THE UNIVERSITY *of* EDINBURGH

Monitoring the spread of antibiotic resistance in wastewater

Hannah Catherine Lepper

THE UNIVERSITY *of* EDINBURGH

Thesis submitted for the degree of Doctor of Philosophy

The University of Edinburgh
July, 2022

Declaration

I declare that this thesis and the analysis described in it are my own composition, except where explicitly stated. This work has not been submitted for any other degree or professional qualification.

Hannah C Lepper

July, 2022

Abstract

Background: Antibiotic resistant bacterial infections are causing a growing amount of morbidity and mortality. Effective control and prevention relies on good data on the current burden of antibiotic resistance (ABR). Traditional ABR surveillance from phenotypic, passive, hospital-based testing may not adequately represent the resistome of the general population. Wastewater metagenomics has been proposed as a new type of surveillance to overcome this limitation. It generates rich, quantitative information on the bacterial species and resistance genes of a whole community. Large wastewater metagenomic datasets are now available to monitor and explore drivers of ABR in the community. However, questions remain about how to collect, analyse, and interpret these novel datasets. In this thesis, I aimed to 1) address key unknowns in wastewater data, including sources of resistance, environmental resistance dynamics, and what statistical models describe the distribution of the data well, and 2) investigate global and local patterns in wastewater resistance and identify potential community and hospital drivers.

Methods: I used a systematic review to find evidence in the literature for dissemination of ABR from hospitals to wastewater. I next developed a compartmental transmission model to investigate environmental resistance dynamics and its impact on human ABR levels. I implemented a multi-response statistical model to correlate hospital-based surveillance (EARS-Net) data with resistance gene abundance in sewage samples from around the world analysed with metagenomics by the Global Sewage Surveillance Project. Finally, I used a paired sampling design and multiple statistical methods to compare the resistome of sewage from hospitals, communities, and wastewater treatment plants (WWTPS) in Scotland. I also investigated the links between ABR in humans and antibiotic consumption in the modelling and data analysis chapters.

Results: I found increasing evidence in primary research that resistant bacteria and resistance genes can be disseminated from hospital patients to wastewater and into natural water sources. Modelling the dynamics of ABR in an environmental reservoir indicated that the environment can theoretically influence human ABR levels as much as or more than an animal reservoir, and mitigate intervention impacts. Combining EARS-Net and sewage metagenomic data indicated that some types of ABR are positively correlated in sewage and hospitals (such as aminoglycosides), but many are not (such as vancomycin and aminopenicillins). The paired sampling study demonstrated that hospital and community sewage resistomes are distinct, and WWTPs mostly reflect community sewage resistomes. I found mixed evidence for an impact of antimicrobial consumption on human ABR levels. Overall, the impact of antibiotic consumption at the population level appears to be small in these datasets.

Conclusions: Wastewater metagenomics is a valuable way of monitoring ABR in the community. It can indicate the composition of the reservoir of ABR in the general population and what drives it. However, hospital rather than mixed municipal effluent may need to be collected to monitor clinical resistance patterns. To make the most of this new source of data more flexible modelling frameworks that account for

wastewater metagenomics specific factors such as high dimensionality and overdispersion. Comparing resistance patterns in hospitals to community sewage implied that patients and/or the hospital environment may present unique and strong selection pressures for resistance. Finally, we also show that differential antibiotic consumption alone cannot explain the observed patterns in resistance abundance on the national or international level.

Lay summary

Antibiotic resistance is the ability of bacteria to escape the effects of antibiotics, making infections harder or impossible to treat. To prevent illness and death from resistant infections, we need to stop the bacteria from spreading and treat infections with the right kind of antibiotics. My thesis is about how we can use data from sewage samples to understand why resistant bacteria spread and what kinds of resistance are most common.

Many healthy people in the community carry resistant bacteria in their gut without knowing, but these bacteria could cause serious infections later or spread to other people. We can use samples of sewage to learn more about resistance in the community. Sewage samples are easy to collect, but contain bacteria from the faeces, urine and skin of thousands of people. The profile of all the different kinds of bacteria and resistance genes in a sample can be identified by looking at the DNA in a sewage sample. Analysing all the DNA of a sample in this way is known as metagenomics. By combining sewage and metagenomics, we have a very rich source of information on the resistance profile (or resistome) of the community.

In this thesis I asked: what is in the community resistome, and how and why does the community resistome differ between countries? Secondly, what sampling and data analysis methods should be used to make the most of sewage metagenomics?

I reviewed existing studies of resistant bacteria and resistance genes from hospitals in wastewater. These studies showed that resistant bacteria can be carried from hospital patients and environments to wastewater.

I next studied sharing of resistant bacteria between humans, animals, and the environment with a computational model. The results showed that the environment should be considered as an additional source of resistance for humans, and it can make preventive measures like reducing antibiotic usage in animals less effective.
I also compared the resistance levels of patients in hospitals with community sewage in countries across Europe. Other studies have also made this comparison, but here I used a statistical model that made use of more information about both datasets. Hospital and sewage resistance patterns reflected each other in some types of resistance but not in other types.

I went on to explore more about how the resistance profile of hospital and community sewage might be different by collecting sewage samples from hospitals, groups of households, and from the whole community (including hospitals). For the statistical analysis I used three different methods to make our results more robust. Groups of households and the whole-community sewage had similar resistomes, and the hospital sewage resistome had more types of resistance genes.

I have shown in this thesis that sewage metagenomics is a valuable way of monitoring resistance in the community. Wastewater can also be used to look at hospital resistance, especially in samples of sewage from directly outside hospitals. However,

we also demonstrate that the resistome of patients may be quite different to the resistome of healthy community members. Finally, we have shown that using flexible data analysis methods that reflect the structure of sewage metagenomics better is important to make the most of wastewater metagenomics.

Acknowledgements

I would first like to thank my supervisors, Bram, Mark, and Luke, for giving me the opportunity to do this PhD, and for all of their essential help and guidance throughout the projects.

Thank you to the Novo Nordisk Foundation for funding this work from the grant 614EPI R83810. In addition, thank you to the Academy of Medical Science for their grant to Meghan Perry (SGL016_1086), a portion of which funded data collection for Chapter 5.

Thank you also to Meghan Perry for taking a chance on my study design idea and making it possible, and for many helpful comments on the implementation and interpretation. Thanks Jarrod Hadfield, for technical expertise in model design and implementation. Thanks also to my annual review panel chair, Helen Stagg, for additional support and a new perspective on the projects.

And thank you to the many other people who helped with study design, data collection, and analysis: Scottish Water, especially Mark McEwan, David Wills, and Moray Simon; NHS Lothian, especially Carol Philips and Barbara Moore; and DTU, especially Frank Aarestrup, Hanne Nielsen, and Patrick Munk.

Thank you to the other members of Epigroup past and present for technical and academic advice and moral support, and for being my friends. Thanks to Bryan Wee for help with genomic analysis, and to Leonie Maier and Alex Morgan for comments on sections of this thesis and idea generating discussions. Thanks also to Olga Dluzniewska for working with me to do some additional analysis of the Chapter 5 data during her internship in June, 2022. And the others: Roo Cave, Feifei Zhang, Dishon Muloi, Jordan Ashworth, Shengyuan Zhao, Thomas Dalhuisen, Stefan Rooke, and Lu Lu – thank you all!

Finally, thank you to my family and friends for their love and support all the way to the end, especially my partner Mike, my parents Patricia and Clive, and my sister Ellamae.

Publications

A previous version of Chapter 3 is published as preprint on bioRxiv:

Hannah C. Leper, Mark E.J. Woolhouse, and Bram A.D. van Bunnik. The role of the environment in transmission of antimicrobial resistance between humans and animals: a modelling study. bioRxiv, May 2022. Doi: 10.1101/2022.05.19.492687

Abbreviations

ABR	Antibiotic resistance
ARG	Antibiotic resistance gene
AST	Antimicrobial susceptibility testing
ATC	Anatomical Therapuetic Chemical Classification
CAESAR	Central Asia and Eastern European Surveillance of Antimicrobial Resistance network
CDC	Centre of Disease Control
CI	Confidence intervals
CLSI	Clinical & Laboratory Standards Institute
DDD	Defined Daily Dose
DGGE	Denaturing gradient gel electrophoresis
DNA	Deoxyribo nucleic acid
EARS-Net	European Antimicrobial Resistance Surveillance Network
ECDC	European Centre of Disease Control
ELPD	Expected Log Pointwise Predictive Density
ESAC-Net	European Surveillance of Antimicrobial Consumption Network
ESBL	extended spectrum ß-lactam
EUCAST	European Committee on Antimicrobial Susceptibility Testing
FAST	Fourier amplitude sensitivity test
FPKM	Fragments per kilobase million
GDP	Gross domestic product
GLASS	Global Antimicrobial Resistance and Use Surveillance System
GPS	Global positioning system
GSSP	Global sewage surveillance project
HDI	Human development index
HIC	High income countries
HWW	Hospital wastewater
KMA	*K*-mer alignment
LIC	Low income country
LMIC	Lower and middle income countries
MALDI-TOF	Matrix-assisted laser desorption/ionization - time of flight
MLST	Multi-locus sequence typing
MRSA	Methicillin resistant *S. auerus*

MWW	Municipal wastewater
NIH	National Institute of Health
OBD	Occupied Bed Days
PCR	Polymerase chain reaction
PECO	Population, Exposure, Comparison, Outcome
PFGE	Pulsed-gel electrophoresis
qPCR	Quantitative polymerase chain reaction
ReLAVRA	Latin American and Caribbean Network for Antimicrobial Resistance Surveillance
SD	Standard deviation
UI	Uncertainty intervals
VRE	Vancomycin-resistance enterococci
WAIC	Widely Applicable Information Criterion
WHO	World Health Organisation
WWTP	Wastewater Treatment Plant

List of Figures

List of Tables

Table of Contents

1 General introduction

1.1 What is antibiotic resistance?

Antibiotic resistance (ABR) is the ability of bacteria to tolerate or prevent the effects of the group of compounds used to treat bacterial infections known as antibiotics. Antibiotics are a staple of modern medicine, reducing mortality from infectious diseases throughout the 20th century (Cohen, 2000), as well as making a range of life-saving surgeries and immune-compromising chemotherapies possible (Smith & Coast, 2013). However, since the discovery of penicillin in 1928, bacterial resistance has repeatedly emerged soon after the introduction of nearly every new antibiotic (J. Davies & Davies, 2010; Ventola, 2015). Bacteria themselves produce antibiotics and other small molecules to compete with their bacterial neighbours, destroying others nearby or inhibiting their growth (Granato et al., 2019). This means that bacteria have already been in an arms race with each other to develop new antibiotics and gain resistance for millennia (D'costa et al., 2011). Although there is much research into new antibiotics, there is a declining discovery rate (Hutchings et al., 2019), and doubt that current resistance can be reversed (Andersson & Hughes, 2010): resistance may be a permanent feature of infection medicine.

Antibiotic resistance is also a serious threat to global public health. Recent estimates suggest that 0.911 – 1.71 million deaths attributable to resistant infections occurred in 2019 (Murray et al., 2022). Forecasting estimates of the future burden of ABR are highly uncertain due to a lack of reliable data, particularly in Lower and Middle Income Countries (LMICs) (de Kraker et al., 2016), but one estimate suggested there could be 10 million deaths as a result of resistant infections in 2050 (O'Neill, 2014). Untreatable infections will also have a knock-on effect on the economy and the standard of living. The World Bank estimates that without a reduction in the current burden of antibiotic resistance, a 1.1 – 3.8% fall in GDP is likely by 2050, as well as losses in livestock production, increases in health care costs, and increases in poverty (World Bank, 2017). In LMICs, the number of people falling into extreme poverty could increase by 6.3 – 26.2 million people by 2050 (World Bank, 2017). A hindrance to combatting ABR is a lack of good quality, non-resource-intensive, heterogeneous, information-rich surveillance data on resistance levels in the general population – metagenomic analysis of wastewater has been proposed as a means to fill in these gaps (Aarestrup & Woolhouse, 2020; Pruden et al, 2021).

In this thesis, I investigate the use of wastewater metagenomics for surveillance of antibiotic resistance. The general introduction will describe how and why bacteria gain resistance and spread, the epidemiology of ABR in humans, animals, and the environment, and the use of surveillance, particularly wastewater metagenomics, for understanding and monitoring resistance. I will then summarise my aims, objectives, and methodology.

1.2 Mechanisms, emergence, and spread of resistance in bacteria

Several routes to antibiotic tolerance or resistance are found in bacteria. Firstly, bacteria can break down or modify the antibiotic molecule so it is no longer effective (Blair et al., 2015). For example, enzymes such as carbapenemase hydrolyse β-

lactam antibiotics before peptidoglycan synthesis (the antibiotic target) is disrupted (Llarrull et al., 2010). Secondly, bacteria can reduce the intracellular concentration of an antibiotic (Blair et al., 2015). Reducing cell wall permeability can decrease uptake of an antibiotic, or increasing expression of efflux pumps can remove antibiotics in the cell (Blair et al., 2015). Overexpression of efflux pumps that target antibiotics such as tetracycline are particularly common in Gram-negative bacteria (X. Z. Li et al., 2015). Thirdly, the target of an antibiotic can be altered (Blair et al., 2015). For example, fluoroquinolone destroys cells by binding to topoisomerases and preventing their action to unwind supercoiled DNA (Aldred et al., 2014). Resistance to fluoroquinolones is often caused by mutations to the topoisomerases which mean that the antibiotic can no longer bind (Aldred et al., 2014).

Resistance can be gained or intrinsic. Intrinsic resistance is a particular problem in Gram-negative bacteria, which have lower cell wall permeability than Gram-positive bacteria, and a greater number of efflux pumps (Cox & Wright, 2013). Gained antibiotic resistance comes in two types: mutations and acquired genes (J. Davies & Davies, 2010; Munita & Arias, 2016). Mutation can alter the structure of an antibiotic target, and can also adapt existing genes to generate new resistance determinants (J. Davies & Davies, 2010). Transmission of gene mutations is vertical, i.e. these resistance genes are spread through bacterial replication. Gene acquisition, on the other hand, occurs when a bacteria takes in DNA from another organism through horizontal gene transmission (Munita & Arias, 2016; Peterson & Kaur, 2018). A gene can be introduced into a new bacterial cell by conjugation, when plasmids are transferred between bacteria on specialised pili; through transduction, when genes are incorporated into the genome of a bacteriophage and transmitted to another cell; or through transformation, when bacteria take up free DNA from the substrate around them (Peterson & Kaur, 2018). Conjugation is thought to be the most common form of horizontal gene transfer of resistance genes (von Wintersdorff et al., 2016), although transduction and transformation also appear to be frequent, especially in the environment (Calero-Cáceres et al., 2019; Peterson & Kaur, 2018). Horizontal gene transfer is an ancient microbial process, allowing many bacteria to gain access to a diverse repertoire of genes (the pangenome) which can provide adaptive benefits to the local environment (Brockhurst et al., 2019).

Acquired and intrinsic antibiotic resistance mechanisms can give bacterial lineages a selective advantage in the presence of antibiotics. With increasing use of antibiotics in medical and veterinary practice over the last 100 years, selection pressure for antibiotic resistance has also increased (J. Davies & Davies, 2010). For example, antibiotic consumption in humans is thought to have increased by 65% between 2000 and 2015 alone – this increase is mostly driven by usage trends in LMICs, but HICs are still the majority consumers of antibiotics (Klein et al., 2018). In addition, antibiotics are increasing in concentration in the environment, potentially creating more selection pressure (Polianciuc et al., 2020). Rapid selection of resistance in response to exposure to antibiotics has been documented extensively with in vitro and in vivo experiments (Palmer & Kishony, 2013), and studies have also found that antibiotic exposure is a risk factor for resistant infections and in humans (Chatterjee et al., 2018). It is therefore crucial that antibiotics are not overused, in order to reduce unnecessary increases in selection pressures. In addition to problems of overuse, suboptimal usage

of antibiotics, such not treating early or failure to finish a treatment course, has been observed to select for resistance (Pradipta et al., 2018). In addition, use of antibiotics may lead to 'bystander selection', or selection pressure for resistance in bacteria that are not the therapeutic target, such as commensal or potentially pathogenic gut bacteria (Morley et al., 2019). Co-selection may also occur, where direct selection for a gene conferring resistance to an antibiotic or metal that is at locally elevated concentrations can also select for other antibiotic or metal resistance genes. This can be through linkage, when the genes are nearby on the chromosome, on the same mobile genetic element, or part of the same operon (Baker-Austin et al., 2006). In addition this can occur when a single resistance gene provides multiple resistance or cross resistance, such as the DsbA–DsbB system, with increases efflux of several metals and antibiotics (Baker-Austin et al., 2006). Co-selection and bystander selection are particular concerns for microbial communities exposed to multiple anthropogenic pollutants, such as fresh water microbiomes (Imran et al., 2019). To reduce selection pressure from overuse and suboptimal use of antibiotics, policies to guide antibiotic prescribing in humans and livestock have been introduced (Charani et al., 2021; Majumder et al., 2020). Testing the effectiveness of these interventions at population-level is challenging, but some studies have shown that improved antibiotic prescribing has been followed by a population level decrease in resistance prevalence, e.g. (Aliabadi et al., 2021).

The spread of antibiotic resistance also involves on transmission of the resistant bacteria between hosts. The mode of transmission of resistant bacteria depends on the bacterial and host species. As a wide variety of bacterial species carry resistance, there is also a variety of transmission modes: contact, such as faecal-oral transmission of *E. coli* or sexual transmission of gonorrhoea; airborne, such as tuberculosis; droplet, such as meningococcus; vector borne, such as Lyme disease; and vehicular, such as via food or environmental fomites (Doron & Gorbach, 2008). Once an invading resistant bacterial colony has been transferred to a new host, it must be able to establish a population for transmission to be complete. The complex microbial ecology in the host microbiome can affect establishment (Kim et al., 2017). For example, the mammalian gut is a dense microbial community, with competition for space and nutrients (Kim et al., 2017). Resistance genes may confer a cost on their host, such as reducing the speed of replication for bacteria carrying large plasmids (Andersson et al., 2007). If the newly introduced resistant bacteria are not competitive, a colony may not be established. Exposing the host to antibiotics could make establishment of new resistant infections more likely by selecting for resistance, and compensating for the cost of resistance (N. G. Davies et al., 2019). At the same time, the mechanism of antibiotic resistance can influence the fitness of resistant bacteria in a microbial community; for example, those that increase intercellular concentrations of antibiotics such as efflux pumps may lead to depletion of nearby competing cells that are sensitive (Bottery et al., 2016). Establishment of infection is also affected by host immunity (Wheatley et al., 2021). The interplay of transmission modes, microbial ecology, and host immunology result in complex dynamics with multiple theoretical outcomes.

1.3 Occurrence of antibiotic resistance in humans, animals, and the environment

The number of resistant infections and the rates of resistance (measured as proportion of bacterial isolates that are resistant) is increasing. Estimates of the current burden of resistance in humans include: 33,110 deaths attributable to resistant bacterial infections in Europe in 2015 (Cassini et al., 2019); 35,900 deaths attributable to infections from resistant bacteria and *Candida* in 2019 in the US (CDC, 2019); and 1.27 million deaths attributable to resistant bacterial infections worldwide in 2019 (Murray et al., 2022). The estimated proportion of bacterial isolates that are resistant depends on the bacterial species, type of resistance phenotype (e.g. tetracycline vs ampicillin resistance), site of infection (e.g. blood-stream or urinary tract infection), origin of infection (hospital- or community-acquired), country of study, and the methodology used to generate the estimate. For example, methicillin-resistant *Staphylococcus aureus* (MRSA), estimated to have caused 150,000 resistant infections in Europe in 2015 (Cassini et al., 2019), has an overall prevalence of less than 5% in Europe but may be as common as 80% in countries in Africa and the Middle East, according to the Global Burden of Disease study (Murray et al., 2022). The WHO Global Antimicrobial resistance and use Surveillance System (GLASS) report, on the other hand, reports that the proportion of patients with a blood-stream infection caused by MRSA was 33% in LMICs compared to 15% in High Income Countries (HICs) (WHO, 2021a). Similarly, fluoroquinolone-resistant *E. coli* was found in <30% of isolates in Europe and North America, but up to 80% in southern Asia (Murray et al., 2022). Globally, the proportion of patients with community-acquired ciprofloxacin-resistant *E. coli* is around 30%, rising to around 45% in patients with hospital-acquired infections (CDC, 2019).

Current estimates of the spread of resistance focus on invasive infections. However, commensal and asymptomatically carried bacteria in healthy human gut microbiome also contain resistance genes (Lamberte & van Schaik, 2022). Theoretically, these bacteria could contribute to transmission of resistance to pathogenic bacteria and cause infections in the host, but little is known about the actual health risk attributable to a resistance gene in a healthy gut microbiome (A Zhang et al., 2021). Estimates of the prevalence of resistance in healthy gut communities is rare, but include a few studies of resistance in commensal *E. coli* isolates from community members: 72% of isolates resistant to ampicillin in LMICs between 1989 and 2020 (Nji et al., 2021); 44% of isolates resistant to ampicillin in China in 2009 (B. Li et al., 2014); and 45% of isolates resistant in the Paris area in 2010 (Massot et al., 2016). Collections of metagenomic data from faecal microbiomes from healthy individuals are also illuminating; for example, demonstrating that tetracycline resistance genes are the most common kind of antibiotic resistance gene (ARG) in healthy human guts, and that gut metagenomes from China had the highest resistance gene abundances (Feng et al., 2018).

Resistant bacteria are found in humans, animals, and the environment. Many bacterial species, such as *E. coli*, can be exposed to antibiotics and gain resistance in more than one of these groups, and can be transmitted between groups (McEwen & Collignon, 2018; Robinson et al., 2016). Accordingly, humans, animals, and the

environment have shared microbiomes and resistomes, and an increase in one group may lead to increases in resistance in the other groups (McEwen & Collignon, 2018; Pehrsson et al., 2016). Therefore, the concept of One-Health is crucial in ABR research. The One-Health framework is based on the recognition that human health is connected to environmental and animal health, and approaches that work in multiple sectors are needed (Robinson et al., 2016). In the case of ABR, it suggests that the occurrence and drivers of resistance must be captured in all three compartments to understand patterns in resistance prevalence, and to inform interventions (WHO, 2021b).

Antibiotic resistance is also a health issue in wild and domestic animals. Although it has been estimated that 73% of antibiotics are used in food animals, there are few datasets available to estimate occurrence of resistance in livestock (Van Boeckel et al., 2017). Monitoring data from the European Centre of Disease Control (ECDC) suggests that some resistance rates to some, though not all, antibiotics in livestock are increasing (Food & Authority, 2022). The level of resistance also varies by the livestock species studied. Sampling commensal *E. coli* from healthy animal faeces, pan-susceptibility was more common in isolates from pigs and cows than from broilers and turkey (Food & Authority, 2022). In LMICs, antibiotic resistance in livestock is also increasing, and resistance levels are also higher in chickens than in pigs and cattle (Van Boeckel et al., 2019). However, there are fewer national surveillance programmes for livestock in LMICs than HICs despite a greater livestock population in these areas, making these estimates uncertain (Van Boeckel et al., 2019). Antibiotic resistance has also been observed in wild animals (Allen et al., 2010). Wild animals such as migratory wild birds, may contribute to the spread of antibiotic resistance caused by anthropogenic pollution, but the impact of resistance in wild animals on human health is poorly understood (Allen et al., 2010; Dolejska & Literak, 2019).

The final One-Health component is antibiotic resistance in the environment. Antibiotics and resistant bacteria from human and animals are disseminated to the environment, most notably through wastewater and animal manure (Larsson & Flach, 2022; Polianciuc et al., 2020; Woolhouse & Ward, 2013). Antibiotic resistance genes have even been found in air, water, and soils close to regions of human activity as well as pristine environments (Van Goethem et al., 2018). Some cases of environment to human or animal transmission have been observed (Leonard et al., 2015). The environment may present a potent resistance selection opportunity for bacteria, as it contains a diverse community of bacteria and repertoire of resistance genes (Huijbers et al., 2019; Larsson & Flach, 2022). However, there is a limited understanding of the contribution of environmental resistance to human resistance (Bürgmann et al., 2018). If the environment can act as a reservoir of resistance, there could be dynamic consequences for human and animal epidemiology which need to be understood.

1.4 Monitoring ABR: current approaches and limitations

Surveillance is an essential component of controlling the spread of infectious diseases, including antibiotic resistance. According to the WHO, it is a cornerstone not only for understanding ABR burden but also for informing policies and prevention strategies (WHO, 2021a). Traditionally, government health departments have monitored

antibiotic resistance with surveillance programmes that are hospital-based, and collect reports of the phenotypic susceptibility testing of a set of drug-bug combinations from routine, case-based surveillance (i.e. passive data collection) (Ashley et al., 2018; Tacconelli et al., 2018). This type of surveillance is crucial for guiding prescribing practice, both at the time it is carried out for a specific patient, and when it is collated on the level of a specific hospital or region (Tacconelli et al., 2018). As well as national surveillance programmes, three international surveillance networks have been collating traditional surveillance data for over 20 years: the European Antimicrobial Resistance Surveillance Network (EARS-Net), the Central Asia and Eastern European Surveillance of Antimicrobial Resistance network (CAESAR), and the Latin American and Caribbean Network for Antimicrobial Resistance Surveillance (ReLAVRA) (WHO, 2021a). In addition to these, over the last five years the WHO has set up a new global surveillance network, the Global Antimicrobial Resistance and Use Surveillance System (GLASS), which collects yearly data from hospitals and outpatient susceptibility testing for around 100 countries (WHO, 2021a). As this surveillance indicates numbers of serious infections, they are essential in current estimates of ABR burden, e.g. (Cassini et al., 2019; Murray et al., 2022; O'Neill, 2014).

However, traditional surveillance approaches are also limited. Data from these surveillance systems may be unreliable. They are resource intensive, requiring medical microbiology laboratories and staff, which has led to patchy data in low-resource, Low Income Countries (LIC) and Lower and Middle Income Countries (LMIC) settings (Ashley et al., 2018; Diallo et al., 2020; Iskandar et al., 2021). Although guidelines such as the European Committee on Antimicrobial Susceptibility Testing (EUCAST) and Clinical & Laboratory Standards Institute (CLSI) exist to standardise the microbiological testing, there is much heterogeneity in methodology between countries and even between laboratories within a country (Tacconelli et al., 2018). For example, EARS-Net only mandated that EUCAST guidelines be followed in 2019 – prior to this different countries may have used different guidelines (European Centre for Disease Prevention and Control, 2020). Heterogeneity in testing methods is even greater in LICs and LMICs (Iskandar et al., 2021). Therefore, there is a lack of comparability of the results between countries, years, and drug-bug combinations (Iskandar et al., 2021; Tacconelli et al., 2018).

In addition to unreliability, there is also a question about how representative and informative these data are. Inpatient hospital samples may not be representative of resistance rates in the community. This is especially the case in invasive blood-stream infection monitoring, where for a patient to be sampled and tested, they must be seriously ill (Tacconelli et al., 2018). Moreover, the resistance patterns in samples of invasive isolates may not be representative of the resistance among all bacteria carried by a patient – i.e., these datasets do not help relate infection to carriage (Diallo et al., 2020). The use of susceptibility testing of isolates means that the genetic basis for resistance is not recorded, resulting in a loss of information about the mechanisms and relatedness of resistance in different samples (Diallo et al., 2020; McArthur & Tsang, 2017). By focusing on a set of the most currently clinically relevant drug-bug combinations, traditional surveillance systems may also fail to capture the whole landscape of emerging resistance (Diallo et al., 2020; McArthur & Tsang, 2017). Clinic- and hospital-based surveillance systems are also not compliant with the One-Health

framework, as they do not integrate monitoring across human, animal, and environment sectors (Hu et al., 2018).

Some surveillance programmes have taken steps to address some of these limitations, for example WHO GLASS, which collects data from multiple sample types, including blood-stream, genital and urinary tract (WHO, 2021a). WHO GLASS also asks for the origin of infection to be collected (hospital, community, or unknown), and for information to be given on testing standards and external quality assessments used (WHO, 2021a). In addition, a new surveillance scheme is currently being set up as part of the GLASS technical module on One-Health to fill gaps in data on resistance in animals, the environment, and asymptomatic community carriers (WHO, 2021b). The extended-spectrum beta-lactamase *E. coli* Tricycle Project will collect samples from hospitalised patients, healthy community members, food animals, municipal, slaughterhouse, and wet market sewage samples, and natural water samples (WHO, 2021b).

1.5 Monitoring ABR: Wastewater surveillance and metagenomics

Researchers and policy makers have suggested that wastewater samples may be an attractive new source of population-level surveillance for ABR. Wastewater surveillance has been used for decades to monitor the spread of infectious diseases with high numbers of unreported cases in the community, including enteroviruses such as polio (Asghar et al., 2014), *Salmonella typhi* (Diemert & Yan, 2019), and more recently for SARS-CoV2 (Fernandez-cassi et al., 2021). For antimicrobial resistance surveillance, sewage sampling could address several of the limitations of traditional surveillance: it is unbiased, representing all people in a municipal area whether they are healthy or not; it can represent the whole human faecal microbiome, including pathogenic and commensal bacteria that may be carrying resistance; it is easy and cheap to collect in all countries; it is not individual-level, so there are fewer data sensitivity issues; and it is more One-Health compatible by integrating the environment (Aarestrup & Woolhouse, 2020; Hendriksen et al., 2019; Pruden et al., 2021). In addition, sewage sampling for surveillance can have multiple uses, including monitoring risks of antibiotic resistance in the environment on human health, and the impacts of anthropogenic pollution on environmental health (Huijbers et al., 2019).

Metagenomics has also been proposed as a tool for wastewater surveillance of antibiotic resistance (Aarestrup & Woolhouse, 2020; Hendriksen et al., 2019). In the metagenomic approach, all DNA in a sample is extracted and sequenced without culturing, after which the collection of sequenced read fragments can be compared to reference databases of bacterial, archaeal, and eukaryotic genomes (Wooley et al., 2010). The read collection can also be searched for hits to resistance genes to capture a quantitative profile of resistance gene abundance in a sample (Bengtsson-Palme et al., 2017; de Abreu et al., 2021; Rooney et al., 2021). Metagenomics was developed as a way to assess the whole microbial community in a sample, particularly an environmental one, in which the majority of bacterial species are not culturable so are not suited to whole genome sequencing (Bengtsson-Palme et al., 2017; Wooley et al., 2010). It can provide rich information on the taxonomies and resistance genes of species in samples which represent whole microbial communities, which is particularly

useful for exploratory analyses and characterising intra-species interactions (Bengtsson-Palme et al., 2017; Wooley et al., 2010). In addition, metagenomic data can be used for investigating whole community dynamics, which is particularly important for understanding the microbial ecology of the environmental resistance (de Abreu et al., 2021; Wooley et al., 2010).

There are now an increasing number of wastewater surveillance datasets, including several metagenomics ones. The largest is the Global Sewage Surveillance Project, which has applied metagenomic analysis to samples from over 100 countries between 2016 and 2018 (*Global Sewage Surveillance Project*, 2020; Hendriksen et al., 2019). A key finding of this study was that the global sewage resistome forms two clusters: one with developed countries, including Europe, North America, and Australasia, and another other with developing countries in South America, Africa, and Asia (Hendriksen et al., 2019). Other international datasets of wastewater surveillance include (Pärnänen et al., 2019) and (Huijbers et al., 2020), who applied large panels of qPCR wastewater across Europe. These studies found that there was correspondence between the abundance of resistance genes in sewage and the level of resistance in clinical surveillance data (Huijbers et al., 2020; Pärnänen et al., 2019). Other research groups have sampled widely across single countries, including China (Su et al., 2017), Germany (Alexander et al., 2020), and the UK (Ludden et al., 2017), or within a single city, including Copenhagen in Denmark (Brinch et al., 2020), Stockholm (Kwak et al., 2015) and Gothenburg (Hutinel et al., 2019) in Sweden, New Delhi in India (Lamba et al., 2018), and more. In addition, studies are now sampling from individual effluent points, particularly hospitals (Korzeniewska et al., 2013; Perry et al., 2021).

Wastewater surveillance with metagenomics is a new technology and comes with unknowns. An important question is how well the human faecal microbiome is represented in wastewater systems. Other non-human sources such as pharmaceutical effluent may also contribute resistant material (Guardabassi et al., 1998), and microbiological processes within wastewater may alter the resistome (McLellan & Roguet, 2019; Vandewalle et al., 2012). How much of an effect these have on the composition of sewage is not well understood. Some studies have shown that sewage resistomes mostly represent the human microbiome (Newton et al., 2015), whereas others have highlighted that environmental bacteria as well as faecal bacteria dominate the sewage microbial composition (Vandewalle et al., 2012). Another unknown is which human microbiomes are represented in the sewage. Many wastewater treatment plants (WWTPs) capture hospital effluent as well as domestic effluent, but the size of the contribution of hospitals to the WWTP resistome is still being investigated (Buelow et al., 2018). Moreover, although wastewater can be found all over the world, wastewater systems in LMICs and HICs may vary greatly, with poorer wastewater infrastructure and fewer people connected to the main sewers in LMICs (Nadimpalli et al., 2020), introducing a risk of additional non-human contamination and that the whole community may not be represented in sewage samples in these countries. Epidemiological data such as antibiotic consumption could therefore have different relationships with resistomes in community members, patients, and mixed municipal sewage.

Metagenomic analysis leads to a trade-off between quantification of a wide breadth of resistance genes and loss of information on the bacterial host a resistance gene has come from, and detection of some kinds of resistance gene (Bengtsson-Palme et al., 2017). A lack of connection between the resistance gene and its bacterial host results in some loss of information in comparison to culture-based whole genome sequencing methods (Aarestrup & Woolhouse, 2020). In addition, it means that we cannot tell whether there are multiple copies of a resistance gene in a sample because there are several resistant bacteria, or because there are several copies within a few bacteria. Bacterial resistance that arises from up-regulation of efflux, or changes to the sequence of a chromosomal gene such as *gyrA* that prevent antibiotic binding, are not always included in reference databases, particularly ResFinder, which focuses on acquired resistance genes (Bengtsson-Palme et al., 2017). In addition, point mutations to chromosomal genes giving resistance will only be detected if a read overlaps with this region, making detection of these genes particularly low sensitivity (Bengtsson-Palme et al., 2017; Rooney et al., 2021). These factors must be taken into account when analysing and interpreting wastewater data.

Wastewater metagenomics is also a noisy source of data. There are many opportunities for randomness to influence the resulting resistance gene read hit counts. Rare bacteria and resistance genes might not be captured in a single sample of wastewater (Bengtsson-Palme et al., 2017; de Abreu et al., 2021). Wastewater sampling can capture different numbers of people by chance, and estimating the true number of people captured for normalisation is challenging (Isaksson et al., 2022). The process of sample storage, DNA extraction and amplification introduces further opportunities for some fragments of bacterial genomes to be over- or underrepresented in comparison to the sample's composition (Bengtsson-Palme et al., 2017; de Abreu et al., 2021; Ko et al., 2022). Finally, the position of a read within a gene sequence can influence the accuracy of read calling – a read from a non-variable region of two very similar genes could be assigned to either, depending on the mapping algorithm (Bengtsson-Palme et al., 2017; de Abreu et al., 2021). The impact of these sources of noise are overdispersion and zero-inflation in read hit counts (Bengtsson-Palme et al., 2017). Protocols for wastewater metagenomics, from sampling to read calling, can be established to minimise and standardise this noise between datasets, and these pipelines are increasingly frequently used (de Abreu et al., 2021).

Appropriate analysis methods are needed to take advantage of sewage metagenomic data by dealing with high dimensionality, noise, and overdispersion in metagenomic data. However, many studies of wastewater currently use simple analysis such as bivariate correlation, descriptive cluster methods, and linear regression on transformed read counts. These techniques risk loss of information by not considering the whole resistance profile, loss of power by not accounting for explainable variance, and are too non-specific to represent mechanistic theories about human and sewage resistance patterns. Some examples of other types of analysis of metagenomic data exist (Duarte et al., 2021; Hendriksen et al., 2019; Peng et al., 2016), but further innovation is needed in this area.

1.6 Aims and objectives

Overall aim of thesis: Understanding the spatial and temporal factors that affect the global spread of ARGs in humans and the urban wastewater environment.

Thesis objectives:

1) To investigate the role that the environment might play in the spread and emergence of ABR;
2) To identify epidemiological datasets that best represent the occurrence of ARGs in sewage;
3) To perform analysis of the spatial and temporal distributions of ARGs in sewage to inform the design of surveillance systems;
4) Using wastewater metagenomic data, develop predictive models of the distribution of resistance genes in space and time.

1.7 Thesis outline

In this thesis, I investigated the sources and spread of resistance in wastewater, and the uses of wastewater metagenomics for surveillance of resistance.

In Chapter 2, I used a systematic scoping review to investigate dissemination of resistance from hospitals to wastewater, and evaluate the quality of evidence and knowledge gaps in this area.

In Chapter 3, I used a mathematical model to investigate the transmission of resistance between humans, animals, and the environment, and comment on the consequences of an environmental reservoir of resistance for human resistance epidemiology and prevention strategies.

In Chapter 4, I used a statistical multi-response linear model to estimate the correlations between sewage and clinical resistance levels using data from EARS-Net, Global Sewage Surveillance Project, and comment on the relationship between antibiotic consumption and resistance abundance.

In Chapter 5, I used a paired sampling design and metagenomic analysis wastewater in Scotland and multiple statistical methods to investigate the relationship between hospital and community wastewater resistomes.

In Chapter 6, I will discuss the findings of Chapters 2 – 5 and make concluding remarks.

2 Dissemination of hospital-associated antibiotic resistance to wastewater: a systematic scoping review

2.1 Abstract

Background: Hospital wastewater is a potential reservoir of clinical antibiotic resistance (ABR). This reservoir presents a challenge to using wastewater for surveillance of resistance in the community. However, it is not yet known how well hospital wastewater represents the patient resistome, or how much downstream municipal effluent represents the hospital effluent resistome. In this chapter I aimed to describe the evidence for dissemination of hospital-associated ABR to hospital and municipal wastewater.

Methods: I used a systematic scoping review. I downloaded articles about resistant bacteria or resistance genes and hospital and municipal wastewater from PubMed, Scopus, and Web of Science. Data extraction was conducted on articles that claimed they had demonstrated evidence in favour of or against dissemination of ABR from a hospital source into hospital or municipal wastewater. Data extracted included the route of dissemination, study conclusion, statistical and laboratory methods, and types of resistant bacteria studied.

Results: A total of 1454 unique studies were downloaded from databases, of which 201 studies were included in any data extraction. A number of studies claimed to find evidence of dissemination of resistance from hospital effluent to municipal wastewater (40 studies) and a few for hospital patients to hospital wastewater (9). Among 61 studies with conclusive findings, most used medium resolution typing methods: gene presence (e.g. PCR) for resistance typing (26) and fingerprint (e.g. PFGE) or phenotypic methods for bacterial species typing (17 studies each). 128 used some kind of statistical method, but conclusive studies were more likely than inconclusive to have used statistics (73.4% vs. 59.3%). Few conclusive studies used control samples from the community (8 effluent and 2 healthy human samples).

Conclusions: There is evidence from the literature that resistant bacteria can disseminate from hospital patients to hospital and municipal wastewater. However, there are many medium or low evidence quality studies, as well as conflicting results, which hinder interpretation. To fully disentangle the community resistome from the hospital resistome in municipal wastewater, we need more high evidence quality studies that use high resolution typing, statistical methods, and targeted sampling designs including controls.

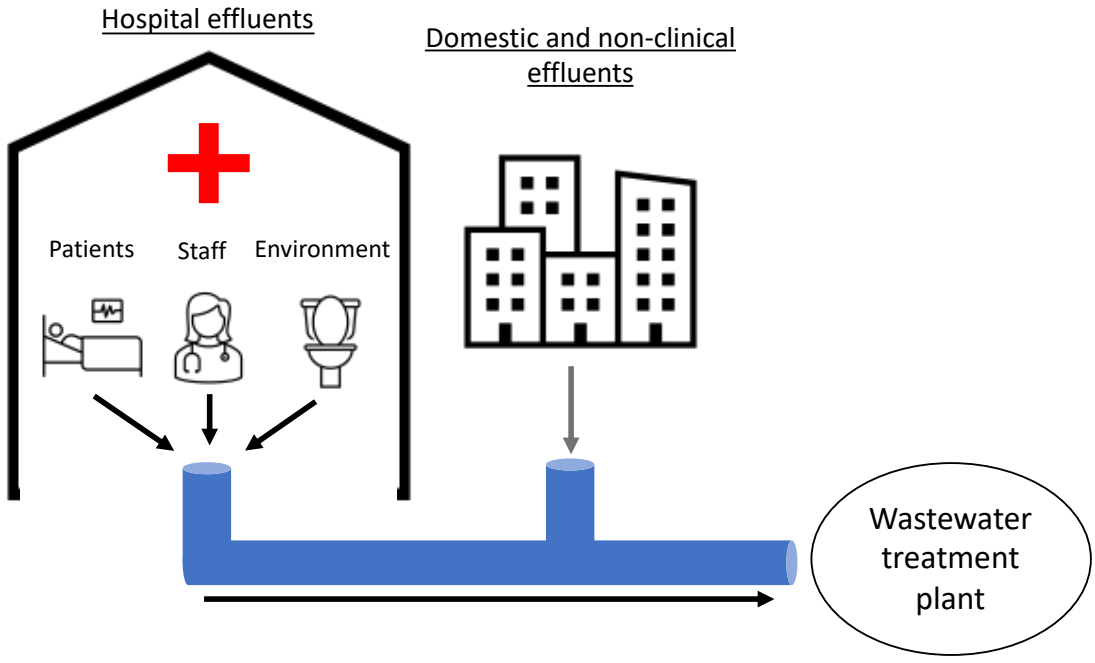
2.2 Introduction

Antibiotic resistance in wastewater is a growing area of research. As a reservoir of the human resistome (Hendriksen et al., 2019; Newton et al., 2015), it has been proposed as a medium for monitoring resistance in the human population (Aarestrup & Woolhouse, 2020; Hendriksen et al., 2019; Pärnänen et al., 2019; Pruden et al., 2021). However, wastewater systems collect effluents from multiple sources as well as the community, including hospitals and industrial premises (Pruden et al., 2013), abattoirs (Nguyen et al., 2021) and other large buildings like offices and schools. Wastewater samples therefore represent multiple resistome signatures. For sewage surveillance to be an effective method for monitoring resistance in the community or in hospitals, the contribution of each of these sources to resistance in a mixed effluent sample must be understood.

The contribution of hospitals is a particular concern, because hospitals are a focal point for bacterial infections and antibiotic usage. Resistant bacteria therefore can emerge in hospitals that are not common in the community. There is concern that municipal wastewater downstream from hospitals could have elevated resistance levels, or an altered resistance profile, which reflects the hospital rather than the community (Hassoun-Kheir et al., 2020; Hocquet et al., 2016). This has implications for both surveillance and the risks to human and environment health of wastewater: clinical effluent resistomes may alter sewage surveillance findings and may contain resistant bacteria that are of public health concern. Therefore, there has been interest in comparing hospital, municipal, and community effluents (Gao et al., 2022; Hassoun-Kheir et al., 2020; S. Zhang et al., 2020).

Figure 2.1: Potential routes of dissemination of hospital-associated resistant bacteria to municipal sewage

Routes that I aimed to study in this review highlighted in black, other related routes in grey.



Within hospitals, resistant bacteria can be found in patients, staff, and the environment, and transmission between these sources has also been observed or suspected (Constantinides et al., 2020; Tacconelli & Cataldo, 2008). This suggests there are potentially multiple sources of resistance in the hospital effluent (Fig. 2.1). It also seems possible that patient and hospital environment resistomes could be characterised by different resistant bacteria, for example more environmental bacteria thriving in the hospital environment (Hassoun-Kheir et al., 2020). To understand how hospital or municipal wastewater can be used for surveillance of patients or community members, the contribution of these different sources to hospital effluent also needs to be investigated.

A growing number of studies have compared resistance in hospital patients, environments, and effluents, and there is a need to synthesise the results of these studies. Evidence from a research field can be collated and synthesised in systematic reviews. Previous systematic reviews have investigated resistant bacteria in WWTPs (Pazda et al., 2019), and community compared to hospital effluents (Hassoun-Kheir et al., 2020; S. Zhang et al., 2020). However, there is lack of reviews considering different hospital to wastewater transmission routes, and of broad scope reviews that consider multiple resistant bacteria, bacterial genes, and study designs.

In this chapter, I conducted a systematic scoping review of the literature on the dissemination of hospital-associated ABR in wastewater. Using broad search terms and screening criteria, I identified studies that sampled wastewater and hospital resistance sources, and made one or more conclusions about dissemination of resistant bacteria. I aimed to describe the evidence for dissemination of ABR from hospital sources to hospital or municipal wastewater. My objectives were to: 1) systematically review the literature to identify studies of the dissemination of hospital-associated ABR to wastewater; 2) summarise the evidence for different hospital-based sources of resistance wastewater; and 3) appraise the characteristics and quality of the methodologies of these studies.

2.3 Methods

2.3.1 Eligibility criteria

I adapted the PECO (Population, Exposure, Comparison, Outcome) framework for environmental health evidence reviews (Collaboration for Environmental Evidence, 2013) to define the following criteria for assessing the relevance of a study to the research questions, described in Table 2.1.

I did not make a comparison group necessary so that the research question and eligibility criteria were applicable to a range of study designs, i.e. both direct and indirect studies of the exposure. Municipal wastewater is defined as wastewater containing all domestic effluent or a mixture of domestic and other effluent types, treated or untreated. Clinical wastewater includes effluent from hospitals, treated or untreated. I also included static sewage such as latrines and septic tanks.

Based on this framework and the objectives of the study, I selected the following criteria for a study to be ineligible in the evidence synthesis: 1) not available in English; 2) not about antibiotic resistance in bacteria; 3) not primary data analysis published in a peer-reviewed journal (i.e., excluded reviews and conference abstracts); 4) observational study design (i.e. no experimental studies such as those that use an experimental wastewater system); 5) no human wastewater of any kind sampled directly or indirectly; 6) only samples from one wastewater source type with no comparison or exposure group studied (i.e. hospital or municipal wastewater samples only). The aim of these criteria was to select studies that have appropriate sampling to study any of the routes of dissemination of ABR from clinical sources into wastewater shown in Fig. 1. For example, studies that measured resistance in hospital patients or environment and hospital wastewater, or studies that measured resistance in municipal wastewaters up- and downstream of a hospital, would be included.

Table 2.1 PECO-style framework for each research question

PECO question element	Description
Population	ABR genes or bacteria in hospital or municipal wastewater
Exposure	ABR genes or bacteria in hospital sources (wastewater, patients, environment)
Comparison	ABR genes or bacteria in municipal wastewater without hospital impact, or none (population and exposure samples only)
Outcome	A claim that the study results provide evidence in support of or against directional transfer of ABR genes or bacteria from the exposure source to one of the populations of interest.

The eligibility of studies according to criteria 1 – 6 was initially assessed in a title and abstract screen. If an abstract was not available, the full text was used.

2.3.2 Sources of information and search strategy

I searched the following electronic databases: PubMed, Scopus, and Web of Science. The search terms used for each database can be found in Table 2.2. Search terms were identified through studying search terms of other similar systematic reviews (Chatterjee et al., 2018; Greig et al., 2015) and tested using PubMed. I aimed to retrieve articles that were about both ABR in bacteria and wastewater. I used a date restriction between 1900 and the 21st July, 2021.

2.3.3 Data extraction

Included studies were split into two categories for data extraction. I conducted a full data extraction from studies with an evidence statement in support of dissemination of ABR from clinical environments to wastewater. An example of an evidence statement

Table 2.2: Database search terms

Database	Search terms
PubMed	(sewage OR sewer OR sludge OR slurry OR "waste water" OR effluent) AND (antimicrobial drug resistance[MeSH Terms] OR antimicrobial drug resistances[MeSH Terms] OR antibiotic resistance, bacterial[MeSH Terms] OR antibiotic resistance, microbial[MeSH Terms] OR drug resistance, microbial[MeSH Terms] OR antimicrobial drug resistance[MeSH Terms] OR antimicrobial drug resistances[MeSH Terms] OR resistome) AND (hospi* OR patient* OR clinic*) AND ("1900"[Date - Publication] : "2021/07/21"[Date - Publication])
Scopus	TITLE-ABS-KEY ("antimicrobial resistan*" OR "antibiotic resistan*" OR "drug resistan*" OR "multidrug resistan*" OR resistome) AND TITLE-ABS-KEY (sewage OR sewer OR sludge OR slurry OR "waste water" OR effluent) AND TITLE-ABS-KEY (hospi* OR clinic* OR patient*) AND LOAD-DATE < 20210722
Web of Science	(TS = ("antimicrobial resistan*" OR "antibiotic resistan*" OR "drug resistan*" OR "multidrug resistan*") OR TI = ("antimicrobial resistan*" OR "antibiotic resistan*" OR "drug resistan*" OR "multidrug resistan*") OR FT = ("antimicrobial resistan*" OR "antibiotic resistan*" OR "drug resistan*" OR "multidrug resistan*") OR SU = ("antimicrobial resistan*" OR "antibiotic resistan*" OR "drug resistan*" OR "multidrug resistan*") OR WC = ("antimicrobial resistan*" OR "antibiotic resistan*" OR "drug resistan*" OR "multidrug resistan*")) AND (TS = (sewage OR sewer OR sludge OR slurry OR "waste water" OR effluent) OR TI = (sewage OR sewer OR sludge OR slurry OR "waste water" OR effluent) OR FT = (sewage OR sewer OR sludge OR slurry OR "waste water" OR effluent) OR SU = (sewage OR sewer OR sludge OR slurry OR "waste water" OR effluent) OR WC = (sewage OR sewer OR sludge OR slurry OR "waste water" OR effluent)) AND (TS = (hospi* OR clinic* OR patient*) OR TI = (hospi* OR clinic* OR patient*) OR FT = (hospi* OR clinic* OR patient*) OR SU = (hospi* OR clinic* OR patient*) OR WC = (hospi* OR clinic* OR patient*)) Date range selected: 01/01/1900 – 21/07/2021

might be “Our results suggest that resistant bacteria in hospital effluents are carried into municipal wastewater” or “Our results show that resistance in WWTP influent is

impacted by hospital effluents." For studies that did not make this claim of this kind, I used partial data extraction. The data that was extracted in these two groups is presented in Table 2.3.

