear age using cubic splines, including offset of exposure time from 2013 until either death or 2021. Outcome variable is the number of days with acute health service use during the exposure period. Data set based on all admissions to Norwegian hospitals between 2008-2021 for patients older than 60 years in 2013. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

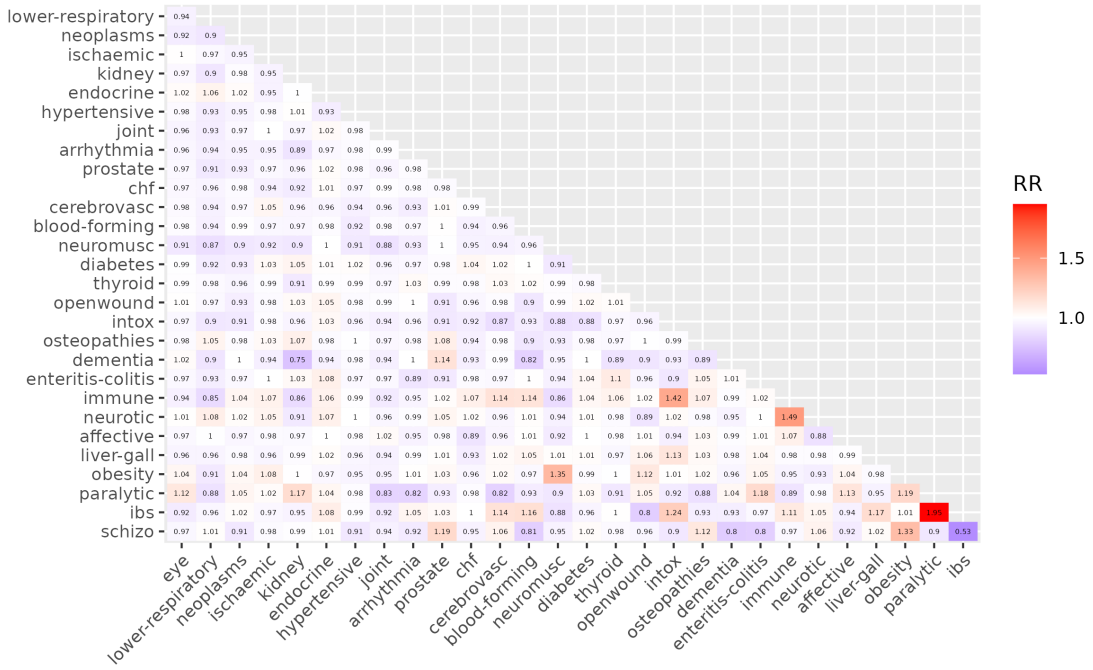


Figure 14: Pairs of diagnoses associated with increase (red) or decrease (blue) in number of days with acute health service use per patient above the age of 60. Numbers can be interpreted as the relative increase in number of days with acute health service. Results are based on log-linear Poisson model, with covariates being prevalence of underlying diagnoses, pairwise interactions between the diagnoses, sex of patients, non-linear age using cubic splines, including offset of exposure time from 2013 until either death or 2021. Outcome variable is the number of days with acute health service use during the exposure period. Data set based on all admissions to Norwegian hospitals between 2008-2021 for patients older than 60 years in 2013. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

5.3.2 General practitioner

Compared to the previous plot, the interactions in Figure 15 are mainly centered around dementia. Thus patients with dementia who also have the diagnoses linked to it in the plot, are prone to have more days with visits to their general practitioner. Immune diseases, obesity and paralytic have similarly central placements. The strongest interactions are present in dementia patients who also have been diagnosed with hyperplasia of prostate (32%) or obesity (30%), as well as patients with a combination of immune diseases and paralytic syndromes (29%). Another thing to note is that more diseases in this plot have no lines connecting them, indicating no strong increase in GP use specifically associated with combinations including those.

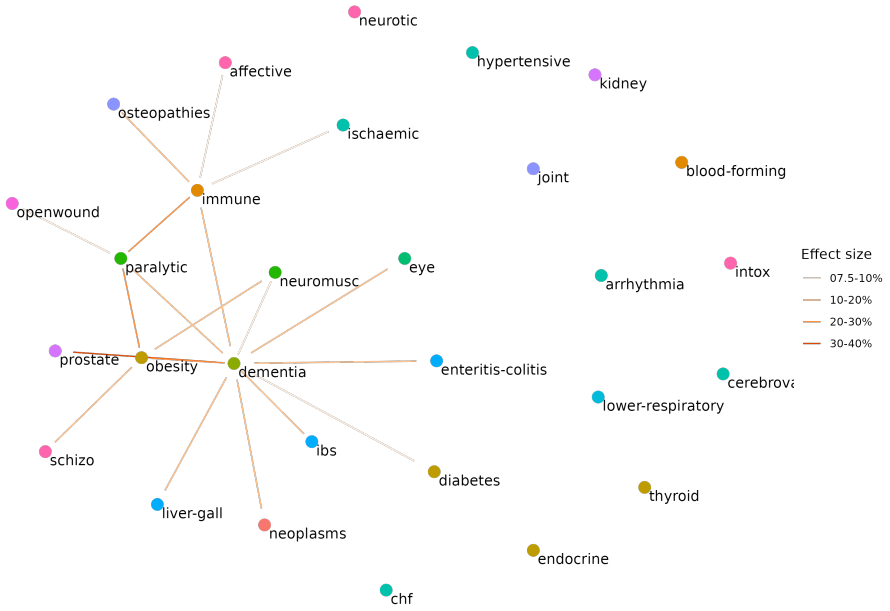


Figure 15: Network map showing pairs of diagnoses associated with increase in number of days with visits to general practitioner per patient above the age of 60. Lines between two diagnoses corresponds to at least 7.5% increase for patients with history of both diagnoses between 2008 - 2011. The interactions are incidence-rate ratios based on log-linear Poisson model, with covariates being prevalence of underlying diagnoses, pairwise interactions between the diagnoses, sex of patients, non-linear age using cubic splines, including offset of exposure time from 2013 until either death or 2021. Outcome variable is the number of days with visits to general practitioner during the exposure period. Data set based on all admissions to Norwegian hospitals between 2008-2021 for patients older than 60 years in 2013. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

