Antibiotic use
Human consumption patterns and effect on bacteria

—
Pål Haugen
A dissertation for the degree of Philosophiae Doctor

Human drivers

Animal consumption

Human consumption

Outpatient consumption
Hospital consumption

Antibiotic production waste

Agriculture
Aquaculture

Prophylaxis
Treatment

Growth

Selective pressure

Intrinsic resistance

Horizontal gene transfer

Antibiotic producing organisms

Level of Antimicrobial resistance

Naturally occurring drivers
Antibiotic use

Human consumption patterns and effect on bacteria

Pål Haugen

A dissertation for the degree of Philosophiae Doctor

UiT
The arctic university of Norway
Faculty of health sciences
Department of pharmacy

October 2014
# Table of contents

Acknowledgements................................................................................................................. 5  
List of papers............................................................................................................................. 7

Terminology and abbreviations ................................................................................................. 8

Introduction ............................................................................................................................... 10  
Antibiotic use............................................................................................................................. 10

Box 1 Antibiotic profiles of European countries based on ECDC 2011 data............................ 12

Surveillance of antibiotic consumption....................................................................................... 13

Predictors of antibiotic consumption......................................................................................... 15

  *Patient.................................................................................................................................. 15

  Prescriber................................................................................................................................. 16

  Geography............................................................................................................................... 18

Summary.................................................................................................................................... 19

Variety of methods in pharmacoepidemiology ......................................................................... 20

  Summary................................................................................................................................. 22

Antibiotic consumption and resistance in bacteria .................................................................. 23

Types of resistance..................................................................................................................... 23

  *Intrinsic............................................................................................................................... 23

  Acquired................................................................................................................................. 23

  Inter- and intracellular transport of genetic elements............................................................ 25

The cost and reversal of resistance......................................................................................... 27

Summary.................................................................................................................................... 28

Aims.......................................................................................................................................... 30

Overall aim............................................................................................................................... 30

  *Paper I................................................................................................................................. 30

  Paper II................................................................................................................................. 30

  Paper III............................................................................................................................... 30

Material and Methods............................................................................................................... 31

Computer software .................................................................................................................. 31

Data sources ............................................................................................................................. 31

Variables used .......................................................................................................................... 32

Paper I Multivariate methods and ordination......................................................................... 34

Paper II Regional difference in outpatient antibiotic consumption...................................... 35
Paper III Maintenance of genetic elements imposing a fitness cost ........................................36

Summary of papers and main results ................................................................. 37
Paper I .................................................................................................................. 37
Paper II ............................................................................................................... 39
Paper III ............................................................................................................. 40

Discussion ........................................................................................................ 41
Methodological considerations ........................................................................... 43
Limitation of studies ......................................................................................... 46
Implications of the studies ................................................................................ 47

Conclusions ....................................................................................................... 49
Future aspect ..................................................................................................... 50

References ........................................................................................................ 51
Acknowledgements

“Everything here is so cold
Everything here is so dark”

[Mayhem, The freezing moon, De mysteriis dom satanans [CD], Deathlike silence, Oslo 1994]

Tromsø certainly can be both a bit cold and a bit dark at certain time intervals. Luckily, several people have contributed to lighten spirits in the daily PhD life during my stay at department of pharmacy.

I have had the pleasure of working close with a multitude of experts. Researchers in epidemiology, molecular biology, evolution, mathematical modelling, clinical pharmacy and statistics have joyfully helped me and joined in all my outbursts of scientific discussions.

I am especially grateful to my two supervisors Raul Primicerio and Lars Småbrekke. Raul’s infectious enthusiasm and positive attitude have elevated the statistical discussions in my office to rock concert volumes (ask my neighbouring offices, they know when Raul is supervising me). Without Raul’s positive attitude and statistical knowledge this project would have been much, much duller.

Lars has a keen eye for details (which I tend to lack) and has been an exceptional motivator for finishing the projects. Besides guidance through epidemiology, his enthusiasm for new methodology and care for the tired PhD student has been of great value. The close contact and friendship has been much appreciated.

Pål Johnsen guided me through my first project and the cooperation with Pål and Irina was a great experience and is a fond memory. Thank you for delightful discussions and expert supervision.
The collaboration with Gunnar Skov Simonsen and Anne-Sofie Furberg has been much appreciated. Your rapid, detailed responses and insightful comments have truly helped me in developing the projects.

Vidar Sørum has been a keen partner for discussions, methodological considerations, (a multitude of) frustrations and general tomfoolery. Thank you for a long lasting friendship and your patience with my lengthy explanations of methodology.

Kaare, Marit, Klaus, Nils, Beate, June, Kjell and Frode and all members of the MMPE research group: thank you for the discussions and quick responses on all of my questions.

Dear colleagues at the department of pharmacy, I wish you all the best, thank you for four fantastic, yet challenging years.

Finally, the importance of the work being put into creating, updating and developing the R language for statistical computing cannot be underestimated. This is an “easy” programming language that is free to use for anybody and is continuously updated by enthusiasts. Open source software has become an important part of my scientific thinking; I wish I had the knowledge to implement such programs in an even broader manner than I have done. To all developers of R, thank you, and keep up the good work!

Tromsø, September 2014

Pål Haugen
List of papers

Paper I

Haugen P, Primicerio R, Simonsen GS, Furberg AS, Småbrekke L. Antibiotic consumption profiles identified from a prescription database using multivariate analysis [manuscript]

Paper II

Haugen P, Simonsen GS, Primicerio R, Furberg AS, Småbrekke L. Outpatient antibiotic use in Norway depends on municipality population size [manuscript]

Paper III

**Terminology and abbreviations**

The word antibiotics may lack precision in some circumstances; in this thesis antibiotic refers to compounds used against infections caused by bacteria.¹

The Defined Daily Dose (DDD) is used as a measure for antibiotic consumption. The World Health Organization (WHO) defines DDD as *“the assumed average maintenance dose per day for a drug used for its main indication in adults”*.²

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>Akaike Information Criteria</td>
</tr>
<tr>
<td>ATC</td>
<td>Anatomical Therapeutic Chemical classification system</td>
</tr>
<tr>
<td>CA</td>
<td>Correspondence Analysis</td>
</tr>
<tr>
<td>CCA</td>
<td>Constrained Correspondence Analysis</td>
</tr>
<tr>
<td>CDC</td>
<td>Centres for Disease Control and prevention</td>
</tr>
<tr>
<td>CRE</td>
<td>Carbapenem-Resistant Enterobacteriaceae</td>
</tr>
<tr>
<td>DAG</td>
<td>Directed Acyclic Graph</td>
</tr>
<tr>
<td>DDD</td>
<td>Defined Daily Dose</td>
</tr>
<tr>
<td>DID</td>
<td>DDD / 1000 inhabitants / day</td>
</tr>
<tr>
<td>ECDC</td>
<td>European Centre for Disease prevention and Control</td>
</tr>
<tr>
<td>ESAC</td>
<td>European Surveillance of Antimicrobial Consumption</td>
</tr>
<tr>
<td>GP</td>
<td>General Practitioner</td>
</tr>
<tr>
<td>HGT</td>
<td>Horizontal Gene Transfer</td>
</tr>
<tr>
<td>MAUP</td>
<td>Modifiable Areal Unit Problem</td>
</tr>
<tr>
<td>NorPD</td>
<td>Norwegian Prescription Database</td>
</tr>
<tr>
<td>OLS</td>
<td>Ordinary Least Squares</td>
</tr>
<tr>
<td>OR</td>
<td>Odds Ratio</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>QR</td>
<td>Quantile Regression</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>US</td>
<td>United States</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
</tbody>
</table>
Introduction