For studies with an evidence statement, I recorded the finding (positive or negative) and the route of ABR dissemination studied. For example, I would record that the evidence claim stated that the authors believed positive evidence was found for dissemination of ABR from hospital wastewater to municipal wastewater. Recording the evidence claim and direction of transfer follows the methods of Muloi et al, 2018, and allowed the study of the full range of different routes identified, making it suitable for a scoping review. Secondly, by recording some data from conclusive and inconclusive studies, I could identify which methods generated conclusive results, therefore what study designs are needed in future.

Table 2.3: Data extraction for each of the evidence groups

	Evidence group	
Data extraction	*Full data extraction (studies with an evidence claim)*	*Partial data extraction (studies without an evidence claim)*
Year of publication	Yes	Yes
Countries sampled (sewage samples only)	Yes	Yes
Route of ABR transfer studied	Yes	NA
Study finding	Yes	NA
Antibiotic resistance phenotype	Yes	Yes
Laboratory typing methods	Yes	No
Statistical analysis methods	Yes	Yes
Hospital sample types	Yes	Yes
Other sample types	Yes	No

For all included studies, I extracted information on the types of antibiotic resistance phenotypes measured in a study. I extracted this data for all included studies because I hypothesised that the antibiotic phenotype measured may influence the study results and evidence statement. For example, some phenotypes such as carbapenem resistance could be being more likely to be detected in hospitals than in community settings. For a resistance to a particular antibiotic to be considered to have been measured, the study needed to either use phenotypic testing of isolates, or detect the presence of a resistance gene that conferred resistance to that antibiotic. When possible, I grouped antibiotics studied into the WHO Anatomical Therapeutic Chemical Classification (ATC) Level 5 groups for comparability between studies and to reduce

the number of antibiotic groupings. Studied that used shotgun metagenomics, whole genome sequencing, or a large PCR panel of resistance genes were recorded as measuring all antibiotic resistance phenotypes.

For studies with an evidence statement, I extracted information on the bacterial group that was measured. I considered a bacterial group to be measured if its presence in a sample was detected by culture, PCR, or metagenomics. Studies that used shotgun or taxonomic metagenomics were recorded as measuring all bacterial groups.

Statistical methods were recorded as any kind of frequentist statistical test (such as regression, t-test, Pearson's correlation, with a p-value reported), Bayesian analysis (such as regression with a Bayesian fitting method), cluster analysis (including clustering on the basis of abundance similarity and sequence similarity, such as principal components analysis or hierarchical clustering of gene sequences), phylogenetics (including maximum likelihood, parsimony, and Bayesian methods), permutations (such as Redundancy Analysis with Monte Carlo permutations), network analysis (such as minimum spanning trees), and machine learning (such as random forest analysis). Both cluster analysis based on genetic sequence similarity and phylogenetics aim to demonstrate the relatedness of strains, but I distinguish between them to capture the use of a model in the latter analysis method. When the type of analysis used was unclear, or the statistical model was described but the fitting method was not, the type of statistics was classed as 'unclear'.

Information on the laboratory methods used for establishing resistance types and bacterial types were collected from conclusive studies. Laboratory methods for resistance typing were broadly classified as phenotypic (disk diffusion, incubation with antibiotic in broth or agar, or incubation with non-antibiotic reagents such as clavulanic acid in the E test), gene presence (PCR, qPCR, or metagenomics without sequence analysis), or gene sequence (whole genome and resistance gene sequencing). Based on Muloi et al, 2018, I describe phenotype testing as low resolution, gene presence testing as medium resolution, and gene sequencing as high resolution, with respect to determining the relationship between resistant bacteria in different samples (Muloi et al., 2018).

Similarly, laboratory methods were classified (in order of resolution) as phenotypic (selective media culture, colony morphological characteristics), bacterial fingerprinting such as PFGE, DGGE and MALDI-TOF, partial sequencing such as MLST, 16/18s, and shotgun metagenomics, and whole genome sequencing.

2.3.4 Evidence quality assessment

I assessed the quality of the evidence collected by recording data on a) appropriateness of laboratory methods for studying transmission, such as the use of high-resolution tools like whole genome sequencing, and b) use of statistics, such as t-tests, regression, cluster analysis, or phylogenetics.

2.4 Results

2.4.1 Duplicates and title and abstract screening

In total, 2083 studies were identified in public databases, which after de-duplication resulted in 1454 unique studies. Fig. 2.2 presents the numbers of studies that were found in each database, and numbers of studies excluded in each step in the eligibility screening. A total of 201 studies were included in the data extraction phase, of which 61 made an evidence claim and had full data extraction, and 140 did not and had partial data extraction.

2.4.2 Study characteristics

Full study characteristics for the 61 full data extraction studies is in Appendix A Table 1.

2.4.2.1 Year and country of studies

The earliest publication year in all included studies was 1973, and half of all studies were published in or after 2017. Fig. 2.3 shows the distribution of source of resistance studies by year.

Fig. 2.4 shows the number of studies found with samples from each country. A total of 70 countries were sampled in studies included in this chapter. The country with samples analysed in the most studies the United Kingdom (19 studies), followed by Germany (15), Spain (15), Portugal (13), and Sweden (12). Although all continents are represented, there is a concentration of studies in Europe.

2.4.2.2 Antibiotic resistance phenotype studied

I extracted the antibiotic resistance phenotypes that were measured in all 201 included studies. The antibiotic phenotypes most frequently studied were penicillins with extended spectrum such as ampicillin and amoxicillin (101 studies), fluoroquinolones such as ciprofloxacin (99), aminoglycosides such as gentamicin (97), tetracyclines (95), and third generation cephalosporins (85). I classified 23 studies as measuring all antibiotic resistance phenotypes. One study did not report the antibiotic resistance phenotypes studied. In addition, six studies reported resistance to antibiotics that could not be assigned to a class that is represented in the WHO ATC index (oleomycin, avilamycin, sulfazoritrim, and monensin). These antibiotics are included in the category “No ATC group found”. Excluding studies that looked at all resistance phenotypes, studies measured 5.9 different ATC Level 5 antibiotic phenotypes on average (range 1 – 21, standard deviation 4.7). The proportion of studies different dissemination routes for each resistance phenotypes is presented in Fig. 2.5. The types of antibiotic resistance phenotype studied were similar in each dissemination route and in inconclusive compared to conclusive studies.

2.4.2.3 Bacteria studied

I extracted data on the bacterial groups measured in studies that made an evidence statement. The most frequently studied types of bacteria were *E. coli* (15 studies),

Enterococcus spp. (13), *Enterobacteriaceae* (7), *Klebsiella pneumoniae* (6), and *Pseudomonas aeruginosa* (6). Three studies did not measure the presence or abundance of any type of bacteria, all of which were qPCR studies. One study investigated all bacteria using shotgun metagenomics. The proportion of studies studying different dissemination routes for each bacterial group measured is presented in Fig. 2.6.

2.4.2.4 Hospital sample types

Among all included studies, 103 collected samples from hospital effluents, 55 collected samples from hospital patients, eight collected samples from the hospital environment, and one collected samples from health care workers.

2.4.2.5 Other sample types

Studies with conclusive findings about dissemination from hospitals to wastewater also often collected samples from other sources. These included natural waters, such as rivers and lakes to which wastewater is discharged; soils where wastewater is used for irrigation; healthy human samples and samples from outpatients; and other effluents that are potentially sources of resistance in municipal wastewater, including community sources, industrial effluents, and slaughterhouses. In total, eight conclusive studies collected community effluents and two collected community clinical samples.

2.4.2.6 Usage of statistics

Of all 201 included studies, 73 did not use any statistics (36.3%). The most frequently used statistical method was frequentist analysis (83 studies). Next were cluster analysis (44 studies), phylogenetics (12), permutation analysis (5), Bayesian analysis (2), machine learning (2), and network analysis (2). The statistical methods used were unclear in two studies.

Among positive evidence claims identified, 40 were in studies with any kind of statistical analysis, whereas 16 were not. Negative evidence claims were only found in five studies, all of which used statistics. 73.4% of studies that made conclusive evidence claims used statistics was (45/61), compared to 59.3% of studies with inconclusive findings (83/140), potentially indicating that use of statistics is important in giving study authors conclusive findings. The proportion of studies with each statistical method or with none by conclusion types is plotted in Fig. 2.7. The use of statistics appears to generate more conclusive results, especially negative ones.

2.4.2.7 Usage of typing methods

Among studies that made an evidence statement about a source of resistance in wastewater, most used gene presence as the highest resolution resistance typing method (26 studies), followed by phenotypic testing (25) and resistance gene sequencing (10). For bacterial typing, an equal number of studies used phenotypic, or fingerprint methods (17 studies each), 16 used partial sequencing, and eight used whole genome sequencing. Three studies did not identify bacterial species. The different conclusions made by studies with each kind of resistance and bacterial typing is presented in Fig. 2.8.

2.4.3 Routes of dissemination of hospital-associated ABR to wastewater

Table 2.4 presents the number of evidence claims in support of and against the different routes of dissemination, as well as the number of these studies with the highest evidence quality (defined as use of statistics plus the highest resolution resistance or bacterial group typing). The dissemination of ABR from hospital wastewater to municipal wastewater received the most attention, with 42 studies claiming evidence about this route.

Table 2.4: Numbers of studies support or against dissemination of hospital-associated ABR to wastewater

All studies in category followed by number of highest evidence quality studies in brackets. HWW: hospital wastewater; MWW: municipal wastewater

Route of dissemination	Number of studies supporting	Number of studies against
Hospital patients to HWW	9 (2)	2 (1)
Hospital patients to MWW	5 (0)	1 (0)
HWW to MWW	40 (8)	2 (0)
Hospital to MWW (via indirect sampling design)	7 (0)	0

2.4.3.1 Hospital patients to hospital wastewater

Nine studies claimed to find evidence in support of dissemination of resistance from hospital patients to hospital wastewater, and two claimed to find evidence against it. In total, 35 studies collected samples from both hospital patients and hospital wastewater, i.e. 31% (11/35) of studies captured in this review with appropriate study design made an evidence claim about this direction of transfer. Studies investigating this route mostly studied bacteria that are human pathogens and frequently cause hospital-associated resistant infections, such as *K. pneumoniae*, *K. oxytoca*, and *A. baumannii* (Fig. 2.6).

A close relationship between resistance in patients and hospital wastewater has been found in several studies across the European (Iversen et al., 2004; Jakobsen et al., 2008; Ory et al., 2016; Popa et al., 2021; Röderová et al., 2016; Seruga Music et al., 2017) and African (Atmani et al., 2015; King et al., 2020) continents. There is high quality evidence in favour of this route of dissemination. For example, Popa et al, 2021 used whole genome sequencing of multi-drug resistant *K. pneumoniae* isolates from patients and effluents of an infectious disease hospital in Romania (Popa et al., 2021). Maximum likelihood phylogenetics indicated that clinical samples and hospital effluent belonged to the same clonal group. However, other high evidence quality studies have found evidence against this route of dissemination. One study used whole genome sequencing study of *K. pneumoniae* in patients in a haematology ward and wastewater

from a UK hospital did not find evidence that resistant bacteria from patients disseminated to hospital wastewater (Ludden et al., 2020).

2.4.3.2 Hospital patients to municipal wastewater

Six studies claimed to find evidence in support of or against dissemination of ABR from hospital patients to municipal wastewater. The study with the highest resolution typing in this group was Oracova et al, 2017, who showed some relatedness in MLSTs of vancomycin resistant *Enterococcus* spp. in hospital patients and municipal wastewater in the Czech Republic (Oravcova et al., 2017). A group of studies compared resistance in national hospital-based surveillance data to municipal wastewater (Karkman et al., 2020; Mbanga et al., 2021; Pärnänen et al., 2019), but none of these conclude that any correlations found are caused by dissemination of ABR from hospitals to wastewater.

2.4.3.3 Hospital environment and health care workers to wastewater

Eight studies were found that sampled both the hospital environment and wastewater, but none concluded that the hospital environment was a source of resistance in the wastewater. Ludden et al, 2020, sequenced *K. pneumoniae* from multiple hospital environments, including equipment and patient areas and found only 2% of these samples were colonised, although they do not comment on the possibility of dissemination of these bacteria to the hospital sewage (Ludden et al., 2020). One study took samples from hospital personnel and patients as well as municipal wastewater and found no resistance to vancomycin in enterococci in the hospital staff, without drawing a conclusion about whether this indicated the role of staff in the sewage resistome (Oravcova et al., 2017).

2.4.3.4 Hospital wastewater to municipal wastewater

The most frequently studied dissemination route studied was the link between hospital wastewater and municipal wastewater (40 out of 61 studies with a conclusions). A range of bacterial species were studied in this group, but the most frequently measured was *E. coli* (Paulshus et al., 2019; Verburg et al., 2019) and *Enterococcus* spp. (Lamba et al., 2018). Six studies used the highest resolution typing methods to investigate this (whole genome sequencing), all of which found positive evidence that resistant bacteria in hospital wastewater are disseminated to municipal wastewater (Cahill et al., 2019; Gouliouris et al., 2019; Ludden et al., 2020; Popa et al., 2021; Pot et al., 2021). Other compelling evidence for this route of dissemination came from studies with samples before and after the hospital wastewater entered the main wastewater stream. One such study in Ireland found similar sequence types were found in the hospital effluent and the municipal wastewater downstream of the hospital entrance point (Cahill et al., 2019).

In contrast to the consistent evidence from studies with sequence comparison, studies that investigated the profile of bacterial species and resistance genes in sewage produced more mixed results. Ng et al, 2017 (Ng et al., 2017) and Kutilova et al, 2021 (Kutilova et al., 2021) conducted metagenomics on the wastewater of urban hospitals and WWTPs in Singapore and the Czech Republic, respectively. Although both found some similarities in the resistome and microbiome of these samples, their conclusions do not agree. Ng et al found more similarity and tentatively suggest a patient-to-

hospital wastewater-to-municipal wastewater route of dissemination as the explanation. Kutilova et al, on the other hand, suggest that patient-derived bacteria in the hospital wastewater do not contribute to the resistance in WWTP influent, likely due to dilution of hospital wastewater, although they do not rule out dissemination. Similarly, two studies in France making use of a large qPCR panel found that the hospital wastewater does not contribute to overall antibiotic resistance abundance in WWTP influent although again they do not rule out hospital dissemination (Buelow et al., 2018, 2020). In general, studies comparing the abundance of resistance genes or resistant bacteria have highlighted that dilution of the hospital signature reduces the overall contribution of hospitals to resistance patterns in municipal wastewater (Buelow et al., 2018, 2020; Kutilova et al., 2021; Paulshus et al., 2019).

2.4.3.5 Indirect studies of hospital impact on wastewater

Finally, studies have looked for an impact of hospitals on municipal wastewater through indirect sampling, i.e. by sampling WWTPs that are and are not impacted by hospital effluents. Many studies that have used this study design have claimed to find evidence in support of a contribution, mostly in Europe (Alexander et al., 2020; Guardabassi et al., 1998; Harris et al., 2014; Linton et al., 1974; Ludden et al., 2017; Novo & Manaia, 2010), and one study in India (Akiba et al., 2015). However, other studies with this design have not found evidence for an overall effect (Buelow et al., 2018, 2020). Alexander et al, 2020 and Ludden et al, 2017 both provide good evidence for an effect by sampling from multiple WWTPs. Alexander et al found an increased abundance of antibiotic resistance genes in hospital-influenced WWTPs, and Ludden et al only found carbapenemase producing Enterobacteriaceae in hospital-influenced WWTPs.

Figure 2.2: Flow diagram of screening process

Numbers of studies included and excluded at each stage of the screening process (importing from database; de-duplication; title and abstract screening; full text screening), plus numbers of studies in each included category.

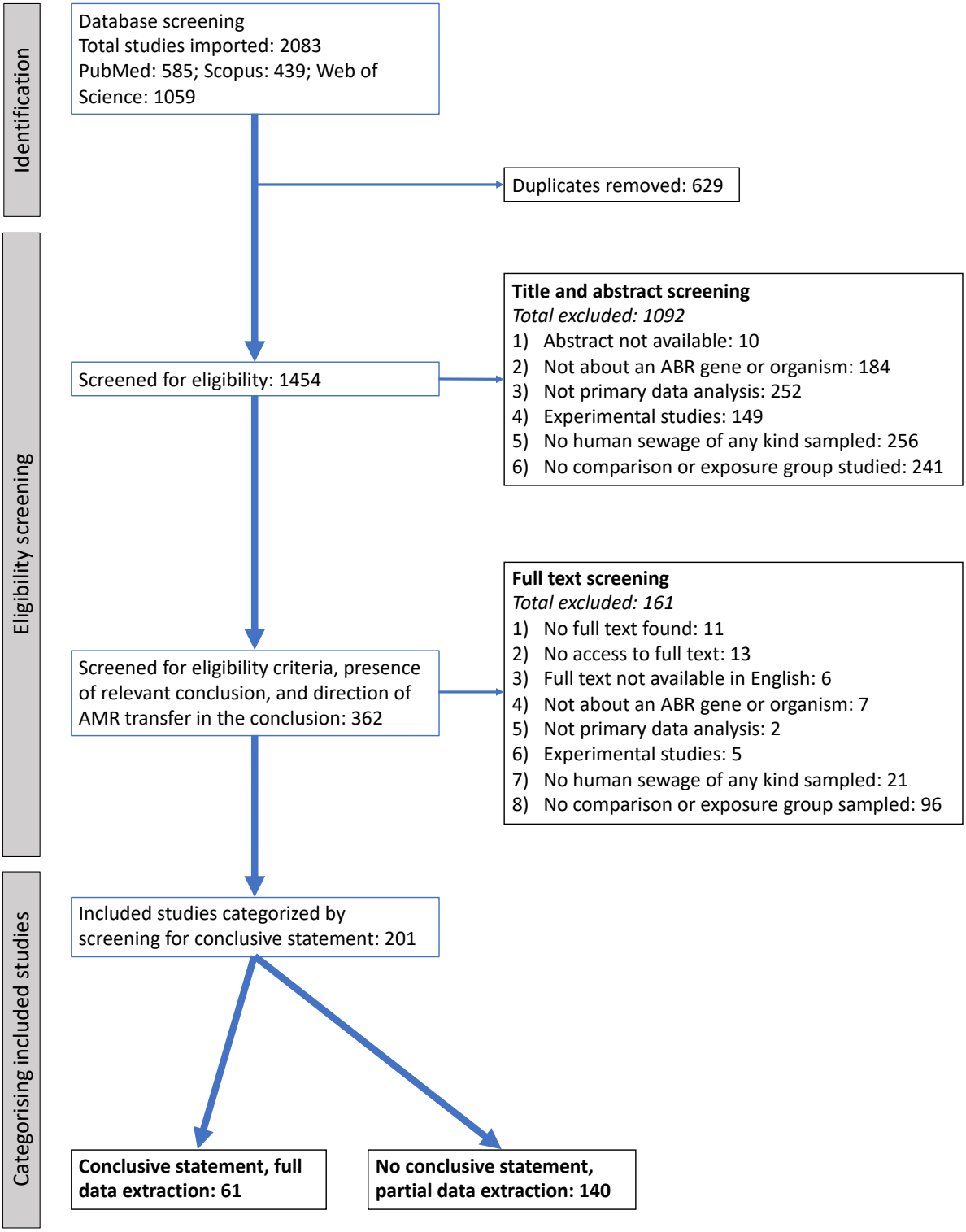


Figure 2.3: Year of included study publication

Year of publication of included studies, coloured by whether they contained an evidence statement about dissemination of ABR from clinical sources to wastewater.

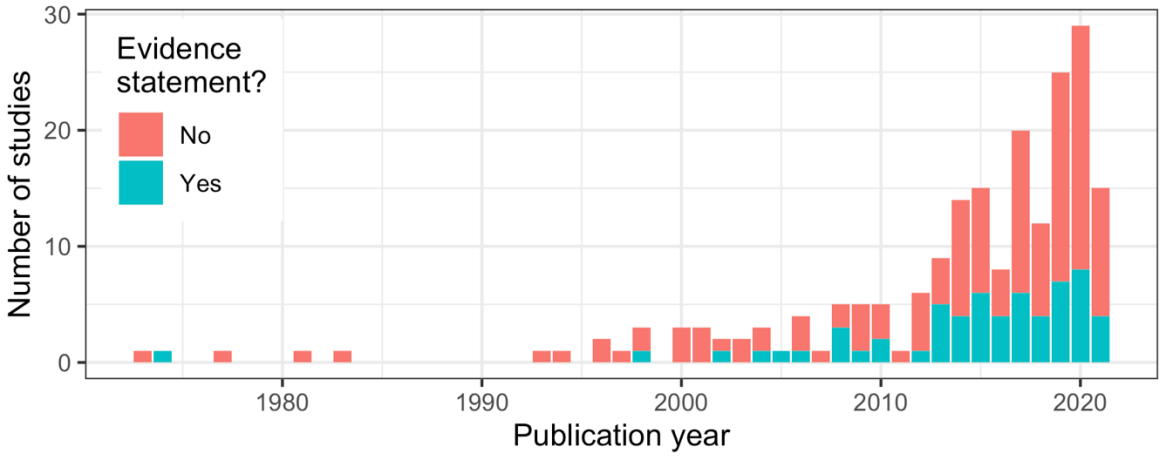
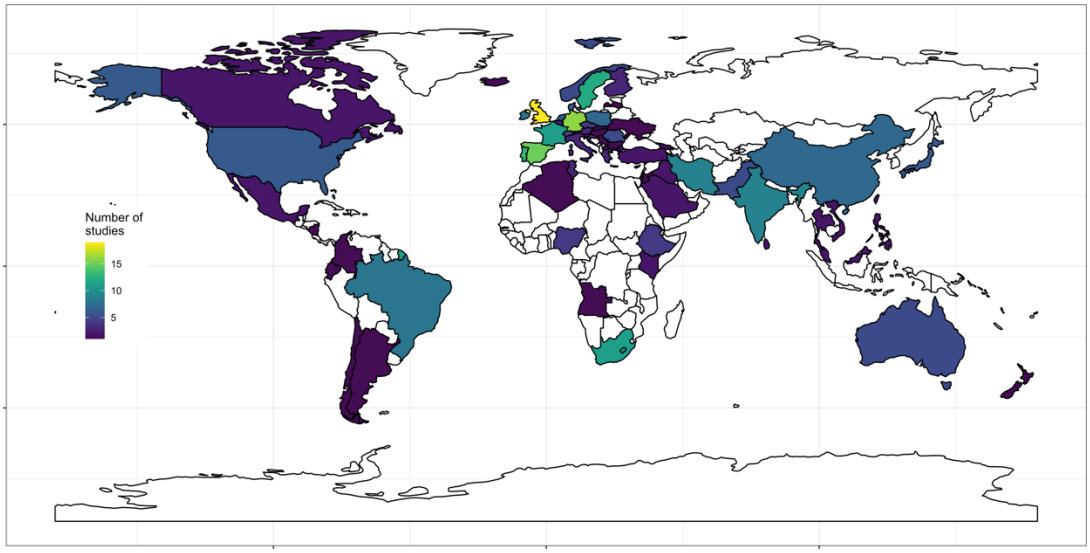


Figure 2.4: Global distribution of included studies

A) Numbers of included studies using a sample from a country. B) The proportion of studies with a positive finding in each country. In both plots, no background colour indicates no relevant studies were found to sample in this region.

A)



B)

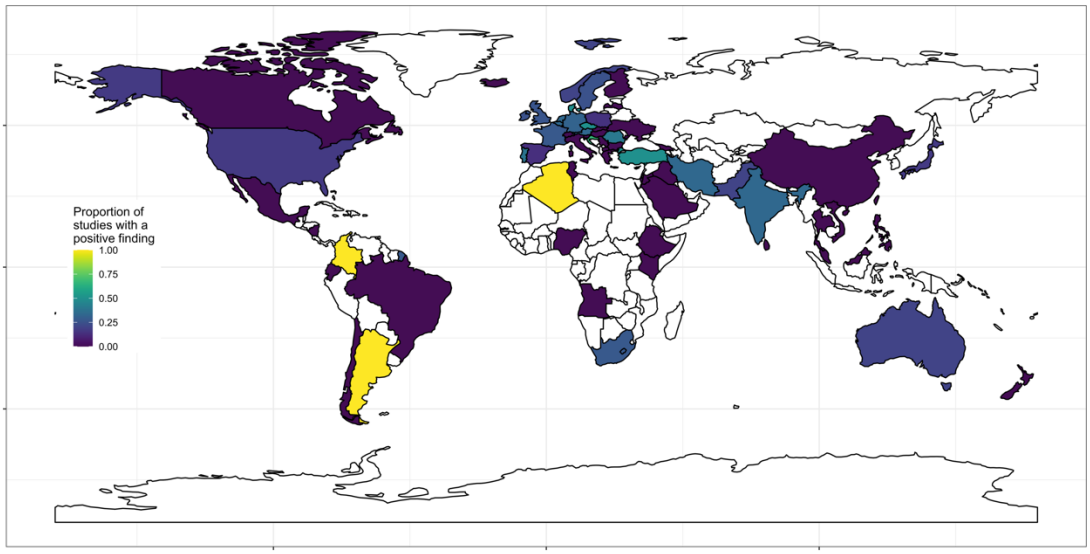


Figure 2.5: Antibiotic resistance phenotypes studied

Proportions of studies with different conclusion types for each antibiotic resistance phenotype studied. Phenotypes occurring in fewer than 5 studies excluded. The number of studies in each category is printed below each bar. HWW: Hospital wastewater; MWW: Municipal wastewater.

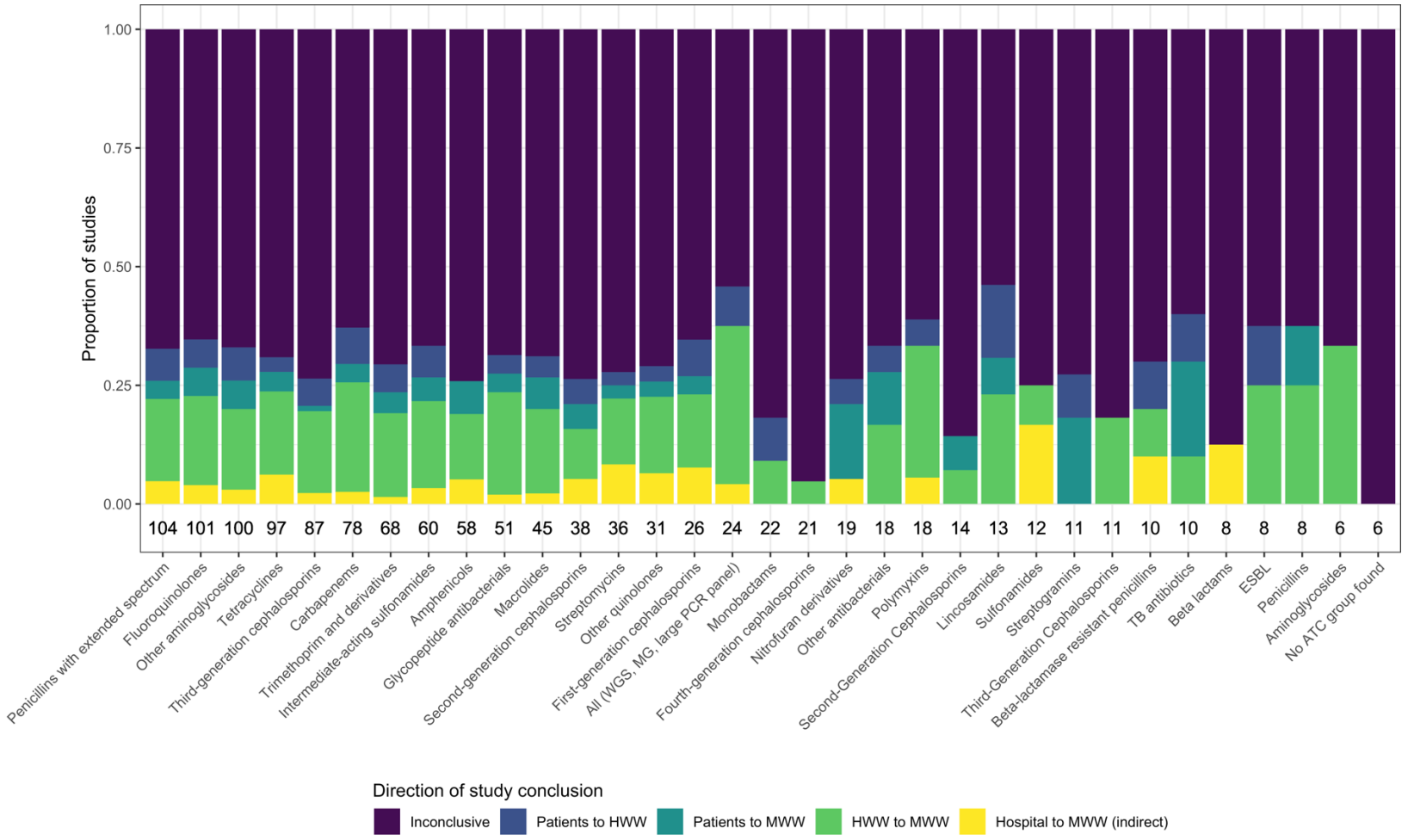


Figure 2.6: Bacterial groups studied

Proportion of studies with different conclusion types for each bacterial groups studied. The numbers of studies in each group are plotted below each bar. HWW: Hospital wastewater; MWW: Municipal wastewater.

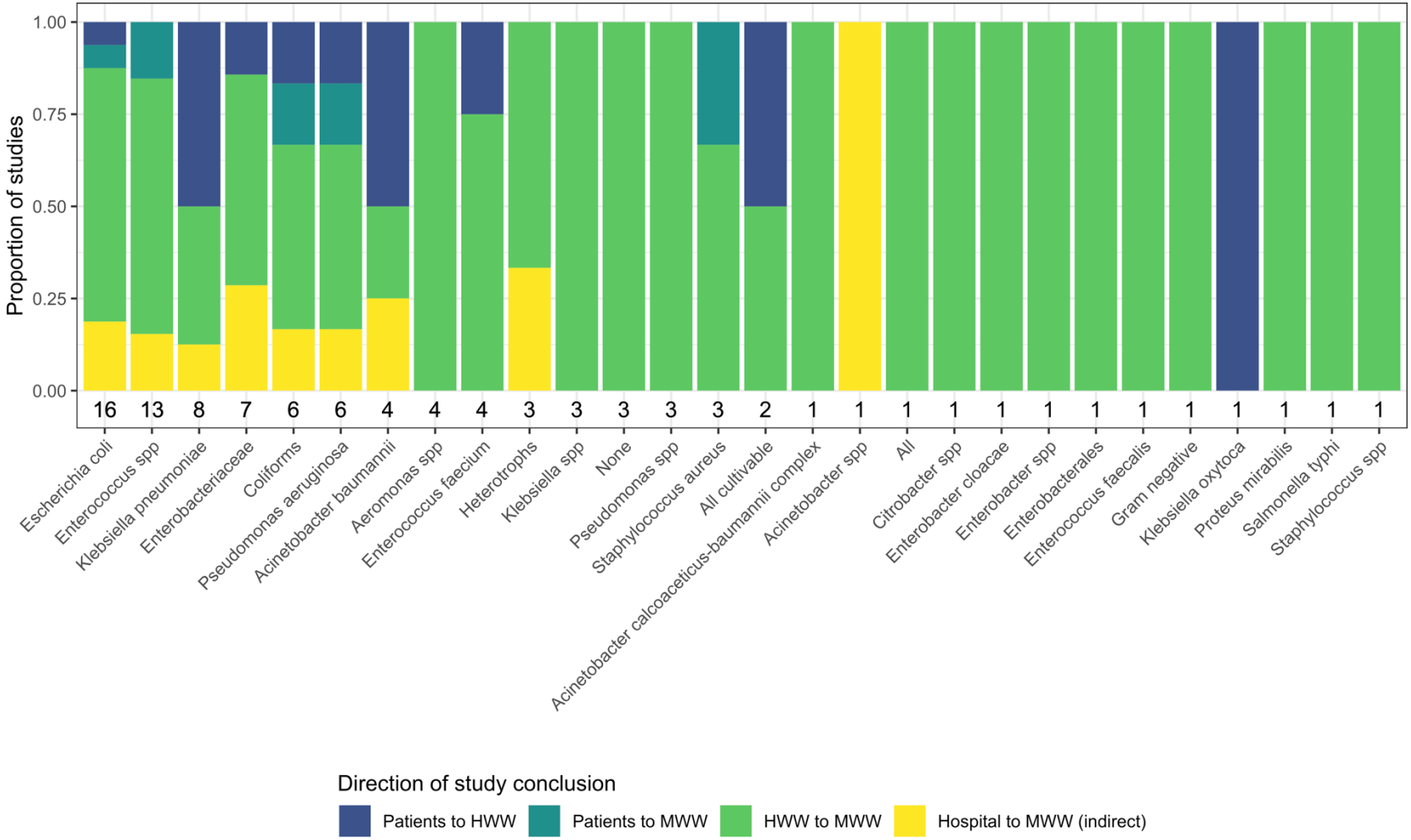


Figure 2.7: Type of statistics used

Proportion of studies with different conclusions for each kind of statistics, or no statistics. The number of studies in each category is printed beneath each bar.

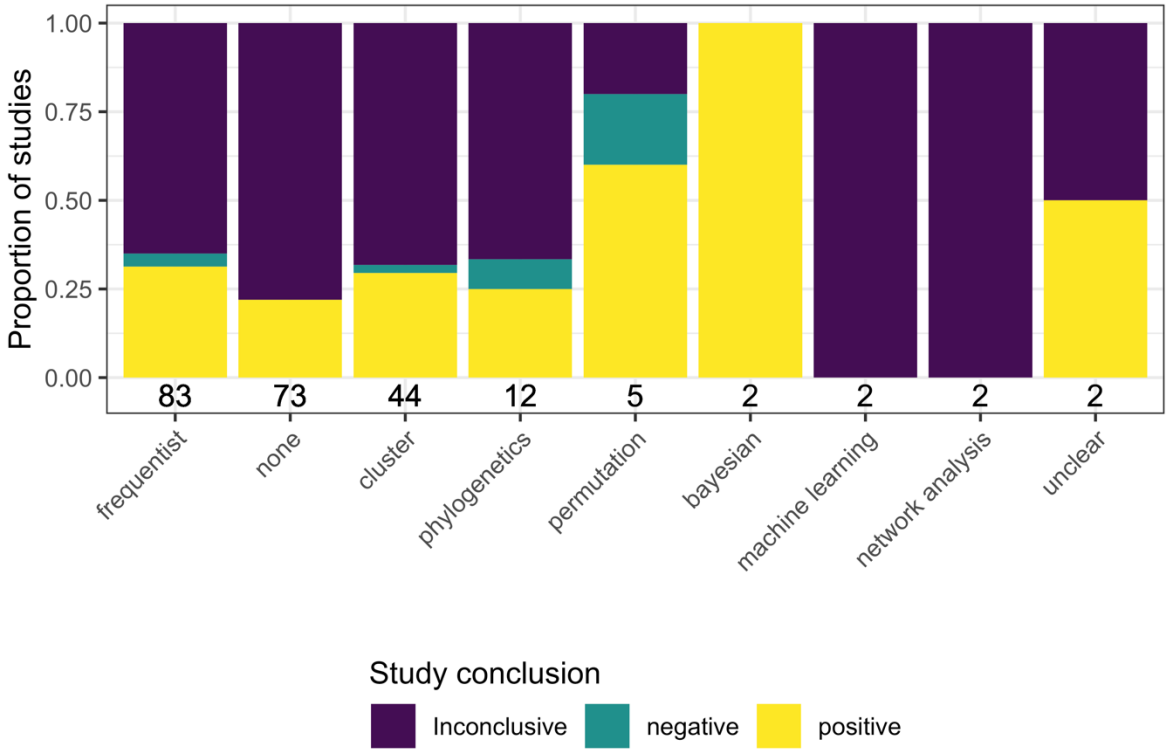
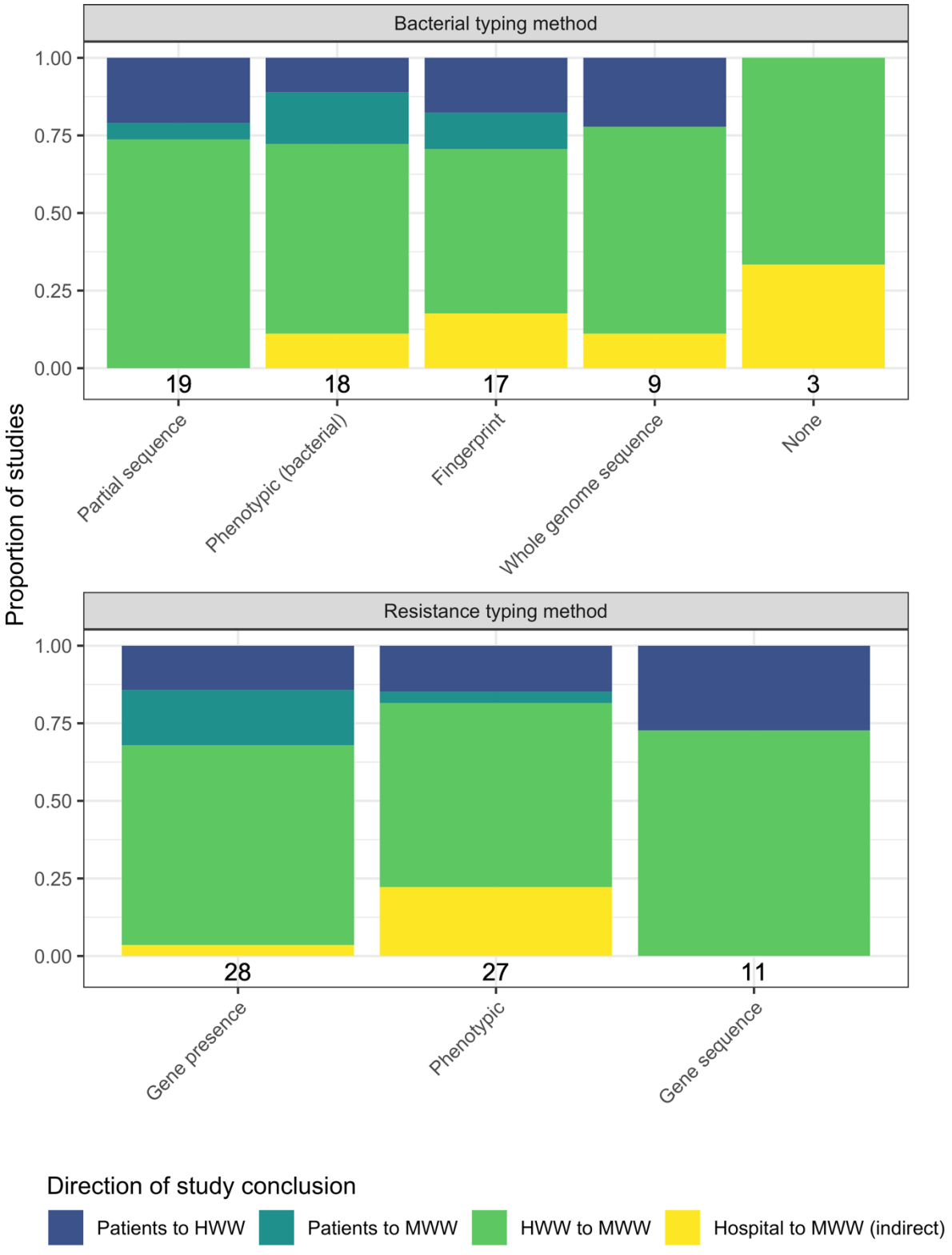


Figure 2.8: Typing methods used

Proportion of studies with different conclusions for each kind of typing. The number of studies in each category is printed beneath each bar.



2.5 Discussion

In this chapter, I used a systematic scoping review to investigate the dissemination of resistance from clinical sources to wastewater. Of 2083 screened studies, data was extracted from 201. There is increasing evidence that ABR in clinical sources, such as hospital patients and effluent, can be a source of resistance in municipal wastewater. However, conflicting results remain, and the contribution of ABR to the overall resistance profile in sewage may be low. There is a need for more high evidence quality studies using high resolution typing, control samples, larger sample sizes (e.g. more than three sampling sites and more than one sample per site), and statistical methods, rather than those with only a descriptive analysis of the phenotypic results of isolates from one or two wastewater samples.

There was some high quality evidence that ABR can disseminate from patient to hospital effluent from studies using partial or whole genome sequencing, multiple sampling, and statistical analysis (Ory et al., 2016; Popa et al., 2021). This is supportive of the use of hospital wastewater itself as a way of non-invasively monitoring the patient resistome. Patient to hospital effluent dissemination has implications for wastewater treatment and for wastewater surveillance. Firstly, it adds to the evidence that hospital wastewater presents a health risk. The prevalence of treatment of hospital wastewater varies around the world, but may particularly be a risk in LMICs where wastewater treatment can be less efficient (Nadimpalli et al., 2020). Secondly, it raises the question of to what extent sewage samples taken downstream of hospitals reflect the hospital resistome instead of the community resistome.

However, there were also some conflicting results on the relatedness of hospital patient and hospital effluent resistant bacteria. Two studies with similar methodologies, one in a Romania (Popa et al., 2021) and the other in the UK (Ludden et al., 2020), produced different results. This may simply show that although resistant bacteria in patients can be disseminated to hospital effluent, a resistant infection in a patient is not guaranteed to be represented in the sewage. To capture patient resistomes, it may be necessary to take multiple samples per hospital and use high sensitivity techniques such as qPCR. At the same time, it may be that this dissemination route could be more common under some conditions than others. For example, the Romanian study was conducted in an infectious disease hospital whereas the UK study was in a haematology ward of a general hospital. Possible differences between these hospital types could lead to the hospital effluent bearing more patient-derived resistant bacteria, such as prevalence of resistant colonisations and infections, site of resistant infections (e.g., blood infections less likely to be represented than gut and urinary tract infections), use of antibiotics, and patient age. Understanding the impact of hospital-related factors on the dissemination of resistance from patients to sewage could inform on the risks of effluents from different hospitals, and the different impacts of these hospitals on downstream municipal sewage. However few if any studies have systematically compared the effluents of different hospitals, although one study of different hospital wards did find some variation (Perry et al., 2021).

Given the evidence for a patient-to-hospital wastewater dissemination route for resistant bacteria, it is interesting that few studies compared resistance in the hospital environment (such as sinks, toilets) and hospital wastewater. It has been shown that hospital wastewater can select for resistance (Hutinel et al., 2021; Kraupner et al., 2021), that the hospital environment is rich in antibiotic residues (Voigt et al., 2019), and that environment to patient transmission of resistant bacteria can occur (Constantinides et al., 2020), suggesting the hospital environment may represent a reservoir and enricher of resistant bacteria. The selection of resistant bacteria within wastewater is an area of concern, as it may increase the amount of resistance in wastewater and therefore the load on WWTPs (Kraupner et al, 2021). There were even fewer studies that collected samples from health care workers as well as sewage. It seems possible that if patients can contribute to resistance in the hospital wastewater, staff could do so as well, although potentially different rates of colonisation and water use in staff may result in this dissemination route being less frequent. If there are resistant bacteria from the hospital staff and environment found in hospital effluents, accounting for this in the analysis of wastewater surveillance may improve patient-level predictions, so this is an important gap in the literature.

The evidence around the dissemination of resistance from hospital effluent to municipal wastewater was mixed. Hospital effluent has been shown to be distinctive and to reflect the patient population (Hassoun-Kheir et al., 2020; Hocquet et al., 2016). The next question is then do these resistant bacteria reach municipal WWTPs, where surveillance samples are usually taken? Studies attempting to address this question have been numerous, as clinically-derived resistant bacteria making up a portion of WWTP influent is a public health concern (Manaia, 2014; Pruden et al., 2013), as well as calling into question the use of municipal wastewater for community resistance surveillance. Whole genome sequencing studies have found related strains in hospital and municipal wastewater (Cahill et al., 2019; Ekwanzala et al., 2019). On the other hand, studies that have used metagenomics (Kutilova et al., 2021; Ng et al., 2017) or large qPCR (Buelow et al., 2018, 2020) panels to characterise the resistome of wastewater suggest a low overall contribution of hospital wastewater to the municipal wastewater resistance signature. On top of this, studies have found some positive correlations between resistance levels in patients and resistance gene abundance in sewage (Karkman et al., 2020; Mbanga et al., 2021; Pärnänen et al., 2019).

In all these studies, interpretation of the results is a challenge. Hospital effluent generally makes up a very small fraction of the volume of WWTP influent (Hocquet et al., 2016), in addition to travelling hundreds or thousands of metres to the nearest WWTP. These factors may be expected to lead to a smaller chance of capturing clinical bacteria in a given sample. Resistance levels in hospitals and hospital wastewater may also correlate with resistance levels in the community because some of these bacteria are circulating in the community. Relatedness found between clinical sources and municipal wastewater may reflect this rather than dissemination of resistance from hospitals to WWTPs. To resolve this, ‘control’ samples from community sites, healthy community members, and outpatients can be used. Some conclusive studies did collect these samples, but not many. In addition, control samples would be most useful when selected to be comparable to the main samples (e.g. regional matching between community and hospital effluent sampling).

The use of any kind of statistical analysis was more common in studies with conclusive results, suggesting that the use of statistics generates more conclusive findings. Sewage data is often noisy, complex, and high dimensional. Making sense of this type of data requires flexible statistical approaches (Bengtsson-Palme et al., 2017; Nguyen et al., 2021). For example, resistance gene counts are likely to be overdispersed, and there may be factors such as physicochemical characteristics that must be accounted for in the analysis. However, more powerful models such as regression, phylogenetics, and machine learning were rare. This was partly due to the low sample size in many studies. Disentangling the multiple sources of resistance in municipal wastewater will require more studies with a combination of greater sample size, control samples, and more powerful statistical techniques.

Few studies had negative findings. This could indicate consensus in the literature that hospital ABR can disseminate to municipal wastewater. However, many other factors could reduce the number of negative findings. There could be a reporting and publication bias against evidence that suggests this dissemination does not happen. Bias against negative or 'null' results is well documented in clinical and social studies, and are thought to also be an issue for the environmental sciences (Bilotta et al., 2014). Authors may also be more cautious to conclude against this route of dissemination if it is considered plausible, especially if the evidence provided by their study is not of high quality. Only studies with statistical analysis had a negative conclusion, which may point to a lack of negative conclusions due to low evidence quality in this field.

The choice to select studies for inclusion based on an evidence statement introduces limitations. Not having many restrictions on study design allowed a broad scope in types of evidence and transmission routes identified. However, I may have excluded studies with relevant data or results if the authors did not discuss them. An evidence claim is also potentially subjective, and some studies that intended to claim evidence may have been missed. In addition to limitations caused by this criterion, I could not quantitatively compare the results of the selected studies because of the heterogeneity of types of data and results included, which is a limitation of having a broad scope. A review focussed on one of the identified ABR transmission routes (such as patients to hospital effluent) could select studies on a sampling design basis, and their quantitative results could then be analysed. Narrow focus systematic reviews of this kind are better suited to a) confidently identify all relevant datasets, and b) thoroughly assess the evidence for such a transmission route.

In conclusion, I found that there is increasing evidence to show that hospital-associated bacteria can disseminate to hospital effluent, and possibly to WWTP influent. Knowledge gaps remain around the overall contribution of hospital effluent to resistomes in municipal wastewater, and how this could vary for different types of hospitals and in different countries. Studies with control samples in the community, robust sampling design, statistics, and high-resolution typing methods are needed to quantify hospital-to-WWTP influent ABR dissemination more precisely. This is particularly important for informing translating the findings of wastewater surveillance to community burdens of resistance.