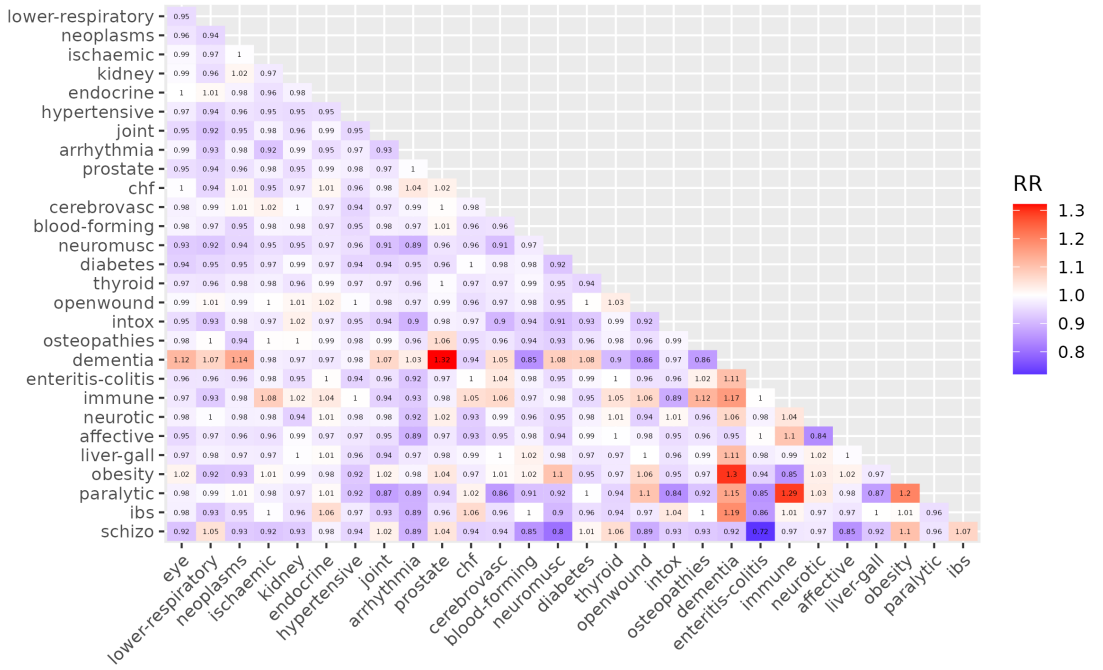


Figure 16: Pairs of diagnoses associated with increase (red) or decrease (blue) in number of days with visits to general practitioner per patient above the age of 60. Numbers can be interpreted as the relative increase in number of days with acute health service. Results are based on log-linear Poisson model, with covariates being prevalence of underlying diagnoses, pairwise interactions between the diagnoses, sex of patients, non-linear age using cubic splines, including offset of exposure time from 2013 until either death or 2021. Outcome variable is the number of days with acute health service use during the exposure period. Data set based on all admissions to Norwegian hospitals between 2008-2021 for patients older than 60 years in 2013. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

5.3.3 Days hospitalized

Figure 15 shows combinations of diagnoses associated with increased number of days hospitalized. Diagnoses with many interactions include obesity, schizophrenia, immune diseases, chronic kidney disease and irritable bowel syndrome. Strong increases in days hospitalized can be found in patients with immune diseases, who also have neurotic/somatoform (33%) or mood affective disorders (30%), irritable bowel syndrome or endocrine diseases. Other strong associations can be found in patients with the following pairs of diagnoses: dementia and hyperplasia of prostate (33%), hyperplasia of prostate and schizophrenia (35%), schizophrenia and kidney diseases (54%), kidney diseases and paralytic syndroms (37%). Note that patients with either disease in the eye, joint disorders, arrhythmia or neoplasms in combination with other diseases, have no increased number of days hospitalized due to the specific combination.

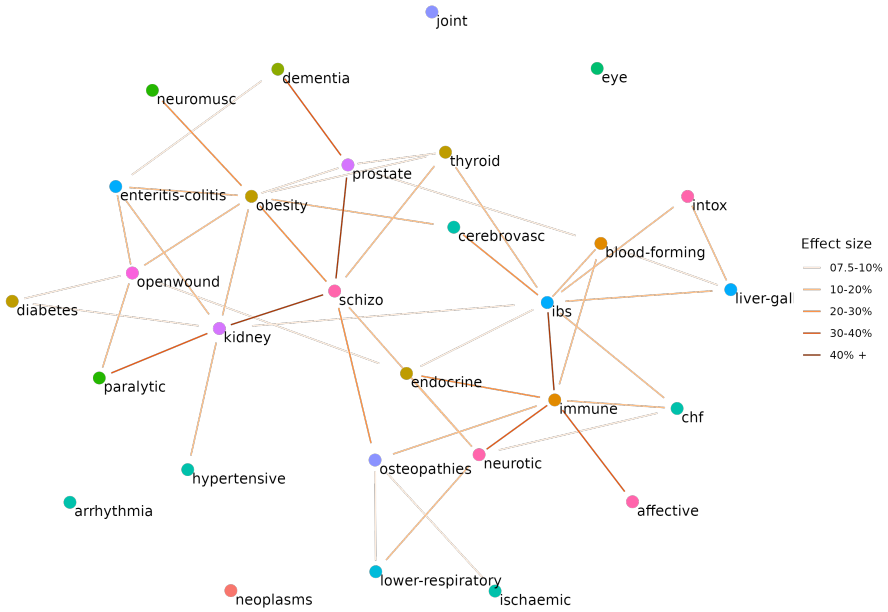


Figure 17: Network map showing pairs of diagnoses associated with increase in number of days hospitalized per patient above the age of 60. Lines between two diagnoses corresponds to at least 7.5% increase for patients with history of both diagnoses between 2008 - 2011. The interactions are incidence-rate ratios based on log-linear Poisson model, with covariates being prevalence of underlying diagnoses, pairwise interactions between the diagnoses, sex of patients, non-linear age using cubic splines, including offset of exposure time from 2013 until either death or 2021. Outcome variable is the number of days hospitalized during the exposure period. Data set based on all admissions to Norwegian hospitals between 2008-2021 for patients older than 60 years in 2013. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

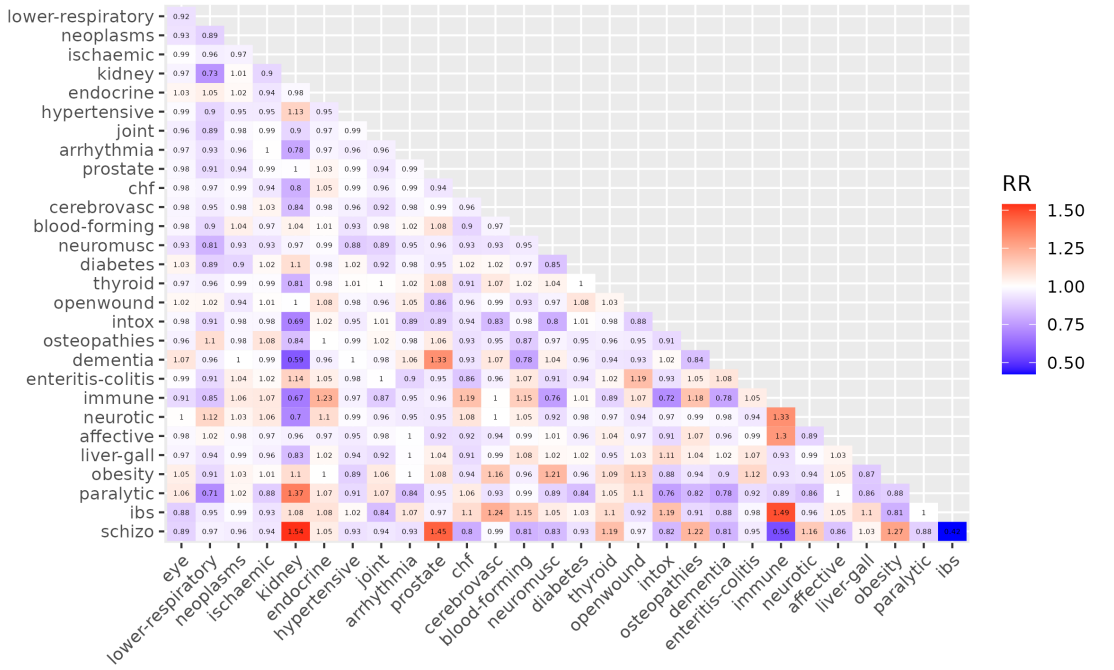


Figure 18: Pairs of diagnoses associated with increase (red) or decrease (blue) in number of days hospitalized per patient above the age of 60. Numbers can be interpreted as the relative increase in number of days with acute health service. Results are based on log-linear Poisson model, with covariates being prevalence of underlying diagnoses, pairwise interactions between the diagnoses, sex of patients, non-linear age using cubic splines, including offset of exposure time from 2013 until either death or 2021. Outcome variable is the number of days with acute health service use during the exposure period. Data set based on all admissions to Norwegian hospitals between 2008-2021 for patients older than 60 years in 2013. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

5.3.4 Out-of-hour general practitioner

Finally, Figure 19 shows results on utilization of out-of-hour general practitioner. Strong associations can be found in schizophrenic patients who also have been diagnosed with substance-induced mental disorders (25%), chronic lower respiratory diseases (26%), cerebrovascular diseases (23%) or obesity (29%). Similarly, patients with obesity who also have dementia (27%), open wounds (21%), neuromuscular diseases (21%), paralytic syndromes (20%) or schizophrenia (29%) have an increased use. Immune diseases also have several interactions, including neurotic/somatoform disorders (22%) and mood affective disorders (15%). Finally, a strong increase, at 55%, is found in patients with irritable bowel disease and paralytic syndromes.