Antibiotic resistance in bacteria is a prime example of the implications of Darwin’s theory of evolution: A selective pressure will change the frequency of genotypes in a population in favour of the organisms that are best adapted to the environment.\textsuperscript{3,4} In presence of antibiotics the resistant phenotypes have an advantage over their susceptible counterparts, and can potentially dominate the environment in relatively short time.\textsuperscript{5} Antibiotic resistance and related genes predates human antibiotic production and use, and resistant bacteria are even found in environments where human impact is improbable.\textsuperscript{6,7} However, evidence for human activities as drivers of increased levels of antibiotic resistance is overwhelming, even though some reports show that the link between consumption and resistance is not always clear-cut.\textsuperscript{3,8-16}

Antibiotic use

Antibiotics are commonly used in agriculture and aquaculture as well as for human consumption. In Germany in 2008, 85\% of the total sales of antibiotics were used for animals. In contrast, Norwegian and Swedish antibiotic sales to animals in 2012 constituted 15\% of total sales.\textsuperscript{11,17,18} In Denmark the sales of antibiotics for animal consumption constituted 70\% of the volume sold in 2012.\textsuperscript{19}

Human antibiotic consumption can be divided into hospital and outpatient use. Of the total human consumption in Norway approximately 80\% is from outpatients.\textsuperscript{17,20}

In Norway the total consumption of antibiotics is relatively stable. There was an increase from 19 DDD / 1000 patients /day (DID), in 2006 to 20 DID in 2013. However, methenamine (Anatomical Therapeutic chemical Classification system (ATC) class J01XX05) is frequently used in Norway for prevention of urinary tract infections. The antibiotic consumption excluding this drug was 16.3 both in 2006 and 2013.\textsuperscript{17}
A major task to combat evolution of antibiotic resistance is to monitor resistance in bacteria. Also, monitoring consumption of antibiotics is important to identify countries, regions or demographic groups with an undesirable consumption. Reports on consumption of antibiotics can be based on sales from wholesalers, reimbursements or prescription databases. If sold antibiotics are consumed is seldom investigated. A global study on adherence to acne treatment revealed adherence as low as 50%.21 Other reports show non-adherence as low as 9% for patients receiving antibiotic treatment.22

To what extent sales of antibiotics reflects the accurate consumption will vary depending on indication, how many and frequency of doses, if the patient feels healthy, possible side effects of the drug and also personality.22,23 Therefore, some uncertainty is connected to measurements of antibiotic consumption. Most data on antibiotic consumption are therefore an estimate of the true consumption.

The European Centre for Disease prevention and Control (ECDC) have published data on consumption of antibiotics in several European countries.24 With a few exceptions the figures only cover outpatient consumption. No decrease in the consumption of antibiotics was found in 22 countries during 2007 – 2011. In 2011 Greece and Belgium were among the highest consuming countries, with approximately 30 DID, whereas Netherlands and Estonia consumed less than half of this, approximately 12 DID.24 A large difference in the ratio of broad-versus narrow spectrum antibiotic use was identified. Sweden and Norway had a ratio of 0.2 and Malta a ratio of 142.7.24 The data on antibiotic use in European countries are presented in more detail in Box1.
Box 1 Antibiotic profiles of European countries based on ECDC 2011 data

The ECDC 2014 report (Table 3.1) presents data on DID for 8 subgroups of antibiotics from 29 countries. In the following, the data have been processed by correspondence analysis to visualize the antibiotic profiles for these countries. Countries that included hospital consumption in their reporting (Cyprus, Iceland, Lithuania, and Slovakia), and countries that reported parts of the national community consumption (Romania and Spain) were excluded from the analysis as their antibiotic profiles were assumed to differ due to sampling design. The weighted averages (centroids) of antibiotic profiles for each country were plotted alongside antibiotic types in a biplot.

The countries with the lowest consumption (smallest filled circles) had a higher relative consumption of other antibiotics (J01X), tetracyclines (J01A) and trimethoprim combinations (J01E). Greece and Malta had a high relative consumption of other betalactam antibacterials (Cephalosporins (J01D)) and the ATC-groups J01 B, R or G. Belgium, Portugal, Denmark and Slovenia had high relative consumptions of penicillins (J01C). Finally, Germany had the lowest relative consumption of penicillins.

Biplot of CA results for data from table 3.1 in *Surveillance of antimicrobial consumption in Europe 2011* authored by the ECDC. The biplot captures 72% of the total variation of the data. Black filled circles represent the centroids for each country. The size of each filled circle reflects relative total consumption. Red triangles indicate the different antibiotic groups.
Surveillance of antibiotic consumption

The European Surveillance of Antimicrobial Consumption Network (ESAC-Net) covers 29 European countries including Norway. Twelve quality indicators are used to assess antibiotic consumption. Among these we find DID for several classes of antibiotics, percentages of total consumption for ATC-J01 sub groups and ratio of broad-spectrum antibiotics versus narrow spectrum antibiotics.\(^{24}\)

In the United States (US), as in most EU countries, there is no national or federal prescription database. However, claims databases from insurance companies, and the Medicaid database contains prescription data on the individual level. The Medicaid database includes 19\% of the US population, but has limitations especially with regards to bias in patient demographics.\(^{25}\)

In 1994 Finland and Denmark established registries for prescriptions, and from 2006 all Nordic countries (Finland, Denmark, Norway, Sweden and Iceland) have a national prescription database. None of the databases includes hospital use or indication of disease.\(^{26}\) A recent review summarized the research based on the Nordic prescription databases. Danish studies were most prevalent (51\% of the studies examined). Only a minor part of the studies (6\%) addressed antibiotics.\(^{27}\)

In Norway several sources of antibiotic consumption are available. At the population level, wholesale statistics exists from 1974 until 2013, and the total sales from all classes of drugs are reported annually from the Norwegian

institute of public health. These data are only suited for ecological studies, as no individual patient information is available.

The Norwegian Prescription Database (NorPD) was established in 2004. All Norwegian prescriptions are registered and are available for research with an identifier for patient and prescriber. Further, aggregated data are publically available through online resources. Indication of disease is not routinely collected for prescriptions, and assessment of appropriateness of prescribing is difficult. It is not reasonable to assume a uniform distribution of infectious diseases between genders or age classes of any population. This is illustrated in differences in antibiotic consumption in age groups and genders in Norway. Therefore, causal inferences from available demographic variables in NorPD are hampered with possible bias in prevalence of disease.

Population surveys can be used to investigate individual patients and their consumption of antibiotics. In Norway, studies have been conducted on respiratory tract infections where prescriber and patient information has been combined with type of antibiotic and indication of disease. Population studies in Tromsø and Bergen have used questionnaires to collect data on drug consumption including antibiotic use.
Predictors of antibiotic consumption

National guidelines are developed to aid prescribers in choice of antibiotic therapy.\textsuperscript{35, 36} The indication and the pathogen in question are decisive of which drug is optimal for treatment. Pathogen characteristics that influence the choice of antibiotic include whether the bacteria are gram-negative or positive and known resistance issues.\textsuperscript{35}

However, variables related to both prescriber and patient also affect treatment. In a study including 17 European countries, the authors found no association between morbidity caused by infectious diseases and overall consumption of antibiotics. In this study the authors adjusted for socioeconomic, demographic and supply variables in their analysis.\textsuperscript{37} Consequently, other variables, besides indication of disease, can affect the amount and type of antibiotics prescribed.\textsuperscript{38}

A study comparing antibiotic prescribing in Germany and France addresses five non-microbiological factors influencing the use of antibiotics; prescriber factors, patient factors, cultural influences, social determinants and regulatory practices.\textsuperscript{39} The authors conclude that France has a higher antibiotic prescribing rate, especially for upper respiratory tract infections. The authors attribute these differences to prescriber and patient attitudes in addition to sociocultural and economic differences.