3 The role of the environment in transmission of antibiotic resistance between humans and animals: a modelling study

3.1 Abstract

Antibiotic resistance can be transmitted between animals and humans directly or indirectly, through transmission via the environment. However, little is known about the contribution of the environment to resistance epidemiology. Here, I use a mathematical model to study the effect of the environment on human resistance levels and impact of interventions to reduce antibiotic consumption in animals. I developed a model of resistance transmission with human, animal, and environmental compartments. I compared the model outcomes under different transmission scenarios, conducted a sensitivity analysis, and investigated the impacts of curtailing antibiotic usage in animals. Human resistance levels were most sensitive to parameters associated with the human compartment (rate of loss of resistance from humans) and with the environmental compartment (rate of loss of environmental resistance and rate of environment to human transmission). Higher environmental transmission also reduced the impact of curtailing antibiotic consumption in animals on resistance in humans. I highlight that environment-human sharing of resistance can influence the epidemiology of resistant bacterial infections in humans and reduce the impact of interventions that curtail antibiotic consumption in animals. More data on resistance in the environment and frequency of human-environment transmission is crucial to understanding the population dynamics of antibiotic resistance.

3.2 Introduction

Antibiotic resistance (ABR) is a One-Health issue, present in a variety of commensal and pathogenic bacteria found in humans, animals and the environment (Robinson et al., 2016; Woolhouse et al., 2015). The potential of the environment for dissemination of ABR has been increasingly recognised, with particular focus on the volume of resistance bacteria in human and agricultural wastewater effluent being discharged into natural waters and soils (Bürgmann et al., 2018; Larsson et al., 2018; Manaia, 2014).

There are many potential routes for resistant bacteria into the environment. Several studies have demonstrated is it likely that resistant bacteria in humans can be transferred to the environment, including rivers (Amos et al., 2015), coastal waters (Leonard et al., 2015), and soils (Palacios et al., 2017). In addition, studied have linked resistant bacteria in animals and their respective environments, such as between wild animals and human-impacted environments (Araujo et al., 2014; Swift et al., 2019), and between livestock and their environment, especially in aquaculture (Cabello et al., 2013; Call et al., 2013). However, the risk that the resistance in the environment poses to humans and animals remains poorly understood (Raven et al., 2019).

Mathematical models are an important tool to study complex dynamics inherent in the emergence and spread of resistance (Knight et al., 2019) and can therefore be used to improve our understanding and combat the spread of ABR in humans, animals and the environment. However, a lack of data and understanding about the burden, selection and transmission of resistant bacteria, especially in animals and the environment, presents a challenge to parameterising models of ABR from a One-Health perspective. Consequently there are few models of resistant bacteria that connect humans, animals and the environment (Niewiadomska et al., 2019).

Some existing studies incorporate an environmental component into transmission models of resistant bacteria in hospitals or farms. Two compartmental models found that reducing or eradicating resistance in a hospital setting was harder to achieve when the environment was explicitly modelled (Kouyos et al., 2011; McBryde & McElwain, 2006). Studies taking the environment into account when modelling the spread of resistance in farms have found environmental parameters were key in dynamics of resistance in the farm (Call et al., 2013; Græsbøll et al., 2014). However, a recent modelling study found that interventions to reduce antibiotic consumption in animals would still be effective when the influence of resistance in animals and the environment is considered (Booton et al., 2021). These findings indicate the need for further exploration of the role of the environment with fully dynamic transmission models.

In this study, I aimed to investigate the importance of the environment in the long term dynamics of resistant bacterial infections in humans, including how it might affect the impact of interventions to reduce resistance in humans. A compartmental of resistance transmission within and between humans, animals and the environment was developed. I use a dynamic environmental compartment, improving on existing models

by allowing us to assess the importance of within-environment processes. My objectives were: 1) to investigate how adding an environmental compartment affects the long-term dynamics of resistance in humans, and the sensitivity of the model to its parameters; and 2) to investigate the impact of interventions to curtail antibiotic usage in animals or environment to human transmission on the prevalence of resistance in humans in this model.

3.3 Methods

3.3.1 Model description

I extended the original model presented in (van Bunnik & Woolhouse, 2017), to include an environmental compartment. Humans and animals gain resistant infection by exposure to antibiotics, or exposure to other humans, animals or environments carrying resistant bacteria. Resistance in the environmental compartment is increased by contact with humans or animals who carry resistant bacteria, or via exposure to antibiotics that have been excreted by humans or animals. The environment is not considered to be any one type of environment, such as water or soil, but rather a summation of these types.

I define the model using a system of coupled ordinary differential equations:

$$\frac{dR_H}{dt} = (1 - R_H) \cdot (\Lambda_H + \beta_{HH} \cdot R_H + \beta_{AH} \cdot R_A + \beta_{EH} \cdot R_E) - \mu_H \cdot R_H \quad (1)$$

$$\frac{dR_A}{dt} = (1 - R_A) \cdot (\Lambda_A + \beta_{AA} \cdot R_A + \beta_{HA} \cdot R_H + \beta_{EA} \cdot R_E) - \mu_A \cdot R_A \quad (2)$$

$$\frac{dR_E}{dt} = \gamma_H \Lambda_H + \gamma_A \Lambda_A + \beta_{HE} \cdot R_H + \beta_{AE} \cdot R_A - \mu_E \cdot R_E \quad (3)$$

R_H and R_A are the fractions of the human and animal population that are infected with resistant bacteria, respectively, and R_E is a measure of the amount of resistant infectious bacteria in the environment. Λ_H is the constant rate at which resistance is gained from exposure to antibiotics in humans, and Λ_A is the equivalent in animals. These are composite variables, taking into account both the amount of antibiotics consumed and the rate at which selection causes resistance in bacteria to arise. μ_H is the reversion rate of humans infected with resistant bacteria to having only sensitive bacteria, and μ_A is the equivalent in animals. This includes the rate of clearance of resistant infection and the rate of death in a fixed-size population. The parameters γ_H and γ_A are scaling parameters determining how much of the antibiotic exposure in humans (Λ_H) and animals (Λ_A) will result in excreted antibiotics selecting for an increase in resistant bacteria in the environment. μ_E is the rate of loss of resistant infectious bacteria from the environment. Transmission within and between the compartments is controlled by the β transmission coefficients, with the subscripts indicating the direction of transmission of each coefficient. For example, β_{HH} is the transmission coefficient between humans, and β_{EH} is the transmission from the environment to humans.

Further details about parameter definitions, units and values ranges can be found in the Appendix B Table 1. Fig. 3.1.A shows a flow diagram representing the movement of infectious resistant material between and within the different compartments. All rates are per capita with respect to the human and animal populations, and per environmental unit with respect to the environment (see next section). I used the steady state solutions of this model, obtained numerically, as I were interested in long-term prevalence. The timestep of the model represents one month. Determination of the time step is discussed in Appendix B (Additional Methods Information).

3.3.2 Capacity for resistance in the environment

Equation (3) represents the environment as an unbounded compartment, in which the amount of resistant infectious material in the environment is in the range 0 - ∞. I consider one "unit" of the environment to be the human infectious potential equivalent. This means that for a value of $R_E = 1$, if the transmission coefficients β_{EH} and β_{HH} were the same, each unit of the environment would transfer resistant material to humans at the same rate that an infected human would to another human. Although theoretically the environment may have some maximum capacity for resistant material, I do not have a way to determine this capacity, so I modelled the environment as an unbounded compartment. For comparison, I also explored a version of the model in which resistance levels in the environment cannot exceed 1. In this model the environmental compartment is specified:

$$\frac{dR_E}{dt} = (1 - R_E) \cdot (\gamma_H \Lambda_H + \gamma_A \Lambda_A + \beta_{HE} \cdot R_H + \beta_{AE} \cdot R_A) - \mu_E \cdot R_E \quad (4)$$

This model assumes that there is no growth or dissemination of resistant organisms within the environment. I also assume that the environment is only exposed to antibiotics that are excreted by humans or animals. The environment may be exposed to antibiotics directly through, for example, the effluent of pharmaceutical industries, but I do not consider this specific case here.

3.3.3 Impact of interventions on resistance in humans

I investigated the impact of two types of interventions on the levels of resistance in the human compartment. Firstly, I looked at interventions to remove antibiotic usage in livestock (reducing Λ_A to 0), and how changes to environmental parameters affect the effectiveness of this intervention. Secondly, I looked at interventions that would reduce the transmission of resistant bacteria from the environment to humans (reducing β_{EH} to 0).

I measured the impact of interventions as the percentage decrease in resistance levels in humans, following (van Bunnik and Woolhouse, 2017). I compare equilibrium values of R_H before (R_H^*) and after the intervention (RI_H^*), to obtain the impact, or percentage decrease in human resistance levels:

$$\omega = 1 - \frac{RI_H^*}{R_H^*} \quad (5)$$

I investigate the impact of reducing β_{EH} and of curtailing antibiotic usage in animals (Λ_A).

3.3.4 Sensitivity analysis

I use the extended version of the Fourier Amplitude Sensitivity Test (FAST) (Saltelli et al., 1999) to analyse the relative influence of each parameter on the value of R_H, the outcome measure of interest. A total sensitivity index for each parameter is calculated based on the variance of R_H over variation in all parameters. The R package fast was used for this analysis (Reusser, 2015).

3.3.5 Parameterisation

Due to a paucity of data about many of the parameters in the model, I aimed to explore a wide range of parameter scenarios in this model. I chose the following transmission scenarios: 1) a baseline, with transmission parameter values similar to those of the original (van Bunnik & Woolhouse, 2017); 2) a balanced transmission scenario, with all transmission coefficients equal; 3) human-driven transmission (i.e., if the subscript H denotes the humans and x denotes any other compartment $\beta_{Hx} > \beta_{xx}$); 4) animal-driven ($\beta_{Ax} > \beta_{xx}$); and finally 5) environment-driven ($\beta_{Ex} > \beta_{xx}$).

I also averaged the results across parameter sets generated randomly using sampling distributions for the three parameters R_H that was most sensitive to (viz. μ_H, μ_E, and Λ_H), to avoid over-reliance on model dynamics that are unusual to a particular combination of parameters rather than generally true of the system. All parameter values and sampling distributions can be found in Appendix B Tables 2 and 3, as well as the methods for obtaining transmission scenario parameters.

3.3.6 Software

Analyses were carried out using Wolfram Mathematica version 11.3 (Wolfram Research Inc., 2018), R 4.1 (R Core Team, 2022), and Julia 1.7 (Bezanson et al., 2017). The code for the model, parameter set generation, and visualisations is available at https://github.com/hannahlepper/animal-human-env-model.

3.4 Results

All analyses were conducted in both bounded and unbounded environmental capacity versions of the model.

3.4.1 Long term dynamics of resistance in humans

3.4.1.1 Prevalence of resistance in humans

For all transmission scenarios, parameter sets were identified that corresponded to the intended target equilibrium human resistance prevalence of 71% in both the bounded and unbounded versions of the model (Appendix B Fig. 1) (target prevalence

used in original study from van Bunnik and Woolhouse, 2017). Fig. 3.1.B shows that the amount of resistance in the environment was influenced by the model structure and the transmission scenario. The highest level of resistance in the environment was in the environment-driven, unbounded version of the model, indicating that an implausibly high level of environmental contamination is not needed for observed human resistance levels.

3.4.1.2 Sensitivity analysis

Model sensitivity results are presented in Fig. 3.1.C. In both bounded and unbounded models, human resistance prevalence was most sensitive to μ_H, the rate of loss of resistance from humans, but relatively insensitive to Λ_A, the antibiotic consumption in animals. The rate of transmission from the environment to humans, β_{EH}, was at least as important as β_{HH} and β_{AH}, rates of transmission to humans from other humans and from animals. Moreover, β_{EH} is more influential than any other transmission parameter in the unbounded model. The rate of loss of resistance from the environment, μ_E, was more important for human resistance levels in the unbounded than the bounded model.

3.4.2 Impact of interventions to reduce resistance in humans

3.4.2.1 Impact of curtailing antibiotic usage in animals

Curtailing antibiotic usage in animals had a small impact on human resistance levels, and the impact was lower when the environment was explicitly modelled or when animals contributed less to resistance transmission (Fig 2). The percentage decrease in human resistance levels achieved *without* an environmental compartment and using the parameters of the original model (the 'baseline transmission scenario) was 3.2%. Simply adding an environmental compartment and keeping other parameters reduced the percentage decrease to 2.8% in the unbounded and 2.9% in the bounded model. The animal-driven transmission scenario had the highest impacts (5.8% decrease in human prevalence), and the human-driven scenario had the lowest (0.064%). In the environment-driven transmission scenario, the environmental capacity was influential: when bounded, the impact was low (0.94%), and increased when unbound (3.2%). Both the environmental structure and the transmission parameters affected the impact of antibiotic usage reduction in animals.

3.4.2.2 Reducing Λ_A vs. reducing β_{EH}

I compared the impact (ω) of reducing either Λ_A (the antibiotic consumption in animals) or β_{EH} (the transmission of resistant material from the environment to humans) (Fig. 3.3). I considered pre-intervention values of 0.1 for each parameter, as well as the impacts in different transmission scenarios. This value was chosen so that the size of the intervention was consistent between transmission scenarios in this model and with the previous model (van Bunnik and Woolhouse, 2017). Interventions to reduce β_{EH} had a greater impact than interventions to curtail Λ_A when transmission was human- or environment-driven, or when transmission was balanced. When livestock dominated transmission or for the baseline parameter set, the impacts of interventions to reduce β_{EH} or Λ_A were similar.

3.4.2.3 Effect of β_{EH} on impact of interventions to reduce antibiotic consumption in animals

I next identified the impact of reducing Λ_A across a range of values for β_{EH} (Fig. 3.4). Increasing β_{EH} decreased the size of the impact of curtailing antibiotic usage in animals in all transmission scenarios (Fig 3.4.A). The peaked shape of the impact size in the environmental transmission scenario is caused by the increase in β_{EH} allowing increasing indirect transmission in animals and humans. This effect is only observed when there is little non-environmental transmission. Fig. 3.4.B shows that the decrease in intervention impact was also observed across the range of pre-intervention values for Λ_A. These results indicate that increasing environmental transmission can reduce the impact of curtailing antibiotic usage in animals.

Figure 3.1: Model structure, environmental resistance levels, and sensitivity analysis

A: flow diagram indicating the model structure. B: R_E values in all transmission scenarios and both model structures. C: Fourier Amplitude Sensitivity Tests (FAST), indicating how much variation in R_H was explained by each model parameter. On the left, FAST for the version of the model in which R_E is bounded to 1. On the right, FAST for the version of the model in which R_E was unbounded.

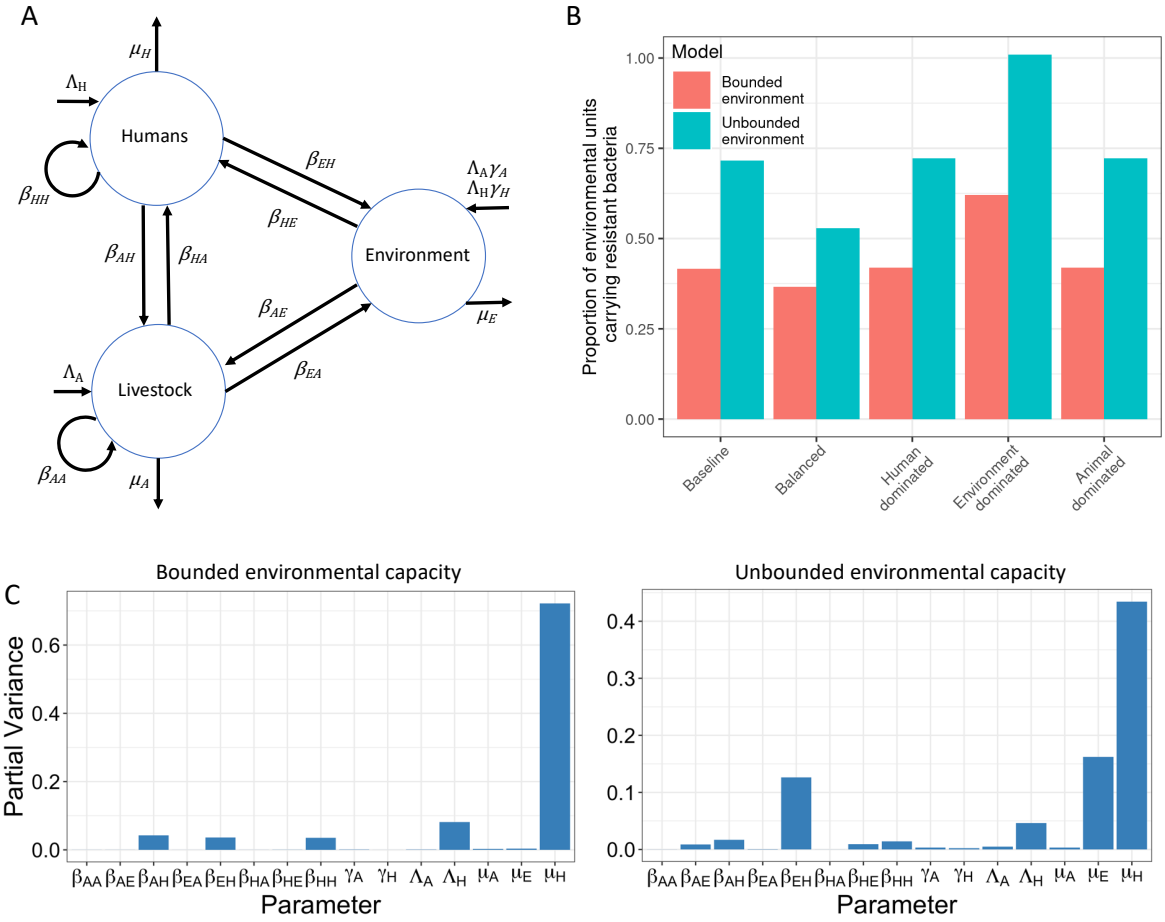


Figure 3.2: Mean impact of reducing antibiotic consumption in animals on human resistance levels across transmission scenarios

Λ_A was reduced from 0.1 to 0 in all scenarios. The green point in the baseline transmission scenario group is the mean impact for the original (van Bunnik and Woolhouse, 2017) model, with no environmental compartment included. Results were averaged for parameter sets with μ_H, μ_E, and Λ_H varied, with error bars indicating standard deviation in results.

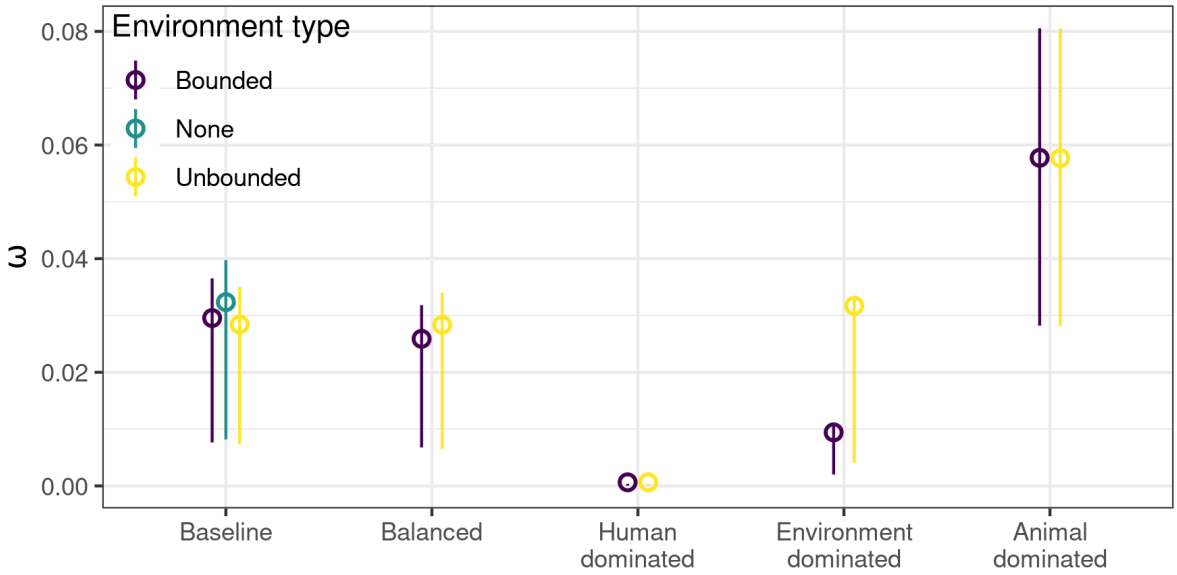


Figure 3.3: Impact of reducing environmental transmission or animal antibiotic consumption on human resistance levels

Violin plots of the impact (proportion decrease in R_H after the intervention) of reducing either β_{EH} or Λ_A in all transmission scenarios and for both model structures. The intervention target was reduced from 0.1 to 0 in each case for consistency.

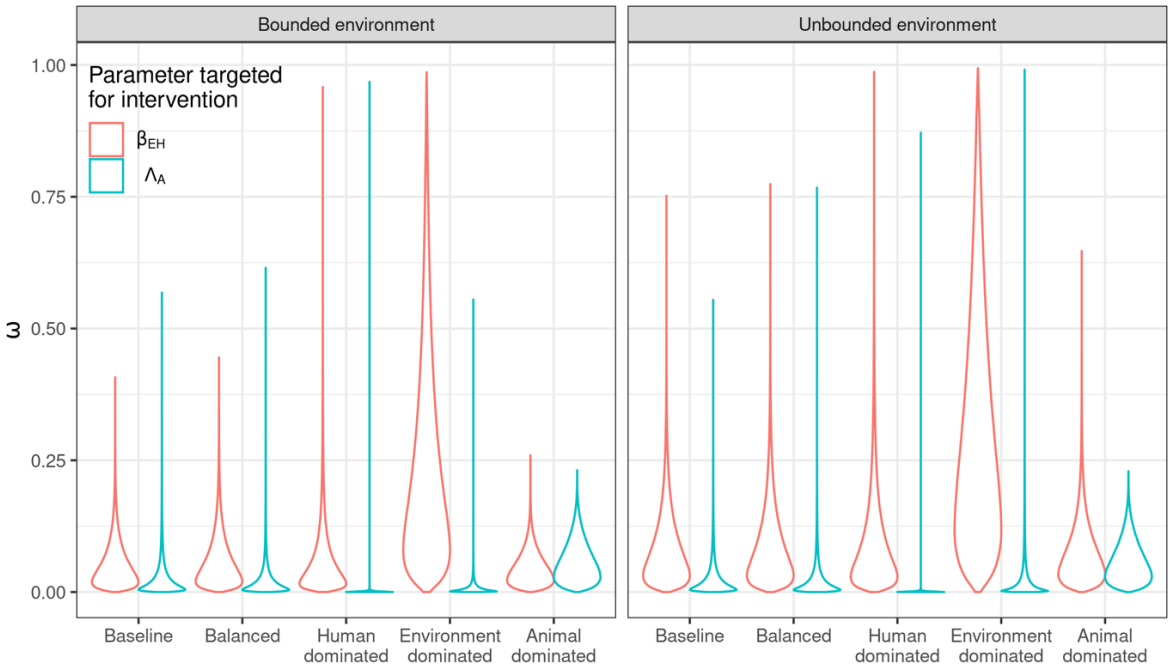
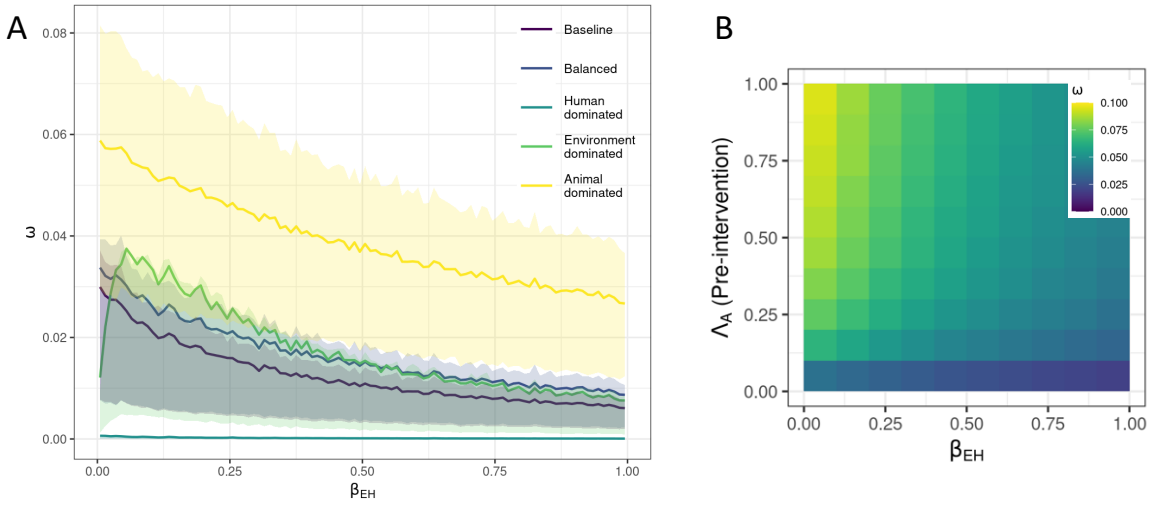


Figure 3.4: Effect of environmental transmission on the impact of decreasing antibiotic consumption in animals on human resistance levels

A: Mean impact of antibiotic decrease in animals on human resistance levels (proportion decrease in human resistance levels) for each transmission scenario with increasing rate of environment to human transmission (β_{EH}). Ribbons indicate 25% and 75% impact quantiles. B: Heatmap of the impact of different pre-intervention values of Λ_A (y axis) against different levels of environment to human transmission, β_{EH} (x axis), for the animal transmission scenario in the unbounded model. The colour of the tiles indicates the average value of the impact of the intervention from 17,000 parameter sets where μ_H, μ_E, and Λ_H were varied.



3.5 Discussion

3.5.1 Key findings

In this Chapter I modelled the transmission of resistant bacteria between humans, livestock animals and the environment, and assessed the impact of interventions that reduce antibiotic consumption in animals or decrease the transmission of resistant bacteria from the environment to humans. I found that human resistance prevalence is sensitive to transmission between humans and the environment. Including an environmental compartment in the model decreased the impact of curtailing antibiotic resistance, and a more transmissible environmental reservoir of resistant bacteria further mitigated the impact of this intervention. Reducing the transmission of resistant bacteria from the environment to humans was found to be a more effective intervention than reducing antibiotic consumption in animals. Overall, these results indicate that resistant bacteria in the environment can influence the prevalence of resistance in humans. The size of environmental influence will depend on the amount and dynamics of resistant bacteria in the environment. Assessing the likelihood of observing these theoretical results in the real world is hindered by a lack of quantified, generalisable data on the types, amount, and degradation of resistance in the environment, and the transmission of resistance between humans, livestock and the environment.

3.5.2 Is curtailing antibiotic usage in animals an effective intervention to reduce human resistance levels?

The greatest observed impact of curtailing antibiotics in animals was a modest 10% decrease in human resistance level in a balanced transmission scenario, and the smallest impact was a <1% in the human-driven transmission scenario. This result provides little theoretical support that curtailment of antibiotics would appreciably decrease resistance in humans in many settings. In contrast, there is some empirical evidence that curtailing antibiotics in livestock could reduce human resistance levels, although from a small set of observational studies (Scott et al., 2018). A study of use of third-generation cephalosporin ceftiofur in broiler rearing in Canada found that resistance in *Salmonella* and *E. coli* was decreased in clinical isolates by 20% and 40%, respectively, after ceftiofur use decreased (Dutil et al., 2010). This real-world population-level effect is greater than these results would predict, and may indicate they are an underestimate, especially with respect to the degree of sharing of resistance between humans and animals. More data-based parameterisation will be crucial to improve the accuracy of One-Health resistance transmission models.

The size of the effect of intervening to reduce antibiotic consumption in livestock varied by transmission scenario (balanced transmission, or transmission driven by either humans, livestock or the environment). Therefore, a key question for assessing the accuracy and relevance of the resulting intervention effect sizes is how realistic are the transmission scenarios? Although transmission of resistance between humans and animals is of great concern, evidence that conclusively demonstrates a case of direct transmission is rare (Muloi et al., 2018; Wee et al., 2020). Accurately parameterising the relationship between resistance in humans and livestock is an

ongoing area of research (Thorpe et al., 2021) which will be crucial for One-Health modelling of resistance. It seems likely that on average across a large human population, human-human transmission is far more common than animal-human transmission and I suggest human-driven scenario to be most relevant for resistance dynamics in the human population.

As the transmission rate from the environment to humans increased, the effectiveness of antibiotic curtailment was decreased. This suggests that the environment can provide a 'back door' transmission route from animals to humans that can reduce the effectiveness of antibiotic curtailment by adding to overall animal-human transmission rates. Using a two-pronged approach by intervening to reduce environmental transmission at the same time could therefore improve the impact of antibiotic usage curtailment. However, the effect of environmental transmission on antibiotic curtailment effectiveness was negligible in the human-dominated transmission scenario (Appendix B Fig. 2), again indicating the importance of transmission setting for this result. It remains unclear if non-human dominated transmission scenarios are realistic, and therefore what the real-world size of this back-door effect might be. There is some evidence that microbiomes in humans, animals and the environment become more shared as interactions become more frequent (Pehrsson et al., 2016), suggesting that transmission scenarios in which humans do not dominate transmission (such as the balances and baseline scenarios) are possible. Further work to quantify environmental resistance concentrations and transmission could improve accuracy of outcome predictions of antibiotic usage interventions. As reducing antibiotic usage in livestock animals is a costly intervention, it is important to ensure optimal implementation.

3.5.3 Could the environment be an effective alternative intervention target?

The rate of transfer of resistant bacteria from environment to humans (β_{EH}) is also a potentially effective intervention target. Human resistance prevalence levels were sensitive to β_{EH} and μ_E, the rate of loss of resistant bacteria from the environment (sensitivity analysis, Fig. 3.1.C), which suggests that interventions to reduce how much resistance humans gain from the environment would be effective. Indeed, the impact of reducing β_{EH} was more effective than antibiotic usage curtailment interventions, although the difference was small in the animal-dominated scenario (Fig. 3.2.A). Interventions that improve sanitation have been proposed to reduce occurrences of transmission of resistance between humans and the environment in informal urban communities in LMICs where there is frequent exposure to resistance bacteria in the environment (Collignon et al., 2018; Nadimpalli et al., 2020). Nadimpalli et al (2020) focus particularly on the potential benefits of improved water and wastewater infrastructure for controlling and preventing ABR transmission, but note that few studies have investigated the impacts of sanitation interventions on ABR.

3.5.4 Should the environment be included in ABR models?

In this model, the environment played an important role in the long-term dynamics of antibiotic resistance levels in humans. Mechanistically, the environment acts as a

reservoir for antibiotic resistance from humans and animals in this model structure. Therefore, parameters that provide more opportunity for transmission to humans were influential in human resistance levels, especially the rate of loss or level of persistence of resistant bacteria in the environment (μ_E). Environmental parameters were also influential in the size of impact of interventions, and these results show that it may be an effective intervention target itself. Existing models that incorporate an environmental component have also highlighted the potentially strong role the environment could play in increasing resistance levels in humans and undermining interventions (Booton et al., 2021; Græsbøll et al., 2014; Kouyos et al., 2011; McBryde & McElwain, 2006). Most models include environment as a constant rather than a dynamic compartment, with the exception of Booton et al, 2021. As I find comparable results to models with constant compartments, this may indicate that models incorporating the environment simply may be enough to account for this additional source of resistant bacteria. On the other hand, the model in Booton et al, 2021, assumes that transmission of resistance (including from the environment) is dependent on exposure to antibiotics and accordingly finds that human antibiotic usage is the most influential parameter for human resistance, downplaying the role of the environment. This contrasting result points to a need for further models that compare the contribution of the environment under different model structures and assumptions. Incorporating the environment into models of ABR spread may be important in understanding ABR prevalence and for evaluating intervention success.

3.5.5 Modelling the environment highlights data needs

The results highlight some key data needs for understanding the importance of ABR in the environment for humans. There are two influential parameters in the model which are difficult to parameterise from existing data: the rate of transfer of ABR from the environment to humans, and the rate of loss of resistance in the human population.

How frequently humans gain resistant bacteria after exposure to an environmental source is unknown. There is evidence that humans can be exposed to resistant bacteria in the environment. For example, one study estimated that the amount of third-generation cephalosporin resistant *E. coli* that humans would ingest during recreational water use in coastal regions in England and Wales poses a risk of infection (Leonard et al., 2015). However it is not clear how often these exposures lead to infection or colonisation (Berendonk et al., 2015). More research that demonstrates a close relationship and epidemiological link between resistant bacteria colonising the environment and humans is needed to understand the frequency of environment-human transmission events. Use of high resolution typing such as whole genome sequencing of, for example, isolates from hospital patients and the hospital environment in longitudinal studies would be ideal for this research.

Studies have provided data on the rate of clearance of resistant infections in humans. A systematic review on methicillin-resistant *S. aureus* (MRSA) and vancomycin-resistant *Enterococcus* (VRE) colonisation found that it takes a period of 88 and 26 weeks on average to clear MRSA and VRE infections, respectively(Shenoy et al., 2014). However, they note that there is considerable methodological heterogeneity in

studies of MRSA and VRE, including varying definitions of clearance, and length of follow-up (Shenoy et al., 2014). The studies also focussed primarily on hospital-associated resistance. Data on resistant bacteria colonisation prevalence and clearance in the community, where the role of exposure to animals and the environment may play a greater role, appear to be rare. Parameterising generalisable One-Health models will therefore be benefitted by more research into resistance in the community.

3.5.6 Limitations

There are some important limitations to this study that should be noted. Firstly, I make simplifying assumptions in the structure and parameterisation of the model. These are suitable to the questions posed in this study, but there are still many complexities in the spread and emergence of ABR in humans, animals and the environment to be explored. Further models should explore the importance of potential complexities, such as heterogeneity of transmission events, separate humans-specific and animal-specific environmental reservoirs, variation in the capacity for resistance in the environment, or the fitness costs to bacteria of carrying resistance in the three populations.

I do not model the dynamics of transmission of resistant bacteria and resistance genes separately, but assume that transmission parameters combine the transmission of both. This is in-keeping with the assumptions of the original model (van Bunnik & Woolhouse, 2017). Resistance genes can be transferred between bacteria via plasmid transfer or bacteriophages, and can also be lost from bacterial lineages. The transmission rates of resistance genes in human population may therefore differ from resistant bacteria, and it is a limitation that I do not capture this in the model. ABR epidemiology and surveillance is usually measured in resistant bacteria so there is little data on the prevalence and transmission rates of specific resistance genes.

Two further important assumptions about resistance in the environment are that I assume that there is no growth of resistant material within the environment, and that all antibiotics secreted into the environment are from human and livestock usage. The dynamics of resistance genes and bacteria in the environment is a complex topic, and although there are potentially environments in which resistance may spread (especially in sewage) much more empirical and modelling research is needed (Bengtsson-Palme et al., 2018; Berendonk et al., 2015). A recent review found that the sources of antibiotics in ground water include excretion from humans and animals (via sewage and manure) but also landfill, aquaculture and industrial sites (Zainab et al., 2020), so not including these sources may limit the accuracy of the results of this model. However the relative contribution of each sources is not well known and may vary from one country to another (Zainab et al., 2020).

3.5.7 Conclusions

This study illustrates the potentially important role of the environment in the epidemiology of resistant bacterial infections in humans. I highlight the need to

consider the role of the environment in the design of ABR control strategies, as it can be influential in human prevalence of resistance, reduce the effectiveness of interventions that curtail antibiotic consumption in animals, and may be an effective intervention target itself via improved sanitation infrastructure. Incorporating the environment into a One-Health model of antibiotic resistance as a dynamic compartment was useful for considering the role of the environment. However, assessing the uncertainty of model predictions is hindered by a lack of data on the types and frequency of resistance in the environment, and the frequency of environment-human transmission events.

4 A multi-response model to combine sewage and hospital antibiotic resistance surveillance data

4.1 Abstract

Background
Quantifying antibiotic resistance gene (ARG) abundance in municipal sewage is an emerging approach to ABR surveillance with the potential to inform our knowledge of resistance in the community. To improve our understanding of drivers of resistance in the community, and the relationship between resistance in hospitals and the community, here we combine hospital and sewage resistance data in a multi-response model. We investigate associations with hospital and community antibiotic usage, and country-level correlations between resistance in hospitals and the community for a range of drug-bug combinations.

Methods
Data sources were European Antimicrobial Resistance Surveillance Network (EARS-Net) for clinical data, the Global Sewage Surveillance Project for metagenomic analysis of wastewater, and the European Surveillance of Antimicrobial Consumption Network (ESAC-Net) for consumption data. We use a series of multi-response models, one for each antibiotic group of clinical interest (aminoglycosides, aminopenicillins, carbapenems, fluoroquinolones, macrolides, third generation cephalosporins, and vancomycin). A multi-response model was used to simultaneously model clinical susceptibility testing results as binomial and resistance gene abundance in sewage as Poisson. We estimated the country-level correlation between clinical and sewage resistance levels using the multi-response model and with Spearman's Rank as a comparison. We also estimated the association between community antibiotic consumption and resistance gene abundance in wastewater.

Results
We found evidence for a positive correlation between clinical and sewage aminoglycoside resistant levels, and no evidence for a correlation in vancomycin resistance. Evidence was mixed for other types of resistance, with correlation estimates sensitive to estimation method and model structure. Likelihood ratio tests and leave-one-out validation indicated that combining sewage and clinical surveillance data with a multi-response model improved model performance for some resistance types (aminoglycosides, aminopenicillins), but was not beneficial for others (macrolides, third generation cephalosporins). There was a positive effect of community antibiotic usage and resistance gene abundance for fluoroquinolone and third generation cephalosporin resistance.

Conclusions
This chapter shows that sewage surveillance data can be useful for supporting clinical predictions for some but not all drug-bug combinations. Sewage surveillance can also shed light on the community drivers of resistance, indicating that primary-case antibiotic usage may lead to increased resistance. Multi-level, multi-response models have advantages over uni-response or univariable methods for combining sewage and clinical data, as they allow inclusion of additional essential covariates and data hierarchies, which could lead to over- and underestimates of correlations if excluded.

4.2 Introduction

Antibiotic resistance (ABR) is a serious and growing threat to global public health (Cassini et al., 2019; Murray et al., 2022; O'Neill, 2014). This threat is compounded by a lack of ABR surveillance, leading to gaps in our knowledge of the distribution and causal factors of ABR, which will be crucial in optimising interventions to prevent and manage resistance infections (Knight et al., 2018a; Tacconelli et al., 2018; WHO, 2019). Currently, the main source of ABR surveillance is from routine data collection of antimicrobial susceptibility testing (AST) of isolates from invasive bloodstream infection samples from hospital medical microbiology laboratories, such as European Antimicrobial Resistance Surveillance Network (EARS-Net, European Centre for Disease prevention and Control (ECDC)). However, this data source narrowly focuses on only a few drug-bug combinations in patients with suspected bacterial infections in highly developed settings (Aarestrup & Woolhouse, 2020; Knight et al., 2018a; Tacconelli et al., 2018). Municipal sewage samples have been posed as an attractive alternative pathway for surveillance of ABR, offering a way of sampling the general population's microbiology easily, inexpensively, and without ethical barriers (Aarestrup & Woolhouse, 2020; Bürgmann et al., 2018; Pruden et al., 2021). Especially when paired with metagenomics, sewage surveillance provides rich and flexible data, as it does not focus on any specific human subpopulation, infection or resistance type (Aarestrup & Woolhouse, 2020; Pruden et al., 2021).

The potential of sewage samples and resistance gene quantification for surveillance of ABR has been demonstrated on a range of scales. Smaller scale studies have focussed on specific cities, (e.g. Copenhagen, Denmark: (Brinch et al., 2020)), or within single hospitals (Perry et al., 2021). Larger scale studies include studies of whole countries, (e.g. China: (Su et al., 2017)), Europe (Huijbers et al., 2020; Pärnänen et al., 2019), and the Global Sewage Surveillance Project (GSP) (Hendriksen et al., 2019), which applied metagenomics to samples from over 100 different countries across the globe. These studies have shown that this method can be used to investigate spatial and temporal variation and identify drivers of resistance.

Combining sewage resistance data with clinical resistance data can provide more information than looking at each independently (Flach et al., 2021; Hutinel et al., 2019; Karkman et al., 2020; Kwak et al., 2015; Pärnänen et al., 2019; Raven et al., 2019). For example, sewage surveillance could be used to complement clinical surveillance, for example through making clinical predictions in countries for which there is no clinical surveillance (Hutinel et al., 2019; Karkman et al., 2020; Kwak et al., 2015). If hospital effluent contributes to the resistome of WWTP influent, sewage may also provide more information about the resistome of hospital patients (Huijbers et al., 2020; Pärnänen et al., 2019; Raven et al., 2019). Municipal sewage can also be used to gain information on the prevalence of resistance in the general population or community, and how it differs from those in hospitals.

The strength of the link between resistance in clinical settings and WWTP influent needs to be established to determine if and how sewage surveillance can be used to monitor hospital and community resistance. Studies addressing this link through

correlating sewage and clinical resistance levels have generally found positive correlations (Huijbers et al., 2020; Hutinel et al., 2019; Karkman et al., 2020; Pärnänen et al., 2019) and, a meta-analysis of studies comparing clinical and sewage resistance levels did find a positive correlation overall (Chau et al., 2022).

However, further research is needed to understand the implications of positive correlations. They could indicate that the resistome of hospital patients are well represented in hospital effluent, or that hospital and community resistomes correlate, or could be due to correlated hospital and community antibiotic usage. The two data sources are also sensitive to different sets of population coverage, sampling and typing methods, and chance, leading to noisy data with many sources of uncertainty. The statistical methods used by existing studies to correlate clinical and sewage data do not take the complex nature of both data sources into account, creating a risk of over- or under-estimating correlation. Methods used include bivariate correlation tests (Chau et al., 2022; Pärnänen et al., 2019), which cannot include other covariates or give more weight observations based on a greater number of samples. Linear regression models that have one dataset as a covariate of the other have also been used (Huijbers et al., 2020; Hutinel et al., 2019; Karkman et al., 2020). These model structures only model the generating process, grouping levels, and explanatory factors of one dataset at a time.

In this chapter, we aimed to investigate community-hospital correlations and drivers of sewage resistance levels using a multi-response linear model. The objectives were a) to build a multi-response model that appropriately models the sampling distribution of both datasets; b) apply this model to assess the correlation between clinical and sewage resistance levels; and c) apply this model to assess the contribution of antibiotic usage to antibiotic resistance levels in the data. We developed a mixed effect generalised linear model with a Poisson and binomial component, and applied it to data from the Global Sewage Surveillance Project, EARS-Net, and the European Surveillance of Antimicrobial Consumption Network (ESAC-Net, ECDC).

4.3 Methods

4.3.1 Datasets

4.3.1.1 Global Sewage Surveillance Data

The collection and metagenomic analysis of the Global Sewage Surveillance Project (GSP) has been described in detail elsewhere (Hendriksen et al., 2019). In brief, 24 hour 1 litre composite samples were taken from wastewater prior to entry to wastewater treatment plants (influent) from urban regions across the world. Samples were frozen to -80^{o}C and transported to the Danish Technical University (DTU), where DNA was extracted using the QIAamp Fast DNA Stool Mini Kit. Pilot sewage samples collected in 2016 were sequenced using Illumina HiSeq4000, and 2017 – 2018 samples were sequenced with NovaSeq6000. To maintain consistency between pilot and main study samples, fragment size, read length, and sequencing depth targets were the same, and the PCR-free Kapa Hyper library prep was used for both sets of

samples. Reads were quality- and adapter-trimmed using BBduk2 (Bushnell, 2014). Hits to genes in the resulting dataset of reads were mapped using KMA (*k*-mer alignment) version 1.2.12 (Clausen et al, 2018) to a collection of gene databases, including ResFinder (Zankari et al., 2012) and Silva (v138) 16/18S (Quast et al., 2013).

Sample collectors also completed a metadata survey form, including information on the location, type and temperature of the wastewater. They used discrete categories to rate the flow, viscosity, and colour of the wastewater, and the socioeconomic status of the population captured in the sample. They also provided information on the transport time, transport temperature, freeze temperature, and pH of the samples.

Resistance gene phenotypes were obtained from the ResFinder database (Zankari et al, 2012). For this analysis, only genes conferring resistance to one of the antibiotics used in ECDC ASTs were considered (see Appendix C Table 1 for a list of the bacterial pathogens and antibiotics tested).

4.3.1.2 ECDC data sources

For clinical resistance data, we used the European Antimicrobial Resistance Surveillance Network (EARS-Net) (ECDC), which reports yearly AST results from 30 European countries in 26 antibiotic group-bacterial species (drug-bug) combinations (Appendix C Table 1). National infectious disease observatories passively collect results of ASTs on isolates from blood or cerebrospinal fluid samples of hospitalised patients with invasive infections (European Centre for Disease Prevention and Control, 2020). AST results are shared with EARS-Net and made publicly available after aggregation to country and antibiotic group level. Laboratory methods and clinical breakpoints vary between medical microbiology labs, but a majority follow European Committee on Antimicrobial Susceptibility Testing (EUCAST) guidelines, and since 2019 it is a requirement for EUCAST to be followed for data to be accepted (European Centre for Disease Prevention and Control, 2020). Therefore, a combination of disk-diffusion, broth microdilution, and MIC tests are used, depending on the drug-bug combination (European Centre for Disease Prevention and Control, 2020).

For antibiotic consumption data, we used the European Surveillance of Antimicrobial Consumption Network (ESAC-Net) (ECDC). This database provides yearly reports of the total daily defined doses (DDD) per 1000 inhabitants per day, per reporting country in the EEA/EU, using the Anatomical Therapeutic Chemical (ATC) classification system (WHO Collaborating Centre for Drug Statistics Methodology, 2022). Antibiotics available on national registers as annual numbers of packages consumed are used to calculate DDDs (European Centre for Disease Prevention and Control, 2020). Only antibacterials for systemic use are included in this dataset (ATC group J01). Consumption data is reported for primary and secondary care use separately.

The EARS-Net and ESAC-Net data are described in detail elsewhere (e.g. McDonell et al, 2017). To briefly describe EARS-Net data, for countries included in this analysis in 2019, the most common drug-bug combination was ampicillin resistant *Enterococcus faecium* (88.5% resistant), and the rarest was carbapenem resistant *E. coli* (0.21% resistant). Time trends vary considerably, but in most cases the proportion

of isolates resistant has increased over time. Briefly describing ESAC-Net data, in hospital in 2019 third generation cephalosporins were used in the greatest quantities, and vancomycin the least (0.17 and 0.043 DDDs per 1000 inhabitants per day on average, respectively). In the community aminopenicillins were used in greatest quantities and carbapenems the least (2.92 and 0.0024 DDDs on average, respectively). Figures showing time trends in EARS-Net and ESAC-Net can be found in Appendix C Fig. 2.

4.3.2 Statistical methods

Two methods were used for correlating the sewage and clinical resistance data: Spearman's Rank, and a multi-response hierarchical linear regression model.

4.3.2.1 Read abundance normalisation

Where resistance gene count data is normalised, the fragments per kilobase per million (FPKM) calculation was used. This measure normalises fragment counts (FC) of each gene in a homology group for a) the average length of the homology group in base pairs (L), and b) the total number of bacterial reads in the sample (B), and applies a scaling factor to facilitate analysis. The formula is:

$$FPKM = FC \cdot \frac{1}{0.001 \cdot L} \cdot \frac{1}{B} \cdot 10^6$$

This normalisation has been used in analyses of the Global Sewage Project dataset (Hendriksen et al., 2019). It allows comparison of fragment counts from genes of different lengths and between samples of different concentrations of bacterial cells.

4.3.2.2 Read homology reduction

Resistance genes in the ResFinder_20190905 database were clustered to 90% sequence identity groups using Usearch (v11.0.7) (Edgar, 2010). Read counts were summed read counts within these cluster groups prior to analysis to account for a) variable read assignment due to low read abundances and the Conclave winner-takes-all strategy used by KMA, and b) to reduce the number of very low or 0 read counts. The resistance phenotype of the group was assumed to be the union of the resistance phenotypes of each group member, i.e., if any gene within the group was recorded as conferring resistance to amoxicillin, the whole group was assumed to have this phenotype. An issue with this strategy may be that if only one gene within a group confers resistance to an antibiotic, we may overestimate the abundance of resistance to that antibiotic.