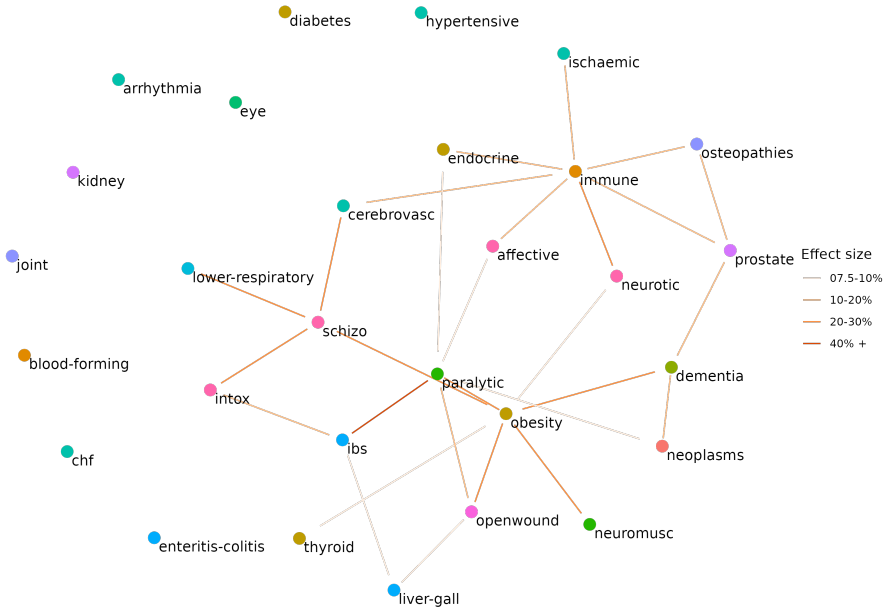


Figure 19: Network map showing pairs of diagnoses associated with increase in number of days with visits to out-of-hour general practitioner per patient above the age of 60. Lines between two diagnoses corresponds to at least 7.5% increase for patients with history of both diagnoses between 2008 - 2011. The interactions are incidence-rate ratios based on log-linear Poisson model, with covariates being prevalence of underlying diagnoses, pairwise interactions between the diagnoses, sex of patients, non-linear age using cubic splines, including offset of exposure time from 2013 until either death or 2021. Outcome variable is the number of days with visits to out-of-hour general practitioner during the exposure period. Data set based on all admissions to Norwegian hospitals between 2008-2021 for patients older than 60 years in 2013. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

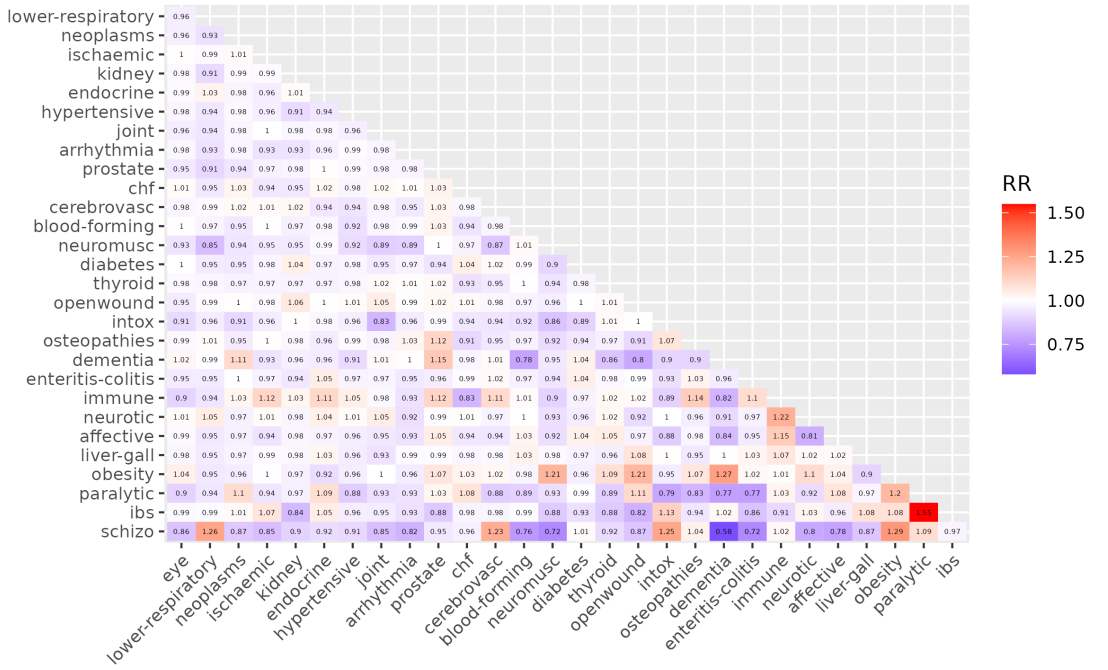


Figure 20: Pairs of diagnoses associated with increase (red) or decrease (blue) in number of days with visits to out-of-hour general practitioner per patient above the age of 60. Numbers can be interpreted as the relative increase in number of days with acute health service. Results are based on log-linear Poisson model, with covariates being prevalence of underlying diagnoses, pairwise interactions between the diagnoses, sex of patients, non-linear age using cubic splines, including offset of exposure time from 2013 until either death or 2021. Outcome variable is the number of days with acute health service use during the exposure period. Data set based on all admissions to Norwegian hospitals between 2008-2021 for patients older than 60 years in 2013. The list of diagnoses are based on the definition of complex multimorbidity by The Norwegian Directorate of Health [33], excluding Amputation and Transplant. Schizophrenia, Mental disorders due to psychoactive substance use, Neurotic and somatoform disorders, and Mood affective disorders are also included.

6 Discussion

6.1 Summary

Descriptive analysis illustrated the potential increase in population, associated with the number of people above the age of 65 doubling in the next 60 years. Furthermore, Figure 9 depicted a rise in the number of diagnoses with age, with almost half of patients above 90 having at least 2 diagnoses. The most common combinations of diagnoses among patients, measured as the prevalence of 2 or 3 diagnoses in patients with at least 1 of them, were found to include hypertensive diseases, arrhythmia, endocrine diseases, ischaemic heart diseases, eye diseases, neoplasms and chronic heart failure.

Various associations were identified, indicating an elevated risk of receiving a specific diagnosis for patients with a history of another diagnosis. For example an increased risk of being diagnosed with paralytic syndromes for patients with cerebrovascular or neuromuscular diseases. Several relations were also found between the psychiatric diseases schizophrenia, mood affective disorders, neurotic/somatoform disorders and substance-induced mental disorders. A two-way relationship was found between diabetes and obesity, indicating a higher risk for patients with one of them of receiving the other. An increased risk was also found for diabetes patients of being diagnosed with chronic kidney diseases. Finally, a higher risk of receiving a diagnosis of immune diseases were found for patients with chronic kidney diseases, joint disorders or neoplasms.