In this thesis I will focus on three predictors of outpatient antibiotic use: Patient, prescriber and geography. Improved understanding of the contribution from these predictors is important to identify possible differences in selection pressure on bacteria, and to address areas or demographic groups with a specific antibiotic consumption pattern. Paper I addresses demographic groups of prescribers and patients, Paper II addresses regional differences and finally Paper III addresses how heterogenic antibiotic pulses affect resistance determinants in bacteria.

Patient

Patient age and gender, income, education and cultural differences are variables that influence antibiotic prescriptions.\textsuperscript{20, 37, 40} A previous Norwegian study showed that female patients had a higher one-year prevalence of antibiotic use
than males (except children below 8 years). There was also a difference in antibiotic profiles between age groups and gender. Females used more broad-spectrum penicillins than males, and males had relatively higher tetracycline consumption compared to females. Finally, high users where found to be among adolescents (predominantly males) and the elderly (no gender difference). In Italy the prevalence of antibiotic prescription showed similar patterns as in Norway, however the overall prevalence was higher in the Italian study. Recently it has been demonstrated that antibiotic consumption increases with age among French outpatients. Females where also found to have higher odds for receiving a prescription of an antibiotic compared to males. Further, when adjusting for type of infection, patients over 45 years had lower odds of being prescribed antibiotics.

Socioeconomic factors such as the proportion of women working and the proportion of the population with higher education have been reported to reduce antibiotic consumption. However, the results depend on the statistical model chosen. An Australian study investigated the effect of patient expectations on the odds for being prescribed a new medication. This study was inconclusive, as patients who expected a new medication and patients that had no opinion both had higher odds of receiving treatment than patients that did not expect a prescription. If the prescriber believed the patient wanted a new prescription the odds increased 10-fold for receiving a prescription. In Spain a study found that prescribers overestimate the patients expectations for receiving a prescription, and prescribe more often than their patients expected. German investigators found that only 10% of patients visiting a physician for the common cold anticipated a prescription of antibiotics. These authors also suggest that physicians overestimate the expectations patients have for receiving antibiotic prescriptions.

**Prescriber**

When treating respiratory tract infections in children, General Practitioner (GP) specialists in Norway had lower odds of prescribing an antibiotic than non-specialists. Age of the prescriber did not influence the odds of prescribing an antibiotic when both age and speciality was included in a logistic regression.
model. However, in the same study specialists had higher odds of prescribing non-penicillin V antibiotics.\textsuperscript{47} In the UK, prescriber age above 45 years increased the amount of prescribed antibiotics.\textsuperscript{48} However, years in practice was not a significant predictor of increased antibiotic prescribing in a study from Ireland.\textsuperscript{49} In Italy, time since certification was found to lower the odds for prescribing an antibiotic.\textsuperscript{43} In Belgium, middle-aged prescribers had higher odds of prescribing broad-spectrum antibiotics than both younger and older prescribers.\textsuperscript{50} Time in practice has also been found to increase the probability for inappropriate antibiotic prescribing as well as being a predictor of high prescribers in both USA and Canada.\textsuperscript{51, 52}

Studies investigating prescriber gender are inconclusive whether and how gender affects prescribing. Prescriber gender can increase odds of prescribing an antibiotic, selecting broad-spectrum alternatives and also being a high antibiotic prescriber. Female prescribers are associated with lower odds of prescribing an antibiotic.\textsuperscript{43} The aforementioned UK study found higher amounts prescribed by male GP’s.\textsuperscript{48} Recently, investigators also found females to have higher odds of prescribing amoxicillin over broader spectrum alternatives for both adults and children, although effect sizes where small for adults.\textsuperscript{50} Norwegian studies, and a US study report no differences between prescriber genders.\textsuperscript{32, 33, 47, 51, 53}

Prescriber age is addressed in a variety of ways across studies. Some studies use age directly, other use time since medical exam or simply the dichotomous version: specialist (presumably older) and non-specialist. Consequently, studies addressing antibiotic prescribing and prescriber age display conflicting results.

Attitudes towards antibiotic prescriptions for respiratory infections among German and French prescribers differ, where the German prescribers are more restrictive in their use of antibiotics compared to the French.\textsuperscript{39} A survey from Wales found that the patient’s clinical characteristics were important for prescribers in choice of antibiotics, as did information on the pathogen in question.\textsuperscript{54} The authors also dichotomised prescribers with regard to their socially responsible prescribing practice; those taking care of the individual patient immediately (higher fluoroquinolone use), or those who try narrow
spectrum choices to prevent future resistance problems (lower fluoroquinolone use).

**Geography**

Where patients live influences both probabilities of receiving antibiotics, type received and amount of antibiotics prescribed. Geographical entities are in themselves hardly the cause of differences in antibiotic consumption, but a proxy for other, unmeasured variables. For instance differences in prevalence of infectious diseases, doctors per thousand inhabitants, income, ease of access to healthcare services, over the counter sales of antibiotics, mortality rate and other demographic variables can influence geographical variation.55

Several authors have studied variation in antibiotic consumption between countries15, 37, 56-60 or regional differences within a country.40, 41, 61-69 An investigation from Germany finds regional variation, however none of the investigated variables predict the observed differences.67 Later attempts, with different aggregation levels, finds areal deprivation as a predictor of regional variation in Germany.65 Amount of antibiotics consumed (measured in DID), both overall and for specific classes was compared between different regions in Switzerland. The investigators found differences between regions, the highest consuming region had 1.5 times the overall antibiotic consumption of north-western Switzerland.70 Consumption of antibiotics are higher during winter, and regions with a high consumption demonstrate higher seasonal fluctuations than lower consuming regions.57, 70 It has been shown that fluctuations in antibiotic use coincide with influenza epidemics, leading to an increased consumption of fluoroquinolones. Although secondary infection of bacteria is possible during influenza outbreaks, increased quinolone use indicates that pathogens other than bacteria can influence the (mis) use of antibiotics.71

Differences between countries are often large, both with respect to type and amount of antibiotics prescribed. A recent comparison of non European Union countries shows a nearly threefold difference in DID between low and high consuming countries. In 2011 Turkey had a total antibiotic consumption of 42.3
DID, in comparison Norwegian total consumption was 20.6 DID (17.2 DID methenamine omitted).\textsuperscript{56, 72}

**Summary**

Evidently, patient, prescriber and geography are not variables influencing antibiotic consumption in isolation. These variables are often considered in combination and seem to interact. Due to differences in choices of explanatory and outcome variables among the aforementioned studies a direct comparison is highly challenging and certainly not straightforward.
Variety of methods in pharmacoepidemiology

During literature search for Paper I and II and the current thesis we identified 34 papers investigating antibiotic consumption at the ecological level using multiple regression (linear, logistic or others). The papers were published between 1999 and 2014 in 23 different journals.\textsuperscript{10, 11, 32, 33, 37, 40, 41, 43, 47, 48, 50, 53, 57, 59, 62, 65, 66, 68, 70, 73-87}

Twenty of the above mentioned papers compared different regression models. The papers not considering model selection in the text used either p-values or p-values in combination with other diagnostic tools (for instance R\textsuperscript{2}) for statistical inference. The methods of model selection of the 20 papers are summarized in Figure 1. P-value interpretation is central to 13 of the studies. Studies relying only on p-values (n=5) typically use either a back- or forward step-wise selection procedure. Some studies (n=4) combine p-values with R\textsuperscript{2} values (or pseudo R\textsuperscript{2}), and a category of others are constructed to cover tests as likelihood ratio test and the Wald test (7 studies).