4.3.2.3 Spearman's Rank Correlations

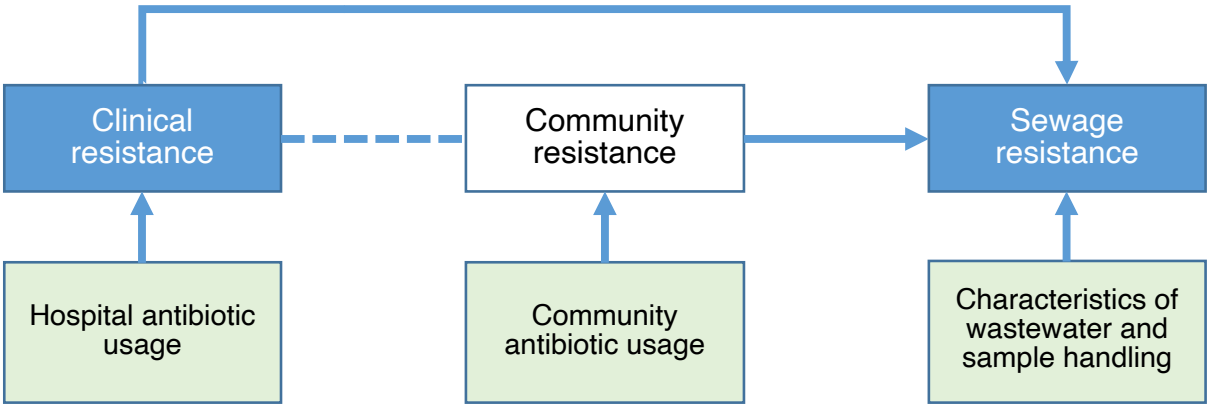
We used Spearman's Rank to investigate the correlations between sewage and hospital data of the same year and country as a simple test of correlation that does not assume normally distributed variables. This test allows comparison with the more complex multi-response model within this study, to aid in assessing its benefits and disadvantages. A separate correlation coefficient was obtained for each drug-bug combination in the ECDC data, selecting sewage abundance data for only those genes

conferring resistance to the same antibiotics as used in ASTs. The unit of observation for the Spearman's Rank tests were resistance levels in each country and year. Resistance levels in the clinical data was proportion of isolates resistant, and FPKM of resistance genes in the sewage data.

In order to have sewage surveillance observations on the same level as the ECDC surveillance observations for calculating the correlation coefficient, we aggregated across gene groups, multiple sampling sites in a single country, and across multiple batches per year. For each year and country, read counts for all phenotypically relevant genes groups, gene group lengths, and total bacterial read count were summed prior to FPKM calculation. This aggregation step results in loss of information and a reduction in the number of observations. In addition, the correlation test itself is limited as it can only look for correlation between two variables and offers no way to understand why they may correlate. To address these issues whilst still estimating a correlation, we next applied a multi-response (a.k.a. multivariate) generalised linear mixed effects model.

4.3.2.4 Multi-response linear model

We used the following conceptual model:



Solid arrows indicate a hypothesised causal relationship, and dotted arrows indicate correlation. Therefore, the model structure assumes hospital and community antibiotic usage has a direct effect on the population in which is it used, and only indirect effects on the other population. It also assumed sewage resistance levels are directly impacted by community resistance, clinical resistance (through hospital wastewater effluent) and by characteristics of the wastewater site and the way the sample was handled. We do not make a causal hypothesis about the relationship between resistance in hospitals and the community, only hypothesise that there may be a correlation. A full directed acyclic graph representing the model structure can be found in Appendix C Fig. 1.

To explain the model structure, we will first show how a multi-response linear model is specified. Multi-response models use a multi-response normal distribution on the residuals of each dataset, allowing estimation of a covariance matrix describing residual variance and between-dataset covariance. For example, for $j \in (1, \ldots, J)$

measurements of $i \in (1, \ldots, N)$ observational units (e.g. sampling site) that are normally distributed, the linear model would be:

$$\boldsymbol{y}_i = \boldsymbol{X}_i\mathbf{B} + \boldsymbol{\epsilon}_i$$
$$\boldsymbol{\epsilon}_i \sim MVNormal(0, \Sigma)$$

Where $\boldsymbol{y}_i$ is a vector of J measurements for observation i, $\boldsymbol{X}_i\mathbf{B}$ is the linear predictor (predictor variables multiplied by the estimated model coefficients), $\boldsymbol{\epsilon}_i$ is a vector of size J of the errors or residuals for observation i, and Σ is the covariance matrix.

Some features of our datasets cannot be addressed with this model structure: neither the EARS-Net data nor the sewage data are normally distributed; we do not have the same units of observation in each dataset; and we have different covariates and hierarchies that we want to apply to each dataset. Therefore, we made some adaptations to this model.

We can use alternative sampling distributions and link functions to transform residuals to a normal distribution (a generalised linear model, or GLM). The ECDC data are AST results of with a binary outcome, so are binomial. As the sewage dataset of read abundance is count data, a Poisson or negative binomial distribution might be appropriate. Although combining different sampling distributions in one multi-response model is rare, one example includes a model combining binomial and Poisson species observation data in a spatially autoregressive multi-response species distribution model (Pacifici et al., 2017). For comparability, we therefore also chose to also use a Poisson distribution for this model with additional accounting for over-dispersion. Chapter 5 will investigate other distributions for read count data. The model can be written:

$$\theta_i = \lambda_i + \epsilon_i$$
$$y_{i,j=1} \sim Pois\left(e^{\theta_{i,j=1}}\right)$$
$$y_{i,j\geq 2} \sim Bin\left(logit(\theta_{i,j\geq 2}), N_i\right)$$

In this model, θ_i is the latent variable for each observation. λ_i is the linear predictor, and includes an offset term for the Poisson model to account for gene length:

$$\lambda_i = \boldsymbol{X_{i,j}}\mathbf{B_j} + \boldsymbol{Z_{i,j}}\boldsymbol{U_j} + \ln(L_i) + \ln(B_i)$$

Where $\boldsymbol{X_{i,j}}$ is a vector of covariates for observation i and dataset j, $\mathbf{B}_j$ is the vector of fixed effect coefficients for dataset j, $\boldsymbol{Z_{i,j}}$ is a vector of membership status of observation i, dataset j, to groups being modelled with random effects, and $\boldsymbol{U_j}$ is the vector of random effects for these groups. This notation is based on Hadfield, 2010. L_i is the gene group length and B_i is the bacterial read count of observation i (which are set to 1 if the observation is from the clinical dataset). N_i is the number of isolates tested for observation i of the clinical dataset ($j \geq 2$).

The residual term (ϵ_i) for this model is specified with two components representing the residual variance due to country and resistance measure (α), and the residual variance due to overdispersion (η_i):

$$\epsilon_i = \alpha_{c,m} + \eta_i$$

The residual structure of the model is based on the model structure described in Hadfield, 2010. In addition to the variance expected due to the sampling distribution (e.g., in the Poisson model, the expected variance of the residuals would be e^{λ_i}), we also estimate residual overdispersion (σ_j^2).

$$\eta_i \sim Normal\left(0, \sigma_j^2\right)$$

The model structure has hierarchical residual errors ($\alpha_{c,m}$) for each country ($c \in (1, \dots, C)$) and each resistance measure ($m \in (1, \dots, M)$). These errors have the hyperprior:

$$\boldsymbol{\alpha_c} \sim MVNormal(0, \Sigma_m)$$

Here $\boldsymbol{\alpha_c}$ is a vector of size M of the deviations of country c for each resistance measure m. Σ_m is the covariance matrix of country-level deviations from the linear predictor (residuals) for each resistance measure. This hyperprior therefore allows group-specific variance, or partitioned variance, and between-group correlation. We chose to specify variance not just for the two datasets (ECDC and sewage surveillance), but also across different bacterial species within the ECDC data, considering each as a separate resistance measure. For example, vancomycin resistance measures include probability of resistance in clinical isolates of *E. faecium*, probability of resistance in *E. faecalis*, and vancomycin resistance gene abundance in sewage samples ($M = 3$). We can therefore estimate country-level correlations in resistance rates between bacterial species, and correlations between bacterial resistance rates and sewage resistance gene abundance.

The covariance matrix is the key element for addressing the aim of this chapter. It is composed of σ_m^2, the between-country variance in measure m, and the between-country covariances between the measures, e.g. $\sigma_{m=1,m=2}$,. After converting this covariance into correlation, we compare it to the Spearman's Rank correlation coefficient obtained previously. Both measure country-level correlations between resistance measures. Two important differences are a) the unit of observation is changed, and b) in the multi-response model, other sources of variance in the data are also accounted for.

A separate model was fit for each antibiotic group, for a total of 7 models (aminoglycosides, aminopenicillins, carbapenems, fluoroquinolones, macrolides, third generation cephalosporins, and vancomycin). In each model, only susceptibility test results, genes conferring resistance to, and usage rates of the antibiotic group of interest were used. Table 4.1 summarises the numbers of observations and groups in each model.

Table 4.1: Numbers of groups and observations in each model

Antibiotic group	Number of gene groups	Number of countries with both sewage and ECDC data	Number of bacterial species	Total number of clinical observations	Total number of gene observations
Aminoglycosides	71	26	6	1583	8236
Aminopenicillins	48	26	3	969	5568
Carbapenems	21	22	4	841	2121
Fluoroquinolones	16	26	4	1013	1856
Macrolides	43	25	1	259	4988
Third generation cephalosporins	30	26	3	857	3330
Vancomycin	10	26	2	632	1110

Uninformative priors were used throughout (Appendix C Table 3). The models were scripted using R v3.6 (R Core Team, 2022) and the MCMCglmm package (Hadfield, 2010), with an adaptation to enable combining binomial and Poisson sampling distributions (courtesy of Jarrod Hadfield). Example code is provided in Appendix C.

4.3.2.5 Other variables included in wastewater model

Primary care antibiotic usage rate was included as a fixed effect. The 90% homology gene group was included as a random effect to capture between-gene group variation in abundance, and to deal with the varying number of observations per gene group. Sampling batch was included as a random effect as another potential source of variation. Whether or not a sample was frozen to -70^{o}C (or lower) was included as a fixed effect, as this was stipulated in the GSP protocol, and because flash freezing is considered the gold standard for preserving microbiological samples (Song et al., 2016).

Other variables in the wastewater portion of the model were selected to take environmental factors into account. We included the flow rate category and the pH of the sample as factors in the model. The impact of environmental factors on sewage resistomes will be further explored in Chapter 5. Although flow rate was recorded as an ordered categorical variable, we included it with a single slope coefficient for model simplicity.

4.3.2.6 Other variables included in clinical model

Two fixed effects were included in the clinical portion of the model. Hospital antibiotic usage rate was included as an interaction effect, with a separate coefficient for each bacterial species. In each drug-bug combination, there was auto-correlation between time points, non-linear time trends, so we also included the resistance proportion in the previous year as a fixed interaction effect, with a separate coefficient for each bacterial species.

4.3.2.7 Comparison models

Uni-response models were also constructed to indicate if there was any improvement to model fit by incorporating the correlation terms. For the sewage data model, this meant that there was no variance partitioning for between-country variance, only a random intercept effect for each country. In the model of the ECDC data, there was still variance partitioning in the between-country variance, but only for bacterial species. In addition, we included models without primary care antibiotic usage to investigate the impacts of this factor on resistance gene abundance and correlation estimates.

4.3.2.8 Model likelihood comparison

For model goodness-of-fit comparison, we used a likelihood ratio test. First we calculated the likelihood ratio, $ll_0 - ll_1$, where ll_0 is the log-likelihood of the reduced model and ll_1 is the log-likelihood of the full model. We calculated the model log-likelihood from the MCMC chain iteration with the lowest reported model deviance for multi-response and uni-response models. The sum of the log likelihood of both uni-response models was used as ll_0. The likelihood ratio test is then performed using the χ^2 distribution with the number of free parameters added as the degrees of freedom. For the model structure described here we used m, the number of bacterial species and therefore the number of correlation terms that were added. Therefore, we are testing whether adding correlation terms between the datasets improved the model likelihood.

4.3.2.9 Model validation

Model validation was performed to assess the ability of the model to predict clinical resistance levels to countries not included in the training dataset. For this, we fit the model on training sets of the data with each country excluded, and then predicted clinical results of the excluded country with the resulting fitted model. This leave-one-out type of validation is used in assessing the predictive ability of a similar model in Karkman et al, 2020. We then compared the error on predictions for different countries, and the error for different drug-bug combinations for the multi-response model and the univariate clinical model.

4.4 Results

4.4.1 Data

4.4.1.1 Metagenomic data

A full description of the Global Sewage Surveillance Project data can be found elsewhere (Hendriksen et al., 2019). Briefly, for samples and genes included in this analysis, the average FPKM of genes conferring resistance to each antibiotic group was: macrolide resistance genes (RGs), 1177.1, standard deviation (SD) 4840.9; aminopenicillin RGs, 138.7, SD 857.9; fluoroquinolone RGs, 108.9, SD 476.9; carbapenem RGs, 58.2, SD 223.4; 3rd generation cephalosporin RGs, 51.1, SD 215.5; aminoglycosides RGs, 35.8, SD 176.5; and vancomycin RGs, 6.6, SD 53.9. The five

most abundant resistance genes were: *erm(B)*, *mph(E), mef(A)*, *msr(D)*, and *erm(F)* (average FPKMs 17032.5, 15674.9, 5909.3, 5182.0, and 2689.3, respectively).

4.4.1.2 Sample characteristics

Of 126 wastewater samples included in the analysis, 106 (84.1%) were frozen to -70^{o}C or lower prior to shipping. Most samples were rated as having the highest flow rate category (76/126, 60.3%). The average pH of samples included in this analysis was 7.30 (standard deviation = 0.51).

Table 4.2: Effect of primary care antibiotic usage on ARG abundance in sewage

Rate ratio estimate is the posterior mode, and uncertainty intervals are highest posterior density intervals.

Antibiotic group	**Rate ratio (95% uncertainty intervals)**
Aminoglycosides	1.04 (0.90 – 1.18)
Aminopenicillins	1.19 (0.92 – 1.43)
Carbapenems	1.11 (0.98 – 1.40)
Fluoroquinolones	1.38 (1.05 – 1.94)
Macrolides	1.16 (0.96 – 1.42)
Third generation cephalosporins	1.36 (1.06 – 1.76)
Vancomycin	0.34 (0.07 – 1.62)

4.4.2 Model results

We judged the strength of evidence for an association in the results below on the basis of the uncertainty intervals of the rate ratio posterior; if 1 is not within the 95% uncertainty, we take this as evidence for an association. If 1 is within the range but close to the edge, we take this as borderline evidence.

4.4.2.1 Effect of primary care antibiotic usage rates on ARG abundance on sewage

There was evidence for a positive effect of 3rd generation cephalosporins and fluoroquinolone usage rates in the community on ARG abundance (Table 4.2). There was borderline evidence for positive effects of macrolide, carbapenem and aminopenicillin usage, and no evidence for an effect for aminoglycosides or vancomycin usage.

4.4.2.2 Effect of hospital antibiotic usage on resistance prevalence in clinical isolates

Positive associations between hospital antibiotic usage rates and clinical resistance levels were observed for the following drug-bug combinations: aminoglycosides and *K. pneumoniae*; aminopenicillins and *E. faecalis*; 3rd generation cephalosporins and *Pseudomonas aeruginosa*; fluoroquinolones and *Pseudomonas aeruginosa*; and carbapenems and *Acinetobacter spp.*, *E. coli*, and *K. pneumoniae* (Appendix C Table 3).

4.4.2.3 Effects of sample characteristics on ARGs

The flow category assigned to the stream of wastewater on collection generally did not correlate with ARGs, except for a negative association with carbapenem resistance genes (Table 4.3). The sample being frozen to -70°C or lower had no effect on the abundance of ARGs. The pH of the wastewater sample had a borderline positive association with resistance genes for aminopenicillin, carbapenems, and third generation cephalosporins.

Table 4.3: Effect of environmental factors on resistance gene abundance in sewage

Rate ratio estimate is the posterior mode, and uncertainty intervals are highest posterior density intervals.

Resistance group	Sample flow rate category	Sample flash frozen	Sample pH
Aminoglycosides	0.97 (0.90 – 1.05)	1.06 (0.80 – 1.36)	0.98 (0.88 – 1.11)
Aminopenicillins	1.02 (0.94 - 1.09)	1.08 (0.86 - 1.36)	1.09 (0.97 - 1.19)
Carbapenems	0.88 (0.78 - 0.99)	1.06 (0.68 - 1.63)	1.14 (0.97 - 1.36)
Fluoroquinolones	1.09 (0.98 - 1.18)	1.05 (0.81 - 1.46)	1.00 (0.88 - 1.15)
Macrolides	1.00 (0.94 - 1.05)	1.02 (0.89 - 1.25)	1.01 (0.93 - 1.08)
Third generation cephalosporins	1.01 (0.90 - 1.12)	1.08 (0.71 - 1.41)	1.14 (0.97 - 1.32)
Vancomycin	0.85 (0.49 - 1.54)	0.92 (0.28 - 4.35)	0.69 (0.26 - 1.37)

4.4.2.4 Effect of previous time point on clinical isolate resistance levels

There was a positive association between resistance prevalence in clinical isolates of a bacterial species at consecutive time points for all drug-bug combinations except for carbapenem resistance in *E. coli* and vancomycin resistance in *Enterococcus faecalis* (Appendix 2.2).

4.4.2.5 Correlations between resistance measures.

The relationship between sewage and clinical resistance measures, and between bacterial species resistance measures, varied by antibiotic group, bacterial species, correlation method and model structure (Fig. 4.1, Appendix C Fig. 3). Summaries of the posteriors for all resistance measure covariance matrices can be found in Appendix C Fig. 3.

Aminoglycoside resistance in hospitals and sewage appeared to positively correlate on the country-level (Fig. 4.1, Fig. 4.2, Appendix C Fig. 3.1). Correlation was particularly strong for *Acinetobacter spp.*, *K. pneumoniae* and *P. aeruginosa*. The correlation estimates were unaffected by removing primary care antibiotic usage and were higher using the multi-response model method compared to the Spearman's Rank correlation (Appendix C Fig. 3.1.B).

Aminopenicillin resistance measures did not appear to have strong positive correlations (Fig. 4.1, Appendix C Figure 3.2). Model results indicated that there was

no relationship for *E. faecalis* or *E. coli*, and potentially a negative one for *E. faecium*. This contrasts to the Spearman's Rank results, which suggested slightly positive relationships for *E. faecium* and *E. coli.* Removing primary care antibiotic usage did not affect correlation estimates.

Evidence for a positive corelation for carbapenem resistance was mixed (Fig 1, Appendix C 3.3). Model results indicated that there was no positive association for any species. Spearman's Rank, however, suggested positive correlations for *Acinetobacter spp*. and *P. aeruginosa*. Removing primary care antibiotic usage made positive correlation in the multi-response model more likely (Appendix C Fig. 3.3.B).

There was also mixed evidence for a positive correlation in Fluoroquinolone resistance (Fig. 4.1, Appendix C Fig. 3.4). Model results indicated no positive correlation even though Spearman's Rank did (Fig. 4.1). Removing primary care antibiotic usage led to strong positive correlation values, indicating the potential importance of this variable for fluoroquinolone resistance (Appendix C Fig. 3.4.B).

Macrolide resistance appeared positively correlated (Fig. 4.1, Appendix C Fig. 3.5). Multi-response models with and without primary care antibiotic usage, and the Spearman's Rank, agreed on a positive correlation (Appendix C Fig. 3.5.B), although there was wide uncertainty in the multi-response estimates.

There were wide uncertainty intervals for third generation cephalosporin resistance correlation estimates correlate (Fig. 4.1, Appendix C Fig. 3.6). Correlation estimates did not agree across model structures and Spearman's Rank, although there was some evidence of positive correlation for *E. coli* (Fig. 4.1).

Vancomycin resistance contrasts to other resistances types examined, appearing to have no correlation or even possibly negative correlation (Fig. 4.1). Results of multi-response models and Spearman's Rank suggested correlations below 0 but not significantly so (Appendix C Fig. 3.7.B).

Comparing the correlation between bacterial species within the clinical data, it is notable fluoroquinolone resistance had positive covariance for all bacterial species combinations, but none were positive for aminopenicillins (Appendix C Fig. 2.3). Resistance in *K. pneumoniae* and *P. aeruginosa* were most often positively correlated with other bacterial species, each having a positive covariance in 8/13 between species comparisons. However, uncertainty intervals were wide for correlation estimates between bacterial species.

4.4.2.6 Other sources of variance

Further sources of variance are provided in Appendix C Table 4.

4.4.2.7 Model likelihood comparisons

We compared the likelihood of models with and without a between-dataset correlation term (Table 4.4). All models except for macrolide and third generation cephalosporin resistance had improved likelihood after adding between-dataset correlation.

Table 4.4: Likelihood ratio tests of multi-response compared to univariate models

Log-likelihood obtained from the parameter set with the lowest deviance. *P*-value indicates result of a likelihood ratio test, probability of achieving the observed log-likelihood ratio if the two models had the same log-likelihood.

Antibiotic group	**No. variance parameters added (degrees of freedom)**	**Log-likelihood ratio**	***p*-value**
Aminoglycosides	6	19.83	<0.01
Aminopenicillins	3	2.09	<0.01
Carbapenems	4	10.21	<0.01
Fluoroquinolones	4	15.36	<0.01
Macrolides	1	-14.16	0.99
Third generation cephalosporins	3	-6.19	0.99
Vancomycin	2	6.23	0.00

4.4.2.8 Model validation

The greatest model error was for predicting carbapenem resistance in *E. coli* for countries excluded in the training set, for which the average error was 574 test results incorrectly predicted, making predictions 18 percentage points away from the observed prevalence of resistance in countries outside of the test dataset (Fig. 4.3). Generally, errors for the univariate model and the multi-response model were similar, although the multi-response model had modestly lower errors for aminopenicillins and vancomycin resistance in all species. There were also some increases in error in multi-response models compared to univariate ones, particularly for carbapenem resistance.

The greatest range of errors for each country (Appendix C Fig. 4) was for fluoroquinolone and third generation cephalosporins models, whereas errors were similar across the board for the vancomycin model. Generally Romania or Malta, for which there was the least data, were predicted worst. On average the worst predicted country across all test sets of data was Romania (13.6 percentage points away from true percent of isolates resistant), and the best predicted was the Republic of Ireland (6.5 percentage points away).

Figure 4.1: Comparison of sewage-clinical resistance correlation estimates for Spearman's Rank and multi-response models

For each drug-bug combination, the correlation estimate from a Spearman's Rank test (left half of circle) and the correlation posterior mode from the multi-response model (right half of circle). The correlation value is printed as text, and the colour reflects the degree of correlation, with blue for negative correlation, white for no correlation, and red for positive correlation. The star indicates that the uncertainty intervals (95% confidence intervals for Spearman's Rank and 95% highest posterior density for multi-response models) did not include 1, and therefore provide some evidence for an association.

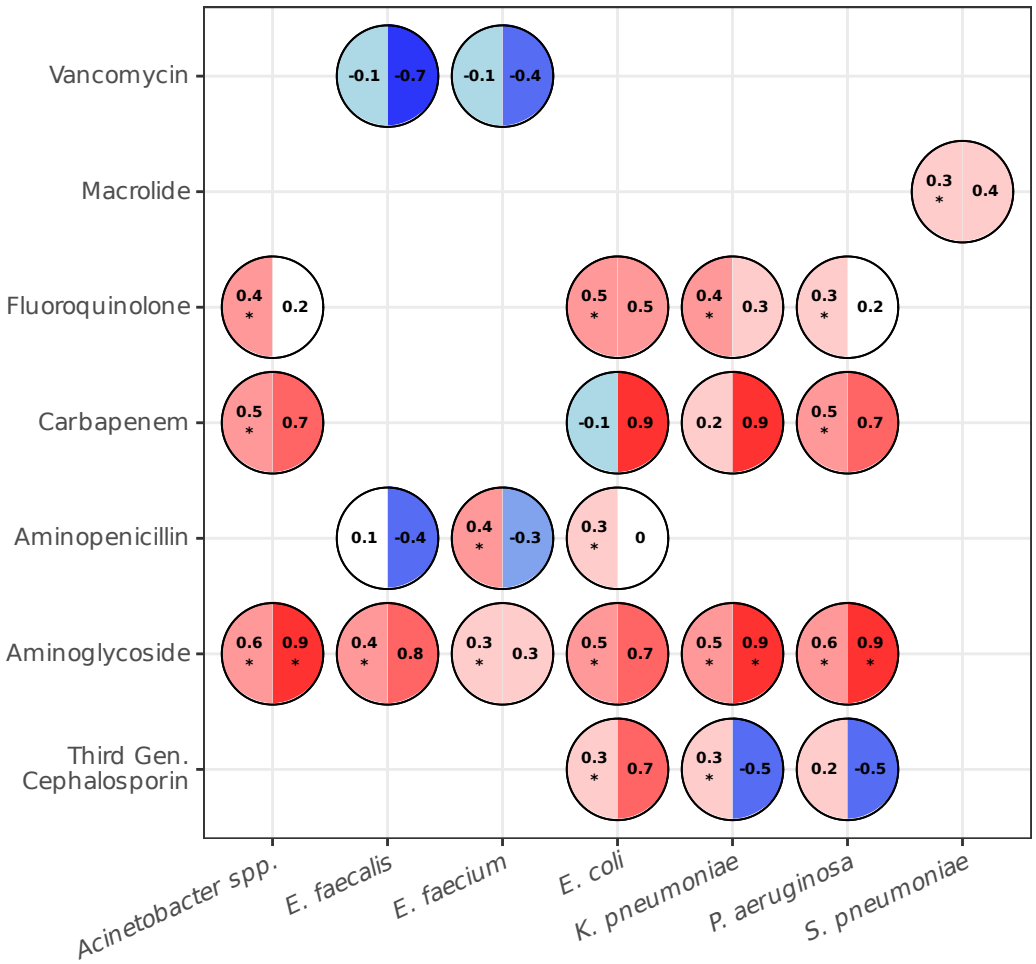


Figure 4.2: Resistance to aminoglycosides in clinical and sewage data

On the left, the average proportion of isolates resistant to aminoglycosides (gentamicin or tobramycin) in 2018 in European countries. On the right, the average FPKM of resistance genes conferring resistance to gentamicin or tobramycin in European countries in 2018. The colour of the country area indicates the resistance level, or grey indicates no data was collected in GSSP or EARS-Net.

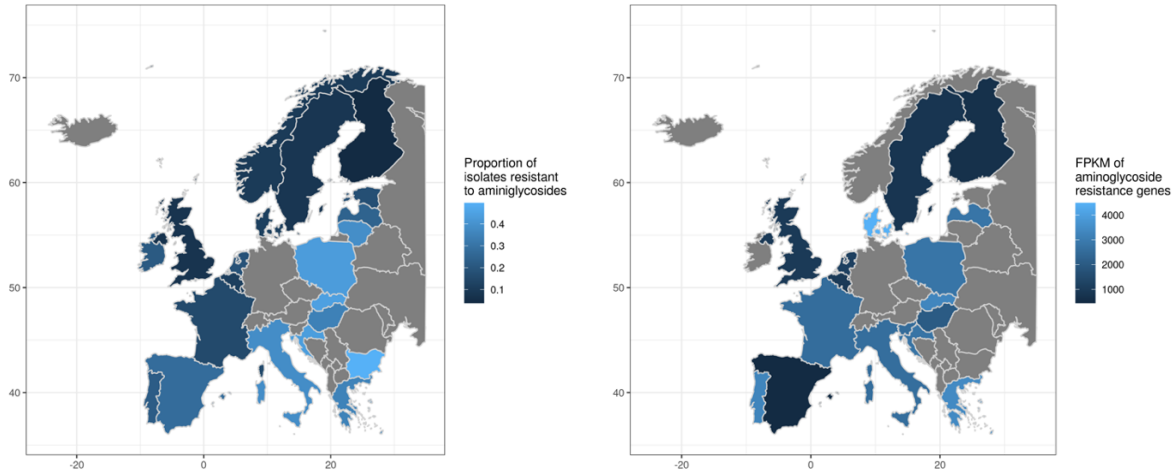
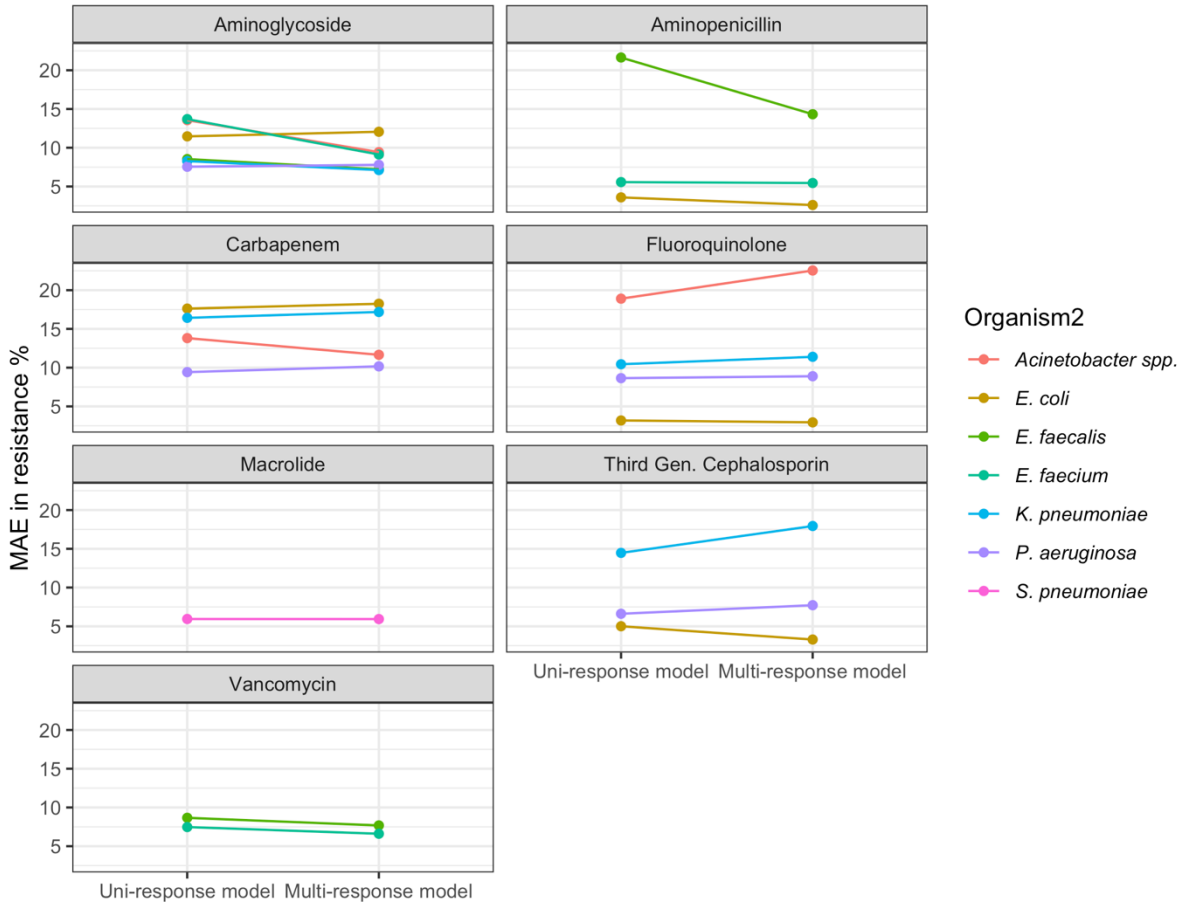


Figure 4.3: Predictive ability of the multi-response and univariate models to new countries

Mean absolute error in the predicted proportion of isolate results for countries not included in test sets. Multi-response models including all the data and univariate models of only clinical data were run on test sets with each country excluded.



4.5 Discussion

In this chapter, we applied a multi-response model to integrate clinical and sewage resistance surveillance data. There was evidence for positive correlations between clinical and sewage resistance levels for aminoglycoside resistant bacteria, but no strong evidence for correlations for other antibiotic groups. However, correlation estimates were uncertain and sensitive to model structure and correlation estimation method. The relationship between community antibiotic usage and sewage resistance level varied by antibiotic, with strongest evidence for a positive effect for fluoroquinolones and third generation cephalosporins. We suggest that these results demonstrate that sewage surveillance can be used for exploring drivers of resistance in the community as well as in clinical settings, but that some drug-bug combinations (especially aminoglycoside resistance) may be better suited to sewage surveillance than others (macrolide or vancomycin resistance). More research that disentangles hospital and community contributions to WWTP influent is needed interpretation of sewage surveillance data.

The most robust evidence for a positive correlation between clinical and sewage surveillance data was for resistance to aminoglycosides, which was consistently found across model structures. A country-level positive correlation between EARS-Net data and resistance gene abundance in sewage was also found in the two other studies, further suggesting this finding is robust (Karkman et al., 2020; Pärnänen et al., 2019) (see Table 4.5 for a comparison of the methods and results of these studies). This positive correlation could indicate that aminoglycoside resistance genes that are abundant in hospital patients and sewage are carried to the WWTPs, i.e. the correlation is due to the signature of the hospital sewage in the WWTPs. Current evidence suggests that the contribution of hospital effluent to the resistome of WWTPs is present but small (Buelow et al., 2018; Gundogdu et al., 2017; Hutinel et al., 2019; Kutilova et al., 2021; Paulshus et al., 2019; Verburg et al., 2019), although aminoglycoside resistance has not been studied separately from other resistance classes in studies comparing hospital and municipal wastewater. The proportion of municipal sewage that has come from hospital effluent by volume is small (estimated to be about 1% based on flow rates of hospital and total municipal effluent) so it is not surprising that the signal would be diluted (Buelow et al., 2018; Verburg et al., 2019). In addition, bacteria and resistance genes may degrade within the wastewater network, and in some European countries hospital wastewater is also treated prior to discharge into the main sewer line (Kumari et al., 2020). These features could further dilute and alter the composition of the hospital effluent-derived resistance genes in WWTPs but are not well understood.

Positive correlations may alternatively indicate that there is sharing of aminoglycoside resistant bacteria between hospitals and the community. The sharing of resistance may be due to hospital-acquired infections spreading within the community, or community-acquired infections being detected in hospitals, or both. Whether hospitals or the communities are the main sources of resistant infection is an ongoing area of research, with a recent modelling paper suggesting that most infections are likely to come from the community (Knight et al, 2018b). In the case of aminoglycoside

resistance, there is evidence that it can be acquired in hospitals (El-Mahdy et al., 2018) and in the community (Roldan-Masedo et al., 2019), but data indicating the proportion of aminoglycoside resistant infections that are community or hospital acquired are lacking.

Table 4.5: Comparison of other EARS-Net studies and results by antibiotic and bacterial species

The data, methods, and whether evidence was claimed for a positive association. NT: not tested.

Data and methods comparisons		
	Pärnänen et al., 2019	**Karkman et al., 2020**
Clinical data source	EARS-Net	EARS-Net; Central Asian and Eastern European Surveillance of Antibiotic Resistance Annual Report, 2018 (CAESAR)
Sewage resistance gene data source	qPCR of 10 European countries	Global Sewage Surveillance Project (2016 data only)
Correlation method	Spearman's Rank	Generalised linear regression model (beta family)
Evidence for positive associations		
Aminoglycosides	Yes: *E. coli*; *K. pneumoniae*; *P. aeruginosa* No: *Acinetobacter* spp., *E. faecium*	Yes: *E. coli*
Aminopenicillins	NT	Yes: *E. coli*
Carbapenems	NT	NT
Fluoroquinolones	Yes: *E. coli*, *K. pneumoniae*, *P. aeruginosa* No: *Acinetobacter* spp.	Yes: *E. coli*
Macrolides	No: *S. pneumoniae*	NT
Third generation cephalosporins	NT	Yes: *E. coli*
Vancomycin	No: *E. faecium*	NT
Beta-lactams	Yes: *E. coli*, *K. pneumoniae*, *S. aureus* No: *Acinetobacter* spp., *P. aeruginosa*, *E. faecium*	NT

Another explanation for positive correlations is that there are similar drivers of resistance in both hospital and community sewage. For example, community- and hospital-acquisition of aminoglycoside resistance could both be independently elevated due to, for example, correlated hospital and community antibiotic usage. To assess this we compared the correlation estimates of models with and without primary care antibiotic usage rates included, and compared the resulting correlation estimate

posterior, as the correlation might be expected to decrease if this was the case. Aminoglycoside correlation estimates appeared to be insensitive to primary care antibiotic usage, suggesting this explanation may not play a strong role in this finding. Of the explanations of positive correlations considered (hospital effluent signatures, hospital-community resistome sharing, or correlated antibiotic usage), it is not possible to quantify the contribution of each to the correlations estimates from the available data. Further research that samples from multiple hospital and community sites is needed to disentangle hospital and community signatures from municipal wastewater resistomes.

For aminopenicillin, fluoroquinolone, and third generation cephalosporin resistance, Spearman's Rank and previous studies (Huijbers et al., 2020; Karkman et al., 2020; Pärnänen et al., 2019) found positive correlations, whereas the full multi-response model did not. The simpler methods that only include one year of data or can match observations on year and country may be able to find higher correlation estimates, although resistance levels were relatively stable over the 2016 – 2018 period in both datasets (Appendix C Figure 3). More explanatory factors and hierarchies in the data were incorporated into the multi-response model, and these may have led to the between-dataset correlations explaining a smaller proportion of variance in either dataset. For example, including primary care antibiotic usage decreased the correlation value in the fluoroquinolone model. This factor was not included in previous studies. Overall, the differences in the correlation estimates point to the sensitivity of the results to the modelling framework. Comparison of the results of model structures (Spearman's Rank, multi-response and uni-response, with and without antibiotic usage) provides more information on how best to combine these two datasets and adds robustness to the results. Studies which do not take this approach risk over- or underestimating correlation values.

In macrolide and vancomycin resistance models, both this study and a previous study agree (Pärnänen et al., 2019) on there being no or a small degree of correlation. There was also no evidence from this study for a correlation in carbapenem resistance, but we do not find any studies estimating a country-level correlation of carbapenem resistance levels in clinical and sewage samples, although two studies on a single hospital level found some evidence of a positive association in hospital isolates and effluent (Flach et al., 2021; Perry et al., 2021). Arguably the strongest evidence for no association is for vancomycin resistance, as this result was most consistent between correlation methods. It is notable that hospital usage of vancomycin is increasing but community usage is rare (Appendix C Fig. 2), and it was a rare type of resistance in the sewage (average 6.6 FPKM) and in clinical data, although the proportion of *E. faecium* isolates that are resistant to vancomycin is increasing (Appendix C Fig. 2). Therefore, vancomycin resistance gene abundance may be particularly hard to detect in wastewater due to dilution. This dilution would lead to any clinical factors being hard to detect in sewage samples. Low-sensitivity methods like sewage metagenomics may not be suited to monitoring these types of resistance.

There are several factors that may decrease the strength of the link between resistance clinical samples and municipal sewage. A key question to answer is how

closely related the microbiomes of blood and cerebrospinal fluid (represented in EARS-Net) to gut and urinary tract (represented in sewage surveillance) are. It has been observed that different sites in the human body have distinct microbiomes (Lloyd-Price et al., 2017), and that the resistome of oral and stool samples differ (Carr et al., 2020), but studies that compare the resistomes of blood to other different body sites are lacking. Differences in the resistome between these sites could lead to lower correlations between clinical and sewage surveillance. In other words, the clinical data source may explain the low correlation. A lack of correlation may also point to a lack of onward transmission of hospital-acquired resistant bacteria in the community. This finding would be of some clinical importance, indicating that antibiotic stewardship and infection prevention control measures would have greater knock-on positive effects for some resistant bacteria than others. On the other hand, it may indicate that the EARS-Net data does not represent the community resistance levels well. Studies using EARS-Net data such as (Cassini et al., 2019) assume that their results can be generalised to the whole population, but few studies have compared community samples with EARS-Net data. In addition, the blood-gut and hospital-community relatedness in resistant bacteria differ between bacterial species. Sewage surveillance will be more effective for clinical predictions for those bacterial species that have higher relatedness to patient blood specimens.

Positive correlations may be obscured in the dataset. Wastewater is a new source of surveillance, and the impact of non-human sources of resistance in the sewage and wastewater passage on the composition of the sewage resistomes should be considered. Wastewater influent at the water treatment plant is a mixture of wastewater from multiple sources, and there has been speculation that animal gut (Raven et al., 2019) and industrial effluents (Guardabassi et al., 1998) also contribute to the resistome. In addition, features within the sewer pipes of the wastewater network may affect the microbial composition. Some features may encourage the growth of some bacteria and the decomposition of others, including warm temperatures and nutrients, exposure to antibiotic residues, heavy metals and disinfectants, and fluctuating concentrations of oxygen (McLellan & Roguet, 2019). This will alter the composition of the microbiome and the resistome within the sewage compared to human microbiomes. Indeed, studies have found that the wastewater is dominated by resident bacteria such as *Acinetobacter* and *Aeromonas* (Guo et al., 2019; Vandewalle et al., 2012). There is also evidence that exposure to antibiotic residues in hospital wastewater could select for resistance in *E. coli* (Kraupner et al., 2021). Accordingly, it has been shown that the wastewater microbiome and resistome is related to, but distinct from, human samples (Newton et al., 2015; Pehrsson et al., 2016; Raven et al., 2019).

Differences in the microbiology of different sites in the body and wastewater will reduce the similarity of the municipal sewage resistome to that of humans, and will add noise to the data. Analysis of sewage data may therefore be at risk of underestimating drivers of resistance, or overestimating the association with human risk factors that correlate with environmental variables, such as latitude and GDP. However, in this study few environmental factors were associated with resistance gene abundance, which could indicate a limited influence. Experimental research that investigates the

impacts of wastewater passage is crucial to confirm which environmental factors need to be considered. Multi-response models will be in important tool for building flexible model structures that estimate the impact of many sources of variance and the impacts of co-factors and environmental factors on sewage resistance.

Primary-care antibiotic usage was positively associated with resistance gene abundance in fluoroquinolone and third generation cephalosporin models, suggesting the usage of these antibiotics in the community could drive selection for resistant infections. This result contrasts the previous analysis of the Global Sewage Surveillance Project, which found no association between resistance gene abundance and antibiotic usage on a global scale (Hendriksen et al., 2019). Positive associations between antibiotic consumption and EARS-Net data have been found in some (McDonnell et al., 2017) but not all (Collignon et al., 2018) studies. On a smaller scale, there is some evidence for community prescriptions influencing resistance selection from studies of non-hospital use of fluoroquinolones and co-amoxiclav and community-acquired UTIs (Kahlmeter et al., 2003; Vihta et al., 2018). In this analysis, we split models by antibiotic group, to allow for differences in the relationship between consumption and resistance. This may have improved our ability to detect positive effect estimates. Finding an association between sewage metagenomics and a community, population-level factor is promising for further research that looks for community drivers of resistance using wastewater surveillance.

Some additional limitations of sewage metagenomics should be considered in interpretation of our results. Metagenomics has low sensitivity which may mask rare resistance genes, such as ones that are usually hospital-associated. Resistance gene data from metagenomics does not indicate the bacterial host of the resistance gene, which limits comparison to culture-based clinical surveillance data, although research is ongoing to infer the host which can mitigate this limitation of metagenomics in future studies (e.g. Munk-Welford et al, *in prep*.). Sampling was focussed on large cities which may not represent the rest of the country well, and there were more sampling sites and time points for some countries than others. EARS-Net also doesn't have full population coverage, with contributing countries reported an average of 67.8% population coverage, and 70% of countries reported high degree of geographical representativeness in 2019 (European Centre for Disease Prevention and Control, 2020). The degree of population overlap between these two datasets is unknown, but if overlap is low this could decrease correlation estimates. Finally, as both sewage surveillance and EARS-Net are population-level measures, this is an ecological study. Applying analysis of ecological data to draw conclusions on risk factors for individuals is often inappropriate (Grimes & Schulz, 2002).

In conclusion, this study provides evidence for a link between aminoglycoside resistance in hospitals and communities. It also demonstrates that multi-response models are powerful and flexible tools for wastewater metagenomic analysis, allowing more efficient use of the data, and more investigation of influential factors, hierarchies, and sampling distributions in both datasets than uni-response methods. We show that sewage surveillance data will be useful for supporting clinical predictions for some drug-bug combinations (such as aminoglycoside and fluoroquinolone resistance), but

not all (such as macrolide and vancomycin resistance). Sewage surveillance can shed light on the community drivers of resistance, indicating that primary-case antibiotic usage may lead to increased resistance in the community.

4.6 Acknowledgements

I would like to acknowledge Jarrod Hadfield for his help formulating the multi-response model and for adapting the MCMCglmm package to accommodate it.

5 A cross-sectional metagenomics study of 8 wastewater networks in Scotland: hospital and community resistomes are distinct

5.1 Abstract

Background: Metagenomic analysis of sewage generates rich information on the human resistome and microbiome, and wastewater surveillance programmes are becoming more widespread. However, sewage samples can combine effluents from many sources, including communities and hospitals. Little is known about the contribution of each source, limiting the interpretation of sewage metagenomics. Here, I aimed to assess the contribution of communities and hospitals to sewage samples, and whether resistomes in patients and community members differ. We applied metagenomics to sewage samples from eight wastewater networks in Scotland, including hospitals, communities, and their connected wastewater treatment plants (WWTPs).

Methods: I selected four WWTPs with a district hospital in their catchment area and matched them to WWTPs without hospitals in their catchment area, taking population size, rural-urban composition, and geographical distance from each other into account. Sewage was sampled from all WWTPs, district hospitals, and community sites (in non-hospital catchment areas). For each catchment I collected community and hospital antibiotic prescription data. Samples were sequenced using Illumina NovaSeq and reads were mapped using the ResFinder database. I used multi-level zero-inflated negative binomial regression models, cluster analysis, and source-attribution random forests to compare resistance in hospitals, communities, and WWTPs, and estimate resistance-prescription associations.

Results: PCoA analysis showed two fully separate resistome groups, with hospital resistome in one and communities and WWTPs in the second. Hospitals had the greatest abundance (average fragments per kilobase million 1461.2, 122.4, and 141.1 for hospitals, communities, and WWTPs, respectively) and richness (average 197.4, 111.6, and 119.7 different genes, respectively) of resistance genes. Even so, the contribution of hospitals to WWTP resistomes was small: resistance gene abundance in WWTPs was not increased by a significant proportion by hospital presence (effect estimate: 0.84, 95% uncertainty intervals: -0.24 – 1.91), and random forest models predicted a low chance that the source of WWTP resistomes was a hospital (5.6% without and 13.9% with hospitals in their catchments). I found no evidence of an impact of prescriptions on sewage resistance abundance.

Conclusion: This chapter suggests WWTP resistome represents effluents from communities more closely than hospitals. Furthermore, it suggests hospital and community/WWTP sewage resistomes are distinct, implying there are different drivers of selection or transmission of antibiotic resistance (ABR) in these groups. Sampling from hospitals effluent will be needed to determine the hospital resistome, as it is not captured in WWTP influent.

5.2 Introduction

Antibiotic resistant infections are thought to be a leading cause of mortality globally (Murray et al., 2022), and are growing in number (Cassini et al., 2019). Although health-care associated resistant infections are the main cause of morbidity and mortality, colonisations of resistant bacteria in the general population are thought to be a major reservoir of invasive infections (Hendriksen et al., 2019). Quantifying hospital and community resistance patterns can inform when and where resistant infections are acquired, which is essential information in effective intervention design (Knight et al, 2018a). However, characterising community carriage is hindered by a lack of data.

Wastewater surveillance and metagenomics has been proposed as a means of monitoring resistance patterns in the general population (Aarestrup & Woolhouse, 2020; Hendriksen et al., 2019). Its advantages over traditional, clinical isolate surveillance methods include convenience of collection and greater representation of the healthy general community microbiomes (Aarestrup & Woolhouse, 2020; Miłobedzka et al., 2022). When paired with metagenomics, it further offers the benefit of being agnostic, providing information on a wide variety of resistance genes and bacterial species (Aarestrup & Woolhouse, 2020; Miłobedzka et al., 2022).

However, wastewater is a composite sample source. It has usually been collected from raw influent to wastewater treatment plants (WWTPs) (Hendriksen et al., 2019; Huijbers et al., 2020; Hutinel et al., 2019; Kwak et al., 2015; Pärnänen et al., 2019; Raven et al., 2019; D. Zhang et al., 2021). In Scotland at least, WWTP influent represents household and office building sewage as well as effluents from schools, hospitals, laboratories, cleaning industries, and surface water (Scottish Water, n.d. a). Despite mixing these sources, WWTP influent has been shown to be a good representation of the human microbiome (Newton et al., 2015; Pehrsson et al., 2016), and resistance levels in WWTPs and clinical samples have been shown to positively correlate (Huijbers et al., 2020; Karkman et al., 2020; Pärnänen et al., 2019). However, effluents coming directly from hospitals have been shown to have greater resistance abundance and diversity than WWTP and community influents (Hassoun-Kheir et al., 2020), and to reflect within-hospital clinical activity (Cai et al., 2022; Perry et al., 2021) and outbreaks (Flach et al., 2021). This raises a question: are resistance genes detected in WWTP influent likely to have come from a hospital rather than a community source? To interpret how wastewater represents resistance in the general population, the relative size of the contribution of hospital and community effluents to resistance in the composite wastewater needs to be quantified. To quantify the relative contributions of hospitals and communities, we need to know how different their resistance profiles are, and how similar WWTP resistance is to each source.