Furthermore, several combinations of diagnoses were found to be associated with increase or decrease in health service utilization. Combinations leading to increased use of acute specialized health services include obesity in combination with either schizophrenia or neuromuscular diseases, a combination of irritable bowel syndrome and paralytic syndromes, and immune diseases in combination with neurotic/somatoform disorders or substance-induced mental disorders. The next outcome, the number of days of visits to general practitioner, were found to increase for patients with obesity who also have been diagnosed with dementia, hyperplasia of prostate or paralytic syndromes, as well as patients with both paralytic syndromes and immune diseases. Additionally, having dementia in combination with a range of diagnoses, indicate increase in visits to general practitioner. The third measure, number of days hospitalized, were found to increase for immune disease patients who also have irritable bowel syndrome, mood affective mental disorders or neurotic/somatoform disorders. Additionally, the following pairs also led to increased number of days hospitalized: dementia and hyperplasia of prostate, hyperplasia of prostate and schizophrenia, schizophrenia and chronic kidney disease, chronic kidney disease and paralytic syndromes. Finally, use of out-of-hour general practitioners increased for patients with both irritable bowel syndrome and paralytic disorders. The diagnoses immune disease, obesity and schizophrenia also led to increased use when in combination with some other diagnoses.

6.2 Interpretation of results

Before proceeding with interpreting the results, it is essential to have a clear understanding of what is modelled. First and foremost, we are looking at medical conditions which have been diagnosed at a hospital. Thus we do not have information about when the condition first appeared within the patient. Some diseases may be latent for years before a diagnosis is set. Furthermore, suppose that a person has some latent diseases which have not been diagnosed, and are not impacting their life. Suppose then that they experience a change in symptoms which causes them to visit a doctor, and are sent to a hospital. That hospital admission may lead to a thorough examination where several diagnoses may be made, each of which may or may not be serious. This illustrates that the analysis does not necessarily show how individual diseases are connected, as it may be influenced by how the health service system is organised.

The descriptive analysis illustrates the challenge of the aging population. Considering both the forecast from SSB in Figure 8, and Figure 9 illustrating that the number of diagnoses increases with age, then the average number of diagnoses in the population will increase. This means that the challenges related to multimorbidity will be present in a larger fraction of patients. These

challenges include complications in treatment, coordination of medication and increased risk of developing other diseases.

Next we consider the most common pairs and triplets. Note that hypertensive diseases, also known as high blood pressure, is the 2nd most common diagnosis in the data set, meaning that a lot of patients have been diagnosed with the condition. The four most common pairs and the three most common triplets include hypertensive diseases, which indicates that the conditions eye diseases, arrhythmia, endocrine and ischaemic may be related to high blood pressure. This is consistent with known relations, where high blood pressure is a risk factor of eye disease such as hypertensive retinopathy, arrhythmias like atrial fibrillation, and ischaemic heart disease [41, 42, 43].

6.2.1 Relative risk

This brings us to the question of whether there is a change in risk of being diagnosed with a condition, associated with having an underlying diagnosis. The results from model 2 indicated several such associations, illustrated in Figure 11. Consider first the increase in risk of paralytic syndromes associated with cerebrovascular and neuromuscular diseases. Cerebrovascular diseases, for example stroke, affect the blood vessels supplying the brain, which can lead to damage or dysfunction in certain areas of the brain. This may lead to paralysis in various body parts, depending on the location and severity of the damage [44]. Neuromuscular disease, on the other hand, for example amyotrophic lateral sclerosis (ALS), include conditions affecting the nerves or muscles, which also may lead to muscle weakness and paralysis [45].

Another relation indicated in the model, is the two way interaction between diabetes and obesity. There is a strong association between obesity and diabetes, where obesity is the most significant risk factor for diabetes [46]. Obesity may lead to a reduced ability to use insulin, which is a hormone for regulating blood sugar levels, referred to as insulin resistance. This can lead to type 2 diabetes over time. As mentioned above, a limitation of basing the model on diagnoses set at hospitals, is that there is no information about the order in which diagnoses are set. Furthermore, the model indicates that diabetes leads to an increased risk of chronic kidney diseases, also known as chronic renal failure. Diabetes is known to cause poor blood circulation through small blood vessels, and the kidney is especially sensitive. The association is well-established: "Diabetes is considered the commonest cause of end-stage renal disease" [47]. Diabetes can also lead to a wide range of other complications, including heart attacks, stroke and eye diseases. Our model does not indicate strong associations from diabetes to other conditions, apart from obesity and chronic kidney disease. This may be an indication that diabetes is well controlled among the population in Norway.

Furthermore, the model indicates that chronic kidney diseases and joint disorders are associated with an increased risk of being diagnosed with immune diseases. The kidneys are the main organs responsible for clearing metabolic waste products and toxins, and are contributing to maintaining the balance (homeostasis) of the immune system. In chronic kidney diseases, the immune system can become severely compromised, which can cause immune diseases [48]. Additionally, certain joint diseases, for example rheumatoid arthritis (RA), are considered autoimmune diseases. These will mistakenly attack the body's tissues, especially the joints, which in turn can lead to the development of other immune diseases [49].

6.2.2 Health service utilization

We will now proceed to discuss how specific multimorbidites can affect health service use, i.e. the results of model 2. We will interpret the results to the best of our ability, but the interpretations are not straight forward, which we will see. First we will go through the results of model 2 for each of the outcomes; acute contacts with specialized health services, visits to general practitioner and out-of-hour general practitioner, as well as days hospitalized.

Starting with the increase in number of acute contacts with specialized health services, the strongest associations are between immune diseases and the psychiatric neurotic/somatoform disorders and

substance-induced mental disorders. Immune diseases have potential for weakening the immune system, making the patient more susceptible to infections. Some mental disorders can be associated with negative behaviors and lifestyle factors which may increase the effect of these infections, leading to an increase in frequency of contacts with acute health services. Additionally, both immune diseases and mental disorders have potential for episodes of exacerbations or flare-ups with severe symptoms, which may lead to hospitalizations. Finally, both immune diseases and mental disorders often include medication as part of treatment, which may cause interactions or side effects. Specifically substance-induced mental disorders may be complicated when combined with immune diseases, as drugs may increase the severity of inflammations. Another association found in the model is between schizophrenia and obesity. One potential reason for this can be antipsychotic-associated weight gain, in other words weight gain associated with medication prescribed for schizophrenia [50]. Additionally, both schizophrenia and obesity may require comprehensive and integrated approaches at treatment, where patients suffering from both may have challenges adhering to strict treatment regiments or managing lifestyle factors. Finally, patients who have paralysis in addition to irritable bowel syndrome may have severe gastrointestinal complications due to the paralysis potentially impacting bowel function. Paralysis may also impact the ability to swallow, which in combination with IBS may lead to nutritional deficiencies and dehydration.

Next we will analyse the associations leading to increased load on general practitioners. First, dementia has a lot of relations with other diagnoses. This may indicate some difficulties related to specifically dementia. One example is that patients with dementia may have a decreased ability to adhere to treatment regiments, and have trouble communicating both pain and symptoms. Therefore, regular GP visits may be required to closely monitor complications associated with other diagnoses. Another reason may be that family and caregivers may want to seek advice from the general practitioners, to discuss challenges and tips. Hyperplasia of prostate and obesity seem to have the strongest association with dementia. Another strong relation from the model is the combination of immune diseases and paralysis.