None of the papers used information criteria, such as Akaike’s Information Criterion (AIC), for model selection. None of the papers considered Directed Acyclic Graphs (DAG) for identification of confounders. These methods are used and described in further details, in Paper II and in the discussion of the thesis. Five papers were considered to have efficient methods for model selection, and avoided using only p-values and R\textsuperscript{2} as selection criteria. Typically these papers used multilevel models and tests based on the likelihood function.
Figure 1 Venn diagram of the distribution of model selection techniques from 20 papers conducting multivariable regression addressing different aspects of antibiotic consumption and resistance. Studies only using p-values rely on stepwise backward elimination or forward selection. R² values are used to explain how much of the variation in the dataset the model explains. Other selection techniques include tests like likelihood ratio and the Wald test.
Summary

Studies on antibiotic consumption that include patient, prescriber and geographical variables are not conclusive. Model selection procedures in pharmacoepidemiology are frequently sub-optimal. In order to efficiently address the predictors patient, prescriber and geography several statistical methods are available, and methodology is under constant development. Implementation of DAG and AIC into studies concerning antibiotic consumption will strengthen analysis, and possibly lead to studies that are more uniform in their conclusions.
Antibiotic consumption and resistance in bacteria

Bacteria have accumulated a broad range of genetic traits encoding resistance against a multitude of antibiotics, and these traits are found worldwide.5,88-90 The problem of resistant bacteria has been known for decades, 91 and genome analyses show that changes in bacterial genomes coincide temporally with an increasing use of antibiotics for humans and livestock.92-94

Bacteria resistant towards antibiotics are by no means restricted to hospitals, agricultural areas or aquaculture where antibiotic use is frequent, but are also found in pristine environments.8, 9, 92, 94, 95

Types of resistance

Intrinsic

Already at the time of penicillin discovery it was reported that bacteria varied in their susceptibility.96 Some bacteria are multidrug resistant through natural cell functions. Efflux pumps, which reduce the concentration of antibiotics within the cell, typically mediate this resistance. These efflux pumps can be targeted towards specific compounds or multiple drugs.97 Simultaneously the protective layer provided by the membrane of bacteria act alone or synergistically with the effects of efflux pumps. For instance gram-negative bacteria have a higher degree of protection due to the outer membrane. In addition to these mechanisms basic cellular functions in bacteria can aid in resisting antibiotics. Intrinsic resistance is usually not considered to be a function that is easily transferred between bacteria.98

Acquired

Bacteria acquire resistance genes either through mutations or transfer of genetic information by Horizontal Gene Transfer (HGT).99

Mutations

Mutations in bacteria can render them less sensitive towards antibiotics. Alteration of drug targets and modification of efflux pumps are examples of resistance due to mutations in protein coding genes.99
**Horizontal gene transfer**

Bacteria exchange genes, even crossing the genus level, by a multitude of pathways.\(^\text{3, 99, 100}\) HGT allows bacteria to rapidly acquire complex genetic changes compared to the slower process of mutations.\(^\text{101}\) Resistance towards single or multiple compounds can be acquired in a single step, letting bacteria take evolutionary leaps. The success of HGT elements in a bacterial population (or in ecosystems) is dependent not only on the mechanism of dispersal, but also on the fitness advantage for the receiving bacteria.\(^\text{102}\)

HGT is a topic of great interest in scientific literature and excellent reviews have been published in top ranking journals. Three of these form the basis for the following brief overview.\(^\text{102-105}\)

**Three modes of horizontal gene transfer**

Natural transformation is the process of uptake and integration of exogenous DNA.\(^\text{103}\) Although the mechanisms for uptake differ between gram positive and gram negative bacteria, the processes are similar.\(^\text{103, 106}\) The DNA may originate from destroyed cells, phages or active excretions of genetic material and can be found inside bacterial hosts (in faeces, blood saliva etc.), or in external environments such as soil, or in water.\(^\text{102}\) It has been demonstrated that bacteria also can use fragmented, damaged, and even ancient non-bacterial DNA.\(^\text{107}\) During natural transformation the competent bacteria taking up the DNA is the active part, in contrast to conjugation and transduction where the receiving cell is (more) passive.\(^\text{102}\)

Conjugation occurs when bacteria transfer DNA (plasmid or integrated conjugative elements) with cell interactions and is dependent on physical contact.\(^\text{104}\) Two bacterial cells connect through a pilus.\(^\text{103}\) Although other modes of plasmid transport between cells exist, conjugation is most common. The receiving cell can have some degree of protection against plasmid transfer through exclusion at the surface and restriction enzymes.\(^\text{102}\)

As plasmids were common in bacteria prior to widespread antibiotic use, it is possible that human antibiotic use has led to resistance determinants being acquired by pre-existing plasmids.\(^\text{108, 109}\)
Transduction is transfer of bacterial DNA through a bacteriophage. Phages can incorporate host (bacterial) DNA and transfer it as double stranded DNA. This process is independent of contact between donor and recipient. However, this form of transfer is mostly limited to related species as bacteriophages normally have relatively narrow host ranges.\textsuperscript{104}

**Inter- and intracellular transport of genetic elements**

While bacteriophages, plasmids and other Mobile Genetic Elements (MGE) are spreading intercellularly, there are MGEs that are spreading intracellularly and are transported between cells by the aid of an intercellular MGE.\textsuperscript{104, 105}

Integrons are intracellular genetic elements that have the ability to acquire gene cassettes and possess a system to express these genes (promoter). Gene cassettes are small genetic segments that are integrated and transcribed within the content of integrons.\textsuperscript{110} Integrons rely on intercellular MGE (for instance a plasmid) to facilitate transport between cells and on transposons for intracellular mobilization (between genome and plasmid).\textsuperscript{105} A visualization of plasmid and transposon mediated integron mobilization within and between cells is displayed in Figure 2.
**Figure 2** Overview of plasmid and transposon mediated mobilization of integrons. Integrons possess the ability to capture (and express) gene cassettes. When integrated in a transposon the integron can relocate between the chromosome and plasmids. The plasmid can be transferred to a new cell where the integron either can remain on the plasmid and express incorporated gene cassettes or, through the transposon, relocate to other parts of the genome. MGE (transposons and plasmids) connected to a integron are coloured in light blue, transport between cells in red, dotted lines represents possibilities of gene cassette capture.
While natural transformation, conjugation and transduction are the classical modes of HGT other mechanisms exist. Outer membrane vesicles have the ability to transfer DNA in Acinetobacter baylyi. Other structures called nanotubes have been demonstrated to facilitate transport of plasmids.111, 112

**The cost and reversal of resistance**

Antibiotics act as selective agents in three ways: First, they select for already existing, naturally occurring resistant genotypes or spontaneous mutations. Secondly, they can act as a mediator for bacterial heterogeneity by increasing genetic variability and finally they can act as signalling molecules.113