To address this question, previous studies have compared the resistance profile of hospital effluent and WWTP influent (Paulshus et al., 2019; Verburg et al., 2019), or WWTPs with and without hospitals in their catchment area (Buelow et al., 2018). Generally, these studies have concluded that the hospitals have only a small or no contribution to downstream resistance (Buelow et al., 2018; Paulshus et al., 2019;

Verburg et al., 2019). However, there few studies capturing more than one wastewater network, or collecting a sample of community-only wastewater for comparison. Lack of sample size and a comparison group limits the power of their analysis and the generalisability of their conclusions.

In this study, we collected wastewater samples from eight WWTPs, four hospitals, and four community sites in a paired sampling design. Our aim was to investigate how the resistomes of hospital and community sewage differ, and how this impacts WWTP resistomes. My objectives were to 1) compare the resistome and microbiome profiles of hospitals, communities, and WWTPs; 2) model the abundance of resistance genes in different site types and estimate the influence of antibiotic prescriptions and environmental factors; and 3) estimate the contribution of hospital effluent to the WWTP influent resistome. We applied metagenomics to each sample, obtaining a measure of the abundance of acquired resistance genes and bacterial species at each site. We also collected antibiotic prescription data for the hospitals and communities sampled. I used a combination of cluster, statistical, and source attribution methods to analyse the data. The combination of multiple study sites, community samples, metagenomics, paired sampling design, and robust analysis methods improve our power over previous studies to discern differences and drivers of the resistome in each sewage sample type.

5.3 Materials and methods

5.3.1 Sample site selection

I selected eight WWTPs for sampling. I aimed to select four WWTPs with a district hospital in their catchments, and to pair them to four further WWTPs that did not have a district hospital in their catchment but served similar population sizes, locations, and urban level. Table 5.1 describes the sites and their locations.

I also selected a sampling site from each of the four district hospitals in the catchment areas. There were several access points to the wastewater network in each hospital. In Border's General and St John's hospitals, there was an access point that represented the whole hospital so we sampled from this point. In Hairmyres and Victoria hospitals there was not one point for the whole hospital so I chose sites to a) sample the intensive care units, where many antibiotics are used, and b) sample a large a proportion of the hospital beds.

Finally, we selected a sampling site within the community for each WWTP without a hospital in the catchment area. Using water network maps of the drainage area, I selected a small area where the access points represented a collection of houses in a residential area that was the same straight-line distance from the WWTP as the hospital was from the matched WWTP. Sample collectors (from Scottish Water) then chose an access point from the small area on the basis of safety and convenience.

5.3.2 Sample collection

All samples were collected by Scottish Water. WWTP and hospital samples were taken using autosamplers, which continuously take small samples to make up two 1L total samples representing a 24-hour period. At community sites access points were smaller so the autosamplers could not be used. Instead, spot samples of 333ml were taken over 15 minutes (i.e., 3 samples at five-minute intervals) to make up two 1L samples.

Samples were taken from all sites within a pair over the same 24-hour period, i.e., autosamplers set up on the same day at both WWTPs and the hospital, and the community spot sample and collection of autosamples took place the next day. Samples were collected from the Kirkcaldy and Stirling catchment areas in March, 2020, and again in November, 2020. Galashiels, Hawick, East Calder, and Kinneil Kerse catchment areas were sampled in November, 2020. Philipshill and Allers catchments were sampled in December, 2020.

Table 5.1: Selected WWTP names, locations, and area descriptions

The population equivalent is an estimate of the population served, provided by Scottish Water. The distance between sites is measured as the straight-line distance between GPS co-ordinates.

Pair	WWTP name	Population equivalent	Distance between sites (m)	Upstream site	Area description
1	Galashiels	28,983	3785	Borders General Hospital	Town in rural area
	Hawick	17,642	3257	Hawick community	Town in rural area
2	Kirkcaldy	62,057	1376	Victoria Hospital	Small city
	Stirling	70,481	1876	Stirling community	Small city
3	East Calder	113,254	5021	St John's Hospital	Mix of town and industrial estate
	Kinneil Kerse	45,019	4926	Grangemouth community	Mix of town and industrial estate
4	Philipshill	61,244	2036	University Hospital Hairmyres	Large town
	Allers	61,024	2243	East Kilbride community	Large town

Metadata from the site was also recorded by sample collectors: the temperature and pH of the sample, the weather, time of sampling (spot sample) or sample collection

time (autosamplers), the colour of the water and the flow rate (in litres per second at WWTPs, and on a 1 – 5 scale at hospitals and community sites).

All samples were transported to the University of Edinburgh and frozen to -70°C on the same day they were sampled. Frozen samples were packed in polystyrene and shipped to the National Food Institute, Danish Technical University (DTU) after at least 48 hours at this temperature. Transport to DTU took 24 hours or less.

5.3.3 DNA extraction and sequencing

Samples were processed upon arrival at DTU, after thawing. The same procedure was used as for the Global Sewage Surveillance project (GSSP) (Hendriksen et al., 2019; Munk-Welford et al., 2022). Briefly, 250ml of sewage pellet was generated by centrifuging the samples. DNA was extracted from the samples using the QIamp Fast DNA Stool mini kit and an optimised protocol described in (Knudsen et al., 2016). DNA was then shipped on dry ice for sequencing. After shearing to 300bp target fragment size, PCR-free Kapa Hyper library preparation was used. Sequencing was done using Illumina NovaSeq and 150bp paired-end sequences.

5.3.4 Read trimming and mapping

Quality- and adapter-trimming followed the previously used protocol for the GSSP (Munk-Welford et al., 2022), using BBduk2 (Bushnell, 2014). Briefly, common adapters were removed and low-quality base pairs were trimmed from the 3' end of the reads, using a Phred score of Q20, or a 1% error rate. For resistance genes, reads were then mapped to ResFinder (Zankari et al., 2012) using KMA with default alignment parameters (Clausen et al., 2018). For taxonomic calls, Kraken2 (v. 2.0.7-beta) (Wood & Salzberg, 2014) was used with default setting against a custom database of representative bacterial, viral, protozoan and human sequences using MGmapper_makedb.pl (Petersen et al., 2017). Using a representative database allows us to identify a clade for sequence fragments from genomes that do not have a high quality genome in the reference databases (NIH, 2019). This is particularly important for microbiomes from environmental samples, where many bacterial species may not be known (Bengtsson-Palme et al., 2017).

5.3.5 Prescription data

5.3.5.1 Community prescription data

I obtained community prescription data in an information requestion from Public Health Scotland. Prescriptions of interest were those made by community sources (mainly general practices but also dentists and some other sources) to households with a post code sector within the catchment area of all WWTPs in the 6 months prior to sampling. I collected data on the monthly daily defined doses (DDDs) for every antibiotic under the Anatomical Therapeutic Chemical Classification (ATC) code J01. Non-zero values of less than five were categorised as 0 – 5 DDDs by NHS Lothian for data privacy. I estimated the population size of the catchment area by finding the Data Zones (defined by the Scottish Government) that overlapped with

the catchment and taking the sum of the 2019 population estimates for these Zones, which are based on extrapolation from the 2011 Scottish census (Lowe, 2020).

5.3.5.2 Hospital prescription data

Hospital prescription data was obtained from the pharmacists at each hospital via the digital drug dispensing recording programme, known as JAC. For hospitals, I acquired the monthly net number of doses of antibiotics (ATC code J01) dispensed from the hospital pharmacy in the 6 months prior to sampling. I calculated the DDDs using the preparation strength, number of doses given, and mode of administration from the hospital data, and the dose factor provided by the WHO ATC index (WHO Collaborating Centre for Drug Statistics Methodology, 2021). I normalised the DDDs by the estimated monthly number of occupied bed days (OBDs) for each hospital, obtained from Public Health Scotland (WHO Collaborating Centre for Drug Statistics Methodology, 2021). This provides the quarterly OBDs, from which an average monthly OBD was estimated.

5.3.6 Sequence data normalisation methods

For resistance gene abundance, fragments per kilobase per million reads (FPKM) (Chapter 4, page 56) was used. This measure normalises the number of read hits to a reference gene by the number of base pairs in the gene and the total number of bacterial fragments in a sample, to allow meaningful comparison between different genes and samples. For bacterial species abundance, the relative abundance was used, which is the sum of read hits for genes in a specific bacterial species divided by the total number of bacterial species reads in the sample.

5.3.7 Homology cluster grouping of resistance genes

I use the same method as in Chapter 4 to group resistance genes in the ResFinder database with >90% sequence similarity into clusters, taking the sum of read hits to all genes in each cluster. Normalisation was then conducted on the summed counts. Also, in the same way as in Chapter 4, I assume that the resistance phenotype of a cluster is the combination of the resistance phenotypes of all resistance genes within the cluster.

5.3.8 Resistome and microbiome cluster analysis

I used principal coordinate analysis (PCoA) to assess the distance between the resistomes and microbiomes of each sample. The Bray-Curtis distance was calculated for all sample pair combinations, using the FPKM to measure the abundance of each resistance gene and the relative abundance for bacterial species. Relative abundance and centre-log ratio transformation (CLR) are recommended measures for normalising bacterial species read counts from microbiome data (Gloor et al., 2017). The variances of site type group on these coordinates were also estimated. The strength of the influence of each resistance gene and bacterial species along the axis explaining the most variation for distinguishing hospital and non-hospital samples was obtained by

correlating abundance with the axis. This and all following analysis was done using R (4.1) (R Core Team, 2022).

5.3.9 Source attribution random forests

I used a source-attribution random forests approach to investigate whether the WWTPs resemble hospitals or communities more closely, following the methods from a recent study of gut metagenomes in humans and livestock species (Duarte et al., 2021). Random forests are a supervised classification ensemble algorithm, in which multiple decision trees are built using random subsets of the observations and the variables. For each encountered observation, each tree gets a weighted 'vote' for a classification, resulting in a probability for each possible classification. Random forests are thought to be a useful method for analysis of genomic datasets with a large number of genes and a small number of samples (Chen & Ishwaran, 2012), being less susceptible to over-fitting than, for example, logistic regression (Matsuki et al., 2016). I trained three flat random forests, using 1) the hospital and community resistome data only to predict WWTP resistome classification, 2) all data to predict hospital/community/WWTP, and 3) all data to predict hospital/not hospital classification. Due to low sample size, I did not subset data for training. I included the union of the 50 most abundant resistance genes in each site type, resulting in 80 genes in total. I used the same training parameters as (Duarte et al, 2021): a 10 x 10 cross validation grid for re-sampling, and the number of variables selected for each tree was the square root of the number of features. Model performance was assessed with accuracy and the Kappa measure. All were implemented using the packages caret (Kuhn, 2008) and randomForest (Liaw & Wiener, 2002).

5.3.10 Linear mixed effects models structures

As in Chapter 4, I used the gene length of the gene group and the total bacterial fragment count as offsets in models of the read hit counts to account for the count related properties of the data whilst also normalising.

To select the linear model family, I compared the Widely Applicable Information Criterion (WAIC) and Expected Log Pointwise Predictive Density (ELPD) (Gelman et al., 2014) of Poisson and negative binomial models with and without zero-inflation. Zero-inflated models are mixture models which account for two different random processes simultaneously generating zeroes, leading to a greater number of zeroes than would be expected if only one process was acting. In this case, the two sources of zeroes are a) the chance that the gene is not present in the site, and b) the chance that the gene was present but undetected by the sampling and sequencing process. The model is parameterised with the zero-inflation parameter, z_i, which is the binomial probability of 'true' zeroes. We were interested in Poisson or negative binomial distributions for the positive observations to take overdispersion in the count data into account. Both zero-inflation and overdispersion have been recognised as features of count data from microbial communities that can impact the power of the results if not taken into account, especially for metagenomic analysis (Jonsson et al., 2019; X. Zhang et al., 2017).

All models were implemented using brms (Bürkner, 2017) and flat, uninformative priors were used – the package default random effect parameters, and $Normal(0, 100)$ for fixed effects.

5.3.10.1 Model structure 1: Impact of hospital presence on resistance gene abundance

In this model I included only observations from WWTPs to ask if having a hospital in a catchment area increased the resistance gene abundance in WWTP influent. The observational unit was a read hit count for a resistance gene group in a sample. Two hierarchies are present in the data which may be a source of variation: the pairing, i.e. the group that a WWTP was matched to, and the catchment area of the WWTP. Which of these hierarchies would be more important was not known a priori. Therefore I compared the WAIC of two models with each hierarchy as a random intercept effect (and an extra model with neither) and used the structure with the best fit. A second random intercept effect used 70% sequence homology clusters to account for differences in the abundance of different gene groups. A 70% clustering was used because most groups had only one gene in them in the 90% clustered groups. Presence of a hospital in a WWTP's catchment area was used as a binary fixed effect.

5.3.10.2 Model structure 2: association between upstream and downstream resistance gene abundance

I assessed the link between community and hospital sites and their downstream WWTPs by correlating the abundance of resistance genes at these sites with those at WWTPs. The random effect structure was the same as model 1. I also included a term for the FPKM of the resistance gene at the upstream site. I hypothesised that community sites would have more similar resistance gene abundance to WWTPs, and therefore have a higher association value. I tested for this by estimating separate fixed effects for the hospital/WWTP and community/WWTP associations. I also allowed for resistance abundance to be affected by the distance between the upstream site and the downstream WWTP, as longer distances may lead to greater degradation and dilution. To do this I used the straight-line distance in kilometres between each site as a fixed effect.

5.3.10.3 Model structure 3: impact of prescription rates on resistance abundance in sewage

I tested the association between prescription rates and resistance abundance using a third set of models. Separate models were used to look at hospital and community prescriptions for a) all antibiotic types, b) amoxicillin prescriptions and resistance genes, c) carbapenem prescriptions and resistance genes, and d) vancomycin prescriptions and resistance genes. For the carbapenem model I looked at hospital prescription only. The use of different antibiotic groupings followed methods in (Perry et al., 2021), to allow for the resistance-prescription relationship to vary depending on whether antibiotics are typically used in the community and the hospital (amoxicillin), or only in hospital settings (carbapenems and vancomycin). In all the hospital prescriptions models, only the hospitals and their four connected WWTPs were included.

I estimated the impact of direct selection by matching each resistance genes with the DDDs of the antibiotic to which it confers resistance. Resistance gene groups often confer resistance to multiple antibiotic groups. To address this, for each resistance gene I took the sum of the DDDs of all antibiotic groups (at ATC level 4) that the gene group confers resistance to before normalisation. For example, for a gene conferring resistance to sulfonamides and tetracyclines I would sum the DDDs of both antibiotic groups.

The models included random intercept effects for the 70% gene clusters, a random intercept effect for the sample site to account for repeat measurements, and a fixed effect for the type of site (i.e., hospital, community, or WWTP). The impact of prescriptions was assessed through fixed effect terms for the antibiotic prescription rate matched to the resistance gene, and a second term for the total rate of prescription of any antibiotic in the area, log(x+1) transformed. In vancomycin and carbapenem models, prescription rate was scaled. Including the total rate of prescription allowed us to estimate the indirect selection effects of the total amount of prescriptions on selection for all resistance genes. I hypothesised that the type of site (WWTP, community or hospital) could affect the degree of association. For example, we might expect a closer association between hospital prescriptions and resistance levels in hospitals than WWTPs. This was accounted for by including an interaction term between matched antibiotic prescriptions and site type.

5.3.10.4 Model structure 4: impact of environmental factors

A fourth set of models were used to assess the impact of environmental factors on resistance gene abundance in sewage. The random effect structure included intercept effects for sample site and resistance gene cluster, and there was a fixed effect for site type. I then added either sample pH, sample temperature, weather category (sun, cloud, rain), flow rate (WWTP samples only), or sample collection time. A separate model for each environmental variable was used, similar to a univariate analysis. The association between each environmental variable was also estimated using Spearman's Rank, rank-biserial correlation, Kruskal-Wallis, or Goodman-Kruskall-γ.

5.3.10.5 Model structure 5: impact of distance on net change

A fifth mixed effects model structure was used to estimate the impact of distance between an upstream site and the downstream site on the net change in the FPKM of a resistance gene. Net change for each resistance gene in each WWTP catchment area was calculated as the downstream FPKM minus the upstream FPKM. Distance was measured as a straight line between the GPS co-ordinates of the upstream site and the WWTP, and included in the model as a fixed effect with an interaction for upstream site type (hospital or community). The random effects structure included intercepts for WWTP catchment area and gene cluster.

5.4 Results

5.4.1 Metagenomic data

The average number of reads mapped to reference databases was 46,635,892 (standard deviation 16,549,909). The proportion of read counts mapped to human and bacterial genomes is reported in Appendix D Table 1.

A total of 15,115 different bacterial species were detected in the samples. The most abundant bacteria by CLR were *Flavobacterium* sp. ($3.8x10^{-15}$), *Acinetobacter wuhouensis* ($3.3x10^{-15}$), *Kinneretia asaccharophila* ($3.2x10^{-15}$), *Acidovorax konjaci* ($3.2x10^{-15}$), and *Pseudomonas* sp. 31-12 ($3.1x10^{-15}$).

A total of 360 resistance gene groups were detected in the samples. There were 20,890 read hits to resistance genes on average per sample (range 302 – 91,249). Resistance genes made up 0.043% of bacterial read counts on average (range 0.0016 – 0.23). The most abundant resistance gene groups by FPKM were *sul1* (67.91), *blaOXA-233* (40.91), *aph(6)-Id* (15.44), *msr(E)_1* (14.42), and *ere(A)_5* (14.16).

Table 5.2: Comparison of total abundances and diversity indexes for resistance genes and bacterial species by site type

Total abundance and Shannon index for resistance genes calculated using FPKM. Total abundance for bacterial species calculated using total bacterial read counts, and Shannon estimated using relative abundances of bacterial species.

Site type (number of samples)	**Resistance genes**			**Bacterial species**		
	Total abundance	*Richness*	*Shannon*	*Total abundance*	*Richness*	*Shannon*
Hospitals (5)	1461.22 (659.01)	197.4 (21.56)	3.43 (0.31)	5169.19k (1603.05k)	14842.80 (70.77)	5.79 (0.82)
WWTPs (9)	141.09 (95.36)	119.7 (29.4)	3.22 (0.53)	4068.90k (1151.81k)	14842.22 (84.90)	5.60 (1.31)
Communities (5)	122.42 (97.01)	111.6 (31.71)	3.46 (0.21)	4714.25k (2401.15k)	14866.80 (38.17)	6.57 (0.91)

The relative proportion of the FPKM of resistance genes in each sampling site is plotted in Fig. 5.1. Here I compared the resistome of repeated samples from Victoria Hospital, Stirling community, Kirkcaldy WWTP, and Stirling WWTP. Generally, these were similar, apart from an increase in the proportion of the sample representing the tet(A) gene in Kirkcaldy WWTP, March 2022 sample. Many factors could explain this, such as sampling conditions, seasonal effects, and changes to population movements due to the COVID-19 prevention measures. However, with only two time points I was not able to investigate and account for drivers of this difference. Therefore, the March 2022 Kirkcaldy WWTP sample was removed from statistical analysis. However, results were similar overall whether this sample was or was not included.

The bacterial and resistance gene counts and diversity measures according to site type (hospitals, community, and WWTP) can be found in Table 5.2. Hospitals had a higher abundance of resistance genes and bacterial species than community or WWTP sites. Hospitals also had highest richness in resistance genes. Shannon diversity in resistance genes and bacterial species was similar across WWTPs, communities and hospitals.

5.4.2 Site characteristics

A full table of the site metadata is included in Appendix D Table 2. Sample conditions were broadly similar. The sample pH was 6.8 on average (standard deviation, SD: 0.52). Sample temperature varied the most by site type. Hospitals had a higher temperature of 16.2^{o}C (SD 2.1) compared to 10.2^{o}C in community samples (SD 1.6) and 11.8^{o}C in WWTPs (SD 1.2).

5.4.3 Prescription data

In the three months prior to sampling, the average of the total prescribed DDDs per capita from community healthcare providers post code sectors within the drainage areas sampled was 2.02 (SD 0.91). The Stirling drainage area had the highest per capita prescriptions (January – March, 2020, 3.52 DDDs per capita) and Kinneil Kerse had the lowest (September – November, 2020, 1.04 DDDs per capita). The antibiotics with the highest community DDD per capita in the three months before sampling were lymecycline (0.31, SD 0.16), doxycycline (0.29, SD 0.16), amoxicillin (0.24, SD 0.15), flucloxacillin (0.19, SD 0.01), and clarithromycin (0.16, SD 0.09).

For hospitals, I calculated the DDDs of the net antibiotic doses dispensed from pharmacies to wards per occupied bed days (OBD) in the three months prior to sampling. In total, net 1.15 DDDs per OBD were issued to hospital wards (SD 0.39). Hairmyres Hospital had the highest net antibiotics issued per OBD (October – December, 2020, 1.77 DDDs per OBD), and Victoria Hospital had the lowest (September – November, 2020, 0.76 DDDs per OBD). The antibiotics with the highest net DDDs per OBD in the hospitals were amoxicillin (0.16, SD 0.12), flucloxacillin (0.14, SD 0.06), amoxicillin and beta lactamase inhibitor, such as co-amoxiclav (0.14, SD 0.07), clarithromycin (0.13, SD 0.12), and doxycycline (0.10, SD 0.04).

5.4.4 Cluster analysis

Fig. 5.2.A. shows the results of the PCoA of distances between the FPKM of resistance genes for each site, with the centroids and variance by PCoA axis for the different types of site displayed. Hospitals form a distinct cluster with non-overlapping variances with other groups, whereas the centroids and variances for the community and WWTP sites are similar. This indicates that the resistome of hospitals is distinguishable from community and WWTP sites, but WWTP and community sewage resistomes are not distinguishable from each other. There was no clustering by catchment area or pair group. I extracted the twenty resistance genes with the greatest correlations with PCoA axis 1, which discriminated between hospital-type and community-type

compositions (Appendix D Table 3). These results show that several carbapenemases are enriched in hospital sample sites.

Fig. 5.2.B. shows the results of the PCoA of the distances between the relative abundance of bacterial species read counts, with centroids and variances by site type group. The hospital sites microbiomes had a distinct composition from WWTP site microbiomes, with community microbiome sharing some compositional features of each. I extracted the twenty bacterial species with the greatest correlations with PCoA axis 2, which discriminated most strongly between hospital and WWTP-type compositions (Appendix D Table 3). Some human-related bacterial genera such as *Klebsiella* and *Bifidobacterium* are enriched in hospital samples, whilst environment-related bacteria such as *Acidovorax* are enriched in WWTP samples.

Table 5.3: Model comparison using expected log pointwise predictive density and Widely Applicable Information Criterion

Higher ELPD indicates a better fit, and lower WAIC indicates a better fit. ELPD: expected log pointwise predictive density; WAIC: Widely Applicable Information Criterion.

Model	ELPD	WAIC
Poisson	-313714	729144
Zero-inflated Poisson	-279254	651291
Negative binomial	-14241	28357
Zero-inflated negative binomial	-14245	28358

5.4.5 Random Forest

I first trained a random forest model using only hospital and community site data to identify hospital vs. community sample site differences. The model accuracy was 99.0% and the Kappa was 97.7%. Fig. 5.3.A shows the classification predictions of the model for WWTP site data. The proportion of trees that voted for 'hospital' classification of resistomes from WWTP sites when there was a hospital in the catchment area was 13.9% on average, with wide variation between sites (SD 7.3). WWTPs without hospitals had a lower average proportion of votes for "Hospital" classification (5.6%) and lower variation between sites (SD 1.5). This could indicate that the resistomes of WWTPs with hospitals in their catchment area were more similar to hospitals than those without, although the wide variance in classification votes for WWTPs with hospitals mean caution is needed in interpreting this result.

In the second model, I trained a random forest model using all data (Fig. 5.3.B) to identify differences in all site type. The accuracy of this model was 71.2%, and the Kappa was 53.6%. The lower accuracy of this model may partially be due to the uneven number of samples from different site types (5 community, 5 hospital, and 9 WWTP samples). The most voted for classification of each sample did correspond to the true sample type, with most uncertainty in distinguishing community resistomes

from WWTP resistomes: an average 18.4% (SD 7.9) of votes were given to WWTP classification when predicting a community sample.

Table 5.4: Random intercept, shape, and zero-inflation parameter estimates for selected models

Presents posterior means and 95% uncertainty intervals in brackets. SD: standard deviation; WWTP: wastewater treatment plant.

Model	SD cluster	SD WWTP captured	Shape	Zero-inflation
Model structure 1 (impact of hospitals on resistance abundance)	3.14 (2.81 – 3.51)	0.69 (0.36 – 1.39)	0.72 (0.67 – 0.87)	0.02 (0.00 – 0.05)
Model structure 2 (association between resistance abundance at upstream sites and WWTPs)	2.77 (2.47 – 3.09)	0.71 (0.37 – 1.41)	0.82 (0.73 – 0.92)	0.01 (0.00 – 0.03)
Model structure 3 (association between community prescriptions and sewage resistance)	2.80 (2.54 – 3.11)	-	0.28 (0.27 – 0.30)	0.01 (0.00 – 0.03)
Model structure 3 (association between hospital prescriptions and sewage resistance)	3.17 (2.84 – 3.54)	-	0.33 (0.29 – 0.37)	0.04 (0.00 – 0.09)

5.4.6 Linear mixed effects models

5.4.6.1 Model family goodness of fit comparisons

The fit of the data to Poisson, negative binomial, and zero-inflated distributions was compared. I used two goodness of fit measures to compare model performance: the ELPD, estimated using leave-one-out cross-validation; and the WAIC. The values of the measures for each model family are presented in Table 5.3. Assuming a difference of more than three between goodness of fit measurements indicates one fit was better than the other, negative binomial models were a better fit than Poisson, so I eliminated Poisson as a potentially model structure. Zero-inflation did not improve fit, but only slightly worsened fit according to ELPD and was indistinguishable from the fit of the model without zero-inflation according to WAIC. As there is a theoretical basis for zero-inflation, I included zero-inflated negative binomial distribution was used in all following models.

5.4.6.2 Impact of hospital effluent on resistome of wastewater treatment plants

The impact of hospitals on WWTPs was looked at with two models. The first compared the abundance of resistance genes in WWTPs with and without a hospital in their catchment area (model structure 1). The second looked at the association between

resistance abundance in hospitals and WWTPs, and between community sites and WWTPs, to see which upstream-site type is more closely correlated with the mixed WWTP samples (model structure 2).

In model 1, there was no significant of impact of a hospital in the catchment area of a WWTP on resistance gene abundance, with an effect size of 0.84 (posterior mean), -0.24 – 1.91 (95% uncertainty intervals). Posterior means for random effects, shape parameter, and zero-inflation parameter for models 1 and 2 can be found in Table 5.4. The overall model fit for this and all following models was assessed with Bayes R^2 (Gelman et al, 2019), presented alongside MCMC diagnostics in the Appendix D Table 4.

Table 5.5: Associations between prescriptions and resistance levels

Presented as the posterior mean of the estimated association and the 95% uncertainty intervals. If an estimate was not estimated a dash (-) is used.

Wastewater source	**Antibiotic group**	**Prescription source**	
		Hospital	**Community**
WWTP	All	0.24 (-1.86 - 2.35)	0.68 (-0.35 - 1.69)
	Amoxicillin	0.57 (-10.71 - 12.31)	0.48 (-4.32 - 5.33)
	Carbapenem	0.22 (-0.46 - 0.99)	-
	Vancomycin	-0.32 (-3.60 - 3.43)	0.55 (-1.08 - 2.63)
Hospital	All	1.33 (-0.90 - 3.56)	-0.67 (-1.71 – 0.37)
	Amoxicillin	-2.16 (-15.47 - 7.73)	-4.33 (-11.16 - 2.49)
	Carbapenem	0.21 (-0.40 - 0.87)	-
	Vancomycin	0.26 (-0.84 – 1.41)	0.19 (-1.96 - 2.32)
Community	All	-	0.52 (-0.59 - 1.63)
	Amoxicillin	-	-4.44 (-4.32 - 5.33)
	Carbapenem	-	-
	Vancomycin	-	1.16 (-1.08 - 2.63)
All	Total	-0.76 (-2.29 – 0.56)	0.19 (-0.56 - 0.89)

In model structure 2, the association between community sewage and WWTP sewage was positive (0.76, 0.56 – 0.97). The association between hospital and WWTP sewage was also positive (0.49, 0.40 – 0.59). The interaction term estimating the difference between the community-WWTP and hospital-WWTP associations was negative (-0.27, -0.48 – -0.07), indicating that the hospital-WWTP was lower than the community-WWTP association. There was no evidence for an effect of distance between the sampling site and WWTP (-0.16, -0.58 – 0.30). The FPKM of resistance genes upstream vs. downstream sites is plotted in Fig. 5.4.

5.4.6.3 Associations between antibiotic prescriptions and sewage resistance levels

The association between prescriptions and resistance levels was assessed using model structure 3. Seven separate models were used to estimate the associations between community prescriptions and hospital prescriptions and sewage resistance for all antibiotic types, as well as amoxicillin, carbapenem, and vancomycin groups

individually. Prescription data FPKM of resistance genes are plotted in Fig. 5.5. All associations between resistance abundance and prescription rates can be found in Table 5.5. Random effects, shape and zero-inflation parameters for models with all prescription types included are reported in Table 5.4. The numbers of observations, sampling sites, and gene groups in each model is in Appendix D Table 5.

There was no association between resistance abundance in wastewater samples and any prescription data (Table 5.5). These results do not provide strong evidence for direct or indirect selection from antibiotic consumption.

In model structure 3, I also tested for the difference in the resistance abundance by site. Hospitals also had a higher resistance abundance than communities or WWTPs in both models. In the community prescription model, the hospital association estimate was 3.29 (3.02 – 3.56) and the WWTP association estimate was -1.32 (-3.41 – 0.71) with community as the reference category. In the hospital prescription model, hospital was the reference category and the WWTP association estimate was -0.27 (-0.52 - -0.03).

5.4.6.4 Associations between environmental variables and sewage resistance levels

Separate model formulae with the same basic structure were used to test the impact of pH, sample collection time, sample temperature, and flow rate on the abundance of resistance genes in the sewage. The effect estimates of these models can be found in Table 5.6. There was a positive association between resistance abundance and sample pH and temperature. Samples collected earlier in the day were found (though with some uncertainty) to have higher resistance gene read counts. These results suggest that environmental factors could have an impact on resistance abundance. Correlations between environmental factors are recorded in Appendix D Table 6. Notably, stream temperature and sample time were negatively correlated. This may be due to hospital samples being warmer and collected earlier in the day.

Table 5.6: Associations between environmental variables and sewage resistance levels

Separate models used for each variable. The association estimate is the posterior mean and standard deviation of the association between the environmental variable and resistance gene abundance.

Environmental variable	**Association estimate (95% uncertainty)**
pH	0.30 (0.07 – 0.54)
Sample collection time	-0.30 (-0.63 - 0.03)
Sample temperature	0.12 (0.03 – 0.22)
Flow rate	0.00 (0.00 – 0.00)
Weather: Sunny Cloud Rain	 Reference category 0.05 (-0.73 – 0.83) -0.49 (-0.78 – 0.19)

The association value for the flow rate was unexpectedly precise, so I checked the distribution of the two datasets (Appendix D Figure 1). I also tried a simple linear regression of the association between the FPKM and the flow rate, which returned the same association and confidence intervals (0.00, 0.00 – 0.00). The small number of flow rate observations may account for unusual confidence range.

5.4.6.5 Impact of distance and site type on net change in resistance gene abundance

The average net change in resistance gene FPKM between an upstream site and a WWTP was -1.80 (SD: 16.5). There was no association between the distance between an upstream and downstream site and the net change in resistance gene FPKM, whether the upstream site was a hospital (-0.39, -1.92 – 1.11) or a community site (-0.06, -1.90 – 1.77). Upstream site type did not affect the net change (effect of upstream site being a hospital -2.89, -10.81 – 4.83). These results do not provide evidence that the size of a wastewater network influences net change in the concentration of resistance genes. Appendix D Figure 2 shows the net change in resistance genes, stratified by upstream site type.

Figure 5.1: Relative abundance of resistance gene groups in each sampling site

Relative abundance was calculated based on FPKM. Sample names on the X axis were generated as location name, site type, month, and then year of sampling. The top 20 most abundance gene groups are plotted, all others are grouped into the 'Other' category for visual clarity.

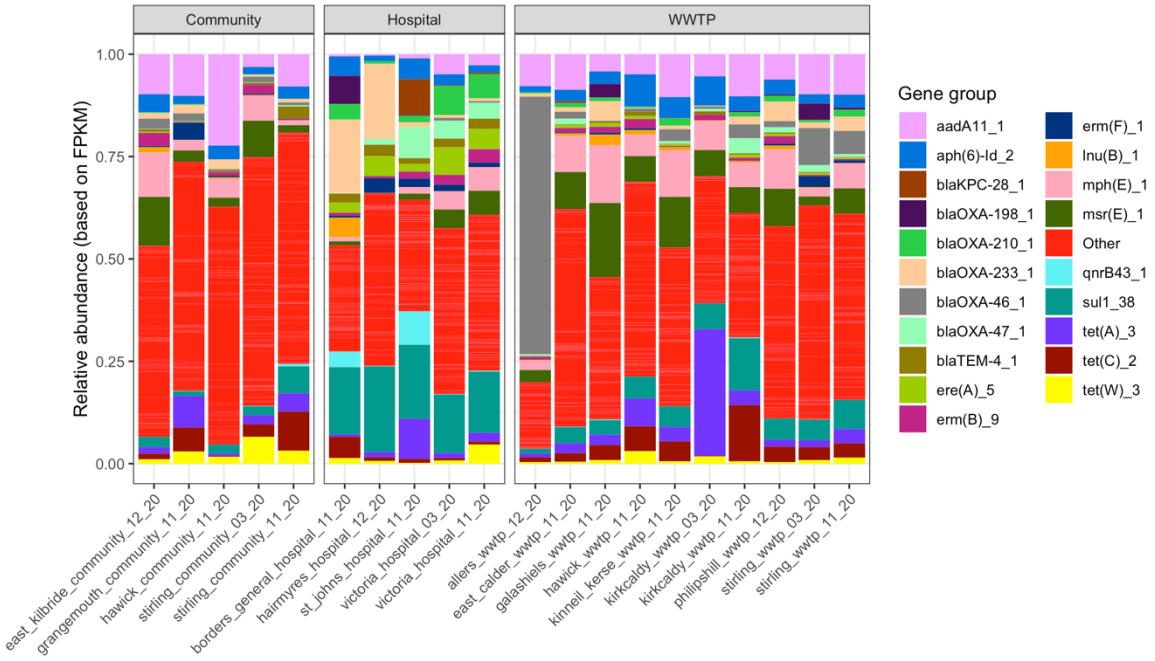


Figure 5.2: Principal Coordinates and Principal Components Analysis of the resistome and microbiome, grouped by site type

The colour of the dots indicates the site type (hospital, community, or WWTP). Empty points indicate hospitals and WWTPs with hospitals, and filled points indicate community samples and WWTPs without hospitals. Dot shape indicates the pair grouping of the sample (e.g., all circular points are from the same pair). Ellipses width and height indicates the variance of PCoA1 and PCoA2, respectively, for each group, and the angle of the ellipse indicates the covariance between the two axes by group. A) the PCoA of the resistome (measured in FPKM), and B) of the microbiome (measured in relative abundance), using Bray-Curtis distances.

A) B)

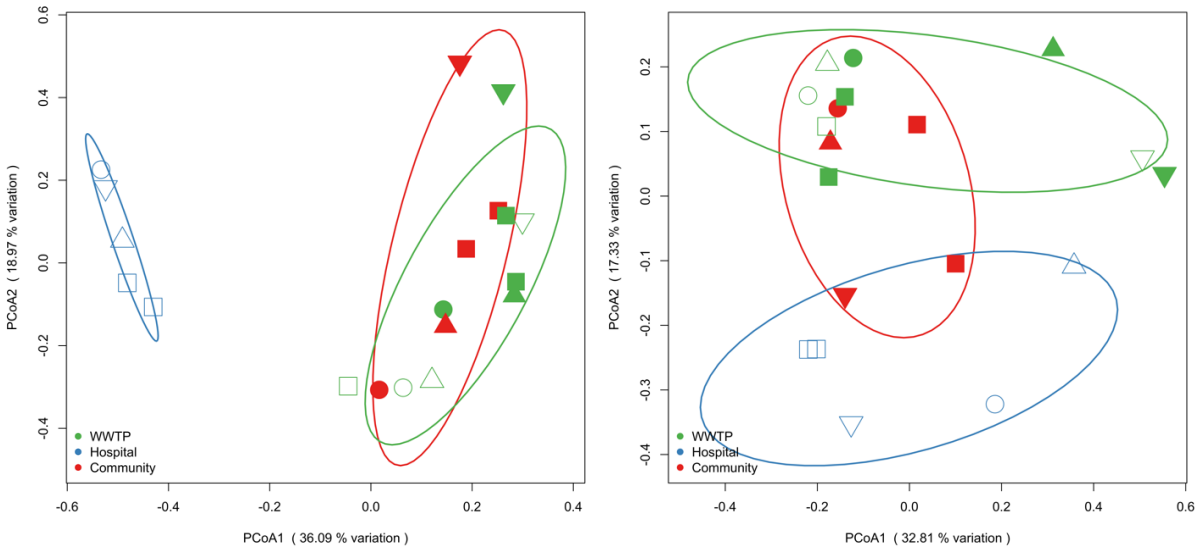


Figure 5.3: Classification predictions of samples resistomes from random forest models

A) the proportion of trees voting for 'hospital' site type classification predictions of WWTP sample resistomes with and without a hospital in their catchment area from model 1 (trained using hospital and community samples only). B) the proportion of votes from trees for the site type classification of each sample from model two, which was trained using all data, grouped by the true source of the sample.

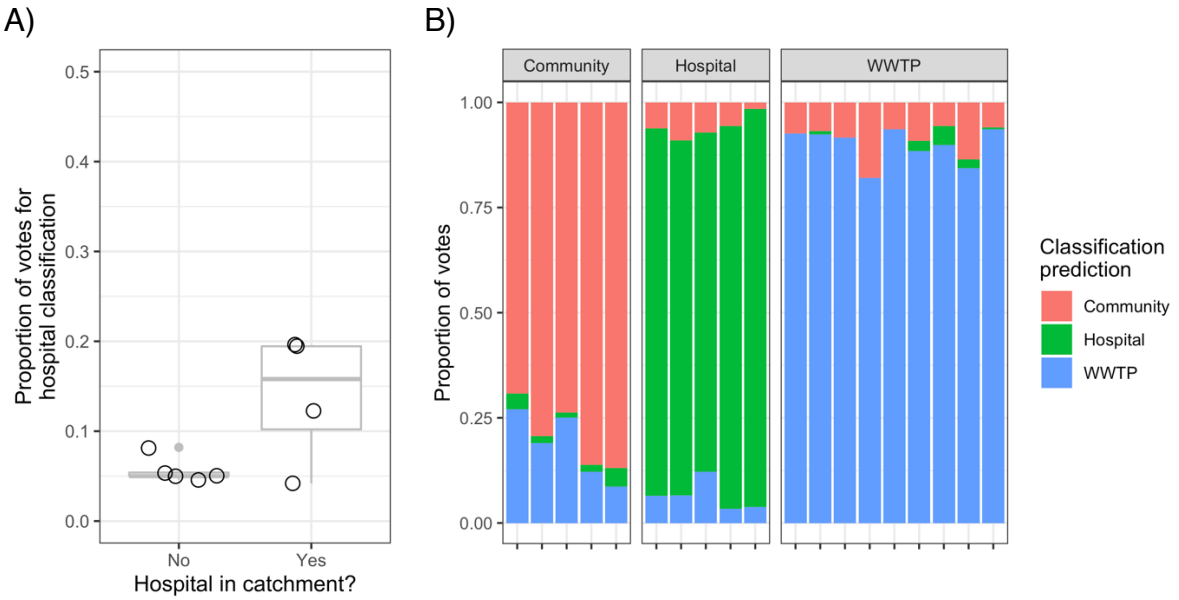


Figure 5.4: Resistance gene abundance in hospital and community samples vs. WWTP samples

FPKM of each resistance gene in hospital (left) and community (right) samples against the FPKM in connected WWTP sites. One observation is excluded from the community sample panel to so that both plots could have the same y axis for comparison (values of the excluded point: FPKM in WWTP sample = 135.1, FPKM in upstream site = 6.6). A log scale is used on the y axis.

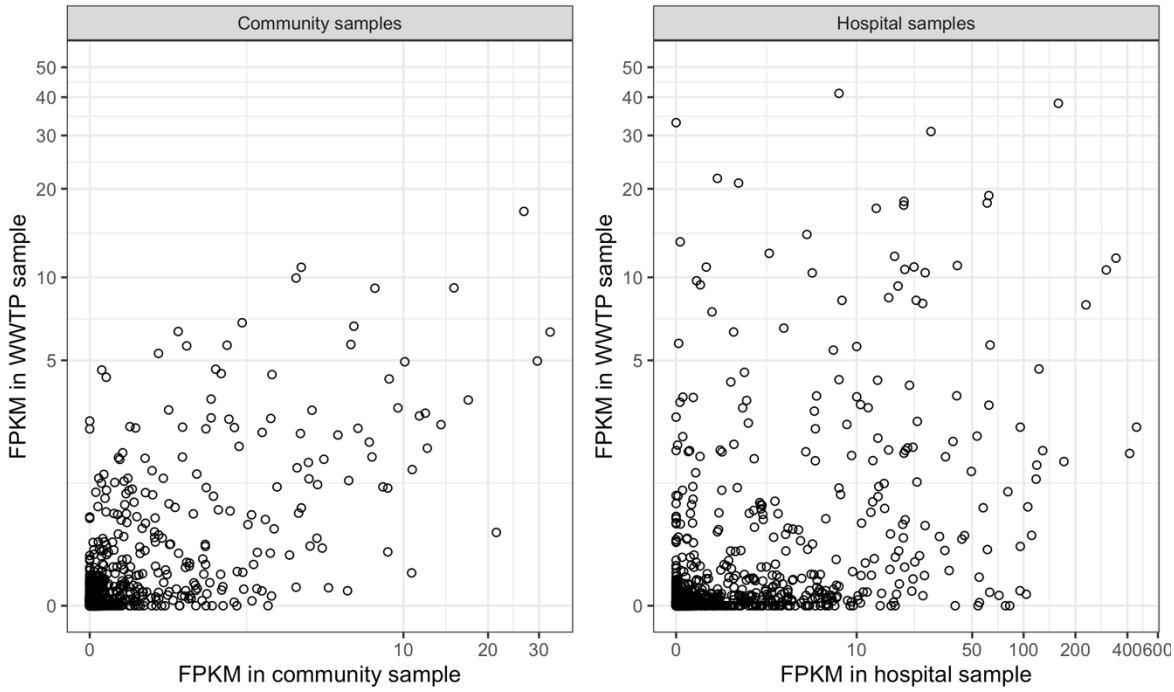
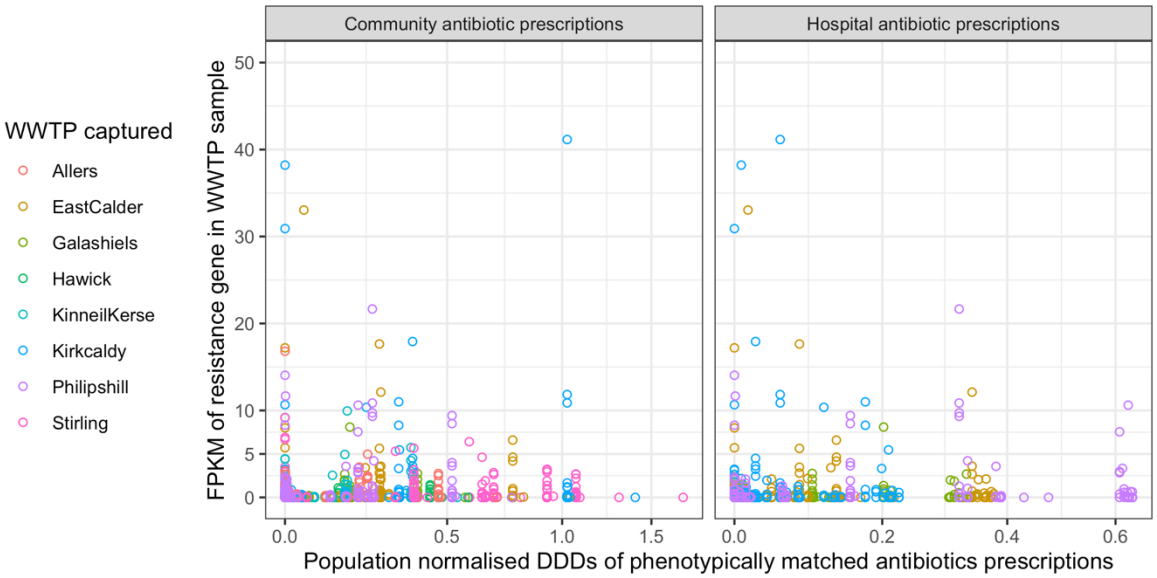


Figure 5.5: Antibiotic prescription rates and sewage resistance gene abundance

Phenotypically-matched antibiotic prescription rates against the FPKM of resistance genes in WWTPs, for community (left) and hospital (right) prescriptions.



5.5 Discussion

In this study we applied a paired study design and metagenomics to wastewater samples from hospitals, WWTPs, and communities in Scotland. The results show that hospitals had a distinctive resistome compared to WWTPs and communities, with a greater abundance and diversity of resistance genes. However, having a hospital in a catchment area did not increase the abundance of resistance in WWTP influent, and the WWTP resistome was unlikely to have the hospital as its primary source, suggesting the hospital effluent contribution to WWTP influent resistance is small. There was no evidence for an impact of prescription rates on resistance levels in this dataset, indicating that selection for resistance from antibiotic consumption was not detectable in the timeframe of this data.

5.5.1 Community and hospital sewage resistomes differed

The resistome of the hospital sewage was distinct from community sewage and WWTP influent. This distinction was due to greater abundances of the resistance genes and by a greater number of unique resistance genes. The greater number and diversity of resistance genes was not reflected by greater numbers or diversity of bacterial species in hospital wastewater, suggesting that more bacteria carried any and/or carried a greater number of resistance genes in hospital than community wastewater. Hospital wastewater from around the world has been found to have a greater abundance and diversity of resistance genes or resistant bacteria than municipal wastewater (Hassoun-Kheir et al., 2020; Hutinel et al., 2019; Korzeniewska et al., 2013; Kwak et al., 2015; Lamba et al., 2018; Paulshus et al., 2019; Verburg et al., 2019), although not in every case (Gundogdu et al., 2017; D. Zhang et al., 2021). However, few other studies have demonstrated that community effluent only and composite WWTP influent samples have similar resistome compositions. This result indicates that it is not just that hospital effluent and WWTP influent are distinct, but that effluents from community households and hospitals differ.

Many factors that could affect resistance gene abundance differ between hospital and community sewage. Hospitals are a focal point of patients with bacterial infections, co-morbidities, and recent exposure to antibiotics and other drugs, which may lead to higher rates and diversity of resistance in the patient microbiome. Hospital patients could be gaining resistant infections and colonisations during their stay in hospital, which are then carried through toilets, sinks and showers to the hospital wastewater.

The hospital environment itself may further impact the human to environment composition of the sample and itself select for resistant bacteria. Environmental properties that may influence microbiological growth have been shown to differ between hospital and municipal effluents, such as higher levels of suspended solids, ammonia and nitrate, and a lower biodegradable fraction (Khan et al., 2021; Majumder et al., 2021), as well as warmer temperatures and stagnation (Kizny Gordon et al., 2017). Indeed, I found that hospital wastewater was on average 6^{o}C warmer than community and WWTP wastewater. Hospitals are large building, and it has also been shown that the water network within large buildings provide opportunities for biofilm

formation (Hocquet et al., 2016; Kizny Gordon et al., 2017). In support of this theory, I found hospital and WWTP microbiomes clustered separately in a PCoA (Fig. 5.1), and that the separation was driven by greater abundance of human gut bacteria in hospital and a greater abundance of environmental bacteria in WWTP samples. The hospital sewage microbiome was also found to cluster separately from community and WWTP samples in previous studies in two studies in The Netherlands (Buelow et al., 2018; Verburg et al., 2021), and to cluster separately from WWTP but with community effluent in a study in Spain (Quintela-Baluja et al., 2019).

In addition to impacts on the microbial composition, the hospital environment may influence resistomes through exposure to higher concentrations of antibiotics, other drugs, and antiseptics that are excreted or disposed of into the wastewater system leading to selection for resistance (Khan et al., 2021; Majumder et al., 2021),(Hassoun-Kheir et al., 2020; Hocquet et al., 2016). Studies have investigated colonisation of the hospital environment and patients and found cross-compartment clonal groups of resistant bacteria, indicating patient-environment transmission (Constantinides et al., 2020; Feng et al., 2020). Moreover, experimental evolution studies have found that exposing bacteria to hospital wastewater selects for resistance more strongly than WWTP wastewater (Hutinel et al., 2021; Kraupner et al., 2021). Resistant bacteria that are colonising any part of the wastewater system, from sink drains to pipe junctions within the building, may be represented in the resistome of the hospital effluent. Overall, these results are consistent with the hospital water environment acting as a reservoir and enricher for resistant bacteria.