One of the strongest relations found by the model associated with increased number of days hospitalized is the combination of irritable bowel syndrome (IBS) and immune diseases. This may be due to several factors, for example that some immune diseases may weaken the immune system and lead to infections in various organs, including the gastrointestinal tract, which is also the organ affected by IBS. This may lead to severe complications. Additionally, both IBS and immune diseases may have periods of flare-ups or exacerbations, so the combination can potentially increase both risk and severity of these. Other strong relations are for schizophrenic patients who also have chronic kidney disease or hyperplasia of prostate, and patients having immune diseases combined with mood affective- or neurotic/somatoform disorders.

Finally, several combinations affecting load on out-of-hour general practitioners are centered around immune diseases, obesity and schizophrenia. The strongest association is the combination of paralytic syndroms and IBS. This may potentially be explained similarly to the association on acute contacts with specialized health services, for example that they may have severe gastrointestinal complications.

Note that all models indicate various combinations leading to decrease in health service use. As mentioned, this may be due to several factors, including what is referred to as selection or collider bias [51]. In our case, by conditioning on the survival of the patients, we also condition on the severity of the diseases. Patients with a combination of 2 severe diagnoses, will probably have a low survival rate, and therefore our model is primarily based on patients with milder versions of those diseases.

While we have interpreted the results of the analysis, it is important to note that we are not health professionals or doctors. However, the results of our analysis appear to coincide with established knowledge in the fields. This reinforces our confidence that our models are accurate.

6.3 Implications

We will now discuss some potential implications of the models. Suppose a tool is developed which takes into account a range of underlying diseases, and predicts the conditions with the highest risk of being diagnosed. This would be a powerful tool for the general practitioner, who is responsible for coordinating all diseases in their patients. For each patient, the general practitioner is presented with potential conditions the patient is more likely to develop, as well as symptoms to be especially aware of. This has potential for earlier detection of diseases, reducing human error and lessening the load on general practitioners. Furthermore, out-of-hour general practitioners and acute specialized health services can have similar use of the model, as well as being presented with complications which needs to be monitored in potential hospitalizations.

Additionally, the model may have several use cases in further research, especially a model predicting health service use associated with combination of conditions. It may for example be used for finding associations that produce higher load, which can be further researched. Considering the aging population in the world, this model can also be used for predicting the added health care use resulting from the older population. Recall the plot of the forecast of the number of people over the age of 65 from SSB data in the descriptive analysis, estimated to double in the next 60 years. Combining this with the model can give a prediction of the average increase in load for health service providers like general practitioner and acute admissions, helping policy makers and others in their decision-making process. Older people tend to have more health problems, including conditions related to frailty. Consider for example this study [52], showing benefits of orthogeriatric care for older patients with hip-fracture. Perhaps the organizations around single-diseases in hospitals is not the most suitable option for the future, and alternative organisation structures focusing on interdisciplinary treatments should be considered [28, 29].

6.4 Comparison with existing literature on data and MM

A range of studies have explored various ways of visualizing the relationship between diagnoses. One study have for example modelled common conditions associated with type 2 diabetes, and created a visual representation of diseases central in the multimorbidity network [16]. Another study used partial correlation combined with network analysis to categorize conditions into sub-groups [14]. In comparison to prior research we have demonstrated how the relative risk of being diagnosed with a chronic diagnosis, given an underlying diagnosis, can be modelled according to Poisson regression. We have also presented a clear and visually intuitive representation of the results using graph drawing by force-directed placement.

A few studies have researched how multimorbidity is associated with health care use, but are mostly limited to the prevalence of multimorbidity, or the association from larger clusters [23, 24, 11]. We have demonstrated a way to include each specific diagnosis, and demonstrated a model of the increased/decreased load on various measures of health service utilization associated with the prevalence of specific combinations of diagnoses. Additionally, we have also here visualized the relations in a clear and intuitive way using graph drawing by force-directed placement.

6.5 Strengths and limitations

A major strength of our study is the quality of the data set available to us, which contains complete information about all admissions to hospitals in Norway during the period 2008-2021, as well as the number of days of utilization for each patient of general practitioner, out-of-hour general practitioner, acute contacts with specialized health services and days hospitalized. Additionally, we have utilized Poisson models for their robustness when implemented with large data sets, which results in very small p-values. Another strength of our study is the novel way of visualizing relative risk between several diagnoses, as well as combinations of diseases associated with increased health service use. Finally, we are including a wide range of diagnoses simultaneously, which can provide deeper insight into their relationships.

There are also some limitations to consider. Because we are considering diagnoses set at hospitals, we do not have information about the exact point in time the disease appeared in the patients. If for example a patient is admitted to a hospital and diagnosed with obesity and diabetes, we have no information of the chronological order of the physical diseases. Furthermore, we can investigate the association between the prevalence of diagnoses and health care use. However, several factors complicates the causal interpretation of the results. One factor is that certain diagnoses may lead to a higher risk of patients being admitted to nursing homes, and for the rest of the data period they are treated there instead of at regular health services. Likewise, patients with complex multimorbidities may stop visiting general practitioners, since all future follow-up occurs at the hospital. One potential complication of these factors is that some combinations of diagnoses indicates decreased health service utilization, which may not be straight forward to account for. Furthermore, as mentioned earlier, we need to be aware of collider biases, since patients with several serious conditions have a low survival rate.

6.6 Further Research

The results of the analysis presented in this thesis have strengthened our belief in using Poisson regression for modelling relative risk of diseases, and increase in health service use. We have provided a basic interpretation of the results of the analysis, but a more thorough interpretation by health professional and doctors should be made. Additionally, we can see several potential ways of further expanding the model, for example by investigating associations between 3 or more diagnoses. This may lead to more accurate results, and potentially uncovering hidden multimorbid relationships generating increased stress on health services. Another potential direction for expansion is extrapolating the model using the forecast of age distribution of the population in Norway, presented by SSB [40]. Considering the predicted distribution of age, it is possible to predict the expected number of days of health service use, for example in the year 2100. Considering the ageing population predicted to impact the western world, a measure of the extra load on health services associated with this increase may be a powerful planning tool for health professionals and policy makers. Finally, other techniques could further strengthen the model, making the results easier to interpret. One possible directions for this are experimental emulation [53], which refers to emulating a physical experiment using models. Another direction is sibling analysis [54], where data from siblings are used to control for various factors.

7 Conclusion

Through this project, we have utilized registry data to evaluate several aspects associated with multimorbidity. Our primary objectives were to investigate the associations between chronic diseases, in addition to how specific combinations of these relate to health service use. To address this, we have demonstrated several models. We developed a model for the increase or decrease in risk of being diagnosed with specific diseases for patients with various underlying diseases, as well as models for the change in health care utilization associated with specific combinations of chronic diagnoses. These models have their limitations, but they have provided already established associations within the field, as well as some associations that may have potential for further studies.

Bibliography

[1] World Health Organization WHO. 'Ageing and health'. In: (2022). URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health.

[2] *Recommendations — Multimorbidity: clinical assessment and management — Guidance — NICE*. eng. Publisher: NICE. Sept. 2016. URL: https://www.nice.org.uk/guidance/ng56/chapter/Recommendations (visited on 27/02/2023).

[3] Bruno Pereira Nunes et al. 'Multimorbidity and mortality in older adults: A systematic review and meta-analysis'. en. In: *Archives of Gerontology and Geriatrics* 67 (Nov. 2016), pp. 130–138. ISSN: 01674943. DOI: 10.1016/j.archger.2016.07.008. URL: https://linkinghub.elsevier.com/retrieve/pii/S0167494316301388 (visited on 05/06/2023).