Acquired antibiotic resistance often comes with a fitness cost for the bacteria. This can be a cost of harbouring the gene itself, from a plasmid, or the fitness cost of an integron, or combinations of these. The disadvantage of this cost is counteracted by the presence of antibiotics. Antibiotics allow the resistant bacteria to outcompete their susceptible counterparts, which suffer a fitness reduction induced by the antibiotics.114, 115

The concentration of an antibiotic that visually inhibits bacterial growth after incubation is referred to as the Minimum Inhibitory Concentration (MIC). This concentration is used to define bacteria as susceptible or resistant.116, 117 However, selection of resistant bacteria starts at concentrations lower than the MIC. Sub-MIC concentrations allow bacteria to grow, but at a slower rate than in antibiotic free environments, and can thereby act as a selective force promoting resistant genotypes.113

A review on reversal of resistance highlights that reduced antibiotic consumption does not always imply rapidly reduced prevalence of resistant bacteria.115 The authors present investigations where reduced levels of resistance are found 118,119 alongside studies that demonstrate no effect.120 One study even found increased prevalence of antibiotic resistance after reduced consumption.121 Even if genetic elements encoding antibiotic resistance impose a fitness cost bacteria harbouring these can still persist. Theoretical models and
Experimental evolution suggest that reduced fitness can be mitigated through compensatory mutations (as opposed to reversal to non-susceptibility), or episodes of periodic selection.\textsuperscript{114,122,123} We have addressed periodic selection and fitness costs related to integrons in Paper III where we combined results from evolutionary experiments with theoretical modelling to investigate conditions that favour MGE’s harbouring antibiotic resistance determinants.

**Summary**

Evidently, several factors affect antibiotic resistance. It is not the scope of the presented work to cover all areas, but rather to put the presented papers into a broader framework. Illustrating the general mechanisms affecting antibiotic resistance demands a balance between the principle of parsimony and enough details to encapsulate the known dynamics. Figure 3 summarizes factors involved in the evolution, spread and persistence of antimicrobial resistance. Each factor depicted in Figure 3 has a number of components or sub-levels that are omitted from display. The topics addressed by study I, II and III are highlighted. Human outpatient antibiotic consumption, selective pressure from antibiotics and horizontal gene transfer are addressed in detail in the three papers.
Figure 3 Flowchart depicting how human activity and naturally occurring phenomena affect the level of antibiotic resistance. The figure is based on a number of references from reviews and original reports, mainly references 6, 15, 102, 113, 115, 124-126. Research topics addressed by paper I - III are coloured in red.
Aims

Overall aim
Describe the antibiotic drug consumption in Norway, and address how heterogenic antibiotic environments can favour bacterial resistance.

Paper I
Identify patterns in Norwegian antibiotic prescriptions and identify prescriber and patient characteristics that are associated with specific antibiotic profiles

Paper II
Determine if regional differences in outpatient antibiotic prescriptions can be explained by a south-north axis at municipality and county level in Norway.

Paper III
Investigate to what extent horizontally acquired resistance traits pose a fitness cost for bacteria, and how periodic selection can favour these resistance traits.
Material and Methods

In the current work three analytical approaches were used.

First, a multivariate analysis was used to summarize prescription data and identify patterns of consumption in the entire Norwegian population.

Secondly, quantile regression was used to investigate the effect of a south-north axis and population size at three different aggregation levels. For model selection DAG and AIC were used.

Finally, we investigated periodic selections that favoured maintenance of an acquired resistance determinant. This was done through a mathematical model that was parameterized with data from laboratory experiments.

These methods are selected on the basis of the research question and the structure of the data available. The methods and data sources are summarized in Table 1 and presented in further detail below.

Computer software

Central to paper I through III is the software and computer language R. This highly flexible computing environment has been used in several versions during the projects, including additional packages. R version 3.0.3 “Warm Puppy” is the latest addition being used. Further packages such as “Vegan”, “quantreg”, “deSolve”, “rgl”, “diagram” and “VennDiagram” have been essential tools to solve the different analytical tasks and graphical presentations. Besides R, SPSS has been used, mainly for organizing data, and finally the online tool DAGitty was used for construction of the DAG model used.

Data sources

Papers I and II are based on data from NorPD. This database is a compulsory registry where all dispensed prescriptions in Norway are registered with information on patient, prescriber, type of prescription, type of drug, amount of drug, the pharmacy and date.
**Variables used**

In Papers I and II variables from NorPD are combined with official statistics on Norwegian demographics (provided by Statistics Norway) and an open source database on geographical locations of administrative centres in Norway.\textsuperscript{136, 137} We used data from 2004 until 2010 and included all prescriptions from ATC class J 01. In accordance with previous studies, we excluded methenamine from the analysis.\textsuperscript{20} Patient age, gender, municipality of residence, prescriber age, prescriber gender and type and amount of drug dispensed were used in the two studies. For a detailed overview see Table 1 and Papers I and II. In Paper III we used data from laboratory experiments (measures of relative fitness) and existing literature to parameterize a mathematical model. We simulated antibiotic pulses at different time intervals and presented median simulation results. ANOVA with post hoc t-test with Bonferroni corrections were used to measure difference in relative fitness between experimental groups of bacteria.
### Table 1 Overview over statistical methods, data sources and variables used in Papers I - III

<table>
<thead>
<tr>
<th>Methods</th>
<th>Origin of Data</th>
<th>Outcome variables</th>
<th>Explanatory variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correspondence analysis</td>
<td>NorPD</td>
<td>DID for 8 antibiotic groups&lt;sup&gt;*&lt;/sup&gt;</td>
<td>Age&lt;sup&gt;*&lt;/sup&gt; and gender</td>
</tr>
<tr>
<td>Permutation tests</td>
<td>Statistics Norway</td>
<td></td>
<td>Patient and prescriber</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Type&lt;sup&gt;®&lt;/sup&gt;, amount and year</td>
</tr>
<tr>
<td><strong>Study 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariable regression</td>
<td>NorPD</td>
<td>Aggregated DID for all antibiotics at three different aggregation levels, data</td>
<td>Residence and age&lt;sup&gt;•&lt;/sup&gt;</td>
</tr>
<tr>
<td>Quantile regression, directed</td>
<td>Statistics Norway</td>
<td>where age adjusted</td>
<td>Patient</td>
</tr>
<tr>
<td>acyclic graphs, Akaike</td>
<td>Norwegian geographical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>information criteria</td>
<td>coordinates</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Amount and year</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Location and population size</td>
</tr>
<tr>
<td><strong>Study 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathematical modelling</td>
<td>Simulations</td>
<td>Results from simulations</td>
<td>Experimental treatment</td>
</tr>
<tr>
<td>Differential equations</td>
<td>Laboratory experiments</td>
<td>Relative fitness of bacterial populations</td>
<td></td>
</tr>
<tr>
<td>Anova / t-test</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>•</sup>Age was estimated by subtracting year of birth from year of dispensing.

<sup>®</sup>ATC codes were used to aggregate antibiotics into 8 groups (see Paper I for a detailed description).

<sup>•</sup>Age was estimated as in Paper I, however age was solely used for a direct age adjusting of consumption at municipality and county level.
**Paper I Multivariate methods and ordination**

Prescriptions were excluded from the raw data if they were connected to institutions, included methenamine (ATC class J01XX05), if age, or gender, for patient and prescriber were missing or non-sense. Multiple entries on date, patient ID and drug type were summarized and duplicates removed, negative DDDs were excluded prior to aggregation.

Before running the multivariate analysis, we assembled ATC codes into 8 antibiotic groups. These 8 groups were used as response variables and were grouped accordingly: Tetracyclines, omitting doxycycline (J01AA), doxycycline (J01AA02), phenoxymethylpenicillin (J01CE), extended spectrum penicillins (J01CA), Trimethoprim (J01E), macrolides (J01FA), quinolones (J01M).