5.5.2 Hospital resistomes made a small contribution towards resistance in WWTP influent

The results indicate that the contribution of hospital effluent is small, and the primary source of resistance genes in wastewater is likely to be the community. This result is consistent with previous research, which has found no increase in resistance gene abundance in WWTP influent when a hospital is present (Buelow et al., 2018), and WWTP influent resistance to be more similar to community than hospital effluent (Gundogdu et al., 2017; Kutilova et al., 2021; Paulshus et al., 2019; Quintela-Baluja et al., 2019; Verburg et al., 2019).

I investigated hospital effluent contribution by estimating the gene-level association between resistance gene abundance in WWTPs and in hospitals or communities upstream, and using a source-attribution random forest model to estimate WWTP resistome source likelihood. Hospital resistance gene abundance was positively associated with resistance gene abundance in the downstream WWTP, suggesting that there is a link between hospitals and WWTPs. Bacterial isolates in WWTP influent have been found to be related to those of hospital isolates (Gouliouris et al., 2018), but a comparison of WWTP isolates to hospital *and* community isolates is lacking. Random forest models indicated that each site type (hospital, community, or WWTP) had a recognisable resistome, but that a signature of the hospital in the resistome of WWTP influent with a hospital in its catchment area was not detectable. If the hospital signature is not strongly detectable, the positive association between hospital and

WWTP resistance abundance may indicate that the hospital resistome is linked to the local community resistome, which is represented in the WWTP influent. In addition, it suggests that although hospital wastewater contains high levels of resistance, treating before discharge into the main sewage network may not appreciably decrease the load of resistant bacteria that need to be removed by the WWTP. However, a contribution of hospital effluent to resistance in WWTP influent cannot be completely ruled out.

A small contribution of hospital effluent is not surprising when we consider the volume of wastewater from hospitals compared to all the other sources entering a WWTP. Resistance genes were more than 10 times more abundant in hospital water than community water, but after dilution this increase may be negligible. Previous studies have estimated hospital effluent to be about 1% of the total volume of wastewater processed by a WWTP each day (Buelow et al., 2018; Verburg et al., 2019) and suggested this as a key factor in the low impact of hospital wastewater on WWTP resistance. In addition, human bacteria and their genomes may be degraded within the wastewater network, potentially decreasing the number of resistant bacteria from hospital effluent in a WWTP influent sample (Pehrsson et al., 2016; Verburg et al., 2019). Against this hypothesis, I found no relationship between the net change in resistance gene abundance between an upstream and downstream site and distance. In other countries, hospital effluent is treated prior to discharge into the main sewage network (Khan et al., 2021; Majumder et al., 2021), which would further decrease the hospital signature in WWTP influent. Further research is needed to quantify dilution and degradation of the hospital signal, including: deeply sequenced metagenomics of sewage to reduce noise (e.g., 100s of millions of reads per sample rather than 10s of millions), studies correlating the volume of hospital wastewater contributing to WWTP influent and resistance gene abundance, and studies along the length of a wastewater pipe to assess rates of degradation, possibly in experimental settings. On top of dilution and degradation, there may be other sources of resistance genes in WWTP influent than hospitals and communities. WWTP influent in Scotland includes wastewater from swimming pools, laboratories, and laundries, for example (Scottish Water, n.d. a), although most surface water is not mixed with domestic sewage until treatment (Scottish Water, n.d. b). Drinking water may also be a source of resistant bacteria that could be represented in WWTP influent, although the contribution is probably small (Piotrowska & Popowska, 2014). Additional sources of resistance genes may make the contribution of hospital effluent to WWTP resistance smaller and increase the uncertainty that a resistance gene in WWTP influent is from a hospital patient or community member.

5.5.3 Communities have a local resistome

Resistance gene abundance at an upstream site was positively associated with the abundance of the same gene at the WWTP influent. Finding that the resistance gene abundance in a community sample representing only a few household is predictive of abundance in the downstream WWTP, despite WWTPs receiving effluents from thousands of households and wastewater metagenomics having low sensitivity is interesting. It indicates that there could be a community-wide resistome in common

within a drainage area, i.e. that different community samples within a catchment area may also potentially correlate. Positive associations between hospital and WWTP wastewater has been found for bacterial species abundance (Buelow et al., 2018; Verburg et al., 2021) and resistance rates (Paulshus et al., 2019), but few if any studies have compared resistance gene abundance in these WWTPs and community sites. Studies of a whole WWTP drainage area, including different community sewage sampling sites and resistance hotspots like hospital effluent, could be used to investigate how resistance varies within a community. Despite finding a resistance gene level association between communities and their downstream WWTP, I did not find any distinct groupings by catchment area in the cluster analysis. This may be an indication that the overall profile and diversity is similar across the communities sampled. All of the communities included here are within relatively close geographic range, in the same country, and all in built up areas (although of different sized towns), which may lead to similar resistance gene profiles.

5.5.4 No impact of antibiotic consumption on resistance abundance in detectable in sewage

No relationships were found between prescription rates and resistance gene abundance in wastewater samples. Although antibiotic usage drives selection for resistant bacteria, there is mixed evidence for a positive association between resistance abundance in sewage and antibiotic consumption rates in the population. Previous studies (including Chapter 4 in this thesis) have found that associations with prescription data or with antibiotic residues and sewage resistance levels depend on the antibiotic studied and the modelling framework used, and that the effect is often small (Hendriksen et al., 2019; Perry et al., 2021). The dynamics of acquisition, colonisation, and transmission of resistant bacteria or resistance genes are complex. Aggregation, low sensitivity, and sampling error in wastewater metagenomics may make it difficult to detect small effect sizes and individual-level effects. These limitations of wastewater metagenomics have been considered previously considered (Aarestrup & Woolhouse, 2020; Miłobedzka et al., 2022). More sensitive resistance gene typing methods such as qPCR may be more powerful for discerning effects of antibiotic consumption. The prescription data used in this study is also aggregated at population level, potentially further masking any effects. Finally, this study is a snapshot of one time and over a relatively small geographic region, with similar prescription rates. Positive associations have been found in longitudinal datasets over a greater area (Hendriksen et al., 2019), possibly allowing investigation of this relationship over a greater range of prescription rates that has higher detectability.

In addition to the possibility that low sensitivity and data aggregation prevented detection of a small association, other factors not accounted for in the model may have masked an association. It may be possible that an unaccounted demographic or health-care related factor that affects resistance gene abundance may decrease the association through confounding. An important factor that was not considered in this analysis was area deprivation. However, greater levels of local deprivation have been shown to lead to an increase in resistance rates in hospitals (Alividza et al., 2018), but also to be positively associated with prescription rates in the UK (Lambourg et al.,

2022) and Scotland (Covvey et al., 2014), so would be more likely to lead to a higher rather than a lower association estimate in this data. An in-depth analysis of the prescriptions, demographics, and hospital resistance levels in this data will be needed to further address confounding. An alternative explanation is reverse causality if prescriptions were less frequently given due to higher resistance levels in an area, which could reduce the association estimate. Community and hospital prescribing guidelines in NHS boards in Scotland are updated every 6 months, taking into account recent resistance rates in clinical samples from hospitals and the community (Carol Philips, personal communication), so there is a potential route to reverse causality. However, the time lag on this process (i.e., time from use of an antibiotic driving a noticeable increase in resistance to a reduction in prescriptions) is unknown. Reverse causality has not to my knowledge been previously observed in sewage resistance data although it has been considered in cross-sectional population level resistance and prescription data (Jit et al., 2020). Longitudinal sampling designs with community demographic data can be used to detect confounding and reverse-causality.

There was no impact of total antibiotic usage on resistance abundance, indicating that there was no evidence for indirect selection. Direct selection is when a resistance gene is selected for because it confers a growth advantage during exposure to an antibiotic. Resistance genes that are near the directly selected gene (for example in the same genome or plasmid) but confer resistance to a different antibiotic may be indirectly benefitted and increase in prevalence in the bacterial population. Previous studies have not found an association between total antibiotic resistance and resistance gene abundance in wastewater (Hendriksen et al., 2019; Perry et al., 2021), and together this may indicate that this effect may not be detectable at a population level.

5.5.5 Environmental factors impact resistance gene abundance

I found a positive effect of higher pH or sample temperature and early sample collection time, although flow rate and the weather had no impact. The wastewater microbiome and resistome are thought to be sensitive to environmental conditions that may influence the growth and degradation of bacteria (Guo et al., 2019; Pärnänen et al., 2019). Previous studies have found that physicochemical properties of wastewater, such as pH, ammonia concentration, and flow rate, may affect wastewater resistance gene abundance (Guo et al., 2019; Harnisz et al., 2020; Yang et al., 2019). In this study there may not have been enough variation to detect an impact of flow rate. There was a negative association with the time that the autosampler was installed, suggesting that samples collected in morning had a greater resistance gene abundance. Diurnal patterns in resistance gene abundance have been found in wastewater, likely impacted by changes water usage throughout the day (Cai et al., 2022; Guo et al., 2019). This finding may also point to a limitation of the autosamplers, as it suggests that the resistome signal is strongly influenced by what had most recently been added and there may therefore be degradation or other changes in composition in the sample during the 24 hour sampling period. Wastewater studies may benefit from collecting samples at the same time or taking separate samples through the day to take diurnal activity into account (Chau et al., 2022).

5.5.6 Conclusion

In this study I demonstrate that hospital and community sewage resistomes differ, with enrichment of the abundance and diversity of resistance genes in hospitals. The influence of hospital-specific drivers of resistance, including patient-environment transmission and exposure to antibiotics, is potentially strong. However, I also find that the contribution of hospitals to resistance abundance in WWTP influence is small, likely due to the high level of dilution of hospital wastewater. I suggest that resistance genes in WWTP samples are therefore primarily from community sources, and WWTP influent is a good indicator of the community resistome. I find no evidence for an impact of prescription data on resistance gene abundance, suggesting this effect is small or non-existent. Wastewater studies would benefit from sampling sites within the wastewater network, including hospitals and communities, and taking longitudinal samples, to further investigate drivers of sewage resistance.

5.6 Acknowledgements

I would like to acknowledge contributions from many people who helped with this project. Meghan also contributed to the design, collection, and interpretation of this analysis. Scottish Water supplied information on different WWTPs, contributed to sample collection design, as well as providing autosamplers and collecting the wastewater samples, including Mark McEwan, David Wills, Moray Simon, John Lee Bar, and Rebecca Campbell. DNA extraction, sequencing, and bioinformatics was done by collaborators at DTU, including Prof. Frank Aarestrup, Dr Hanne Nielsen, Dr Saria Otani, and Baptiste Avot. Dr Bryan Wee from the University of Edinburgh conducted the taxonomic read mapping. Carol Philips, Barbara Moore, and Kelly Legard at NHS Lothian set up the request form for hospital prescription data, and helped in design of prescription data collection. Barry Melia and Aidan Morrison at Public Health Scotland helped in setting up an appropriate community prescription request form and fulfilled it. And finally thanks to many more pharmacists and estate managers who helped us collect wastewater and prescriptions from the hospitals.

6 General discussion

6.1 Summary of results and objectives

Antibiotic resistance (ABR) is a serious threat to global public health. In addition to causing high morbidity and mortality, resistant bacteria can be asymptomatically carried by healthy humans. This silent carriage could contribute to resistance burden but is hard to monitor. Resistant bacteria can also be found in the environment – human- and livestock-derived bacteria in environments impacted by wastewater, agriculture, and landfill are particularly concerning. These sources of resistant bacteria pose a risk to human and environmental health, but also present an opportunity for surveillance of resistance in human populations. In this thesis, I investigated the potential links between the environment and humans and the dynamic consequences of this relationship. I also built statistical models of hospital and municipal wastewater metagenomes to estimate associations with clinical resistance surveillance and antibiotic prescription data, and to inform design of surveillance programmes.

My first objective was to investigate the role that the environment might play in the spread and emergence of ABR. I addressed this objective in Chapter 3 by using a mathematical model to study transmission between humans, animals, and the environment. The results of this model suggested that the environment can theoretically play a role in human ABR epidemiology, and perhaps even a stronger one than animals. However, more data on resistance transmission and persistence, especially in the environment, are needed to contextualise these theoretical results. I also address this objective in Chapter 2, using a systematic scoping review to find evidence of dissemination of ABR from hospitals to municipal wastewater. This work indicated that hospital wastewater can contain resistant bacteria from patients, and untreated and treated municipal wastewater may do so as well. This suggests that wastewater, especially hospital wastewater, could play a role in dissemination of clinically relevant antibiotic resistance into the environment.

My second objective was to identify epidemiological data sets that represent antibiotic resistance genes (ARGs) in sewage. I focused on two sources of data: hospital-based resistance surveillance and antibiotic consumption. In Chapter 4 I used a multi-response model of European Antimicrobial Resistance Surveillance Network (EARS-Net) and Global Sewage Surveillance data to estimate country-level correlations between sewage and hospital resistance levels for different drug-bug combinations. I found that a few, but not all, drug-bug combinations were correlated in hospitals and sewage. In the systematic review in Chapter 2, I also found some examples of positive correlations between hospital-based surveillance and resistance abundance in sewage. Overall, this indicates some support that hospital-based surveillance data can at least partially explain ARG abundance in sewage. Prescription rates from national surveillance data were positively associated with some resistance gene phenotypes, giving some support for this dataset representing ARGs in sewage.

In Chapter 5, I used a paired study design to compare wastewater metagenomes from wastewater treatment plants (WWTPs) with and without hospitals, and hospital contributing to their influent, and community effluents, along with prescription data. Here, I found no associations between prescription data and resistance levels in either hospital effluent or community effluents. This seemingly contradictory finding may

imply that the scale on which sewage and prescription data are collected are important – an association may be found internationally where there is more variation in resistance levels and consumption patterns that are not detectable in relatively homogeneous national data.

My third objective was to use spatial and temporal analysis of ARGs in sewage to inform the design of surveillance systems. This objective was addressed in Chapter 5. Here, differences in the resistome of hospital, community, and mixed municipal effluent highlight the need for single-source sampling to capture the full range of resistance patterns in a WWTP catchment area. The systematic review in Chapter 2 also highlights that hospital effluent can represent the patient resistome, but not in all cases, and that community samples are rare in wastewater studies to date. Together, these findings suggest that hospital effluent sampling could be complimentary to WWTP influent sampling. Further sampling in the community would also be useful at this early stage of WWTP surveillance systems, as much is still unknown about how local differences in resistomes can combine in WWTP influent. In Chapters 4 and 5 I also investigated associations between ARGs and physicochemical properties of the sample, such as pH, sample temperature, and flow rate. I found some associations with pH and sample temperature. These associations could be due to the sewage environment generating opportunities for growth or destruction of bacteria (McLellan & Roguet, 2019), in which case collecting data on sample properties as metadata is key.

The final objective was to develop predictive models of resistance gene abundance. In Chapter 4 I used the multi-response linear model to predict clinical resistance. The results showed that jointly modelling both datasets improved model likelihood but not accuracy of predictions of clinical resistance. Therefore, this model structure can explain existing data but is not yet an improvement for extrapolating to unseen data. In Chapter 5 I used a zero-inflated negative binomial linear model and a random forest analysis to model and predict ARG abundance and wastewater sample type. Model comparison indicated that zero-inflation and negative binomial structures accounted for data variation better than Poisson models, suggesting this is a more appropriate model structure. Also in Chapter 5, I demonstrated that random forest analysis can suit wastewater metagenomic data well, accurately predicting wastewater sample type from resistance gene abundance. In summary, in this thesis I have highlighted the need for other model structures than classical generalised linear models and bivariate correlations, and that comparing the results and performance of multiple model frameworks during analysis is needed.

6.2 Implications of thesis

Wastewater metagenomes must represent human resistomes for this surveillance technique to be useful. In this thesis, I show that the wastewater resistome does reflect the community and hospital resistomes, although not perfectly, supporting the use of metagenomics of wastewater for surveillance. Previous studies have compared hospital, community, and WWTP resistance data (Lamba et al., 2018; Paulshus et al., 2019; Verburg et al., 2019), but few have done so on an international scale, or collected sewage samples from multiple hospital and community sites. Understanding

the relatedness of community and WWTP influent is important as it implies that dilution and wastewater microbiology do not influence or destroy the human resistome signal, lending support for the use of wastewater for surveillance. Despite speculation that the community can act as a reservoir for resistant bacteria, there are few surveillance datasets of resistance in the community, so WWTP influent sampling can fill this important gap. However, Chapters 2, 4, and 5 suggest that hospital effluent resistomes are quite different to WWTP influent. Therefore, wastewater surveillance from WWTPs may be more appropriate for monitoring community rather than hospital resistance. Surveillance of wastewater from hospitals could be valuable for detecting hospital-based outbreaks early, and for identifying environmental reservoirs of resistance in the hospital.

The environment is a mixing pot for resistance from different humans, animals, and resident environmental bacteria (Larsson & Flach, 2022). In this thesis, I show that an environmental reservoir can theoretically have a strong impact on human health (Chapter 3). Few empirical observations of human-environmental health links exist (Bürgmann et al., 2018), although other models of the environment have also highlighted its potential importance (Booton et al., 2021; Call et al., 2013; Græsbøll et al., 2014; McBryde & McElwain, 2006). There is increasing recognition that sanitation and lack of access to clean drinking water is a driver of antibiotic resistance, especially in Lower and Middle Income Countries (LMICs), implying that the environment does indeed represent an important source of resistance or play an important role in transmission of resistance (Bürgmann et al., 2018).

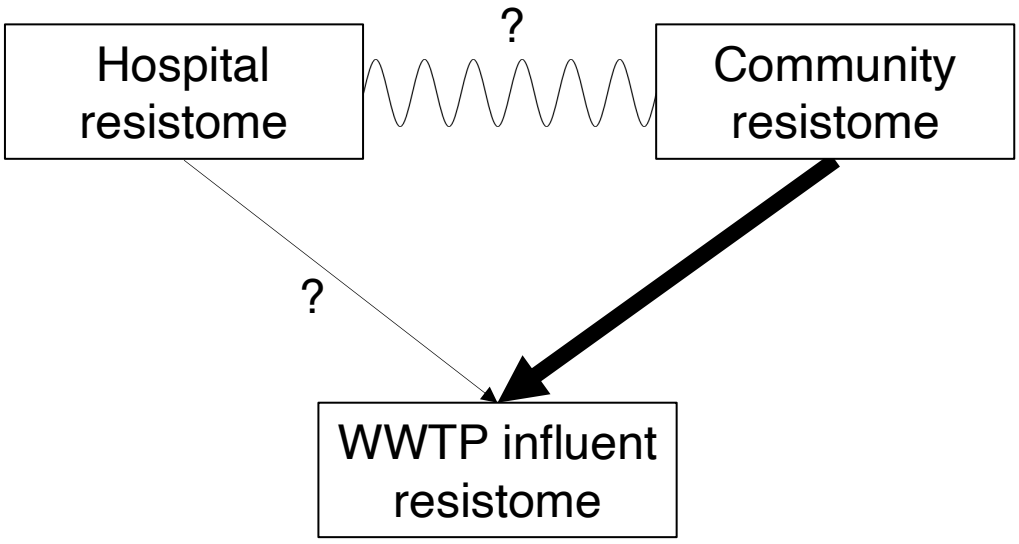
The role of the environment can have implications for clinical settings as well. As is highlighted in Chapter 2, organisms from clinical environments can be found in hospital wastewater and, in Chapter 5, I also describe the distinctiveness of the hospital resistome and how this may reflect not only the patient but also the hospital environment resistome. An increasing number of studies have considered the transmission of resistance between patients and the environment in hospitals, e.g. (Constantinides et al., 2020; Feng et al., 2020). Together these findings highlight unknowns about the dynamics of resistance in the hospital environment: does increased persistence of resistant bacteria in the hospital environment increase incidence of resistant nosocomial infections? How much might the effectiveness of patient-focused interventions around antimicrobial stewardship and infection prevention in hospitals be mitigated by environment-patient resistance transmission? More modelling and empirical work is needed to study the role of the hospital environment.

The results of this thesis have implications for our understanding of ABR in the community. Firstly, it provides some information to address the question: what is the relationship between resistance in hospitals and in the community? The results in this thesis could be explained by hospital resistomes being influenced by but still separate from local community. For example, in Chapter 5 I find that ARG abundance in both hospital and community effluents were positively associated with downstream ARG abundance but the association value was lower for hospital than community effluent. Two non-mutually exclusive mechanisms can explain either of these observations: 1) the hospital resistance profile is correlated with the community resistance profile, and

the community resistance profile is correlated with the WWTP influent; and 2) that the hospital resistome signature is still prevalent enough in the WWTP influent to generate this correlation (Fig. 6.1). Other studies that have found positive relationships between resistance in hospitals and in WWTP influent have variously made both of these interpretations, but few if any consider both at the same time (Buelow et al., 2020; Gouliouris et al., 2019; Pärnänen et al., 2019; Paulshus et al., 2019). If (1) is true, the implication is that hospital resistance patterns should be considered tied to the local community, both when predicting the resistance profile on admission and in interpreting hospital-based surveillance data. If (2) is true, the implication is that analysis of WWTP influent for surveillance must account for this contribution in order to be applicable to the community. In this thesis we can comment further by combining the results of Chapters 4 and 5. Hospital effluent had a distinctive signature but did not appear to leave a strong trace in WWTP influent in Chapter 5. In addition, the phenotype of resistance genes that were highlighted as hospital-type in Chapter 5 (such as *blaTEM* and *blaOXA*) did not have a positive correlation with sewage ARG abundance in Chapter 4. These findings suggest that (1) is playing a greater role than (2), as they suggest the hospital signature degrades (due to dilution) before it reaches the WWTP influent. However, to fully resolve the contribution of (1) and (2) with wastewater sampling, studies that sample from multiple hospital and multiple community effluents within a WWTP catchment area are needed.

Figure 6.1: Relationships between resistance in hospitals, communities, and municipal effluent

Waved line indicates correlation (without assumption about direction of impact) and arrows indicate direct impact. There is some evidence for hospital-community correlations and some evidence for a low impact of the hospital on the WWTP influent resistome, but more precise quantification is needed.



Finally, this thesis has some implications for interventions to reduce antibiotic consumption to mitigate ABR in the community. Antibiotic stewardship is often suggested as a key route for reducing resistance (Aliabadi et al., 2021; Majumder et

al., 2020). The results of this thesis do not provide strong evidence to support the effectiveness of reducing antibiotic consumption. In Chapter 2, the results suggest that curtailing antibiotic usage in animals may be less effective than sanitation-related factors that reduce environment-human transmission. In later chapters, I demonstrate some support for antibiotic consumption driving resistance at a population-level in the community after accounting for time trends and local resistance patterns, but the relationship could only be detected on an international level and not on a national level, suggesting the effect is small. Therefore, there is some empirical and theoretical evidence that the impact of consumption on community resistance may be small overall when compared to other factors, such as environmental transmission and local resistance levels. Most other studies looking at risk factors for resistance levels have focused on resistant infections in hospital settings, e.g. (Chatterjee et al., 2018). Comparing our findings to those of other studies trying to link silent community carriage of extended ß-lactam (ESBL) resistant *Enterobacteriaceae* with risk factors, travel is frequently identified, as well as antibiotic consumption and overcrowding (Karanika et al., 2016; Otter et al., 2019). Overall, the implication is that antibiotic consumption reduction may be effective in the community, but its effectiveness in comparison to other possible interventions, such as sanitation practices, is uncertain. Moreover, the most effective intervention may vary by country, depending on country-level differences in community access to antibiotics and sanitation. Studies that examine the impact of multiple risk factors in the whole gut resistome are needed to give the full picture. Both sewage and individual gut metagenomes can be used in future studies assessing these impacts, identifying both population-level and individual-level effects.

6.3 Future research directions

In this thesis I present evidence that municipal wastewater is more representative of community than hospital effluents. Further studies are needed that compare wastewater and human metagenomes to explore similarities and differences between these two microbial communities. A few existing studies have done this to some extent such as (Pehrsson et al., 2016), but more in-depth analysis of healthy community resistomes in multiple countries will be useful for highlighting the ways to extract 'human' type signatures from wastewater.

I use a very simplified model to study human-animal-environment transmission. More studies are needed that consider the complex nature of human-environment interactions, including different environment types, and the importance of heterogeneous and stochastic transmission. In addition, there is still a lack of mathematical models of ARG abundance in sewage. Some studies have used some modelling approaches to look at decay of ARGs in sewage (Amos et al., 2015), but none have tried to link ARG abundance with human population-level epidemiology in the way that some models of SARS-CoV-2 in wastewater have (Fernandez-Cassi et al., 2021). Future research may try different possible model structures, especially around the sampling distribution to be used to get from a human resistance prevalence to the wastewater ARG abundance. Normalisation for population size will be a key challenge for this modelling. A key unknown for future models will be the distribution

function that links human resistance prevalence to ARG abundance in the sewage, and future modelling efforts should compare the results of different options.

In Chapter 4 I used sewage and national surveillance data from Europe only. A clear next step for developing this work would be to use global sources of data, such as WHO Global Antimicrobial Resistance and Use Surveillance System (GLASS). This model may need to take into account differences in this relationship in Higher Income Countries (HICs) vs. LMICs. In addition, many demographic variables such as human development index (HDI) are collected on the country level and could be incorporated into this model. Previous work has found a link between ARG abundance in the sewage and HDI (Hendriksen et al., 2019), so this would be interesting to explore with this new model structure.

Chapter 5 demonstrates that community wastewater samples can represent the community resistome of a group of households. However, few studies take multiple community samples. A next step could be to repeat the Chapter 5 study design with additional community samples, especially in catchment areas with hospitals, to give further clarity on differences between hospitals and communities, and to demonstrate if there are differences in the resistome within a catchment area.

Finally, I highlight that the hospital effluent resistome is distinctive from the community resistome, and links have been found between patients and hospital effluent. However few if any studies have investigated the similarity of resistance in hospital effluent and the hospital environment. Hospital effluent could be a tool for non-invasively monitoring both the resistome of the patients and the environment, but studies that elucidate this additional link in the chain are needed to validate this.

6.4 Concluding remarks

In this thesis, I demonstrate the usefulness of wastewater metagenomics as a tool for monitoring resistance abundance in the community. Wastewater represents community resistomes more closely than hospital resistomes, but with some similarities to the hospital that still need to be resolved. The wider consequences of resistance in the environment may be strong but need to be studied more empirically. Further research into variation within wastewater catchment areas is required to map resistance in wastewater back to humans.

Bibliography

Untapped savings : water services in the NHS / Audit Commission for Local Authorities and the National Health Service in England and Wales. (1993). In *Untapped savings : water services in the NHS*. HMSO.

Global Sewage Surveillance Project. (2020). https://www.compare-europe.eu/Library/Global-Sewage-Surveillance-Project. Accessed: 07/07/2022.

Aarestrup, F. M., & Woolhouse, M. E. J. J. (2020). Using sewage for surveillance of antimicrobial resistance. *Science*, *367*(6478), 630–632. https://doi.org/10.1126/science.aba3432

Adator, E. H., Narvaez-Bravo, C., Zaheer, R., Cook, S. R., Tymensen, L., Hannon, S. J., Booker, C. W., Church, D., Read, R. R., & McAllister, T. A. (2020). A one health comparative assessment of antimicrobial resistance in generic and extended-spectrum cephalosporin-resistant *Escherichia coli* from beef production, sewage and clinical settings. *Microorganisms*, *8*(6), 1–22. https://doi.org/10.3390/microorganisms8060885

Akiba, M., Senba, H., Otagiri, H., Prabhasankar, V. P., Taniyasu, S., Yamashita, N., Lee, K. ichi, Yamamoto, T., Tsutsui, T., Ian Joshua, D., Balakrishna, K., Bairy, I., Iwata, T., Kusumoto, M., Kannan, K., & Guruge, K. S. (2015). Impact of wastewater from different sources on the prevalence of antimicrobial-resistant *Escherichia coli* in sewage treatment plants in South India. *Ecotoxicology and Environmental Safety*, *115*, 203–208. https://doi.org/10.1016/j.ecoenv.2015.02.018

Al-Nassir, W. N., Sethi, A. K., Li, Y., Pultz, M. J., Riggs, M. M., & Donskey, C. J. (2008). Both oral metronidazole and oral vancomycin promote persistent overgrowth of vancomycin-resistant enterococci during treatment of *Clostridium difficile*-associated disease. *Antimicrobial Agents and Chemotherapy*, *52*(7), 2403–2406. https://doi.org/10.1128/AAC.00090-08

Aldred, K. J., Kerns, R. J., & Osheroff, N. (2014). Mechanism of Quinolone Action and Resistance. *Biochemistry*, *53*, 1565–1574.

Alexander, J., Hembach, N., & Schwartz, T. (2020). Evaluation of antibiotic resistance dissemination by wastewater treatment plant effluents with different catchment areas in Germany. *Scientific Reports*, *10*(1), 8952. https://doi.org/10.1038/s41598-020-65635-4

Aliabadi, S., Anyanwu, P., Beech, E., Jauneikaite, E., Wilson, P., Hope, R., Majeed, A., Muller-Pebody, B., & Costelloe, C. (2021). Effect of antibiotic stewardship interventions in primary care on antimicrobial resistance of *Escherichia coli* bacteraemia in England (2013–18): a quasi-experimental, ecological, data linkage study. *The Lancet Infectious Diseases*, *21*(12), 1689–1700. https://doi.org/10.1016/S1473-3099(21)00069-4

Alividza, V., Mariano, V., Ahmad, R., Charani, E., Rawson, T. M., Holmes, A. H., & Castro-Sánchez, E. (2018). Investigating the impact of poverty on colonization and infection with drug-resistant organisms in humans: A systematic review. *Infectious Diseases of Poverty*, *7*(1), 1–11. https://doi.org/10.1186/s40249-018-0459-7

Allen, H. K., Donato, J., Wang, H. H., Cloud-Hansen, K. A., Davies, J., & Handelsman, J. (2010). Call of the wild: Antibiotic resistance genes in natural environments. *Nature Reviews Microbiology*, *8*(4), 251–259. https://doi.org/10.1038/nrmicro2312

Amos, G. C. A., Hawkey, P. M., Gaze, W. H., & Wellington, E. M. (2014). Waste water effluent contributes to the dissemination of CTX-M-15 in the natural environment. *Journal of Antimicrobial Chemotherapy*, *69*(7), 1785–1791. https://doi.org/10.1093/jac/dku079

Amos, G. C. A., Gozzard, E., Carter, C. E., Mead, A., Bowes, M. J., Hawkey, P. M., Zhang, L., Singer, A. C., Gaze, W. H., & Wellington, E. M. H. (2015). Validated predictive modelling of the environmental resistome. *The ISME Journal*, *9*(6), 1467–1476. https://doi.org/10.1038/ismej.2014.237

Andersson, D. I., Patin, S. M., Nilsson, A. I., & Kugelberg, E. (2007). The Biological Cost of Antibiotic Resistance. In R. A. Bonomo & M. . Tolmasky (Eds.), *Enzyme-Mediated Resistance to Antibiotics* (pp. 339–348). John Wiley & Sons, Ltd. https://doi.org/https://doi.org/10.1128/9781555815615.ch21

Andersson, D. I., & Hughes, D. (2010). Antibiotic resistance and its cost: Is it possible to reverse resistance? *Nature Reviews Microbiology*, *8*(4), 260–271. https://doi.org/10.1038/nrmicro2319

Araujo, S., Henriques, I. S., Leandro, S. M., Alves, A., Pereira, A., & Correia, A. (2014). Gulls identified as major source of fecal pollution in coastal waters: A microbial source tracking study. *Science of the Total Environment*, *470*, 84–91. https://doi.org/10.1016/j.scitotenv.2013.09.075

Asghar, H., Diop, O. M., Weldegebriel, G., Malik, F., Shetty, S., Bassioni, L. El, Akande, A. O., Maamoun, E. Al, Zaidi, S., Adeniji, A. J., Burns, C. C., Deshpande, J., Oberste, M. S., & Lowther, S. A. (2014). Environmental surveillance for polioviruses in the global polio eradication initiative. *Journal of Infectious Diseases*, *210*(Suppl 1), S294–S303. https://doi.org/10.1093/infdis/jiu384

Ashley, E. A., Recht, J., Chua, A., Dance, D., Dhorda, M., Thomas, N. V., Ranganathan, N., Turner, P., Guerin, P. J., White, N. J., & Day, N. P. (2018). An inventory of supranational antimicrobial resistance surveillance networks involving low- and middle-income countries since 2000. *Journal of Antimicrobial Chemotherapy*, *73*(7), 1737–1749. https://doi.org/10.1093/jac/dky026

Atmani, S. M., Messai, Y., Alouache, S., Fernández, R., Estepa, V., Torres, C., Bakour, R., Fernandez, R., Estepa, V., Torres, C., Bakour, R., Fernández, R., Estepa, V., Torres, C., & Bakour, R. (2015). Virulence characteristics and genetic background of ESBL-producing *Klebsiella pneumoniae* isolates from wastewater. *Fresenius Environmental Bulletin*, *24*(1), 103–112. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84928164434&partnerID=40&md5=0c6110406a66097031c7a14f018e02ae

Baker-Austin, C., Wright, M. S., Stepanauskas, R., & McArthur, J. V. (2006). Co-selection of antibiotic and metal resistance. *Trends in Microbiology*, *14*(4), 176–182. https://doi.org/10.1016/j.tim.2006.02.006

Bäumlisberger, M., Youssar, L., Schilhabel, M. B., & Jonas, D. (2015). Influence of a non-hospital medical care facility on antimicrobial resistance in wastewater. *PLoS ONE*, *10*(3), e0122635. https://doi.org/10.1371/journal.pone.0122635

Beattie, R. E., Skwor, T., & Hristova, K. R. (2020). Survivor microbial populations in post-chlorinated wastewater are strongly associated with untreated hospital sewage and include ceftazidime and meropenem resistant populations. *Science of the Total Environment*, *740*. https://doi.org/10.1016/j.scitotenv.2020.140186

Bedenić, B., Siroglavić, M., Slade, M., Šijak, D., Dekić, S., Musić, M. Š., Godan-Hauptman, A., & Hrenović, J. (2020). Comparison of clinical and sewage isolates of Acinetobacter baumannii from two long-term care facilities in Zagreb; mechanisms and routes of spread. *Archives of Microbiology*, *202*(2), 361–368. https://doi.org/10.1007/s00203-019-01750-9

Bengtsson-Palme, J., Hammarén, R., Pal, C., Östman, M., Björlenius, B., Flach, C. F., Fick, J., Kristiansson, E., Tysklind, M., & Larsson, D. G. J. (2016). Elucidating selection processes for antibiotic resistance in sewage treatment plants using metagenomics. *Science of the Total Environment*, *572*, 697–712. https://doi.org/10.1016/j.scitotenv.2016.06.228

Bengtsson-Palme, J., Kristiansson, E., & Larsson, D. G. J. (2018). Environmental factors influencing the development and spread of antibiotic resistance. In *FEMS Microbiology Reviews* (Vol. 42, Issue 1, pp. 68–80). https://doi.org/10.1093/femsre/fux053

Bengtsson-Palme, J., Larsson, D. G. J., & Kristiansson, E. (2017). Using metagenomics to investigate human and environmental resistomes. *Journal of Antimicrobial Chemotherapy*, *72*(10), 2690–2703. https://doi.org/10.1093/jac/dkx199

Berendonk, T. U., Manaia, C. M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F., Buergmann, H., Sorum, H., Norstrom, M., Pons, M.-N., Kreuzinger, N., Huovinen, P., Stefani, S., Schwartz, T., Kisand, V., Baquero, F., & Luis Martinez, J. (2015). Tackling antibiotic resistance: the environmental framework.

Nature Reviews Microbiology, *13*(5), 310–317. https://doi.org/10.1038/nrmicro3439

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A Fresh Approach to Numerical Computing. *ArXiv*, *59*(1), 65–98. https://doi.org/10.1137/141000671

Bilotta, G. S., Milner, A. M., & Boyd, I. L. (2014). Quality assessment tools for evidence from environmental science. *Environmental Evidence*, *3*(1), 1–14. https://doi.org/10.1186/2047-2382-3-14

Blair, J. M. A., Webber, M. A., Baylay, A. J., Ogbolu, D. O., & Piddock, L. J. V. (2015). Molecular mechanisms of antibiotic resistance. *Nature Reviews Microbiology*, *13*(1), 42–51. https://doi.org/10.1038/nrmicro3380

Booton, R. D., Meeyai, A., Alhusein, N., Buller, H., Feil, E., Lambert, H., Mongkolsuk, S., Pitchforth, E., Reyher, K. K., Sakcamduang, W., Satayavivad, J., Singer, A. C., Sringernyuang, L., Thamlikitkul, V., Vass, L., OH-DART Study Group, Avison, M. B., & Turner, K. M. E. (2021). One Health drivers of antibacterial resistance: Quantifying the relative impacts of human, animal and environmental use and transmission. *One Health*, *12*(January), 100220. https://doi.org/10.1016/j.onehlt.2021.100220

Bottery, M. J., Wood, A. J., & Brockhurst, M. A. (2016). Selective conditions for a multidrug resistance plasmid depend on the sociality of antibiotic resistance. *Antimicrobial Agents and Chemotherapy*, *60*(4), 2524–2527. https://doi.org/10.1128/AAC.02441-15

Brinch, C., Leekitcharoenphon, P., Duarte, A. S. R., Svendsen, C. A., Jensen, J. D., & Aarestrup, F. M. (2020). Long-Term Temporal Stability of the Resistome in Sewage from Copenhagen. *MSystems*, *5*(5), 1–10. https://doi.org/10.1128/msystems.00841-20

Brockhurst, M. A., Harrison, E., Hall, J. P. J., Richards, T., McNally, A., & MacLean, C. (2019). The Ecology and Evolution of Pangenomes. *Current Biology*, *29*(20), R1094–R1103. https://doi.org/10.1016/j.cub.2019.08.012

Buelow, E., Rico, A., Gaschet, M., Lourenco, J., Kennedy, S. P., Wiest, L., Ploy, M. C., & Dagot, C. (2020). Hospital discharges in urban sanitation systems: Long-term monitoring of wastewater resistome and microbiota in relationship to their eco-exposome. *Water Research X*, *7*. https://doi.org/10.1016/j.wroa.2020.100045

Buelow, E., Bayjanov, J. R., Majoor, E., Willems, R. J. L. L., Bonten, M. J. M. M., Schmitt, H., & van Schaik, W. (2018). Limited influence of hospital wastewater on the microbiome and resistome of wastewater in a community sewerage system. *FEMS Microbiology Ecology*, *94*(7), fiy087. https://doi.org/10.1093/femsec/fiy087

Bürgmann, H., Frigon, D., Gaze, W. H., Manaia, C. M., Pruden, A., Singer, A. C., Smets, B. F., & Zhang, T. (2018). Water and sanitation: An essential battlefront in the war on antimicrobial resistance. *FEMS Microbiology Ecology*, *94*(9), fiy-101. https://doi.org/10.1093/femsec/fiy101

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Bushnell, B. (2014). *BBMap: A Fast, Accurate, Splice-Aware Aligner*. https://www.osti.gov/biblio/1241166

Cabello, F. C., Godfrey, H. P., Tomova, A., Ivanova, L., Dölz, H., Millanao, A., & Buschmann, A. H. (2013). Antimicrobial use in aquaculture re-examined: Its relevance to antimicrobial resistance and to animal and human health. *Environmental Microbiology*, *15*(7), 1917–1942. https://doi.org/10.1111/1462-2920.12134

Cahill, N., O'Connor, L., Mahon, B., Varley, Á., McGrath, E., Ryan, P., Cormican, M., Brehony, C., Jolley, K. A., Maiden, M. C., Brisse, S., & Morris, D. (2019). Hospital effluent: A reservoir for carbapenemase-producing Enterobacterales? *Science of the Total Environment*, *672*, 618–624. https://doi.org/10.1016/j.scitotenv.2019.03.428

Cai, L., Sun, J., Yao, F., Yuan, Y., Zeng, M., Zhang, Q., Xie, Q., Wang, S., Wang, Z., & Jiao, X. (2021). Antimicrobial resistance bacteria and genes detected in hospital sewage provide valuable information in predicting clinical antimicrobial resistance. *Science of the Total Environment*, *795*, 148815. https://doi.org/10.1016/j.scitotenv.2021.148815

Cai, M., Wang, Z., Gu, H., Dong, H., Zhang, X., Cui, N., Zhou, L., Chen, G., & Zou, G. (2022). Occurrence and temporal variation of antibiotics and antibiotic resistance genes in hospital inpatient department wastewater: Impacts of daily schedule of inpatients and wastewater treatment process. *Chemosphere*, *292*(October 2021), 133405. https://doi.org/10.1016/j.chemosphere.2021.133405

Calero-Cáceres, W., Ye, M., & Balcázar, J. L. (2019). Bacteriophages as Environmental Reservoirs of Antibiotic Resistance. *Trends in Microbiology*, *27*(7), 570–577. https://doi.org/10.1016/j.tim.2019.02.008

Call, D. R., Matthews, L., Subbiah, M., & Liu, J. (2013). Do antibiotic residues in soils play a role in amplification and transmission of antibiotic resistant bacteria in cattle populations? *Frontiers in Microbiology*, *4*(JUL), 1–8. https://doi.org/10.3389/fmicb.2013.00193

Carr, V. R., Witherden, E. A., Lee, S., Shoaie, S., Mullany, P., Proctor, G. B., Gomez-Cabrero, D., & Moyes, D. L. (2020). Abundance and diversity of

resistomes differ between healthy human oral cavities and gut. *Nature Communications*, *11*(1), 1–10. https://doi.org/10.1038/s41467-020-14422-w

Cassini, A., Högberg, L. D., Plachouras, D., Quattrocchi, A., Hoxha, A., Simonsen, G. S., Colomb-Cotinat, M., Kretzschmar, M. E., Devleesschauwer, B., Cecchini, M., Ouakrim, D. A., Oliveira, T. C., Struelens, M. J., Suetens, C., Monnet, D. L., the Burden of AMR Collaborative Study Group. (2019). Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. *The Lancet Infectious Diseases*, *19*(1), 56–66. https://doi.org/10.1016/S1473-3099(18)30605-4

CDC. (2019). *Antibiotic resistance threats in the United States, 2019*. https://doi.org/10.15620/cdc:82532

Charani, E., McKee, M., Ahmad, R., Balasegaram, M., Bonaconsa, C., Merrett, G. B., Busse, R., Carter, V., Castro-Sanchez, E., Franklin, B. D., Georgiou, P., Hill-Cawthorne, K., Hope, W., Imanaka, Y., Kambugu, A., Leather, A. J., Mbamalu, O., McLeod, M., Mendelson, M., Mpundu, M., Rawmson, T. M., Ricciardi, W., Rodriguez-Manzano, J., Singh, S., Tsioutis, C., Uchea, C., Zhu, N., Holmes, A. H. (2021). Optimising antimicrobial use in humans – review of current evidence and an interdisciplinary consensus on key priorities for research. *The Lancet Regional Health - Europe*, *7*. https://doi.org/10.1016/j.lanepe.2021.100161

Chatterjee, A., Modarai, M., Naylor, N. R., Boyd, S. E., Atun, R., Barlow, J., Holmes, A. H., Johnson, A., & Robotham, J. V. (2018). Quantifying drivers of antibiotic resistance in humans: a systematic review. *The Lancet Infectious Diseases*, *18*(12), E368–E378. https://doi.org/10.1016/S1473-3099(18)30296-2

Chau, K. K., Barker, L., Budgell, E. P., Vihta, K. D., Sims, N., Kasprzyk-Hordern, B., Harriss, E., Crook, D. W., Read, D. S., Walker, A. S., & Stoesser, N. (2022). Systematic review of wastewater surveillance of antimicrobial resistance in human populations. *Environment International*, *162*(March), 107171. https://doi.org/10.1016/j.envint.2022.107171

Chen, X., & Ishwaran, H. (2012). Random forests for genomic data analysis. *Genomics*, *99*(6), 323–329. https://doi.org/10.1016/j.ygeno.2012.04.003

Clausen, P. T. L. C., Aarestrup, F. M., & Lund, O. (2018). Rapid and precise alignment of raw reads against redundant databases with KMA. *BMC Bioinformatics*, *19*(1), 1–8. https://doi.org/10.1186/s12859-018-2336-6

Cohen, M. L. (2000). Changing patterns of infectious disease. *Nature*, *406*(August), 726–767.

Collaboration for Environmental Evidence. (2013). *Guidelines for Systematic Reviews in Environmental Management. Version 4.2.*

http://www.environmentalevidence.org/wp-content/uploads/2014/06/Review-guidelines-version-4.2-final.pdf

Collignon, P., Beggs, J. J., Walsh, T. R., Gandra, S., & Laxminarayan, R. (2018). Anthropological and socioeconomic factors contributing to global antimicrobial resistance: a univariate and multivariable analysis. *The Lancet Planetary Health*, *2*(9), e398–e405. https://doi.org/10.1016/S2542-5196(18)30186-4

Constantinides, B., Chau, K. K., Quan, T. P., Rodger, G., Andersson, M. I., Jeffery, K., Lipworth, S., Gweon, H. S., Peniket, A., Pike, G., Millo, J., Byukusenge, M., Holdaway, M., Gibbons, C., Mathers, A. J., Crook, D. W., Peto, T. E. A., Walker, A. S., & Stoesser, N. (2020). Genomic surveillance of *Escherichia coli* and *Klebsiella* spp. in hospital sink drains and patients. *Microbial Genomics*, *6*(7). https://doi.org/10.1099/mgen.0.000391

Covvey, J. R., Johnson, B. F., Elliott, V., Malcolm, W., & Mullen, A. B. (2014). An association between socioeconomic deprivation and primary care antibiotic prescribing in Scotland. *Journal of Antimicrobial Chemotherapy*, *69*(3), 835–841. https://doi.org/10.1093/jac/dkt439

Cox, G., & Wright, G. D. (2013). Intrinsic antibiotic resistance: Mechanisms, origins, challenges and solutions. *International Journal of Medical Microbiology*, *303*(6–7), 287–292. https://doi.org/10.1016/j.ijmm.2013.02.009

D'costa, V. M., King, C. E., Kalan, L., Morar, M., Sung, W. W. L., Schwarz, C., Froese, D., Zazula, G., Calmels, F., Debruyne, R., Golding, G. B., Poinar, H. N., & Wright, G. D. (2011). Antibiotic resistance is ancient. *Nature*, *477*(7365), 457–461. https://doi.org/10.1038/nature10388

Davies, J., & Davies, D. (2010). Origins and evolution of antibiotic resistance. *Microbiology and Molecular Biology Reviews*, *74*(3), 417–433. https://doi.org/10.1128/mmbr.00016-10

Davies, N. G., Flasche, S., Jit, M., & Atkins, K. E. (2019). Within-host dynamics shape antibiotic resistance in commensal bacteria. *Nature Ecology and Evolution*, *3*(3), 440–449. https://doi.org/10.1038/s41559-018-0786-x

de Abreu, V. A. C., Perdigão, J., & Almeida, S. (2021). Metagenomic Approaches to Analyze Antimicrobial Resistance: An Overview. *Frontiers in Genetics*, *11*(January), 1–9. https://doi.org/10.3389/fgene.2020.575592

de Kraker, M. E. A., Stewardson, A. J., & Harbarth, S. (2016). Will 10 Million People Die a Year due to Antimicrobial Resistance by 2050? *PLoS Medicine*, *13*(11), 1–6. https://doi.org/10.1371/journal.pmed.1002184

Diallo, O. O., Baron, S. A., Abat, C., Colson, P., Chaudet, H., & Rolain, J. M. (2020). Antibiotic resistance surveillance systems: A review. *Journal of Global*

Antimicrobial Resistance, *23*, 430–438. https://doi.org/10.1016/j.jgar.2020.10.009

Diemert, S., & Yan, T. (2019). Clinically unreported salmonellosis outbreak detected via comparative genomic analysis of municipal wastewater *Salmonella* isolates. *Applied and Environmental Microbiology*, *85*(10), 1–11. https://doi.org/10.1128/AEM.00139-19

Dolejska, M., & Literak, I. (2019). Wildlife is overlooked in the epidemiology of medically important antibiotic-resistant bacteria. *Antimicrobial Agents and Chemotherapy*, *63*(8), 1–5. https://doi.org/10.1128/AAC.01167-19

Doron, S., & Gorbach, S. (2008). Bacterial Infections: Overview. *International Encyclopedia of Public Health*, *January*, 273–282.