[4] *Presset på allmennlegetjenesten øker*. no. URL: https://www.helsedirektoratet.no/nyheter/presset-pa-allmennlegetjenesten-oker (visited on 27/02/2023).

[5] Preety Das, Chris Naylor and Azeem Majeed. 'Bringing together physical and mental health within primary care: a new frontier for integrated care'. eng. In: *Journal of the Royal Society of Medicine* 109.10 (Oct. 2016), pp. 364–366. ISSN: 1758-1095. DOI: 10.1177/0141076816665270.

[6] Albert Roso-Llorach et al. 'Comparative analysis of methods for identifying multimorbidity patterns: a study of 'real-world' data'. en. In: *BMJ Open* 8.3 (Mar. 2018). Publisher: British Medical Journal Publishing Group Section: Epidemiology, e018986. ISSN: 2044-6055, 2044-6055. DOI: 10.1136/bmjopen-2017-018986. URL: https://bmjopen.bmj.com/content/8/3/e018986 (visited on 09/02/2023).

[7] Ingmar Schäfer et al. 'Multimorbidity Patterns in the Elderly: A New Approach of Disease Clustering Identifies Complex Interrelations between Chronic Conditions'. en. In: *PLOS ONE* 5.12 (Dec. 2010). Publisher: Public Library of Science, e15941. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0015941. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0015941 (visited on 09/02/2023).

[8] Ingmar Schäfer et al. 'Reducing complexity: a visualisation of multimorbidity by combining disease clusters and triads'. en. In: *BMC Public Health* 14.1 (Dec. 2014), p. 1285. ISSN: 1471-2458. DOI: 10.1186/1471-2458-14-1285. URL: https://doi.org/10.1186/1471-2458-14-1285 (visited on 09/02/2023).

[9] Jeffrey L. Birk. 'Depression and multimorbidity: Considering temporal characteristics of the associations between depression and multiple chronic diseases.' en. In: *Health Psychology* 38.9 (). Publisher: US: American Psychological Association, p. 802. ISSN: 1930-7810. DOI: 10.1037/hea0000737. URL: https://psycnet.apa.org/fulltext/2019-21944-001.html (visited on 31/01/2023).

[10] Raffaele Palladino et al. 'Associations between multimorbidity, healthcare utilisation and health status: evidence from 16 European countries'. In: *Age and Ageing* 45.3 (May 2016), pp. 431–435. ISSN: 0002-0729. DOI: 10.1093/ageing/afw044. URL: https://doi.org/10.1093/ageing/afw044 (visited on 09/02/2023).

[11] Liam G Glynn et al. 'The prevalence of multimorbidity in primary care and its effect on health care utilization and cost'. In: *Family Practice* 28.5 (Oct. 2011), pp. 516–523. ISSN: 0263-2136. DOI: 10.1093/fampra/cmr013. URL: https://doi.org/10.1093/fampra/cmr013 (visited on 20/12/2022).

[12] Amy Ronaldson et al. 'Associations between physical multimorbidity patterns and common mental health disorders in middle-aged adults: A prospective analysis using data from the UK Biobank'. eng. In: *The Lancet Regional Health. Europe* 8 (Sept. 2021), p. 100149. ISSN: 2666-7762. DOI: 10.1016/j.lanepe.2021.100149.

[13] Pankush Kalgotra, Ramesh Sharda and Julie M. Croff. 'Examining health disparities by gender: A multimorbidity network analysis of electronic medical record'. en. In: *International Journal of Medical Informatics* 108 (Dec. 2017), pp. 22–28. ISSN: 1386-5056. DOI: 10.1016/j.ijmedinf.2017.09.014. URL: https://www.sciencedirect.com/science/article/pii/S138650561730237X (visited on 31/01/2023).

[14] Valerie Kuan et al. 'Identifying and visualising multimorbidity and comorbidity patterns in patients in the English National Health Service: a population-based study'. en. In: *The Lancet Digital Health* 5.1 (Jan. 2023), e16–e27. ISSN: 2589-7500. DOI: 10.1016/S2589-7500(22)00187-X. URL: https://www.sciencedirect.com/science/article/pii/S258975002200187X (visited on 09/02/2023).

[15] John E. Cornell et al. 'Multimorbidity Clusters: Clustering Binary Data From Multimorbidity Clusters: Clustering Binary Data From a Large Administrative Medical Database'. en. In: *Applied Multivariate Research* 12.3 (2008). Number: 3, pp. 163–182. ISSN: 1918-1108. DOI: 10.22329/amr.v12i3.658. URL: https://amr.uwindsor.ca/index.php/AMR/article/view/658 (visited on 09/02/2023).

[16] Alba Aguado et al. 'MorbiNet: multimorbidity networks in adult general population. Analysis of type 2 diabetes mellitus comorbidity'. en. In: *Scientific Reports* 10.1 (Feb. 2020), p. 2416. ISSN: 2045-2322. DOI: 10.1038/s41598-020-59336-1. URL: https://www.nature.com/articles/s41598-020-59336-1 (visited on 31/01/2023).

[17] Alexandra Prados-Torres et al. 'Multimorbidity patterns: a systematic review'. en. In: *Journal of Clinical Epidemiology* 67.3 (Mar. 2014), pp. 254–266. ISSN: 0895-4356. DOI: 10.1016/j.jclinepi.2013.09.021. URL: https://www.sciencedirect.com/science/article/pii/S0895435613004368 (visited on 09/02/2023).

[18] Nancy E. Schoenberg et al. 'Burden of Common Multiple-Morbidity Constellations on Out-of-Pocket Medical Expenditures Among Older Adults'. en. In: *The Gerontologist* 47.4 (Aug. 2007), pp. 423–437. ISSN: 1758-5341, 0016-9013. DOI: 10.1093/geront/47.4.423. URL: https://academic.oup.com/gerontologist/article/47/4/423/680191 (visited on 17/05/2023).

[19] Panagiotis Kasteridis et al. 'Who would most benefit from improved integrated care? Implementing an analytical strategy in South Somerset'. In: *International Journal of Integrated Care* 15.1 (Jan. 2015). ISSN: 1568-4156. DOI: 10.5334/ijic.1594. URL: http://www.ijic.org/article/10.5334/ijic.1594/ (visited on 18/05/2023).

[20] Louisa Picco et al. 'Economic burden of multimorbidity among older adults: impact on healthcare and societal costs'. en. In: *BMC Health Services Research* 16.1 (Dec. 2016), p. 173. ISSN: 1472-6963. DOI: 10.1186/s12913-016-1421-7. URL: http://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-016-1421-7 (visited on 18/05/2023).

[21] Caroline Bähler et al. 'Multimorbidity, health care utilization and costs in an elderly community-dwelling population: a claims data based observational study'. en. In: *BMC Health Services Research* 15.1 (Dec. 2015), p. 23. ISSN: 1472-6963. DOI: 10.1186/s12913-015-0698-2. URL: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-015-0698-2 (visited on 18/05/2023).

[22] Samuel L Brilleman et al. 'Implications of comorbidity for primary care costs in the UK: a retrospective observational study'. en. In: *British Journal of General Practice* 63.609 (Apr. 2013), e274–e282. ISSN: 0960-1643, 1478-5242. DOI: 10.3399/bjgp13X665242. URL: https://bjgp.org/lookup/doi/10.3399/bjgp13X665242 (visited on 08/02/2023).