Remaining groups were organized in a non-standardized group called “Others”.

For each year (7 years) data were aggregated on the variables Patient age (8 groups), patient gender (2 groups) prescriber age (4 groups) and prescriber gender (2 groups) resulting in a table consisting of 896 rows each with 8 response variables. Due to differences in dosages to children and adults the data table was split into analyses for patients over and under 12 years of age. For children we addressed only 4 groups of antibiotics, phenoxymethylpenicillin, extended spectrum penicillins, macrolides and trimethoprim.

Correspondence Analysis (CA) is a multivariate ordination method that allows inspection of data with multiple outcomes. It has previously been demonstrated as a powerful tool for displaying data from contingency tables and its use in medical research is documented elsewhere.\(^\text{138}\) Foremost, CA allows visualisation of complex tables, but is not intended for statistical inference such as classical hypotheses testing. The two tables aggregated from NorPD data were processed in a CA, and the results presented as biplots of the two principal axes.

We added a permutation procedure to assess the effect of the demographic variables. We stratified data on year and used 10 000 permutations of the data for these tests.
Paper II Regional difference in outpatient antibiotic consumption

Preparation of data was structured as for Paper I, except that we also removed prescriptions with missing information on patient residence. Erroneous municipality codes were manually corrected with information from Statistics Norway on municipality codes. Patient age was used to perform a direct age adjusting on both municipality and county level to avoid possible bias due to different age structure in the administrative levels. DID for all antibiotics was used as the response variable.

Selection of variables for the statistical model were supported by DAG methodology through the browser-based program, DAGitty version 2.0.\textsuperscript{134,139} We considered variables available through NorPD, Statistics Norway, open source data on geographical entities and unmeasured variables. We constructed a causal diagram and chose the main effects from the minimal statistical model for further analysis. After selection of variables we investigated their relationship: DID at municipality level along a south-north axis; DID as a function of number of inhabitants in a municipality and finally the south-north axis and the number of inhabitants. Data structure revealed conditions that called for an alternative to Ordinary Least Squares regression (OLS), and we estimated parameters for our statistical model by using Quantile Regression (QR).\textsuperscript{140,141} We assessed two different models structures with AIC. One model where all variables interacted was compared to a reduced model with the main effects. However, we let year of dispensing interact with all variables in both models to avoid bias due to temporal dependencies.

Finally, after selecting the model with the lowest AIC value we aggregated the geographical variable in three different levels to assess if this affected parameter results. This was used as a measure of MAUP\textsuperscript{142} and the south-north gradient was considered at county level, municipality level and a south-north ranking solely based on latitude.
Paper III Maintenance of genetic elements imposing a fitness cost

Relying on the experimental estimation of the fitness cost of newly acquired class-1 integron we developed a mathematical model adapted from earlier publications.\textsuperscript{114, 143} We used our experimental estimates and additional parameters obtained from the literature to parameterize the model and investigate under which environmental conditions genetic elements posing a fitness cost could be maintained.

We simulated antibiotic pulses at different time points and evaluated how the dynamics of 5 different bacterial populations interacted. We included a population free of class-1 integrons that had a relative fitness of 1. We further included a population with a class-1 integron with one gene cassette encoding antibiotic resistance and a relative fitness of 0.93 (range investigated: 0.91 - 0.95). We allowed for a population with a class-1 integron with two gene cassettes encoding antibiotic resistance towards two different antibiotics (same fitness cost of integrase). Finally we included two populations with one and two gene cassettes for antibiotic resistance, however allowing class-1 integrase to mutate rendering it free of function and fitness cost.

Methodologically the mathematical model of Paper III is comparable to the graphical presentation of DAG in Paper II. Other authors have addressed how causal diagrams are related to compartmental models. These authors argue for diagrams as an important tool to understand the system being studied, and pave the way for statistical analysis of complex systems.\textsuperscript{144} Further, compartmental models are based on diagrams and in well known models such as the Susceptible, Infected, Recovered epidemiological model (SIR). In this model differential equations reflect the flow between the compartments.\textsuperscript{144, 145} Our model in Paper III builds on the same framework as a general SIR model, with alterations of the equations to fit the causal assumption of the system under study. To derive such a system, we created a flow diagram to clarify and illustrate the assumptions made, see Paper III supportive information.
Summary of papers and main results

Paper I
Haugen P, Primicerio R, Simonsen GS, Furberg AS, Småbrekke L. Antibiotic consumption profiles identified from a prescription database using multivariate analysis [manuscript]

The study aimed at describing the Norwegian antibiotic consumption within age and gender groups of patients and prescribers. Following the CA we identified several patterns connected to prescriber and patient variables.

Prescriber age influences choice of antibiotics. Younger prescribers have a higher proportion of prescriptions for phenoxyethylpenicillin. Older prescribers favoured more broad-spectrum alternatives, depending on patient characteristics.

We also found prescriber gender differences within age strata of male patients. Male prescribers had a higher proportion of trimethoprim, quinolones and “other” antibiotics. Female prescribers had a higher proportion of doxycycline and extended spectrum penicillins in their prescriptions to male patients above 60 years. This pattern of differences between prescriber genders is not present for the female patients, but can be observed for younger male patients.

Patient age and gender influences the antibiotic profiles we identified. There is a high proportion of tetracyclines prescribed to young adult patients and a higher proportion of doxycycline to the oldest patients. In general female patients have profiles more closely related to broad-spectrum alternatives. These are extended spectrum penicillins, trimethoprim, doxycycline and other antibiotics. There is an interaction between prescriber age and patient age. Older prescribers have a higher proportion of broad-spectrum alternatives within each patient age stratum.

The results from Paper I demonstrate demographic differences in antibiotic profiles. In the Norwegian population gender and age groups experience different antibiotic profiles. Difference in amount and type of antibiotics between age groups and genders have been described by other authors, both in Norway
and in other countries. Our results highlights that there is an interaction between patient and prescriber demographic variables. Prescribers have different antibiotic profiles within age strata of patients. Further, female and male prescribers have different profiles when prescribing antibiotics to males, but not to women. In Norway this phenomenon has not been described previously.

Our findings imply that there are both patient, and prescriber clusters of antibiotic profiles in the population. Therefore the selection pressure on bacteria is different between and within demographic groups of patients.

We answered the aim of the project by identifying not only individual patient and prescriber variables in prescription patterns, but also demonstrated interaction between these variables.
Paper II

Haugen P, Simonsen GS, Primicerio R, Furberg AS, Småbrekke L. Outpatient antibiotic use in Norway depends on municipality population size [manuscript]

The aim of the paper was to investigate regional differences in antibiotic consumption in Norway.

There is a 6-10 fold variation in antibiotic consumption, measured in dispensed DID between Norwegian municipalities. Both at county and municipality level it appears that the consumption declines with latitude. However, adjusting for population size in municipalities demonstrates that the consumption depends on population size, latitude and the interaction between these. The south-north decline in antibiotic consumption diminishes with increasing municipality population size. The results are consistent between the 20th and the 80th percentile. This indicates similar effects of the variables among municipalities with high and low antibiotic consumption.

Different levels of aggregating municipalities, which give some reassurance towards MAUP effects, do not affect the main results. We cannot however, fully exclude these effects. We found no support for yearly differences when we estimated the effects between year and any of the included variables.