Doud, C. W., Scott, H. M., & Zurek, L. (2014). Role of house flies in the ecology of Enterococcus faecalis from wastewater treatment facilities. *Microbial Ecology*, *67*(2), 380–391. https://doi.org/10.1007/s00248-013-0337-6

Duarte, A. S. R., Röder, T., Van Gompel, L., Petersen, T. N., Hansen, R. B., Hansen, I. M., Bossers, A., Aarestrup, F. M., Wagenaar, J. A., & Hald, T. (2021). Metagenomics-Based Approach to Source-Attribution of Antimicrobial Resistance Determinants – Identification of Reservoir Resistome Signatures. *Frontiers in Microbiology*, *11*(January), 1–17. https://doi.org/10.3389/fmicb.2020.601407

Dupouy, V., Doublet, B., Arpaillange, N., Praud, K., Bibbal, D., Brugère, H., Oswald, E., Cloeckaert, A., Toutain, P. L. P.-L., & Bousquet-Mélou, A. (2016). Dominant plasmids carrying extended-spectrum β-lactamases blaCTX-M genes in genetically diverse *Escherichia coli* from slaughterhouse and urban wastewaters. *Environmental Microbiology Reports*, *8*(5), 789–797. https://doi.org/10.1111/1758-2229.12440

Dutil, L., Irwin, R., Finley, R., Ng, L. K., Avery, B., Boerlin, P., Bourgault, A. M., Cole, L., Daignault, D., Desruisseau, A., Demczuk, W., Hoang, L., Horsman, G. B., Ismail, J., Jamieson, F., Maki, A., Pacagnella, A., & Pillai, D. R. (2010). Ceftiofur resistance in *Salmonella enterica* serovar Heidelberg from chicken meat and humans, Canada. *Emerging Infectious Diseases*, *16*(1), 48–54. https://doi.org/10.3201/eid1601.090729

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Ekwanzala, M. D., Lehutso, R. F., Kasonga, T. K., Dewar, J. B., & Momba, M. N. B. (2020). Environmental dissemination of selected antibiotics from hospital wastewater to the aquatic environment. *Antibiotics*, *9*(7), 1–16. https://doi.org/10.3390/antibiotics9070431

Ekwanzala, M. D., Dewar, J. B., Kamika, I., & Momba, M. N. B. (2019). Tracking the environmental dissemination of carbapenem-resistant *Klebsiella pneumoniae* using whole genome sequencing. *Science of the Total Environment*, *691*, 80–92. https://doi.org/10.1016/j.scitotenv.2019.06.533

El-Mahdy, R., Mostafa, A., & El-Kannishy, G. (2018). High level aminoglycoside resistant enterococci in hospital-acquired urinary tract infections in Mansoura, Egypt. *Germs*, *8*(4), 186–190. https://doi.org/10.18683/germs.2018.1145

European Centre for Disease Prevention and Control. (2020). *Antimicrobial resistance in the EU/EEA (EARS-Net) - Annual Epidemiological Report for 2019*. ECDC. https://doi.org/10.1136/vr.g2500

Feng, J., Li, B., Jiang, X., Yang, Y., Wells, G. F., Zhang, T., & Li, X. (2018). Antibiotic resistome in a large-scale healthy human gut microbiota deciphered by metagenomic and network analyses. *Environmental Microbiology*, *20*(1), 355–368. https://doi.org/10.1111/1462-2920.14009

Feng, Y., Wei, L., Zhu, S., Qiao, F., Zhang, X., Kang, Y., Cai, L., Kang, M., McNally, A., & Zong, Z. (2020). Handwashing sinks as the source of transmission of ST16 carbapenem-resistant *Klebsiella pneumoniae*, an international high-risk clone, in an intensive care unit. *Journal of Hospital Infection*, *104*(4), 492–496. https://doi.org/10.1016/j.jhin.2019.10.006

Fernandez-Cassi, X., Scheidegger, A., Bänziger, C., Cariti, F., Tuñas, A., Ganesanandamoorthy, P., Lemaitre, J. C., Ort, C., Julian, T. R., & Kohn, T. (2021). Wastewater monitoring outperforms case numbers as a tool to track COVID-19 incidence dynamics when test positivity rates are high. *Water Research*, *200*, 117252. https://doi.org/10.1016/j.watres.2021.117252

Flach, C.-F., Hutinel, M., Razavi, M., Åhrén, C., & Larsson, D. G. J. (2021). Monitoring of hospital sewage shows both promise and limitations as an early-warning system for carbapenemase-producing Enterobacterales in a low-prevalence setting. *Water Research*, *200*. https://doi.org/10.1016/j.watres.2021.117261

Food, E., & Authority, S. (2022). The European Union Summary Report on Antimicrobial Resistance in zoonotic and indicator bacteria from humans, animals and food in 2019–2020. *EFSA Journal*, *20*(3). https://doi.org/10.2903/j.efsa.2022.7209

Gao, Y. X., Li, X., Fan, X. Y., Zhao, J. R., & Zhang, Z. X. (2022). Wastewater treatment plants as reservoirs and sources for antibiotic resistance genes: A review on occurrence, transmission and removal. *Journal of Water Process Engineering*, *46*(100), 102539. https://doi.org/10.1016/j.jwpe.2021.102539

Gardner, M., Jones, V., Comber, S., Scrimshaw, M. D., Coello-Garcia, T., Cartmell, E., Lester, J., & Ellor, B. (2013). Performance of UK wastewater treatment works

with respect to trace contaminants. *Science of the Total Environment*, *456–457*, 359–369. https://doi.org/10.1016/j.scitotenv.2013.03.088

Gatica, J., Jurkevitch, E., & Cytryn, E. (2019). Comparative Metagenomics and Network Analyses Provide Novel Insights Into the Scope and Distribution of beta-Lactamase Homologs in the Environment. *Frontiers in Microbiology*, *10*. https://doi.org/10.3389/fmicb.2019.00146

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, *24*(6), 997–1016. https://doi.org/10.1007/s11222-013-9416-2

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., & Egozcue, J. J. (2017). Microbiome datasets are compositional: And this is not optional. *Frontiers in Microbiology*, *8*(NOV), 1–6. https://doi.org/10.3389/fmicb.2017.02224

Goldstein, R. E. R., Micallef, S. A., Gibbs, S. G., He, X., George, A., Sapkota, A. A. R., Joseph, S. W., Sapkota, A. A. R., Rosenberg Goldstein, R. E., Micallef, S. A., Gibbs, S. G., He, X., George, A., Sapkota, A. A. R., Joseph, S. W., & Sapkota, A. A. R. (2014). Occupational exposure to *Staphylococcus aureus* and *Enterococcus* spp. among spray irrigation workers using reclaimed water. *International Journal of Environmental Research and Public Health*, *11*(4), 4340–4355. https://doi.org/10.3390/ijerph110404340

Gouliouris, T., Coll, F., Ludden, C., Blane, B., Raven, K. E., Naydenova, P., Crawley, C., Török, M. E., Enoch, D. A., Brown, N. M., Harrison, E. M., Parkhill, J., & Peacock, S. J. (2021). Quantifying acquisition and transmission of Enterococcus faecium using genomic surveillance. *Nature Microbiology*, *6*(1), 103–111. https://doi.org/10.1038/s41564-020-00806-7

Gouliouris, T., Raven, K. E., Ludden, C., Blane, B., Corander, J., Horner, C. S., Hernandez-Garcia, J., Wood, P., Hadjirin, N. F., Radakovic, M., Holmes, M. A., de Goffau, M., Brown, N. M., Parkhill, J., & Peacock, S. J. (2018). Genomic Surveillance of *Enterococcus faecium* Reveals Limited Sharing of Strains and Resistance Genes between Livestock and Humans in the United Kingdom. *MBio*, *9*(6), e01780-18. https://doi.org/10.1128/mBio.01780-18

Gouliouris, T., Raven, K. E., Moradigaravand, D., Ludden, C., Coll, F., Blane, B., Naydenova, P., Horner, C., Brown, N. M., Corander, J., Limmathurotsakul, D., Parkhill, J., & Peacock, S. J. (2019). Detection of vancomycin-resistant *Enterococcus faecium* hospital-adapted lineages in municipal wastewater treatment plants indicates widespread distribution and release into the environment. *Genome Research*, *29*(4), 626–634. https://doi.org/10.1101/gr.232629.117

Græsbøll, K., Nielsen, S. S., Toft, N., & Christiansen, L. E. (2014). How fitness reduced, antimicrobial resistant bacteria survive and spread: A multiple pig -

multiple bacterial strain model. *PLoS ONE*, *9*(7). https://doi.org/10.1371/journal.pone.0100458

Granato, E. T., Meiller-Legrand, T. A., & Foster, K. R. (2019). The Evolution and Ecology of Bacterial Warfare. *Current Biology*, *29*(11), R521–R537. https://doi.org/10.1016/j.cub.2019.04.024

Grassotti, T. T., Zvoboda, D. de A., Xavier Costa, L. da F., de Araujo, A. J., Pereira, R. I., Soares, R. O., Carniel Wagner, P. G., Frazzon, J., & Guedes Frazzon, A. P. (2018). Antimicrobial Resistance Profiles in *Enterococcus* spp. Isolates From Fecal Samples of Wild and Captive Black Capuchin Monkeys (Sapajus nigritus) in South Brazil. *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.02366

Greig, J., Rajić, A., Young, I., Mascarenhas, M., Waddell, L., & LeJeune, J. (2015). A scoping review of the role of wildlife in the transmission of bacterial pathogens and antimicrobial resistance to the food chain. *Zoonoses and Public Health*, *62*(4), 269–284.

Grimes, D. A., & Schulz, K. F. (2002). Descriptive studies: what they can and cannot do. *The Lancet*, *359*(9301), 145–149.

Guardabassi, L., & Dalsgaard, A. (2004). Occurrence, structure, and mobility of Tn1546-like elements in environmental isolates of vancomycin-resistant enterococci. *Applied and Environmental Microbiology*, *70*(2), 984–990. https://doi.org/10.1128/AEM.70.2.984-990.2004

Guardabassi, L., Petersen, A., Olsen, J. E., & Dalsgaard, A. (1998). Antibiotic resistance in *Acinetobacter* spp. isolated from sewers receiving waste effluent from a hospital and a pharmaceutical plant. *Applied and Environmental Microbiology*, *64*(9), 3499–3502. https://doi.org/10.1128/aem.64.9.3499-3502.1998

Gundogdu, A., Kilic, H., Kurekci, C., & Alp, E. (2017). Insufficient decontamination in Sewage treatment plants induce the risk of artificial selection of extended-spectrum beta-lactamase producing *Escherichia coli*. *Cellular and Molecular Biology*, *63*(9), 80–86. https://doi.org/10.14715/cmb/2017.63.9.15

Guo, B., Liu, C., Gibson, C., & Frigon, D. (2019). Wastewater microbial community structure and functional traits change over short timescales. *Science of the Total Environment*, *662*, 779–785. https://doi.org/10.1016/j.scitotenv.2019.01.207

Hadfield, J. D. (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software*, *33*(2), 1–22. http://www.jstatsoft.org/v33/i02/

Harnisz, M., Kiedrzyńska, E., Kiedrzyński, M., Korzeniewska, E., Czatzkowska, M., Koniuszewska, I., Jóźwik, A., Szklarek, S., Niestępski, S., & Zalewski, M.

(2020). The impact of WWTP size and sampling season on the prevalence of antibiotic resistance genes in wastewater and the river system. *Science of the Total Environment*, *741*. https://doi.org/10.1016/j.scitotenv.2020.140466

Harris, S., Morris, C., Morris, D., Cormican, M., & Cummins, E. (2014). Antimicrobial resistant *Escherichia coli* in the municipal wastewater system: effect of hospital effluent and environmental fate. *Science of the Total Environment*, *468–469*, 1078–1085. https://doi.org/10.1016/j.scitotenv.2013.09.017

Hassoun-Kheir, N., Stabholz, Y., Kreft, J.-U. U., de la Cruz, R., Romalde, J. L., Nesme, J., Sørensen, S. J., Smets, B. F., Graham, D., & Paul, M. (2020). Comparison of antibiotic-resistant bacteria and antibiotic resistance genes abundance in hospital and community wastewater: A systematic review. *Science of the Total Environment*, *743*, 140804. https://doi.org/10.1016/j.scitotenv.2020.140804

Hendriksen, R. S., Munk, P., Njage, P., van Bunnik, B., McNally, L., Lukjancenko, O., Röder, T., Nieuwenhuijse, D., Pedersen, S. K., Kjeldgaard, J., Kaas, R. S., Clausen, P. T. L. C., Vogt, J. K., Leekitcharoenphon, P., van de Schans, M. G. M. M., Zuidema, T., de Roda Husman, A. M., Rasmussen, S., Petersen, B., The Global Sewage Surveillance project consortium, Amid, C., Cochrane, G., Sicheritz-Ponten, T., Schmitt, H., Alvarez, J. R. M., Aidara-Kane, A., Pamp, S. J., Lund, O., Hald, T., Woolhouse, M., Koopmans, M. P., Vigre, H., Petersen, T. N., & Aarestrup, F. M. (2019). Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nature Communications*, *10*(1), 1124. https://doi.org/10.1038/s41467-019-08853-3

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L., & Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ (Online)*, *343*(7829), 1–9. https://doi.org/10.1136/bmj.d5928

Hocquet, D., Muller, A., & Bertrand, X. (2016). What happens in hospitals does not stay in hospitals: antibiotic-resistant bacteria in hospital wastewater systems. *Journal of Hospital Infection*, *93*(4), 395–402. https://doi.org/10.1016/j.jhin.2016.01.010

Hu, Y. J., Cowling, B. J., & Fukuda, K. (2018). What is missing in surveillance for control of antimicrobial resistance? *The Lancet Infectious Diseases*, *18*(6), 597–598. https://doi.org/10.1016/S1473-3099(18)30294-9

Huijbers, P., Flach, C. F., & Larsson, D. G. J. (2019). A conceptual framework for the environmental surveillance of antibiotics and antibiotic resistance. *Environment International*, *130*. https://doi.org/10.1016/j.envint.2019.05.074

Huijbers, P. M. C. C., Larsson, D. G. J. J., & Flach, C.-F. F. (2020). Surveillance of antibiotic resistant Escherichia coli in human populations through urban

wastewater in ten European countries. *Environmental Pollution*, *261*, 114200. https://doi.org/10.1016/j.envpol.2020.114200

Hutchings, M., Truman, A., & Wilkinson, B. (2019). Antibiotics: past, present and future. *Current Opinion in Microbiology*, *51*(Figure 1), 72–80. https://doi.org/10.1016/j.mib.2019.10.008

Hutinel, M., Fick, J., Larsson, D. G. J., & Flach, C.-F. (2021). Investigating the effects of municipal and hospital wastewaters on horizontal gene transfer. *Environmental Pollution*, *276*. https://doi.org/10.1016/j.envpol.2021.116733

Hutinel, M., Huijbers, P. M. C., Fick, J., Åhrén, C., Larsson, D. G. J., & Flach, C. F. (2019). Population-level surveillance of antibiotic resistance in *Escherichia coli* through sewage analysis. *Eurosurveillance*, *24*(37), 1–11. https://doi.org/10.2807/1560-7917.ES.2019.24.37.1800497

Imran, M., Das, K. R., & Naik, M. M. (2019). Co-selection of multi-antibiotic resistance in bacterial pathogens in metal and microplastic contaminated environments: An emerging health threat. *Chemosphere*, *215*, 846–857. https://doi.org/10.1016/j.chemosphere.2018.10.114

Isaksson, F., Lundy, L., Hedström, A., Székely, A. J., & Mohamed, N. (2022). Evaluating the Use of Alternative Normalization Approaches on SARS-CoV-2 Concentrations in Wastewater: Experiences from Two Catchments in Northern Sweden. *Environments - MDPI*, *9*(3). https://doi.org/10.3390/environments9030039

Iskandar, K., Molinier, L., Hallit, S., Sartelli, M., Hardcastle, T. C., Haque, M., Lugova, H., Dhingra, S., Sharma, P., Islam, S., Mohammed, I., Naina Mohamed, I., Hanna, P. A., Hajj, S. El, Jamaluddin, N. A. H., Salameh, P., & Roques, C. (2021). Surveillance of antimicrobial resistance in low- and middle-income countries: a scattered picture. *Antimicrobial Resistance and Infection Control*, *10*(1), 1–19. https://doi.org/10.1186/s13756-021-00931-w

Iversen, A., Kühn, I., Rahman, M., Franklin, A., Burman, L. G., Olsson-Liljequist, B., Torell, E., & Möllby, R. (2004). Evidence for transmission between humans and the environment of a nosocomial strain of *Enterococcus faecium*. *Environmental Microbiology*, *6*(1), 55–59. https://doi.org/10.1046/j.1462-2920.2003.00534.x

Jaimee, G., & Halami, P. M. (2016). High level aminoglycoside resistance in *Enterococcus*, *Pediococcus* and *Lactobacillus* species from farm animals and commercial meat products. *Annals of Microbiology*, *66*(1), 101–110. https://doi.org/10.1007/s13213-015-1086-1

Jakobsen, L., Sandvang, D., Hansen, L. H., Bagger-Skjøt, L., Westh, H., Jørgensen, C., Hansen, D. S., Pedersen, B. M., Monnet, D. L., Frimodt-Møller, N., Sørensen, S. J., & Hammerum, A. M. (2008). Characterisation, dissemination and persistence of gentamicin resistant *Escherichia coli* from a Danish university

hospital to the waste water environment. *Environment International*, *34*(1), 108–115. https://doi.org/10.1016/j.envint.2007.07.011

Jensen, H. S., Sekar, R., Shepherd, W. J., Osborn, A. M., Tait, S., & Biggs, C. A. (2016). Spatial and temporal variability of bacterial communities within a combined sewer system. *MicrobiologyOpen*, *5*(4), 616–625. https://doi.org/10.1002/mbo3.356

Jit, M., Ng, D. H. L., Luangasanatip, N., Sandmann, F., Atkins, K. E., Robotham, J. V., & Pouwels, K. B. (2020). Quantifying the economic cost of antibiotic resistance and the impact of related interventions: Rapid methodological review, conceptual framework and recommendations for future studies. *BMC Medicine*, *18*(1), 1–14. https://doi.org/10.1186/s12916-020-1507-2

Jonsson, V., Österlund, T., Nerman, O., & Kristiansson, E. (2019). Modelling of zero-inflation improves inference of metagenomic gene count data. *Statistical Methods in Medical Research*, *28*(12), 3712–3728. https://doi.org/10.1177/0962280218811354

Kahlmeter, G., Menday, P., & Cars, O. (2003). Non-hospital antimicrobial usage and resistance in community-acquired *Escherichia coli* urinary tract infection. *Journal of Antimicrobial Chemotherapy*, *52*(6), 1005–1010. https://doi.org/10.1093/jac/dkg488

Karanika, S., Karantanos, T., Arvanitis, M., Grigoras, C., & Mylonakis, E. (2016). Fecal Colonization with Extended-spectrum Beta-lactamase-Producing *Enterobacteriaceae* and Risk Factors among Healthy Individuals: A Systematic Review and Metaanalysis. *Clinical Infectious Diseases*, *63*(3), 310–318. https://doi.org/10.1093/cid/ciw283

Karkman, A., Berglund, F., Flach, C.-F. F., Kristiansson, E., & Larsson, D. G. J. J. (2020). Predicting clinical resistance prevalence using sewage metagenomic data. *Communications Biology*, *3*(1), 711. https://doi.org/10.1038/s42003-020-01439-6

Khan, F. A., Hellmark, B., Ehricht, R., Söderquist, B., & Jass, J. (2018). Related carbapenemase-producing Klebsiella isolates detected in both a hospital and associated aquatic environment in Sweden. *European Journal of Clinical Microbiology and Infectious Diseases*, *37*(12), 2241–2251. https://doi.org/10.1007/s10096-018-3365-9

Khan, M. T., Shah, I. A., Ihsanullah, I., Naushad, M., Ali, S., Shah, S. H. A., & Mohammad, A. W. (2021). Hospital wastewater as a source of environmental contamination: An overview of management practices, environmental risks, and treatment processes. *Journal of Water Process Engineering*, *41*. https://doi.org/10.1016/j.jwpe.2021.101990

Kim, S., Covington, A., & Pamer, E. G. (2017). The intestinal microbiota: Antibiotics, colonization resistance, and enteric pathogens. *Immunological Reviews*, *279*(1), 90–105. https://doi.org/10.1111/imr.12563

King, T. L. B., Schmidt, S., & Essack, S. Y. (2020). Antibiotic resistant *Kiebsiella* spp. from a hospital, hospital effluents and wastewater treatment plants in the uMgungundlovu District, KwaZulu-Natal, South Africa. *Science of the Total Environment*, *712*, 135550. https://doi.org/10.1016/j.scitotenv.2019.135550

Kizny Gordon, A. E., Mathers, A. J., Cheong, E. Y. L., Gottlieb, T., Kotay, S., Walker, A. S., Peto, T. E. A., Crook, D. W., & Stoesser, N. (2017). The Hospital Water Environment as a Reservoir for Carbapenem-Resistant Organisms Causing Hospital-Acquired Infections - A Systematic Review of the Literature. In *Clinical Infectious Diseases* (Vol. 64, Issue 10, pp. 1436–1444). Oxford University Press. https://doi.org/10.1093/cid/cix132

Klein, E. Y., Van Boeckel, T. P., Martinez, E. M., Pant, S., Gandra, S., Levin, S. A., Goossens, H., & Laxminarayan, R. (2018). Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(15), E3463–E3470. https://doi.org/10.1073/pnas.1717295115

Knight, G. M., Costelloe, C., Deeny, S. R., Moore, L. S. P., Hopkins, S., Johnson, A. P., Robotham, J. V., & Holmes, A. H. (2018). Quantifying where human acquisition of antibiotic resistance occurs: A mathematical modelling study. *BMC Medicine*, *16*(1), 1–11. https://doi.org/10.1186/s12916-018-1121-8

Knight, G. M., Costelloe, C., Murray, K. A., Robotham, J. V., Atun, R., & Holmes, A. H. (2018). Addressing the Unknowns of Antimicrobial Resistance: Quantifying and Mapping the Drivers of Burden. *Clinical Infectious Diseases*, *66*(4), 612–616. https://doi.org/10.1093/cid/cix765

Knight, G. M., Davies, N. G., Colijn, C., Coll, F., Donker, T., Gifford, D. R., Glover, R. E., Jit, M., Klemm, E., Lehtinen, S., Lindsay, J. A., Lipsitch, M., Llewelyn, M. J., Mateus, A. L. P., Robotham, J. V., Sharland, M., Stekel, D., Yakob, L., & Atkins, K. E. (2019). Mathematical modelling for antibiotic resistance control policy: do we know enough? *BMC Infectious Diseases*, *19*(1), 1011. https://doi.org/10.1186/s12879-019-4630-y

Knudsen, B. E., Bergmark, L., Munk, P., Lukjancenko, O., Priemé, A., Aarestrup, F. M., & Pamp, S. J. (2016). Impact of sample type and DNA isolation procedure on genomic inference of microbiome composition. *mSystems*, *1*(5). https://doi.org/10.1128/mSystems.00095-16

Ko, K. K. K., Chng, K. R., & Nagarajan, N. (2022). Metagenomics-enabled microbial surveillance. *Nature Microbiology*, *7*(4), 486–496. https://doi.org/10.1038/s41564-022-01089-w

Koenraad, P. M. F. J., Jacobs-Reitsma, W. F., Van Der Laan, T., Beumer, R. R., & Rombouts, F. M. (1995). Antibiotic susceptibility of campylobacter isolates from sewage and poultry abattoir drain water. *Epidemiology and Infection*, *115*(3), 475–483. https://doi.org/10.1017/S0950268800058635

Korzeniewska, E., Korzeniewska, A., & Harnisz, M. (2013). Antibiotic resistant *Escherichia coli* in hospital and municipal sewage and their emission to the environment. *Ecotoxicology and Environmental Safety*, *91*, 96–102. https://doi.org/10.1016/j.ecoenv.2013.01.014

Kouyos, R. D., zur Wiesch, P. A., & Bonhoeffer, S. (2011). On being the right size: The impact of population size and stochastic effects on the evolution of drug resistance in hospitals and the community. *PLoS Pathogens*, *7*(4). https://doi.org/10.1371/journal.ppat.1001334

Kraupner, N., Hutinel, M., Schumacher, K., Gray, D. A., Genheden, M., Fick, J., Flach, C. F., & Larsson, D. G. J. (2021). Evidence for selection of multi-resistant *E. coli* by hospital effluent. *Environment International*, *150*. https://doi.org/10.1016/j.envint.2021.106436

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, *28*(5), 1–26. https://doi.org/10.18637/jss.v028.i05

Kumari, A., Maurya, N. S., & Tiwari, B. (2020). Hospital wastewater treatment scenario around the globe. In *Current Developments in Biotechnology and Bioengineering* (Issue January).

Kutilova, I., Medvecky, M., Leekitcharoenphon, P., Munk, P., Masarikova, M., Davidova-Gerzova, L., Jamborova, I., Bortolaia, V., Pamp, S. J., & Dolejska, M. (2021). Extended-spectrum beta-lactamase-producing *Escherichia coli* and antimicrobial resistance in municipal and hospital wastewaters in Czech Republic: Culture-based and metagenomic approaches. *Environmental Research*, *193*(October 2020), 110487. https://doi.org/10.1016/j.envres.2020.110487

Kwak, Y. K., Colque, P., Byfors, S., Giske, C. G., Möllby, R., Kühn, I., Mllby, R., & Khn, I. (2015). Surveillance of antimicrobial resistance among *Escherichia coli* in wastewater in Stockholm during 1 year: Does it reflect the resistance trends in the society? *International Journal of Antimicrobial Agents*, *45*(1), 25–32. https://doi.org/10.1016/j.ijantimicag.2014.09.016

Lamba, M., Gupta, S., Shukla, R., Graham, D. W., Sreekrishnan, T. R., & Ahammad, S. Z. (2018). Carbapenem resistance exposures via wastewaters across New Delhi. *Environment International*, *119*, 302–308. https://doi.org/10.1016/j.envint.2018.07.004

Lamberte, L. E., & van Schaik, W. (2022). Antibiotic resistance in the commensal human gut microbiota. *Current Opinion in Microbiology*, *68*, 102150. https://doi.org/10.1016/j.mib.2022.102150

Lambourg, E., Siani, C., & de Preux, L. (2022). Use of a high-volume prescription database to explore health inequalities in England: assessing impacts of social deprivation and temperature on the prescription volume of medicines. *Journal of Public Health (Germany)*. https://doi.org/10.1007/s10389-021-01691-y

Larsson, D. G. J. J., Andremont, A., Bengtsson-Palme, J., Brandt, K. K., Husman, A. M. de R., Fagerstedt, P., Fick, J., Flach, C.-F. F., Gaze, W. H., Kuroda, M., Kvint, K., Laxminarayan, R., Manaia, C. M., Nielsen, K. M., Plant, L., Ploy, M.-C. C., Segovia, C., Simonet, P., Smalla, K., Snape, J., Topp, E., van Hengel, A. J., Verner-Jeffreys, D. W., Virta, M. P. J., Wellington, E. M., & Wernersson, A.-S. S. (2018). Critical knowledge gaps and research needs related to the environmental dimensions of antibiotic resistance. *Environment International*, *117*(May), 132–138. https://doi.org/10.1016/j.envint.2018.04.041

Larsson, D. G. J., & Flach, C. F. (2022). Antibiotic resistance in the environment. *Nature Reviews Microbiology*, *20*(5), 257–269. https://doi.org/10.1038/s41579-021-00649-x

Lekunberri, I., Balcazar, J. L., & Borrego, C. M. (2018). Metagenomic exploration reveals a marked change in the river resistome and mobilome after treated wastewater discharges. *Environmental Pollution*, *234*, 538–542. https://doi.org/10.1016/j.envpol.2017.12.001

Leonard, A. F. C., Zhang, L., Balfour, A. J., Garside, R., & Gaze, W. H. (2015). Human recreational exposure to antibiotic resistant bacteria in coastal bathing waters. *Environment International*, *82*, 92–100. https://doi.org/10.1016/j.envint.2015.02.013

Li, B., Zhao, Z. C., Wang, M. H., Huang, X. H., Pan, Y. H., & Cao, Y. P. (2014). Antimicrobial resistance and integrons of commensal Escherichia coli strains from healthy humans in China. *Journal of Chemotherapy*, *26*(3), 190–192. https://doi.org/10.1179/1973947813Y.0000000113

Li, X. Z., Plésiat, P., & Nikaido, H. (2015). The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. *Clinical Microbiology Reviews*, *28*(2), 337–418. https://doi.org/10.1128/CMR.00117-14

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*(3), 18–22.

Linton, K. B., Richmond, M. H., Bevan, R., & Gillespie, W. A. (1974). Antibiotic resistance and R factors in coliform bacilli isolated from hospital and domestic sewage. *Journal of Medical Microbiology*, *7*(1), 91–103. https://doi.org/10.1099/00222615-7-1-91

Llarrull, L. I., Testero, S. A., Fisher, J. F., & Mobashery, S. (2010). The future of the β-lactams. *Current Opinion in Microbiology*, *13*(5), 551–557. https://doi.org/10.1016/j.mib.2010.09.008

Lloyd-Price, J., Mahurkar, A., Rahnavard, G., Crabtree, J., Orvis, J., Hall, A. B., Brady, A., Creasy, H. H., McCracken, C., Giglio, M. G., McDonald, D., Franzosa, E. A., Knight, R., White, O., & Huttenhower, C. (2017). Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature*, *550*(7674), 61–66. https://doi.org/10.1038/nature23889

Lowe, P. (2020). *Annual Report of the Registrar General of Births, Deaths and Marriages for Scotland 2019 165th Edition*.

Ludden, C., Moradigaravand, D., Jamrozy, D., Gouliouris, T., Blane, B., Naydenova, P., Hernandez-Garcia, J., Wood, P., Hadjirin, N., Radakovic, M., Crawley, C., Brown, N. M., Holmes, M., Parkhill, J., & Peacock, S. J. (2020). A one health study of the genetic relatedness of *Klebsiella pneumoniae* and their mobile elements in the east of England. *Clinical Infectious Diseases*, *70*(2), 219–226. https://doi.org/10.1093/cid/ciz174

Ludden, C., Reuter, S., Judge, K., Gouliouris, T., Blane, B., Coll, F., Naydenova, P., Hunt, M., Tracey, A., Hopkins, K. L., Brown, N. M., Woodford, N., Parkhill, J., & Peacock, S. J. (2017). Sharing of carbapenemase-encoding plasmids between *Enterobacteriaceae* in UK sewage uncovered by MinION sequencing. *Microbial Genomics*, *3*(7). https://doi.org/10.1099/mgen.0.000114

Majeed, H. J., Riquelme, M. V., Davis, B. C., Gupta, S., Angeles, L., Aga, D. S., Garner, E., Pruden, A., & Vikesland, P. J. (2021). Evaluation of Metagenomic-Enabled Antibiotic Resistance Surveillance at a Conventional Wastewater Treatment Plant. *Frontiers in Microbiology*, *12*(May), 657954. https://doi.org/10.3389/fmicb.2021.657954

Majumder, A., Gupta, A. K., Ghosal, P. S., & Varma, M. (2021). A review on hospital wastewater treatment: A special emphasis on occurrence and removal of pharmaceutically active compounds, resistant microorganisms, and SARS-CoV-2. *Journal of Environmental Chemical Engineering*, *9*(2). https://doi.org/10.1016/j.jece.2020.104812

Majumder, M. A. A., Rahman, S., Cohall, D., Bharatha, A., Singh, K., Haque, M., & Gittens-St Hilaire, M. (2020). Antimicrobial stewardship: Fighting antimicrobial resistance and protecting global public health. *Infection and Drug Resistance*, *13*, 4713–4738. https://doi.org/10.2147/IDR.S290835

Malcolm, W., Seaton, R. A., Haddock, G., Baxter, L., Thirlwell, S., Russell, P., Cooper, L., Thomson, A., & Sneddon, J. (2020). Impact of the COVID-19 pandemic on community antibiotic prescribing in Scotland. *JAC-Antimicrobial Resistance*, *2*(4), 1–4. https://doi.org/10.1093/jacamr/dlaa105

Manaia, C. M. (2014). Antibiotic resistance in wastewater: origins, fate, and risks. *Prävention Und Gesundheitsförderung*, *9*(3), 180–184. https://doi.org/10.1007/s11553-014-0452-3

Masarikova, M., Manga, I., Cizek, A., Dolejska, M., Oravcova, V., Myskova, P., Karpiskova, R., & Literak, I. (2016). Salmonella enterica resistant to antimicrobials in wastewater effluents and black-headed gulls in the Czech Republic, 2012. *Science of the Total Environment*, *542*(Pt A), 102–107. https://doi.org/10.1016/j.scitotenv.2015.10.069

Massot, M., Daubié, A.-S., Clermont, O., Jauréguy, F., Couffignal, C., Dahbi, G., Mora, A., Blanco, J., Branger, C., Mentré, F., Eddi, A., Picard, B., Denamur, E., & Group, T. C. (2016). Phylogenetic, virulence and antibiotic resistance characteristics of commensal strain populations of *Escherichia coli* from community subjects in the Paris area in 2010 and evolution over 30 years. *Microbiology (Reading, England)*, *162*(4), 642–650. https://doi.org/10.1099/mic.0.000242

Matsuki, K., Kuperman, V., & Van Dyke, J. A. (2016). The Random Forests statistical technique: An examination of its value for the study of reading. *Scientific Studies of Reading*, *20*(1), 20–33. https://doi.org/10.1080/10888438.2015.1107073

Mbanga, J., Amoako, D. G., Abia, A. L. K., Allam, M., Ismail, A., & Essack, S. Y. (2021). Genomic Insights of Multidrug-Resistant *Escherichia coli* From Wastewater Sources and Their Association With Clinical Pathogens in South Africa. *Frontiers in Veterinary Science*, *8*(February), 636715. https://doi.org/10.3389/fvets.2021.636715

McArthur, A. G., & Tsang, K. K. (2017). Antimicrobial resistance surveillance in the genomic age. *Annals of the New York Academy of Sciences*, *1388*(1), 78–91. https://doi.org/10.1111/nyas.13289

McBryde, E. S., & McElwain, D. L. S. (2006). A Mathematical Model Investigating the Impact of an Environmental Reservoir on the Prevalence and Control of Vancomycin- Resistant Enterococci. *The Journal of Infectious Diseases*, *193*, 1473–1474.

McDonnell, L., Armstrong, D., Ashworth, M., Dregan, A., Malik, U., & White, P. (2017). National disparities in the relationship between antimicrobial resistance and antimicrobial consumption in Europe: An observational study in 29 Countries. *Journal of Antimicrobial Chemotherapy*, *72*(11), 292–300. https://doi.org/10.1093/jac/dkx248

McEwen, S. A., & Collignon, P. J. (2018). Antimicrobial Resistance: a One Health Perspective. *Microbiology Spectrum*, *6*(2), 255–260. https://doi.org/10.1128/microbiolspec.ARBA-0009-2017

McLellan, S. L., & Roguet, A. (2019). The unexpected habitat in sewer pipes for the propagation of microbial communities and their imprint on urban waters. *Current Opinion in Biotechnology*, *57*, 34–41. https://doi.org/10.1016/j.copbio.2018.12.010

Messai, Y., Atmani, S. M., Alouache, S., Fernández, R., Estepa, V., Torres, C., & Bakour, R. (2015). Virulence characteristics and genetic background of ESBL-producing *Klebsiella pneumoniae* isolates from wastewater. *Fresenius Environmental Bulletin*, *24*(1), 103–112. https://www.researchgate.net/publication/281893715

Mhongole, O. J., Mdegela, R. H., Kusiluka, L. J. M., Forslund, A., & Dalsgaard, A. (2017). Characterization of Salmonella spp. from wastewater used for food production in Morogoro, Tanzania. *World Journal of Microbiology & Biotechnology*, *33*(3), 42. https://doi.org/10.1007/s11274-017-2209-6

Miłobedzka, A., Ferreira, C., Vaz-Moreira, I., Calderón-Franco, D., Gorecki, A., Purkrtova, S., Jan Bartacek, Dziewit, L., Singleton, C. M., Nielsen, P. H., Weissbrodt, D. G., & Manaia, C. M. (2022). Monitoring antibiotic resistance genes in wastewater environments: The challenges of filling a gap in the One-Health cycle. *Journal of Hazardous Materials*, *424*. https://doi.org/10.1016/j.jhazmat.2021.127407

Morales Medina, W. R., Eramo, A., Tu, M., & Fahrenfeld, N. L. (2020). Sewer biofilm microbiome and antibiotic resistance genes as function of pipe material, source of microbes, and disinfection: Field and laboratory studies. *Environmental Science: Water Research and Technology*, *6*(8), 2122–2137. https://doi.org/10.1039/d0ew00265h

Moriarty, E., Nourozi, F., Robson, B., Wood, D., & Gilpin, B. (2008). Evidence for growth of enterococci in municipal oxidation ponds, obtained using antibiotic resistance analysis. *Applied and Environmental Microbiology*, *74*(23), 7204–7210. https://doi.org/10.1128/AEM.00341-08

Morley, V. J., Woods, R. J., & Read, A. F. (2019). Bystander Selection for Antimicrobial Resistance: Implications for Patient Health. *Trends in Microbiology*, *27*(10), 864–877. https://doi.org/10.1016/j.tim.2019.06.004

Muloi, D., Ward, M. J., Pedersen, A. B., Fèvre, E. M., Woolhouse, M. E. J., & van Bunnik, B. A. D. (2018). Are Food Animals Responsible for Transfer of Antimicrobial-Resistant *Escherichia coli* or Their Resistance Determinants to Human Populations? A Systematic Review. *Foodborne Pathogens and Disease*, *15*(8), 467–474. https://doi.org/10.1089/fpd.2017.2411

Munita, J. M., & Arias, C. A. (2016). Mechanisms of Antibiotic Resistance. *Microbiology Spectrum*, *4*(2), 119–127. https://doi.org/https://doi.org/10.1128/microbiolspec.VMBF-0016-2015

Munk-Welford, P., Brinch, C., Møller, F. D., Petersen, T. N., Hendriksen, R. S., Seyfarth, A. M., Kjeldgaard, J. S., Bunnik, B. van, Consortium, G. S., Larsson, D. G. J., Koopmans, M., Woolhouse, M., & Aarestrup, F. M. (2022). Global sewage metagenomics provides unparalleled insight into spatial, taxonomic, and temporal components of antimicrobial resistance. *In Prep.*

Murray, C. J., Ikuta, K. S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., Han, C., Bisignano, C., Rao, P., Wool, E., Johnson, S. C., Browne, A. J., Chipeta, M. G., Fell, F., Hackett, S., Haines-Woodhouse, G., Kashef Hamadani, B. H., Kumaran, E. A. P., McManigal, B., … Naghavi, M. (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet*, *399*(10325), 629–655. https://doi.org/10.1016/s0140-6736(21)02724-0

Nadimpalli, M. L., Marks, S. J., Montealegre, M. C., Gilman, R. H., Pajuelo, M. J., Saito, M., Tsukayama, P., Njenga, S. M., Kiiru, J., Swarthout, J., Islam, M. A., Julian, T. R., & Pickering, A. J. (2020). Urban informal settlements as hotspots of antimicrobial resistance and the need to curb environmental transmission. *Nature Microbiology*, *5*(6), 787–795. https://doi.org/10.1038/s41564-020-0722-0

Nadimpalli, M., Delarocque-Astagneau, E., Love, D. C., Price, L. B., Huynh, B. T., Collard, J. M., Lay, K. S., Borand, L., Ndir, A., Walsh, T. R., & Guillemot, D. (2018). Combating Global Antibiotic Resistance: Emerging One Health Concerns in Lower-and Middle-Income Countries. *Clinical Infectious Diseases*, *66*(6), 963–969. https://doi.org/10.1093/cid/cix879

Newton, R. J., McLellan, S. L., Dila, D. K., Vineis, J. H., Morrison, H. G., Murat Eren, A., & Sogin, M. L. (2015). Sewage reflects the microbiomes of human populations. *MBio*, *6*(2). https://doi.org/10.1128/mBio.02574-14

Ng, C., Tay, M., Tan, B., Le, T. H., Haller, L., Chen, H., Koh, T. H., Barkham, T. M. S., & Gin, K. Y. H. (2017). Characterization of metagenomes in urban aquatic compartments reveals high prevalence of clinically relevant antibiotic resistance genes in wastewaters. *Frontiers in Microbiology*, *8*(NOV), 2200. https://doi.org/10.3389/fmicb.2017.02200

Nguyen, A. Q., Vu, H. P., Nguyen, L. N., Wang, Q., Djordjevic, S. P., Donner, E., Yin, H., & Nghiem, L. D. (2021). Monitoring antibiotic resistance genes in wastewater treatment: Current strategies and future challenges. *Science of the Total Environment*, *783*, 146964. https://doi.org/10.1016/j.scitotenv.2021.146964

Niewiadomska, A. M., Jayabalasingham, B., Seidman, J. C., Willem, L., Grenfell, B., Spiro, D., & Viboud, C. (2019). Population-level mathematical modeling of antimicrobial resistance: A systematic review. *BMC Medicine*, *17*(1), 1–20. https://doi.org/10.1186/s12916-019-1314-9

NIH. (2019). *Prokaryotic RefSeq Genomes Frequently Asked Questions (FAQ)*.

Nji, E., Kazibwe, J., Hambridge, T., Joko, C. A., Larbi, A. A., Damptey, L. A. O., Nkansa-Gyamfi, N. A., Stålsby Lundborg, C., & Lien, L. T. Q. (2021). High prevalence of antibiotic resistance in commensal *Escherichia coli* from healthy human sources in community settings. *Scientific Reports*, *11*(1), 1–11. https://doi.org/10.1038/s41598-021-82693-4

Novo, A., & Manaia, C. M. C. M. (2010). Factors influencing antibiotic resistance burden in municipal wastewater treatment plants. *Applied Microbiology and Biotechnology*, *87*(3), 1157–1166. https://doi.org/10.1007/s00253-010-2583-6

O'Neill, J. (2014). *Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations, 2014*. https://doi.org/10.1038/510015a

Oravcova, V., Mihalcin, M., Zakova, J., Pospisilova, L., Masarikova, M., & Literak, I. (2017). Vancomycin-resistant enterococci with vanA gene in treated municipal wastewater and their association with human hospital strains. *Science of the Total Environment*, *609*, 633–643. https://doi.org/10.1016/j.scitotenv.2017.07.121

Ory, J. J., Bricheux, G. G., Togola, A., Bonnet, J. L., Donnadieu-Bernard, F., Nakusi, L., Forestier, C., & Traore, O. (2016). Ciprofloxacin residue and antibiotic-resistant biofilm bacteria in hospital effluent. *Environmental Pollution*, *214*, 635–645. https://doi.org/10.1016/j.envpol.2016.04.033

Otter, J. A., Natale, A., Batra, R., Tosas Auguet, O., Dyakova, E., Goldenberg, S. D., & Edgeworth, J. D. (2019). Individual- and community-level risk factors for ESBL Enterobacteriaceae colonization identified by universal admission screening in London. *Clinical Microbiology and Infection*, *25*(10), 1259–1265. https://doi.org/10.1016/j.cmi.2019.02.026

Pacifici, K., Reich, B. J., Miller, D. A. W., Gardner, B., Stauffer, G., Singh, S., McKerrow, A., & Collazo, J. A. (2017). Integrating multiple data sources in species distribution modeling: A framework for data fusion. *Ecology*, *98*(3), 840–850. https://doi.org/10.1002/ecy.1710

Palacios, O. A., Contreras, C. A., Munoz-Castellanos, L. N., Gonzalez-Rangel, M. O., Rubio-Arias, H., Palacios-Espinosa, A., & Nevarez-Moorillon, G. V. (2017). Monitoring of indicator and multidrug resistant bacteria in agricultural soils under different irrigation patterns. *Agricultural Water Management*, *184*, 19–27. https://doi.org/10.1016/j.agwat.2017.01.001

Palacios, O. A., Contreras, C. A., Muñoz-Castellanos, L. N., González-Rangel, M. O., Rubio-Arias, H., Palacios-Espinosa, A., & Nevárez-Moorillón, G. V. (2017). Monitoring of indicator and multidrug resistant bacteria in agricultural soils under different irrigation patterns. *Agricultural Water Management*, *184*, 19–27. https://doi.org/10.1016/j.agwat.2017.01.001

Palmer, A. C., & Kishony, R. (2013). Understanding, predicting and manipulating the genotypic evolution of antibiotic resistance. *Nature Reviews Genetics*, *14*(4), 243–248. https://doi.org/10.1038/nrg3351

Pärnänen, K. M. M., Narciso-Da-Rocha, C., Kneis, D., Berendonk, T. U., Cacace, D., Do, T. T., Elpers, C., Fatta-Kassinos, D., Henriques, I., Jaeger, T., Karkman, A., Martinez, J. L., Michael, S. G., Michael-Kordatou, I., O'Sullivan, K., Rodriguez-Mozaz, S., Schwartz, T., Sheng, H., Sørum, H., Stedtfeld, S. R., Tiedje, J. M., Giustina, S. V. D., Walsh, F., Vaz-Moreira, I., Virta, M., & Manaia, C. M. (2019). Antibiotic resistance in European wastewater treatment plants mirrors the pattern of clinical antibiotic resistance prevalence. *Science Advances*, *5*(3), eaau9124. https://doi.org/10.1126/sciadv.aau9124

Paulshus, E., Kühn, I., Möllby, R., Colque, P., O'Sullivan, K., Midtvedt, T., Lingaas, E., Holmstad, R., & Sørum, H. (2019). Diversity and antibiotic resistance among *Escherichia coli* populations in hospital and community wastewater compared to wastewater at the receiving urban treatment plant. *Water Research*, *161*, 232–241. https://doi.org/10.1016/j.watres.2019.05.102

Pazda, M., Kumirska, J., Stepnowski, P., & Mulkiewicz, E. (2019). Antibiotic resistance genes identified in wastewater treatment plant systems - A review. *Science of the Total Environment*, *697*, 134023. https://doi.org/10.1016/j.scitotenv.2019.134023

Pehrsson, E. C., Tsukayama, P., Patel, S., Mejia-Bautista, M., Sosa-Soto, G., Navarrete, K. M., Calderon, M., Cabrera, L., Hoyos-Arango, W., Bertoli, M. T., Berg, D. E., Gilman, R. H., & Dantas, G. (2016). Interconnected microbiomes and resistomes in low-income human habitats. *Nature*, *533*(7602), 212–216. https://doi.org/10.1038/nature17672

Peng, X., Li, G., & Liu, Z. (2016). Zero-Inflated Beta Regression for Differential Abundance Analysis with Metagenomics Data. *Journal of Computational Biology*, *23*(2), 102–110. https://doi.org/10.1089/cmb.2015.0157

Perry, M. R., Lepper, H. C., McNally, L., Wee, B. A., Munk, P., Warr, A., Moore, B., Kalima, P., Philip, C., de Roda Husman, A. M., Aarestrup, F. M., Woolhouse, M. E. J., & van Bunnik, B. A. D. (2021). Secrets of the Hospital Underbelly: Patterns of Abundance of Antimicrobial Resistance Genes in Hospital Wastewater Vary by Specific Antimicrobial and Bacterial Family. *Frontiers in Microbiology*, *12*(September). https://doi.org/10.3389/fmicb.2021.703560

Petersen, T. N., Lukjancenko, O., Thomsen, M. C. F., Sperotto, M. M., Lund, O., Aarestrup, F. M., & Sicheritz-Ponten, T. (2017). MGmapper: Reference based mapping and taxonomy annotation of metagenomics sequence reads. *PLoS ONE*, *12*(5), 1–13. https://doi.org/10.1371/journal.pone.0176469

Peterson, E., & Kaur, P. (2018). Antibiotic resistance mechanisms in bacteria: Relationships between resistance determinants of antibiotic producers,

environmental bacteria, and clinical pathogens. *Frontiers in Microbiology*, *9*(NOV), 1–21. https://doi.org/10.3389/fmicb.2018.02928

Piotrowska, M., & Popowska, M. (2014). The prevalence of antibiotic resistance genes among Aeromonas species in aquatic environments. *Annals of Microbiology*, *64*(3), 921–934. https://doi.org/10.1007/s13213-014-0911-2

Polianciuc, S. I., Gurzău, A. E., Kiss, B., Georgia Ștefan, M., & Loghin, F. (2020). Antibiotics in the environment: causes and consequences. *Medicine and Pharmacy Reports*, *93*(3), 231–240. https://doi.org/10.15386/mpr-1742

Popa, L. I., Gheorghe, I., Barbu, I. C., Surleac, M., Paraschiv, S., Măruţescu, L., Popa, M., Pîrcălăbioru, G. G., Talapan, D., Niţă, M., Streinu-Cercel, A., Streinu-Cercel, A., Oţelea, D., & Chifiriuc, M. C. (2021). Multidrug Resistant *Klebsiella pneumoni*ae ST101 Clone Survival Chain From Inpatients to Hospital Effluent After Chlorine Treatment. *Frontiers in Microbiology*, *11*(January). https://doi.org/10.3389/fmicb.2020.610296

Pot, M., Guyomard-Rabenirina, S., Couvin, D., Ducat, C., Enouf, V., Ferdinand, S., Gruel, G., Malpote, E., Talarmin, A., Breurec, S., & Reynaud, Y. (2021). Dissemination of Extended-Spectrum-beta-Lactamase-Producing *Enterobacter cloacae* Complex from a Hospital to the Nearby Environment in Guadeloupe (French West Indies): ST114 Lineage Coding for a Successful IncHI2/ST1 Plasmid. *Antimicrobial Agents and Chemotherapy*, *65*(3). https://doi.org/10.1128/AAC.02146-20

Pradipta, I. S., Forsman, L. D., Bruchfeld, J., Hak, E., & Alffenaar, J. W. (2018). Risk factors of multidrug-resistant tuberculosis: A global systematic review and meta-analysis. *Journal of Infection*, *77*(6), 469–478. https://doi.org/10.1016/j.jinf.2018.10.004

Pruden, A., Arabi, M., & Storteboom, H. N. (2012). Correlation between upstream human activities and riverine antibiotic resistance genes. *Environmental Science and Technology*, *46*(21), 11541–11549. https://doi.org/10.1021/es302657r

Pruden, A., Joakim Larsson, D. G., Amézquita, A., Collignon, P., Brandt, K. K., Graham, D. W., Lazorchak, J. M., Suzuki, S., Silley, P., Snape, J. R., Topp, E., Zhang, T., & Zhu, Y. G. (2013). Management options for reducing the release of antibiotics and antibiotic resistance genes to the environment. *Environmental Health Perspectives*, *121*(8), 878–885. https://doi.org/10.1289/ehp.1206446

Pruden, A., Vikesland, P. J., Davis, B. C., & de Roda Husman, A. M. (2021). Seizing the moment: now is the time for integrated global surveillance of antimicrobial resistance in wastewater environments. In *Current Opinion in Microbiology* (Vol. 64, pp. 91–99). Elsevier Ltd. https://doi.org/10.1016/j.mib.2021.09.013

Public Health Scotland (2020). *Acute hospital activity and NHS beds information (quarterly)*. https://publichealthscotland.scot/publications/acute-hospital-activity-

and-nhs-beds-information-quarterly/acute-hospital-activity-and-nhs-beds-information-quarterly-quarter-ending-31-december-2019/data-files/. Accessed: 12/08/2021

Puchter, L., Chaberny, I. F., Schwab, F., Vonberg, R. P., Bange, F. C., & Ebadi, E. (2018). Economic burden of nosocomial infections caused by vancomycin-resistant enterococci. *Antimicrobial Resistance and Infection Control*, *7*(1), 1–7. https://doi.org/10.1186/s13756-017-0291-z

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), 590–596. https://doi.org/10.1093/nar/gks1219

Quintela-Baluja, M., Abouelnaga, M., Romalde, J., Su, J. Q., Yu, Y., Gomez-Lopez, M., Smets, B., Zhu, Y. G., & Graham, D. W. (2019). Spatial ecology of a wastewater network defines the antibiotic resistance genes in downstream receiving waters. *Water Research*, *162*, 347–357. https://doi.org/10.1016/j.watres.2019.06.075

R Core Team. (2022). *R: A Language and Environment for Statistical Computing* (3.6.3). R Foundation for Statistical Computing. https://www.r-project.org/

Raven, K. E., Ludden, C., Gouliouris, T., Blane, B., Naydenova, P., Brown, N. M., Parkhill, J., & Peacock, S. J. (2019). Genomic surveillance of *Escherichia coli* in municipal wastewater treatment plants as an indicator of clinically relevant pathogens and their resistance genes. *Microbial Genomics*, *5*(5), 1–9. https://doi.org/10.1099/mgen.0.000267

Reinthaler, F. F., Galler, H., Feierl, G., Haas, D., Leitner, E., Mascher, F., Melkes, A., Posch, J., Pertschy, B., Winter, I., Himmel, W., Marth, E., & Zarfel, G. (2013). Resistance patterns of *Escherichia coli* isolated from sewage sludge in comparison with those isolated from human patients in 2000 and 2009. *Journal of Water and Health*, *11*(1), 13–20. https://doi.org/10.2166/wh.2012.207

Remschmidt, C., Behnke, M., Kola, A., Peña Diaz, L. A., Rohde, A. M., Gastmeier, P., & Schwab, F. (2017). The effect of antibiotic use on prevalence of nosocomial vancomycin-resistant enterococci- an ecologic study. *Antimicrobial Resistance and Infection Control*, *6*(1), 1–8. https://doi.org/10.1186/s13756-017-0253-5

Reusser, D. (2015). *fast: Implementation of the Fourier Amplitude Sensitivity Test (FAST)*. https://cran.r-project.org/package=fast

Rizzo, L., Manaia, C., Merlin, C., Schwartz, T., Dagot, C., Ploy, M. C., Michael, I., & Fatta-Kassinos, D. (2013). Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: a review.