[23] Marina Soley-Bori et al. 'Identifying multimorbidity clusters with the highest primary care use: 15 years of evidence from a multi-ethnic metropolitan population'. eng. In: *The British Journal of General Practice: The Journal of the Royal College of General Practitioners* 72.716 (Mar. 2022), e190–e198. ISSN: 1478-5242. DOI: 10.3399/BJGP.2021.0325.

[24] Sandra H. van Oostrom et al. 'Multimorbidity of chronic diseases and health care utilization in general practice'. In: *BMC Family Practice* 15.1 (Apr. 2014), p. 61. ISSN: 1471-2296. DOI: 10.1186/1471-2296-15-61. URL: https://doi.org/10.1186/1471-2296-15-61 (visited on 20/12/2022).

[25] Petra Hopman et al. 'Health care utilization of patients with multiple chronic diseases in The Netherlands: Differences and underlying factors'. en. In: *European Journal of Internal Medicine* 26.3 (Apr. 2015), pp. 190–196. ISSN: 09536205. DOI: 10.1016/j.ejim.2015.02.006. URL: https://linkinghub.elsevier.com/retrieve/pii/S0953620515000412 (visited on 17/05/2023).

[26] Daniel Ky Chan et al. 'Survey of major chronic iIlnesses and hospital admissions via the emergency department in a randomized older population in Randwick, Australia'. en. In: *Emergency Medicine Australasia* 14.4 (Dec. 2002), pp. 387–392. ISSN: 1742-6731, 1742-6723. DOI: 10.1046/j.1442-2026.2002.00343.x. URL: http://doi.wiley.com/10.1046/j.1442-2026.2002.00343.x (visited on 17/05/2023).

[27] Thomas Lehnert et al. 'Review: Health Care Utilization and Costs of Elderly Persons With Multiple Chronic Conditions'. en. In: *Medical Care Research and Review* 68.4 (Aug. 2011), pp. 387–420. ISSN: 1077-5587, 1552-6801. DOI: 10.1177/1077558711399580. URL: http://journals.sagepub.com/doi/10.1177/1077558711399580 (visited on 17/05/2023).

[28] B. Starfield. 'Comorbidity: Implications for the Importance of Primary Care in 'Case' Management'. en. In: *The Annals of Family Medicine* 1.1 (May 2003), pp. 8–14. ISSN: 1544-1709, 1544-1717. DOI: 10.1370/afm.1. URL: http://www.annfammed.org/cgi/doi/10.1370/afm.1 (visited on 17/05/2023).

[29] Anna J Koné Pefoyo et al. 'The increasing burden and complexity of multimorbidity'. en. In: *BMC Public Health* 15.1 (Dec. 2015), p. 415. ISSN: 1471-2458. DOI: 10.1186/s12889-015-1733-2. URL: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-015-1733-2 (visited on 17/05/2023).

[30] Donna M. Zulman et al. 'Multimorbidity and healthcare utilisation among high-cost patients in the US Veterans Affairs Health Care System'. en. In: *BMJ Open* 5.4 (Apr. 2015). Publisher: British Medical Journal Publishing Group Section: Health services research, e007771. ISSN: 2044-6055, 2044-6055. DOI: 10.1136/bmjopen-2015-007771. URL: https://bmjopen.bmj.com/content/5/4/e007771 (visited on 09/02/2023).

[31] G. Zou. 'A Modified Poisson Regression Approach to Prospective Studies with Binary Data'. en. In: *American Journal of Epidemiology* 159.7 (Apr. 2004), pp. 702–706. ISSN: 0002-9262. DOI: 10.1093/aje/kwh090. URL: https://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kwh090 (visited on 10/06/2023).

[32] Gy Zou and Allan Donner. 'Extension of the modified Poisson regression model to prospective studies with correlated binary data'. en. In: *Statistical Methods in Medical Research* 22.6 (Dec. 2013), pp. 661–670. ISSN: 0962-2802, 1477-0334. DOI: 10.1177/0962280211427759. URL: http://journals.sagepub.com/doi/10.1177/0962280211427759 (visited on 12/05/2023).

[33] Helsedirektoratet. 'Styringsinformasjon til helsefellesskapene [nettdokument]. Oslo: Helsedirektoratet (28. juni 2022, lest 07. desember 2022'. In: *Tilgjengelig fra*. Nov. 2021. URL: https://www.helsedirektoratet.no/rapporter/styringsinformasjon-til-helsefellesskapene.

[34] World Health Organization WHO. *https://icd.who.int/browse10/2016/en*.

[35] George Casella and Roger L Berger. *Statistical inference*. Cengage Learning, 2021.

[36] Ludwig Fahrmeir et al. *Regression: Models, Methods and Applications*. en. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. ISBN: 978-3-642-34332-2 978-3-642-34333-9. DOI: 10.1007/978-3-642-34333-9. URL: https://link.springer.com/10.1007/978-3-642-34333-9 (visited on 28/05/2023).

[37] Eric W. Weisstein. *"Bonferroni Correction." From MathWorld–A Wolfram Web Resource. https://mathworld.wolfram.com/BonferroniCorrection.html*.

[38] Gareth James et al. 'Moving Beyond Linearity'. In: *An Introduction to Statistical Learning: with Applications in R*. Ed. by Gareth James et al. New York, NY: Springer New York, 2013, pp. 265–301. ISBN: 978-1-4614-7138-7. DOI: 10.1007/978-1-4614-7138-7_7. URL: https://doi.org/10.1007/978-1-4614-7138-7_7.

[39] Thomas M. J. Fruchterman and Edward M. Reingold. 'Graph drawing by force-directed placement'. en. In: *Software: Practice and Experience* 21.11 (Nov. 1991), pp. 1129–1164. ISSN: 00380644, 1097024X. DOI: 10.1002/spe.4380211102. URL: https://onlinelibrary.wiley.com/doi/10.1002/spe.4380211102 (visited on 18/05/2023).

[40] SSB. *National population projections*. URL: https://www.ssb.no/en/befolkning/befolkningsframskrivinger/statistikk/nasjonale-befolkningsframskrivinger.

[41] A Yildirir. 'Hypertension and arrhythmia: blood pressure control and beyond'. en. In: *Europace* 4.2 (Apr. 2002), pp. 175–182. ISSN: 10995129. DOI: 10.1053/eupc.2002.0227. URL: https://academic.oup.com/europace/article-lookup/doi/10.1053/eupc.2002.0227 (visited on 29/05/2023).

[42] Jindřich Špinar. 'Hypertension and ischemic heart disease'. In: *Cor et Vasa* 54.11-12 (Nov. 2012), e433–e438. ISSN: 00108650, 18037712. DOI: 10.1016/j.crvasa.2012.11.002. URL: http://e-coretvasa.cz/doi/10.1016/j.crvasa.2012.11.002.html (visited on 29/05/2023).

[43] Samantha Fraser-Bell, Richard Symes and Anagha Vaze. 'Hypertensive eye disease: a review: Hypertensive eye disease: a review'. en. In: *Clinical & Experimental Ophthalmology* 45.1 (Jan. 2017), pp. 45–53. ISSN: 14426404. DOI: 10.1111/ceo.12905. URL: https://onlinelibrary.wiley.com/doi/10.1111/ceo.12905 (visited on 29/05/2023).

[44] Brian S. Armour et al. 'Prevalence and Causes of Paralysis—United States, 2013'. en. In: *American Journal of Public Health* 106.10 (Oct. 2016), pp. 1855–1857. ISSN: 0090-0036, 1541-0048. DOI: 10.2105/AJPH.2016.303270. URL: https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2016.303270 (visited on 30/05/2023).