Recently Norwegian regional differences have been addressed in reports where the difference in total DID and prevalence of prescribing has been reported to differ between counties and health regions.\textsuperscript{17,148} We demonstrate that this effect is to a large extent related to municipality population size. By aggregating consumption data over several municipalities, variation within counties and health regions are lost. It is important to acknowledge that there is substantial variation between municipalities within a county (and health region), and this variation has to be considered in analysis of data.
Paper III


We aimed at determining the fitness cost of newly acquired integrons and model conditions that favoured maintenance of both functional and non-functional integrases in bacterial populations.

Initially, a newly acquired integron in *Acinetobacter baylyi* led to a fitness cost. However, when mutating the integrases experimentally, the cost related to the integron was reduced. This also happened during experimental evolution were bacteria were able to mitigate the cost of the acquired integrons after 30 – 42 days.

The mathematical model predicted loss of the functional integrases in the bacterial population when selection pressure imposed by antibiotic drugs was absent.

The combination of a theoretical model and laboratory experiments showed that antibiotic pulses as far apart as 40 days would favour maintenance of integrases in bacterial populations, under laboratory conditions.

Using a combination of both laboratory and theoretical approaches we conclude that initially costly integrons can be maintained in bacterial populations. We demonstrate that time between selective events (i.e. antibiotic selection) is the key parameter under the conditions explored.
Discussion

The recent report by the WHO on surveillance of antimicrobial resistance highlights several microbes as especially worrying in an international public health perspective. Among these we find the species *Escherichia coli*, *Klebsiella pneumoniae*, *Staphylococcus aureus* and *Streptococcus pneumoniae*. These bacteria may potentially cause lethal infections, and their resistance against several antibiotics is threatening effective treatments worldwide. The Centres for Disease Control and prevention (CDC) list seven bacteria, or groups of bacteria, carrying antibiotic resistance genes that cause particular concern. Among these we find the Carbapenem-Resistant Enterobacteriaceae (CRE) and multidrug resistant *Salmonella* spp. The CDC estimates that CRE are responsible for 600 deaths in the USA annually and that the cost related to drug resistant non-typhoidal *Salmonella* spp. infections are in the range of $365 million annually. Recently investigators found evidence for an increase in the use of antibiotics for acute bronchitis in the US between 1996 and 2010. This was in contrast to advices given by CDC to American practitioners.

Investigators have tried to address the cost of antibiotic resistance on healthcare systems. However antibiotics have become such an integrated part of healthcare that estimating the cost is an overwhelming task. It is not only treatment of infectious diseases that will be affected, because the entire healthcare system relies on the effect of antibiotics.

This illustrates the importance of reducing antibiotic consumption. Reduced levels of consumption will lower selection for resistance determinants in bacterial populations. Reducing and preventing antibiotic resistance is one of the major health related challenges for both international and Norwegian health authorities.

A recent comment in Lancet Infectious Diseases calls for studies of regional antibiotic consumption within countries, and the authors also stress the need for studies investigating determinants for antibiotics consumption. In Europe, information on consumption and resistance is extensive, and a number of studies on antibiotic consumption are available. This is in contrasts to
other regions of the world were certified laboratories for microbiological diagnostics and surveillance systems are scarce.\textsuperscript{89} The Nordic countries, with a general low consumption, have well integrated routines for reporting consumption and resistance.\textsuperscript{11, 29, 154, 155} In Papers I and II we present a comprehensive overview of Norwegian antibiotic consumption related to demographic variables and geography. The methodology used in these manuscripts helped to detect several patterns of consumption that have not previously been described in Norway.
**Methodological considerations**

Whether pharmacoepidemiological studies rely on ecological or individual data they share a common goal: address patterns of consumption in a population to aid in decision making for treatment of disease.\(^{156}\) New methods are constantly being developed, or implemented, in healthcare research to address studies with complex data structure.\(^{139,157}\) Two challenges confront researchers working with the large and complex datasets characteristic of pharmacoepidemiological studies. Firstly how to investigate complex causal relationships, and secondly how to cope with multiple, correlated response variables.

Multivariable regressions are predominant in pharmacoepidemiological investigations of antibiotic consumption.\(^{32,33,47,48,74,82,85}\) The use of p values as inferential tools to choose statistical models are commonly encountered. From my perspective, the choice of a statistical model to explain the underlying data structure is of paramount importance in statistical inference, but overlooked in several published studies. The choice of significance testing in a multiple regression setting is predominant, but not without controversy.\(^{144,158}\) Some authors have argued for the implementation of information criteria instead of traditional hypothesis testing.\(^{159,160}\) Among the available information criteria for model selection the AIC is recommended, although alternatives exists depending on type of analysis and data structure.\(^{159}\) In fact, there are strong arguments against using classical hypothesis testing as a method of statistical inference in multiple regressions. However, the alternative use of information criteria for model selection also comes with a risk of selecting suboptimal models, especially if there is a high degree of uncertainty associated with measurement errors.\(^{144,159}\) Discussions on the topic have moved from the scientific, and philosophical, literature over to blogs and a well-structured post on the topic has been published online.\(^{161}\)

Traditionally, goodness of fit estimates for selection of linear regression models have relied on \(R^2\) values. \(R^2\) measures the amount of variance accounted by the regression model. However, in a multivariable setting \(R^2\) is not favourable as a selection criterion for models. The amount of variance explained increases as
more terms are added to the statistical model, thus selecting for larger models and ignoring the principle of parsimony.162

In addition to the use of information criteria in model selection causal diagrams and DAG have been implemented into medical research when specifying candidate models.139,163 Easily accessible online tools with well-defined user guides exist, alongside more code demanding technical solutions, for the implementation of causal analysis to avoid bias.134,164 Despite the availability of DAG tools this methodology has not been considered in pharmacoepidemiological studies concerning antibiotic consumption. Although methodological examples exist, an extensive search of the literature for combinations of “antimicrobial”, “epidemiology” and “causal diagrams” revealed no pharmacoepidemiological studies focusing on antibiotics and DAG methodology. However, recent studies have implemented interesting new methodology on variable selection within the General Estimating Equations (GEE) setting. These authors used a penalized GEE as a data-mining tool to build statistical models in longitudinal data.165,166

Multivariate analyses rely on statistical tools for investigating data with multiple outcome variables.138,167 The expression multivariate analysis is sometimes misinterpreted in pharmacoepidemiological literature in studies that are in fact multivariable.76,79,80 A multitude of multivariate, ordination techniques exists that allow to summarize relationship among dependent variables. Among these we find Principal Components Analysis (PCA), and CA.168

The databases found in the Nordic countries allow connection of these data to other sources of information. Data dredging a serious concern in studies attempting to address hypothesis testing, especially when the vast amount of data is taken into consideration.159 Multivariate analysis avoids this challenge, by focusing on pattern discovery and characterization rather than hypothesis testing. Large contingency tables can be expressed in a two-dimensional space with the aid of a CA and a biplot, and several types of studies can benefit from these graphical techniques to summarize multidimensional results in one figure.138 Evaluation and implementation of tools for data mining has been called
for in the field of pharmacoepidemiology, both CA and PCA are techniques that can meet some of these demands.\textsuperscript{169}

Recently, in a paper including Italian prescriptions, the authors used network analysis as a data-mining tool on prescription data to investigate which classes of drugs that are co-prescribed.\textsuperscript{170} This investigation shows how patients below 30 years of age have more connections between antibiotics and respiratory and hormonal drugs than older age groups.