[45] Darrell Hulisz. 'Amyotrophic lateral sclerosis: disease state overview'. eng. In: *The American Journal of Managed Care* 24.15 Suppl (Aug. 2018), S320–S326. ISSN: 1936-2692.

[46] Mohini Aras, Beverly G. Tchang and Joy Pape. 'Obesity and Diabetes'. en. In: *Nursing Clinics of North America* 56.4 (Dec. 2021), pp. 527–541. ISSN: 00296465. DOI: 10.1016/j.cnur.2021.07.008. URL: https://linkinghub.elsevier.com/retrieve/pii/S0029646521000712 (visited on 30/05/2023).

[47] P. H. Winocour. 'Diabetes and chronic kidney disease: an increasingly common multi-morbid disease in need of a paradigm shift in care'. en. In: *Diabetic Medicine* 35.3 (Mar. 2018), pp. 300–305. ISSN: 07423071. DOI: 10.1111/dme.13564. URL: https://onlinelibrary.wiley.com/doi/10.1111/dme.13564 (visited on 30/05/2023).

[48] J Tecklenborg et al. 'The role of the immune system in kidney disease'. en. In: *Clinical and Experimental Immunology* 192.2 (Apr. 2018), pp. 142–150. ISSN: 0009-9104, 1365-2249. DOI: 10.1111/cei.13119. URL: https://academic.oup.com/cei/article/192/2/142/6412269 (visited on 30/05/2023).

[49] Allan Gibofsky. 'Overview of epidemiology, pathophysiology, and diagnosis of rheumatoid arthritis'. eng. In: *The American Journal of Managed Care* 18.13 Suppl (Dec. 2012), S295–302. ISSN: 1936-2692.

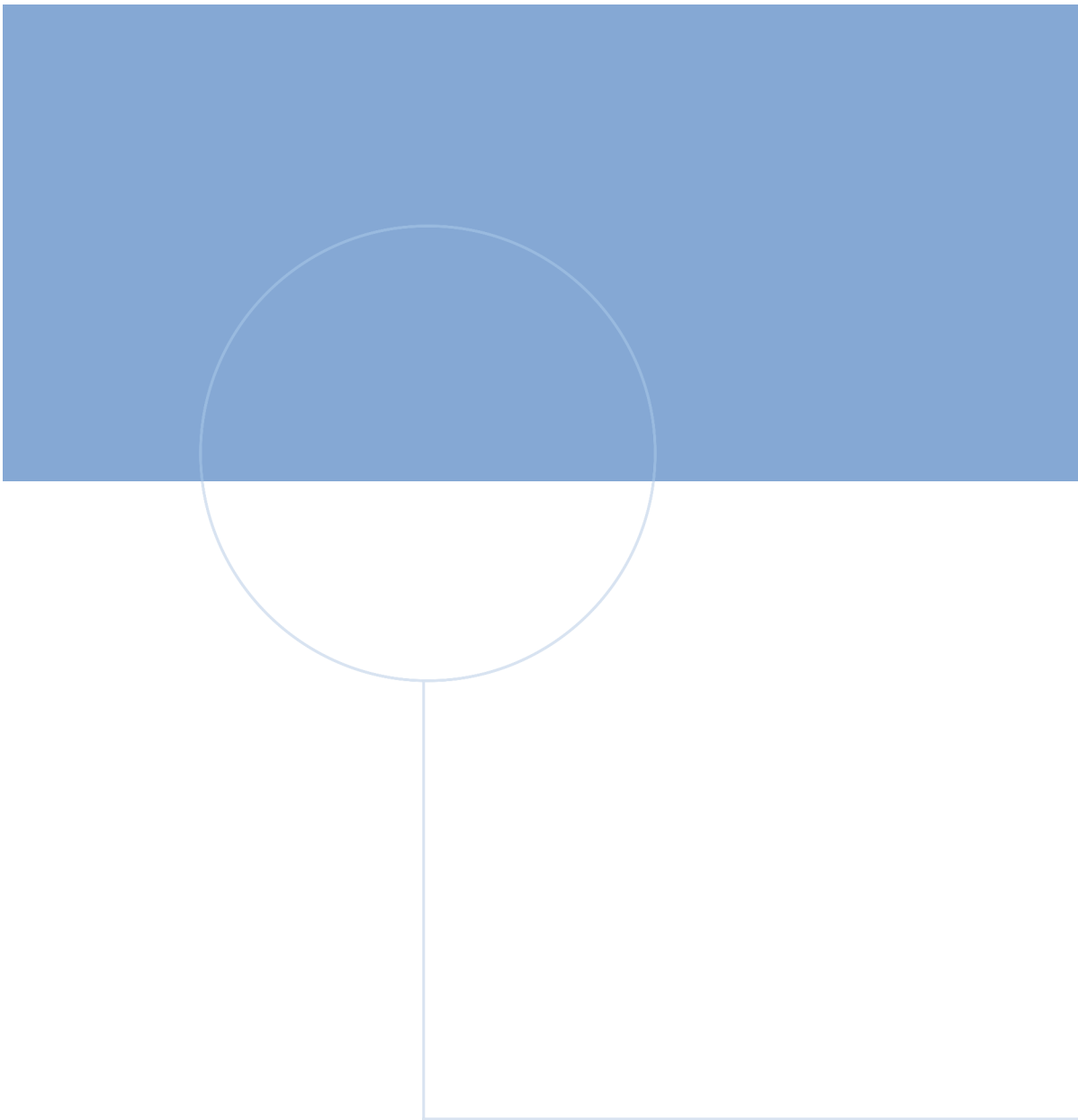
[50] Madhubhashinee Dayabandara et al. 'Antipsychotic-associated weight gain: management strategies and impact on treatment adherence'. en. In: *Neuropsychiatric Disease and Treatment* Volume 13 (Aug. 2017), pp. 2231–2241. ISSN: 1178-2021. DOI: 10.2147/NDT.S113099. URL: https://www.dovepress.com/antipsychotic-associated-weight-gain-management-strategies-and-impact--peer-reviewed-article-NDT (visited on 31/05/2023).

[51] Miguel A. Hernán, Sonia Hernández-Díaz and James M. Robins. 'A Structural Approach to Selection Bias:' en. In: *Epidemiology* 15.5 (Sept. 2004), pp. 615–625. ISSN: 1044-3983. DOI: 10.1097/01.ede.0000135174.63482.43. URL: http://journals.lww.com/00001648-200409000-00020 (visited on 01/06/2023).

[52] Anders Prestmo et al. 'Comprehensive geriatric care for patients with hip fractures: a prospective, randomised, controlled trial'. en. In: *The Lancet* 385.9978 (Apr. 2015), pp. 1623–1633. ISSN: 01406736. DOI: 10.1016/S0140-6736(14)62409-0. URL: https://linkinghub.elsevier.com/retrieve/pii/S0140673614624090 (visited on 31/05/2023).

[53] Miguel A. Hernán and James M. Robins. 'Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available: Table 1.' en. In: *American Journal of Epidemiology* 183.8 (Apr. 2016), pp. 758–764. ISSN: 0002-9262, 1476-6256. DOI: 10.1093/aje/kwv254. URL: https://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kwv254 (visited on 11/06/2023).

[54] Thomas Frisell et al. 'Sibling Comparison Designs: Bios From Non-Shored Confounders and Measurement Error'. In: *Epidemiology* 23.5 (2012). Publisher: Lippincott Williams & Wilkins, pp. 713–720. ISSN: 10443983. URL: http://www.jstor.org/stable/41739651 (visited on 11/06/2023).



NTNU

Norwegian University of
Science and Technology