Whereas regression modelling is frequently used in pharmacoepidemiology, mathematical models are less frequent. A review of the impact of mathematical models in epidemiology has pointed out that articles containing mathematical modelling is among the top cited articles, and that the number of such papers are increasing.\textsuperscript{171} In the antibiotic resistance perspective mathematical modelling is frequent, and several papers address population biology of bacteria, but also transmission dynamics both within and outside hospitals.\textsuperscript{5, 114, 123, 143, 171-173}

In Norway the systematic registration of antibiotic sales to patients has resulted in a large, comprehensive database. Therefore, analysis of outpatient antibiotic consumption provides the richest source of information to address the major antibiotic use, and possibly the main cause of selection pressure on bacteria in the country. The complex character of these data calls for an upgrade of the statistical tools used in pharmacoepidemiological investigations of patterns and causality. Combinations of mathematical modelling and data mining tools such as network analysis and multivariate methods opens for interesting in-depth analyses of these routinely collected data.
Limitation of studies

Paper I and II are based on data from a national prescription database with mandatory entries, and the following limitations should be considered.

Although individual data are available in NorPD, we used aggregated to meet the intended degree of resolution and for ease of interpretation. Further, the aggregation reduced computing time, and reduced response variables into interpretable categories. However, the aggregation reduced information and decreased variation in the data. The patterns that occur in the CA are therefore not from individual prescriptions, but from the groups that were chosen for aggregation.

In paper II we modelled relationships between variables as linear dependencies. The relationship between antibiotic consumption and municipality size indicates that this is a coarse approximation. Given the non-linear response the linear model will capture the sign and magnitude of the relationship but it should not be used for prediction. For prediction a model considering nonlinear terms should be considered. The advantages of linear models are the ease of interpretation, the robustness of estimates and perhaps the familiarity of linear regression. In a paper to the Journal of the American statistical association George E. P. Box comments on the question on normality and linearity:

"Equally the statistician knows, for example, that in true nature there never was a straight line, yet with normal and linear assumptions, known to be false, he can often derive results which match, to a useful approximation, those found in the real world." 174

Compartmental models, DAGs and diagrams are excellent tools for structuring data analysis. These methods allow researchers to get an overview of a study through visualisation of the system. However, there is a risk of oversimplifying very complex systems into a selection of “available variables”. Further, a preliminary study of the expected causal diagram prior to analysis would be beneficial to avoid the researcher adapting the causal model to the data structure.144
Implications of the studies

In Paper I we show how demographic variables shape clusters of different antibiotic profiles in the Norwegian population. In Paper II we show how the volume of antibiotics prescribed vary in space within Norway. Among municipalities the variation in total outpatient antibiotic consumption is substantial. This opens for a scenario where amount of consumed antibiotics vary between demographic groups across municipalities. In an antibiotic resistance perspective this is of great interest, as the selective pressure imposed on bacteria will vary depending both on demographic groups and administrative units. If selection pressure is high within a specific demographic group this effect can be further enhanced through transmission dynamics. Analyses of social networks have shown that contact and possible transmission of disease happens within age-strata of a population, especially when the prevalence of disease is low.\textsuperscript{175} Age is an important factor for creating clusters in social networks; however, this depends on the interaction between nodes.\textsuperscript{176}

Although differences in antibiotic consumption can be found between clusters of the population, differences in antibiotic resistance among demographic groups are not frequently reported. Some studies do however point to age and gender related differences in proportion of antibiotic resistant bacteria. In the US, age of patient has been shown to be a predictor for antibiotic resistance in patients suffering from urinary tract infections. Patient over 65 years had a higher risk for an infection with bacteria resistant towards at least two antibiotics.\textsuperscript{177} In a UK study, male patients with tuberculosis were more likely to have multidrug resistant infections than women.\textsuperscript{178} In Australia it has been shown that indigenous people are at increased risk of having a methicillin resistant \textit{Staphylococcus aureus} infection.\textsuperscript{179} In 2012 the EARS-Net report investigated demographic factors related to resistance in three species of bacteria. They found that males had higher proportion of \textit{E. coli} and \textit{K. pneumonia} resistant towards fluoroquinolones, 3\textsuperscript{rd} generation cephalosporins and aminoglycosides than females. The proportion of resistant \textit{E. coli} increased with age, whereas for \textit{K. pneumonia} the proportion decreased as patient age increased. Combined resistance in \textit{P. aeruginosa} was defined as bacteria resistant towards three or
more antibiotic classes. For these bacteria the oldest patients had the lowest resistance proportions.\textsuperscript{180}

Study III illustrates how pulses of multiple antibiotics can favour multidrug resistance in bacterial populations. Bacteria harbouring integrons can spread within demographic groups with a specific antibiotic profile and thereby experience the necessary heterogeneity required for a functional integrase. If gene cassettes harbouring antibiotic resistance are integrated, selection of antibiotic resistant bacteria within a demographic group will occur. Integrons can harbour several gene cassettes. Selection of one antibiotic can indirectly select for the other acquired gene cassettes, thus effectively maintaining multidrug resistance. This shows how the combination of clusters with specific antibiotic consumption and transmission opens for selective windows that could contribute to the evolution of antibiotic multidrug resistance.

A combination of analysis of resistance patterns and theoretical modelling can answer if our laboratory and theoretical results can have an impact on evolution of resistant bacteria in Norwegian outpatients.
Conclusions

In Paper I we identified patterns among patient and prescriber variables that form clusters of antibiotic consumption and prescribing. We showed how prescriber age influence antibiotic prescription profiles and that there is an interaction between demographic variables for both patient and prescriber.

In Paper II we found a maximum of a 10-fold difference in the outpatient antibiotic consumption when investigating Norwegian municipalities. Consumption in municipalities with small population sizes have decreasing consumption along a south-north axis, this effect is not present for municipalities when population size increases.

In Paper III We found that the fitness cost of integrons are mitigated through bacterial evolution, and that heterogenic environments in the form of antibiotic pulses aids in maintaining a functional integrase.
Future aspect

A natural continuation of Paper I would be to investigate the difference in prescription profiles between age groups of prescribers including indication of disease. Are the differences due to prescriber habits, or can they be explained through differences in diseases treated within the age strata of prescribers?

The observed gender difference between female and male prescribers regarding prescriptions to male patients should be further investigated to identify possible predictors of this difference.

The results from Paper I can also be used to target patient age, or gender, groups with the aim of reducing consumption of specific antibiotic types.

CA could also be implemented into reporting on antibiotic consumption in both national and international reports. ESAC and NORM/NORM-VET reports are areas where such graphic displays can aid in understanding complex tables of consumption.

An interesting development of Papers I to III would be to combine the information obtained into computer simulations. Demographic differences in consumption profiles, regional differences in antibiotic consumption and the effect of periodic selection on bacteria can be combined with existing knowledge on network transmissions. This could aid in increasing the understanding of how resistant bacteria evolve and spread in the outpatient setting.
References


94. Delsol AA, Randall L, Cooles S, et al. Effect of the growth promoter avilamycin on emergence and persistence of antimicrobial resistance in


128. *Chen H. VennDiagram: generate high-resolution Venn and Euler plots* [computer program]. Version 1.6.7; 2014.


133. *Adler D, Murdock D. Rgl: 3D visualization device system (OpenGL)* [computer program]. Version 0.93.996; 2014.


135. *IBM. SPSS, PASW statistics* [computer program]. Version 21.0.0.0; 2012.


166. Blommaert A, Hens N, Beutels P. Data mining for longitudinal data under multicollinearity and time dependence using penalized generalized


Paper I
Paper II
Paper III