Science of the Total Environment, *447*, 345–360. https://doi.org/10.1016/j.scitotenv.2013.01.032

Robinson, T. P., Bu, D. P., Carrique-Mas, J., Fèvre, E. M., Gilbert, M., Grace, D., Hay, S. I., Jiwakanon, J., Kakkar, M., Kariuki, S., Laxminarayan, R., Lubroth, J., Magnusson, U., Ngoc, P. T., van Boeckel, T. P., & Woolhouse, M. E. J. (2016). Antibiotic resistance is the quintessential One Health issue. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *110*(7), 377–380.

Röderová, M., Sedláková, M. H., Pudová, V., Hricová, K., Silová, R., Imwensi, P. E. O., Bardoň, J., & Kolář, M. (2016). Occurrence of bacteria producing broad-spectrum beta-lactamases and qnr genes in hospital and urban wastewater samples. *New Microbiologica*, *39*(2), 124–133. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84971474293&partnerID=40&md5=fd06e2df1c3be4c75773601ccb0b802c

Röderová, M., Sedláková, M. H., Pudová, V., Hricová, K., Silová, R., Imwensi, P. E. O., Bardoň, J., & Kolář, M. (2016). Occurrence of bacteria producing broad-spectrum beta-lactamases and qnr genes in hospital and urban wastewater samples. *New Microbiologica*, *39*(2), 124–133. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84971474293&partnerID=40&md5=fd06e2df1c3be4c75773601ccb0b802c

Roldan-Masedo, E., Sainz, T., Gutierrez-Arroyo, A., Gomez-Gil, R. M., Ballesteros, E., Escosa, L., Baquero-Artigao, F., & Méndez-Echevarría, A. (2019). Risk factors for gentamicin-resistant *E. coli* in children with community-acquired urinary tract infection. *European Journal of Clinical Microbiology and Infectious Diseases*, *38*(11), 2097–2102. https://doi.org/10.1007/s10096-019-03643-7

Rooney, A. M., Raphenya, A. R., Melano, R. G., Seah, C., Yee, N. R., MacFadden, D. R., McArthur, A. G., Schneeberger, P. H. H., & Coburn, B. (2021). Performance characteristics of next-generation sequencing for antimicrobial resistance gene detection in genomes and metagenomes. *BioRxiv*, 2021.06.25.449921. https://www.biorxiv.org/content/10.1101/2021.06.25.449921v1%0Ahttps://www.biorxiv.org/content/10.1101/2021.06.25.449921v1.abstract

Rosenburg Goldstein, R., Kleinfelter, L., He, X., Micallef, S. A., George, A., Gibbs, S. G., & Sapkota, A. R. (2017). Higher prevalence of coagulase-negative staphylococci carriage among reclaimed water spray irrigators. *Science of the Total Environment*, *595*, 35–40. https://doi.org/10.1016/j.scitotenv.2017.03.174

Saltelli, A., Tarantola, S., & Chan, K. P.-S. (1999). A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. *Technometrics*, *41*(1), 39–56. https://doi.org/10.1080/00401706.1999.10485594

Sargeant, J. M., & O'Connor, A. M. (2014). Introduction to Systematic Reviews in Animal Agriculture and Veterinary Medicine. *Zoonoses and Public Health*, *61*(SUPPL1), 3–9. https://doi.org/10.1111/zph.12128

Scott, A. M., Beller, E., Glasziou, P., Clark, J., Ranakusuma, R. W., Byambasuren, O., Bakhit, M., Page, S. W., Trott, D., & Mar, C. Del. (2018). Is antimicrobial administration to food animals a direct threat to human health? A rapid systematic review. *International Journal of Antimicrobial Agents*, *52*(3), 316–323. https://doi.org/10.1016/j.ijantimicag.2018.04.005

Scottish Water. (n.d.). *Trade Effluent : Guide to Services*. https://www.scottishwater.co.uk/-/media/ScottishWater/Document-Hub/Business-and-Developers/Byelaws-and-Trade-Effluent/280521TradeEffluentGuidetoServices.pdf

Scottish Water. (n.d). *Surface Water Policy*. http://www.scottishwater.co.uk/-/media/business/files/connections-documents/developer-services/surfacewaterguidancedoc8ppa4pageshires.pdf?la=en

Seruga Music, M., Hrenovic, J., Goic-Barisic, I., Hunjak, B., Skoric, D., & Ivankovic, T. (2017). Emission of extensively-drug-resistant *Acinetobacter baumannii* from hospital settings to the natural environment. *The Journal of Hospital Infection*, *96*(4), 323–327. https://doi.org/10.1016/j.jhin.2017.04.005

Shenoy, E. S., Paras, M. L., Noubary, F., Walensky, R. P., & Hooper, D. C. (2014). Natural history of colonization with methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant Enterococcus (VRE): A systematic review. *BMC Infectious Diseases*, *14*(1). https://doi.org/10.1186/1471-2334-14-177

Sims, N., Avery, L., & Kasprzyk-Hordern, B. (2021). *Review of wastewater monitoring applications for public health and novel aspects of environmental quality*. Scotland's Centre for Expertise for Waters (CREW).

Smith, R., & Coast, J. (2013). The true cost of antimicrobial resistance. *BMJ (Online)*, *346*(7899), 1–5. https://doi.org/10.1136/bmj.f1493

Song, S. J., Amir, A., Metcalf, J. L., & Amato, K. R. (2016). Preservation methods differ in fecal microbiome stability, affecting suitability for field studies. *mSystems*, *1*(3), 1–12. https://doi.org/10.1128/mSystems.00021-16.Editor

Stark, J. S., Corbett, P. A., Dunshea, G., Johnstone, G., King, C., Mondon, J. A., Power, M. L., Samuel, A., Snape, I., & Riddle, M. (2016). The environmental impact of sewage and wastewater outfalls in Antarctica: An example from Davis station, East Antarctica. *Water Research*, *105*, 602–614. https://doi.org/10.1016/j.watres.2016.09.026

Su, J.-Q., An, X.-L., Li, B., Chen, Q.-L., Gillings, M. R., Chen, H., Zhang, T., & Zhu, Y.-G. (2017). Metagenomics of urban sewage identifies an extensively shared

antibiotic resistome in China. *Microbiome*, *5*(1), 127. https://doi.org/10.1186/s40168-017-0298-y

Swift, B. M. C. C., Bennett, M., Waller, K., Dodd, C., Murray, A., Gomes, R. L., Humphreys, B., Hobman, J. L., Jones, M. A., Whitlock, S. E., Mitchell, L. J., Lennon, R. J., & Arnold, K. E. (2019). Anthropogenic environmental drivers of antimicrobial resistance in wildlife. *Science of the Total Environment*, *649*, 12–20. https://doi.org/10.1016/j.scitotenv.2018.08.180

Tacconelli, E., & Cataldo, M. A. (2008). Vancomycin-resistant enterococci (VRE): transmission and control. *International Journal of Antimicrobial Agents*, *31*(2), 99–106. https://doi.org/10.1016/j.ijantimicag.2007.08.026

Tacconelli, E., Sifakis, F., Harbarth, S., Schrijver, R., van Mourik, M., Voss, A., Sharland, M., Rajendran, N. B., & Rodríguez-Baño, J. (2018). Surveillance for control of antimicrobial resistance. *The Lancet Infectious Diseases*, *18*(3), e99–e106. https://doi.org/10.1016/S1473-3099(17)30485-1

Taucer-Kapteijn, M., Hoogenboezem, W., Heiliegers, L., de Bolster, D., & Medema, G. (2016). Screening municipal wastewater effluent and surface water used for drinking water production for the presence of ampicillin and vancomycin resistant enterococci. *International Journal of Hygiene and Environmental Health*, *219*(4–5), 437–442. https://doi.org/10.1016/j.ijheh.2016.04.007

Thorpe, H., Booton, R., Kallonen, T., Gibbon, M. J., Couto, N., Passet, V., Sebastian Lopez Fernandez, J., Rodrigues, C., Matthews, L., Mitchell, S., Reeve, R., David, S., Merla, C., Corbella, M., Ferrari, C., Comandatore, F., Marone, P., Brisse, S., Sassera, D., Corander, J., & Feil, E. J. (2021). One Health or Three? Transmission modelling of Klebsiella isolates reveals ecological barriers to transmission between humans, animals and the environment. *BioRxiv*. https://doi.org/10.1101/2021.08.05.455249

Valentin, A. S., Santos, S. Dos, Goube, F., Gimenes, R., Decalonne, M., Mereghetti, L., Daniau, C., van der Mee-Marquet, N., Abdoush, H., Alfandari, S., Allaire, A., Aloe, L., Andreo, A., Antoine, E., Aurel, C., Azaouzi, A., Barry-Perdereau, V., Berrouane, Y., Blaise, S., … Vanson, M. L. (2021). A prospective multicentre surveillance study to investigate the risk associated with contaminated sinks in the intensive care unit. *Clinical Microbiology and Infection*, *27*(9), 1347.e9-1347.e14. https://doi.org/10.1016/j.cmi.2021.02.018

Van Boeckel, T. P., Glennon, E. E., Chen, D., Gilbert, M., Robinson, T. P., Grenfell, B. T., Levin, S. A., Bonhoeffer, S., & Laxminarayan, R. (2017). Reducing antimicrobial use in food animals. *Science*, *357*(6358), 1350–1352. https://doi.org/10.1126/science.aao1495

Van Boeckel, T. P., Pires, J., Silvester, R., Zhao, C., Song, J., Criscuolo, N. G., Gilbert, M., Bonhoeffer, S., & Laxminarayan, R. (2019). Global trends in

antimicrobial resistance in animals in low- And middle-income countries. *Science*, *365*(6459). https://doi.org/10.1126/science.aaw1944

van Bunnik, B. A. D., & Woolhouse, M. E. J. (2017). Modelling the impact of curtailing antibiotic usage in food animals on antibiotic resistance in humans. *Royal Society Open Science*, *4*(4). https://doi.org/10.1098/rsos.161067

Van Goethem, M. W., Pierneef, R., Bezuidt, O. K. I., Van De Peer, Y., Cowan, D. A., & Makhalanyane, T. P. (2018). A reservoir of "historical" antibiotic resistance genes in remote pristine Antarctic soils. *Microbiome*, *6*(1), 1–12. https://doi.org/10.1186/s40168-018-0424-5

Vandewalle, J. L., Goetz, G. W., Huse, S. M., Morrison, H. G., Sogin, M. L., Hoffmann, R. G., Yan, K., & Mclellan, S. L. (2012). *Acinetobacter*, *Aeromonas* and *Trichococcus* populations dominate the microbial community within urban sewer infrastructure. *Environmental Microbiology*, *14*(9), 2538–2552. https://doi.org/10.1111/j.1462-2920.2012.02757.x

Varela, A. R., Macedo, G. N., Nunes, O. C., & Manaia, C. M.(2015). Genetic characterization of fluoroquinolone resistant *Escherichia coli* from urban streams and municipal and hospital effluents. *FEMS Microbiology Ecology*, *91*(5). https://doi.org/10.1093/femsec/fiv015

Verburg, I., García-Cobos, S., Leal, L. H., Waar, K., Friedrich, A. W., & Schmitt, H. (2019). Abundance and antimicrobial resistance of three bacterial species along a complete wastewater pathway. *Microorganisms*, *7*(9). https://doi.org/10.3390/microorganisms7090312

Verburg, I., van Veelen, H. P. J., Waar, K., Rossen, J. W. A. A., Friedrich, A. W., Leal, L. H., García-Cobos, S., & Schmitt, H. (2021). Effects of clinical wastewater on the bacterial community structure from sewage to the environment. *Microorganisms*, *9*(4). https://doi.org/10.3390/microorganisms9040718

Vihta, K. D., Stoesser, N., Llewelyn, M. J., Quan, T. P., Davies, T., Fawcett, N. J., Dunn, L., Jeffery, K., Butler, C. C., Hayward, G., Andersson, M., Morgan, M., Oakley, S., Mason, A., Hopkins, S., Wyllie, D. H., Crook, D. W., Wilcox, M. H., Johnson, A. P., Peto, T. E. A., & Walker, A. S. (2018). Trends over time in *Escherichia coli* bloodstream infections, urinary tract infections, and antibiotic susceptibilities in Oxfordshire, UK, 1998–2016: a study of electronic health records. *The Lancet Infectious Diseases*, *18*(10), 1138–1149. https://doi.org/10.1016/S1473-3099(18)30353-0

Voigt, A. M., Faerber, H. A., Wilbring, G., Skutlarek, D., Felder, C., Mahn, R., Wolf, D., Brossart, P., Hornung, T., Engelhart, S., Exner, M., & Schmithausen, R. M. (2019). The occurrence of antimicrobial substances in toilet, sink and shower drainpipes of clinical units: A neglected source of antibiotic residues.

International Journal of Hygiene and Environmental Health, *222*(3), 455–467. https://doi.org/10.1016/j.ijheh.2018.12.013

Von Wintersdorff, C. J. H. H., Penders, J., Van Niekerk, J. M., Mills, N. D., Majumder, S., Van Alphen, L. B., Savelkoul, P. H. M. M., & Wolffs, P. F. G. G. (2016). Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. *Frontiers in Microbiology*, *7*(FEB), 1–10. https://doi.org/10.3389/fmicb.2016.00173

Wee, B. A., Muloi, D. M., & van Bunnik, B. A. D. (2020). Quantifying the transmission of antimicrobial resistance at the human and livestock interface with genomics. *Clinical Microbiology and Infection*, *26*(12), 1612–1616. https://doi.org/10.1016/j.cmi.2020.09.019

Wheatley, R., Diaz Caballero, J., Kapel, N., de Winter, F. H. R., Jangir, P., Quinn, A., del Barrio-Tofiño, E., López-Causapé, C., Hedge, J., Torrens, G., Van der Schalk, T., Xavier, B. B., Fernández-Cuenca, F., Arenzana, A., Recanatini, C., Timbermont, L., Sifakis, F., Ruzin, A., Ali, O., Lammens C., Goossens, H., Kluytmans, J., Kumar-Singh, S., Oliver, A., Malhotra-Kumar, S., & MacLean, C. (2021). Rapid evolution and host immunity drive the rise and fall of carbapenem resistance during an acute *Pseudomonas aeruginosa* infection. *Nature Communications*, *12*(1), 1–12. https://doi.org/10.1038/s41467-021-22814-9

WHO. (2021). *WHO integrated global surveillance on ESBL-producing E. coli using a "One Health" approach: implementation and opportunities*. World Health Organization. https://www.who.int/publications/i/item/who-integrated-global-surveillance-on-esbl-producing-e.-coli-using-a-one-health-approach

WHO. (2021). *Global antimicrobial resistance and use surveillance system (GLASS) report 2021*. World Health Organisation. http://www.who.int/glass/resources/publications/early-implementation-report-2020/en/

WHO. (2019). *Global Antimicrobial Resistance Surveillance System (GLASS) Report Early implementation 2017 - 2018*. World Health Organization. https://www.who.int/publications/i/item/9789241515061

WHO Collaborating Centre for Drug Statistics Methodology. (2021). *ATC/DDD Index 2022*. https://www.whocc.no/atc_ddd_index/. Accessed: 16/12/2021

Williams-Nguyen, J., Bueno, I., Sargeant, J. M., Nault, A. J., & Singer, R. S. (2016). What is the evidence that point sources of anthropogenic effluent increase antibiotic resistance in the environment? Protocol for a systematic review. *Animal Health Research Reviews*, *17*(1), 9–15.

Wolfram Research Inc. (2018). *Mathematica, Version 11.3*. https://www.wolfram.com/mathematica

Wood, D. E., & Salzberg, S. L. (2014). Kraken: Ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*, *15*(3). https://doi.org/10.1186/gb-2014-15-3-r46

Wooley, J. C., Godzik, A., & Friedberg, I. (2010). A primer on metagenomics. *PLoS Computational Biology*, *6*(2). https://doi.org/10.1371/journal.pcbi.1000667

Woolhouse, M. E. J., & Ward, M. J. (2013). Sources of antimicrobial resistance. *Science*, *341*(6153), 1460–1461. https://doi.org/10.1126/science.1243444

Woolhouse, M., Ward, M., Van Bunnik, B., & Farrar, J. (2015). Antimicrobial resistance in humans, livestock and the wider environment. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1670). https://doi.org/10.1098/rstb.2014.0083

World Bank. (2017). *Drug-Resistant Infections: A Threat to Our Economic Future* (Issue March). World Bank Group.

WWT. (2019). *PR19 Challenge Report #5: Water Consumption*. https://waterwise.org.uk/wp-content/uploads/2019/10/WWT-Report-.pdf

Yang, Y., Liu, Z., Xing, S., & Liao, X. (2019). The correlation between antibiotic resistance gene abundance and microbial community resistance in pig farm wastewater and surrounding rivers. *Ecotoxicology and Environmental Safety*, *182*(June), 109452. https://doi.org/10.1016/j.ecoenv.2019.109452

Zainab, S. M., Junaid, M., Xu, N., & Malik, R. N. (2020). Antibiotics and antibiotic resistant genes (ARGs) in groundwater: A global review on dissemination, sources, interactions, environmental and human health risks. In *Water Research* (Vol. 187). Elsevier Ltd. https://doi.org/10.1016/j.watres.2020.116455

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F. M., & Larsen, M. V. (2012). Identification of acquired antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy*, *67*(11), 2640–2644. https://doi.org/10.1093/jac/dks261

Zarfel, G., Galler, H., Feierl, G., Haas, D., Kittinger, C., Leitner, E., Grisold, A. J., Mascher, F., Posch, J., Pertschy, B., Marth, E., & Reinthaler, F. F. (2013). Comparison of extended-spectrum-beta-lactamase (ESBL) carrying Es*cherichia coli* from sewage sludge and human urinary tract infection. *Environmental Pollution*, *173*, 192–199. https://doi.org/10.1016/j.envpol.2012.09.019

Zhang, A. N., Gaston, J. M., Dai, C. L., Zhao, S., Poyet, M., Groussin, M., Yin, X., Li, L. G., van Loosdrecht, M. C. M., Topp, E., Gillings, M. R., Hanage, W. P., Tiedje, J. M., Moniz, K., Alm, E. J., & Zhang, T. (2021). An omics-based framework for assessing the health risk of antimicrobial resistance genes. *Nature Communications*, *12*(1), 1–11. https://doi.org/10.1038/s41467-021-25096-3

Zhang, D., Peng, Y., Chan, C. L., On, H., Wai, H. K. F., Shekhawat, S. S., Gupta, A. B., Varshney, A. K., Chuanchuen, R., Zhou, X., Xia, Y., Liang, S., Fukuda, K., Medicherla, K. M., & Tun, H. M. (2021). Metagenomic Survey Reveals More Diverse and Abundant Antibiotic Resistance Genes in Municipal Wastewater Than Hospital Wastewater. *Frontiers in Microbiology*, *12*. https://doi.org/10.3389/fmicb.2021.712843

Zhang, S., Huang, J., Zhao, Z., Cao, Y., & Li, B. (2020). Hospital Wastewater as a Reservoir for Antibiotic Resistance Genes: A Meta-Analysis. In *Frontiers in Public Health* (Vol. 8, p. 574968). https://doi.org/10.3389/fpubh.2020.574968

Zhang, X., Mallick, H., Tang, Z., Zhang, L., Cui, X., Benson, A. K., & Yi, N. (2017). Negative binomial mixed models for analyzing microbiome count data. *BMC Bioinformatics*, *18*(1), 1–10. https://doi.org/10.1186/s12859-016-1441-7

Appendix A. Dissemination of hospital-associated antibiotic resistance to wastewater: a systematic scoping review

Appendix A Table 1 Summary of included studies

Author and year of publication	Countries sampled	Dissemination route	Evidence in favour or against?	Bacterial group studied	Resistance typing methods	Bacterial typing methods	Statistical methods	Wastewater sample types	Non-wastewater sample types
Ahmad et al, 2014	Pakistan	HWW to MWW	In favour	*Salmonella typhi*	Phenotypic	Phenotypic	None	Untreated hospital and municipal	River
Akiba, et al, 2015	India	Hospital to MWW (indirect)	In favour	*E. coli*	Phenotypic	Fingerprint	Frequentist; Cluster	Treated municipal	-
Alexander et al, 2015	Germany	HWW to MWW	In favour	*Enterococcus spp; Staphylococcus spp; Enterobacteriaceae; P. aeruginosa*	Gene presence	Fingerprint	None	Untreated hospital and municipal, treated municipal	River, groundwater
Alexander et al, 2020	Germany	Hospital to MWW (indirect)	In favour	*E. coli, P. aeruginosa, K. pneumoniae, A. baumannii, and Enterococcus spp*	Gene presence	Fingerprint	Frequentist	Untreated municipal	-
Atmani et al, 2015	Algeria	Patients to HWW; HWW to MWW	In favour	*K. pneumoniae*	Phenotypic; gene presence	Partial sequence	Frequentist	Untreated hospital and municipal	Patients
Azuma et al, 2019	Japan	HWW to MWW	In favour	*E. coli*	Phenotypic	Phenotypic	None	Untreated hospital and municipal, treated municipal	River
Beattie et al, 2020	United States of America	HWW to MWW	In favour	All	Phenotypic	Partial sequence	Permutation	Untreated hospital and municipal, treated municipal	-
Buelow et al, 2018	Netherlands	HWW to MWW	Against	None	Gene presence	Fingerprint	Cluster	Untreated hospital and municipal	-

Butiuc-Keul et al, 2021	Romania	HWW to MWW	In favour	*Pseudomonas spp*	Phenotypic; gene presence; gene sequence	Partial sequence	Frequentist	Untreated hospital and municipal	River, lake
Cahill et al, 2019	Ireland	HWW to MWW	In favour	*Enterobacterales*	Phenotypic; gene presence	Whole genome sequence	None	Untreated hospital and municipal	-
Chávez et al, 2020	Colombia	Patients to MWW	In favour	*P. aeruginosa*	Phenotypic; gene presence	Phenotypic	Frequentist	Untreated municipal	River, patients
Ekwanzala et al, 2019	South Africa	HWW to MWW	In favour	*K. pneumoniae*	Phenotypic	Fingerprint	Frequentist	Untreated hospital and municipal, treated hospital and municipal	River
Eshrati et al, 2020	Iran (Islamic Republic of)	HWW to MWW	In favour	*E. coli*	Phenotypic	Phenotypic	Unclear	Untreated municipal	Patients, livestock
Gibbon et al, 2021	United Kingdom	HWW to MWW	In favour	*Klebsiella spp*	Gene sequence	Whole genome sequence	Phylogenetics; frequentist	Untreated hospital and municipal	River
Gómez et al, 2010	Spain	HWW to MWW	In favour	*Enterobacteriaceae; Enterococcus spp*	Phenotypic	Phenotypic	None	Untreated hospital and municipal	River
Goulouris et al, 2019	United Kingdom	HWW to MWW	In favour	*E. faecium*	Gene sequencing	Whole genome sequence	Cluster	Untreated hospital and municipal, treated municipal	Patients
Guardabassi et al, 1998	Denmark	Hospital to MWW (indirect)	In favour	*Acinetobacter spp*	Phenotypic	Fingerprint	Frequentist	Untreated municipal	-
Gundogdu et al, 2017	Turkey	HWW to MWW	In favour	*E. coli*	Gene sequence	Fingerprint	Frequentist	Untreated hospital, treated and untreated municipal	-

Harris et al, 2014	Ireland	Hospital to MWW (indirect)	In favour	*E. coli*	Phenotypic	-	Bayesian	Untreated hospital and municipal, treated municipal	-
Iversen et al, 2002	Sweden	HWW to MWW	In favour	*Enterococcus spp*	Phenotypic; gene presence	Fingerprint	Cluster	Untreated hospital and municipal, treated municipal	Surface water
Iverson et al, 2004	Sweden	Patients to HWW; HWW to MWW	In favour	*E. faecium*	Phenotypic	Phenotypic	None	Untreated municipal	Community, surface water, livestock, patients
Iweriebor et al, 2015	South Africa	HWW to MWW	In favour	*Enterococcus spp*	Phenotypic; gene presence	Fingerprint	None	Untreated hospital, treated municipal	-
Jakobsen et al, 2008	Denmark	Patients to HWW; HWW to MWW	In favour	*E. coli; Coliforms*	Phenotypic; gene presence	Partial sequence	Frequentist; Cluster	Untreated hospital and municipal	Patients
King et al, 2020	South Africa	Patients to HWW	In favour	*K. pneumoniae and K. oxytoca*	Phenotypic	Phenotypic	None	Untreated hospital and muncipal, treated municipal	Patients, river
Korzeniewska et al, 2013	Poland	HWW to MWW	In favour	*E. coli*	Phenotypic; gene presence	Phenotypic	Frequentist	Untreated hospital and municipal, treated municipal	River, air
Kovacic et al, 2017	Croatia	Patients to HWW	In favour	*A. baumannii*	Gene sequence	Fingerprint	Cluster	Untreated hospital	Patients
Kuhn et al, 2005	Sweden; Denmark; Spain; United Kingdom of Great Britain and	HWW to MWW	In favour	*Enterococcus spp*	Phenotypic	Phenotypic	None	Untreated hospital and municipal	Community, patients, surface water, livestock, farm environment, crops

	Northern Ireland								
Lamba et al, 2018	India	HWW to MWW	In favour	*Enterococcus spp*	Phenotypic; gene presence	Partial sequence	Frequentist	Untreated hospital and municipal	River
Leclercq et al, 2013	France	HWW to MWW	In favour	*Enterococcus spp*	Phenotypic; gene presence	Partial sequence	Frequentist	Untreated hospital and municipal, treated municipal	River
Linton et al, 1974	United Kingdom of Great Britain and Northern Ireland	Hospital to MWW (indirect)	In favour	*Coliforms*	Phenotypic	Phenotypic	None	Untreated municipal	-
Ludden et al, 2017	United Kingdom of Great Britain and Northern Ireland	Hospital to MWW (indirect)	In favour	*Enterobacteriaceae*	Phenotypic	Whole genome sequence	None	Untreated hospital and municipal, treated municipal	-
Ludden et al, 2019	United Kingdom of Great Britain and Northern Ireland	Patients to HWW	Against	*K. pneumoniae*	Gene sequencing	Whole genome sequence	Phylogenetics	Untreated hospital and municipal	Livestock, patients, hospital environment
Müller et al, 2018	Germany	HWW to MWW	In favour	*Gram negative*	Phenotypic; gene presence	Partial sequence	None	Untreated hospital and municipal, treated municipal	River
Narciso-Da-Rocha et al, 2014	Portugal	HWW to MWW	In favour	*Heterotrophs; Coliforms; Enterobacteriaceae; Aeromonas spp; Pseudomonas spp*	Phenotypic; gene presence	Partial sequence	Frequentist; Permutation	Untreated hospital and municipal, treated municipal	-

Ng et al, 2017	Singapore	HWW to MWW	In favour	All	Gene sequence	Partial sequence	Cluster	Untreated hospital and municipal, treated municipal	Suface water
Novo and Manaia, 2010	Portugal	Hospital to MWW (indirect)	In favour	Heterotrophs; *Enterobacteriaceae; Enterococcus* spp	Phenotypic	Phenotypic	Frequentist	Treated and untreated municipal	-
Nunez et al, 2016	Argentina	HWW to MWW	In favour	*Enterococcus* spp	Phenotypic	Phenotypic	Frequentist	Untreated hospital and municipal	-
Oberle et al, 2012	France	HWW to MWW	In favour	*E. coli*	Phenotypic; gene presence	Phenotypic	Frequentist	Untreated hospital and municipal, treated municipal	River
Oravcova et al, 2017	Czechia	Patients to MWW	In favour	*Enterococcus* spp	Phenotypic; gene presence	Partial sequence	None	Untreated municipal	Wildlife, patients, hospital staff, hospital environment
Ory et al, 2016	France	Patients to HWW	In favour	All cultivable bacteria	Phenotypic; gene presence	Partial sequence	Cluster	Untreated hospital	Patients
Paulshus et al, 2019	Norway	HWW to MWW	In favour	*E. coli*	Phenotypic	Phenotypic	Cluster	Untreated hospital and municipal	-
Paulus et al, 2019	Netherlands	HWW to MWW	In favour	None	Gene presence	-	Frequentist	Untreated hospital and municipal, treated hospital and municipal	-
Popa et al, 2021	Romania	Patients to HWW; HWW to MWW	In favour	*K. pneumoniae*	Phenotypic; gene presence; gene sequence	Whole genome sequence	Phylogenetics	Untreated hospital and muncipal, treated hospital and municipal	Patients
Pot et al, 2021	Guadeloupe	HWW to MWW	In favour	*Enterobacter cloacae*	Phenotypic; gene sequence	Whole genome sequence	Phylogenetics; Bayesian	Untreated hospital and municipal, treated municipal	Wildlife samples

Praveenkumarreddy et al, 2020	India	HWW to MWW	In favour	*E. coli*	Phenotypic	Phenotypic	Frequentist	Untreated hospital and municipal, treated hospital and municipal	-
Proia et al, 2018	Belgium	HWW to MWW	In favour	None	Gene presence	-	Frequentist	Untreated hospital and municipal	River
Rahimi and Bouzari, 2015	Iran (Islamic Republic of)	Patients to MWW	In favour	*S. aureus*	Phenotypic; gene presence	Phenotypic	None	Untreated hospital	Patients
Roederova et al, 2016	Czechia	Patients to HWW	In favour	*Enterobacteriaceae*	Phenotypic; gene presence	Fingerprint	Cluster	Untreated hospital and municipal, treated municipal	Patients
Schwartz et al, 2006	Germany	HWW to MWW	In favour	*P. aeruginosa*	Phenotypic	Partial sequence	None	Untreated hospital and municipal	River
Seruga Music et al, 2017	Croatia	Patients to HWW; HWW to MWW	In favour	*A. baumannii*	Phenotypic	Partial sequence	None	Untreated hospital and municipal	Patients, river
Talebi et al, 2008	Iran (Islamic Republic of)	Patients to MWW	In favour	*Enterococcus spp*	Phenotypic; gene presence	Fingerprint	Cluster	Untreated municipal	Patients
Thompson et al, 2013	Australia	HWW to MWW	In favour	*S. aureus*	Phenotypic; gene presence	Fingerprint	Frequentist	Untreated hospital and municipal, treated municipal	-
Tumeo et al, 2008	France	HWW to MWW	Against	*P. aeruginosa*	Phenotypic	Fingerprint	Frequentist	Untreated hospital and municipal	Community building effluent
Varela et al, 2013	Portugal	HWW to MWW	In favour	*Enterococcus spp*	Phenotypic; gene presence	Partial sequence	Frequentist; Cluster	Untreated hospital and municipal, treated municipal	-
Varela et al, 2014	Portugal	HWW to MWW	Against	*Heterotrophs; Coliforms; Enterobacteriaceae;*	Phenotypic	Fingerprint	Frequentist; Permutation	Untreated hospital and municipal, treated municipal	-

				Aeromonas spp; Pseudomonas spp					
Varela et al, 2015	Portugal	HWW to MWW	In favour	*E. coli*	Phenotypic; gene presence	Partial sequence	Permutation	Untreated hospital and municipal, treated municipal	Streams, ponds
Varela et al, 2016	Portugal	HWW to MWW	In favour	*Aeromonas spp*	Phenotypic	Partial sequence	Frequentist; Cluster	Untreated hospital and municipal, treated municipal	-
Verburg et al, 2019	Netherlands	HWW to MWW	In favour	*E. coli; Klebsiella spp; Aeromonas spp*	Phenotypic	Fingerprint	Frequentist	Untreated hospital, nursing home, and municipal, treated municipal	-
Voigt et al, 2020	Germany	HWW to MWW	In favour	*Klebsiella spp, Enterobacter spp, Citrobacter spp, E. coli, Proteus mirabilis, P. aeruginosa; A. calcoaceticus-baumannii complex; S. aureus, E. faecium, E. faecalis*	Phenotypic; gene presence	Fingerprint	Frequentist; Cluster	Untreated hospital and municipal, treated municipal	-
Yang et al, 2009	Taiwan	Patients to MWW	Against	Coliforms	Phenotypic	Phenotypic	Frequentist	Treated and untreated municipal	Patients
Zarfel et al, 2013	Austria	Patients to MWW	In favour	E. coli	Phenotypic; gene presence	Phenotypic	Frequentist	Treated municipal	Patients

Appendix B: The role of the environment in transmission of antibiotic resistance between humans and animals: a modelling study

Appendix B Table 1: Parameter definitions

Parameter	Definition and units
β_{HH}	Per capita rate at which humans acquire antibiotic resistant bacteria as a result of exposure to other humans harbouring resistant bacteria per time step
β_{AA}	Per capita rate at which animals acquire antibiotic resistant bacteria as a result of exposure to other animals harbouring resistant bacteria per time step
β_{AH}	Per capita rate at which humans acquire antibiotic resistant bacteria as a result of exposure to animals harbouring resistant bacteria per time step
β_{HA}	Per capita rate at which animals acquire antibiotic resistant bacteria as a result of exposure to humans harbouring resistant bacteria per time step
β_{HE}	Per environmental unit rate* at which the environment acquires resistant bacteria as a result of exposure to humans harbouring resistant bacteria per time step
β_{EH}	Per capita rate at which humans acquire antibiotic resistant bacteria as a result of exposure to environmental units harbouring resistant bacteria per time step
β_{AE}	Per environmental unit rate* at which the environment acquires resistant bacteria as a result of exposure to animals harbouring resistant bacteria per time step
β_{EA}	Per capita rate at which animals acquire antibiotic resistant bacteria as a result of exposure to environmental units harbouring resistant bacteria per time step
Λ_{H}	Per capita rate at which humans acquire antibiotic resistant bacteria as a result of direct exposure to antibiotics per time step
Λ_{A}	Per capita rate at which animals acquire antibiotic resistant bacteria as a result of direct exposure to antibiotics per time step
γ_H	Proportion of Λ_{H} that reaches the environment as antibiotics (a scalar parameter)
γ_A	Proportion of Λ_{A} that reaches the environment as antibiotics (a scalar parameter)
$\gamma_H\Lambda_H$	Per environmental unit rate* at which the environment acquires resistant bacteria as a result of exposure to a proportion of antibiotics given to humans per time step
μ_H	Per capita rate at which humans with resistant bacteria revert to have only sensitive bacteria per time step
μ_A	Per capita rate at which animals with resistant bacteria revert to have only sensitive bacteria per time step
μ_E	Per environmental unit rate* at which environmental units harbouring resistant bacteria revert to having only sensitive bacteria per time step

Appendix B Table 2: Transmission coefficients

Unbounded model

Parameter	Value					
	Baseline	Balanced	Balanced (low β_{HA})	Environment-driven	Animal-driven	Human-driven
β_{HH}	0.1	0.07432092	0.07432092	0.001	0.001	0.2019663
β_{AA}	0.1	0.07432092	0.07432092	0.001	0.2019663	0.001
β_{HA}	0.1	0.07432092	0.00074321	0.001	0.001	0.2019663
β_{AH}	0.1	0.07432092	0.07432092	0.001	0.2019663	0.001
β_{EH}	0.01	0.07432092	0.07432092	0.1420501	0.001	0.001
β_{EA}	0.01	0.07432092	0.07432092	0.1420501	0.001	0.001
β_{HE}	0.1	0.07432092	0.07432092	0.1420501	0.001	0.2019663
β_{AE}	0.1	0.07432092	0.07432092	0.1420501	0.2019663	0.001

Bounded model

Parameter	Value					
	Baseline	Balanced	Balanced (low β_{HA})	Environment-driven	Animal-driven	Human-driven
β_{HH}	0.1	0.08109928	0.08109928	0.001	0.001	0.20239149
β_{AA}	0.1	0.08109928	0.08109928	0.001	0.20239149	0.001
β_{HA}	0.001	0.08109928	0.00081099	0.001	0.001	0.20239149
β_{AH}	0.1	0.08109928	0.08109928	0.001	0.20239149	0.001
β_{EH}	0.01	0.08109928	0.08109928	0.23084954	0.001	0.001
β_{EA}	0.01	0.08109928	0.08109928	0.23084954	0.001	0.001
β_{HE}	0.1	0.08109928	0.08109928	0.23084954	0.001	0.20239149
β_{AE}	0.1	0.08109928	0.08109928	0.23084954	0.20239149	0.001

Appendix B Table 3: Other parameters

Parameter	Value	
	Fig 1. B	**Fig 2. And 3.**
Λ_H	0.1	Beta(1.7, 15.3) (mean 0.1)
Λ_A	0.1	No intervention: 0.1 or U(0.000001,1.); intervention: 0.0.
γ_H	0.001	0.001
γ_A	0.001	0.001
μ_H	0.1	Beta(1.7, 15.3) (mean 0.1)
μ_A	0.1	0.1
μ_E	0.2	Beta(3, 12) (mean 0.2)

Additional methods information

Methods for finding transmission parameter coefficients

Transmission parameters were chosen by the following method: some parameters were fixed (p_f) whilst the transmission parameters of interest varied (p_v) to reach a human resistance level of 71% (target prevalence in original study, van Bunnik and Woolhouse, 2017):

$$\min_{p_v \in (0,1)} \left| 0.71 - f(p_v, p_f) \right|$$

For example, for the human-driven transmission scenario, rates of transfer of resistance from humans to any other population were varied ($\beta_{HH} = \beta_{HA} = \beta_{HE}$), and all other transmission parameter values were fixed at a low value, 0.001.

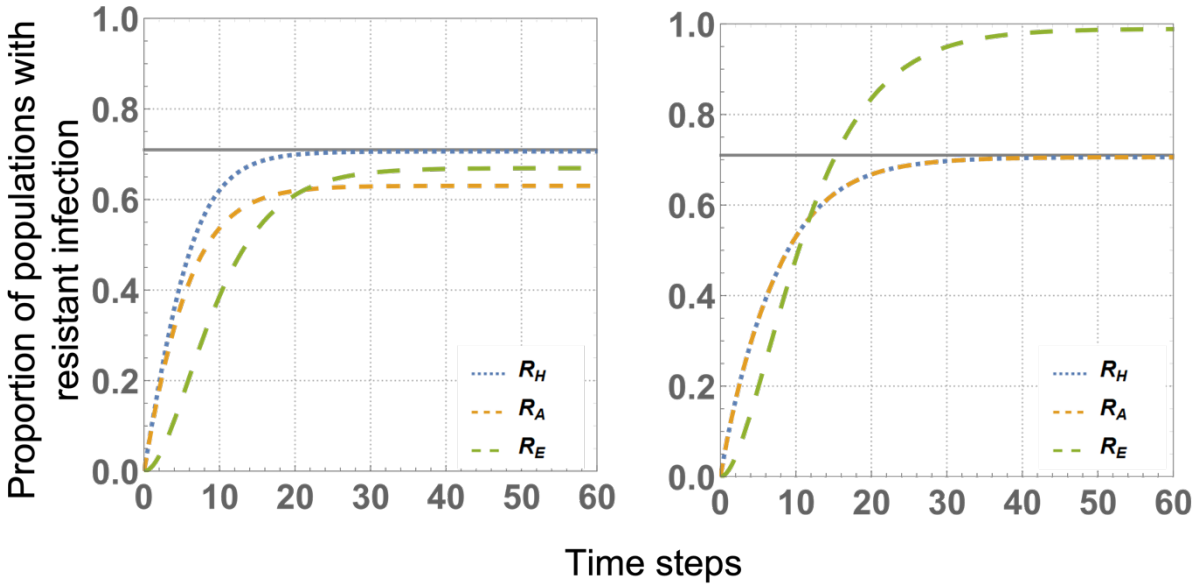
Model timesteps

We selected a value of 0.2 for μ_E in the baseline set of parameters. Zhang et al, 2017 estimated for the rate of loss of various resistance genes from a compost microcosm experiment was 0.0077 per day. We can therefore estimate that the units of our select value is per ~26 days (0.2/0.0077). We intend this model to be used for looking at long term prevalence of resistance in humans, so this estimate of time step is reasonable as it allows us to look over the timescale of years.

To ensure equilibrium values were obtained for all experiments, we initially numerically solved the model to 500 timesteps, and if there was a difference of more than 0.0000001 between the R_H values for the final two timesteps, we solved to 10,000 timesteps.

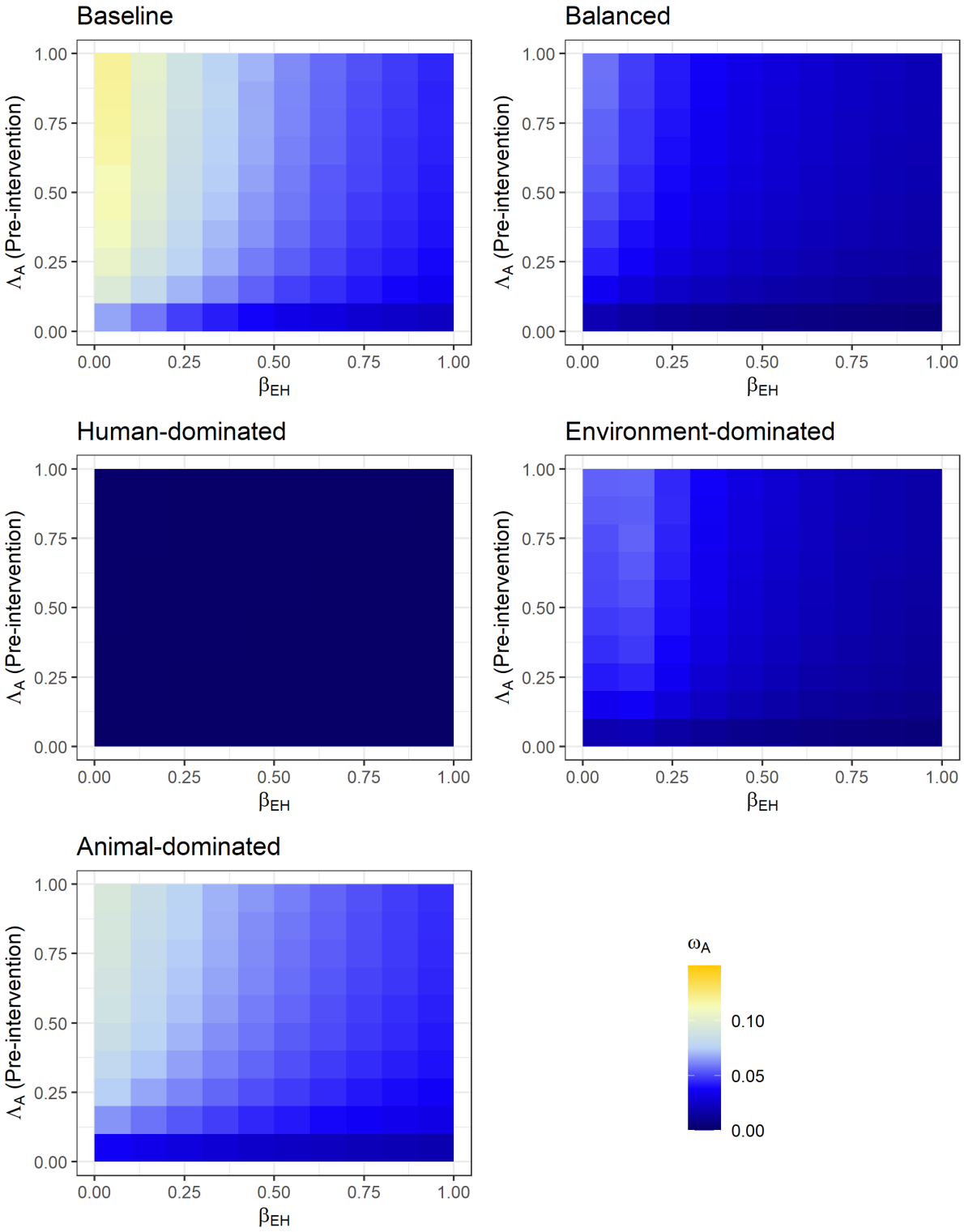
Appendix B Figure 1

Trajectory plot of the fraction of human and animal populations carrying resistant bacteria (R_H, R_E), and the amount of resistant material in the environment (R_E). For bounded environment model (left) and unbounded (right).



Appendix B Figure 2

Heatmaps of the impact of reducing Λ_A, for different pre-intervention levels of Λ_A (Y axis) and β_{EH} (X axis), in all transmission scenarios.



Appendix references

van Bunnik, B. A. D.; Woolhouse, M. E. J. Modelling the Impact of Curtailing Antibiotic Usage in Food Animals on Antibiotic Resistance in Humans. *R. Soc. Open Sci.* **2017**, *4* (4). https://doi.org/10.1098/rsos.161067.

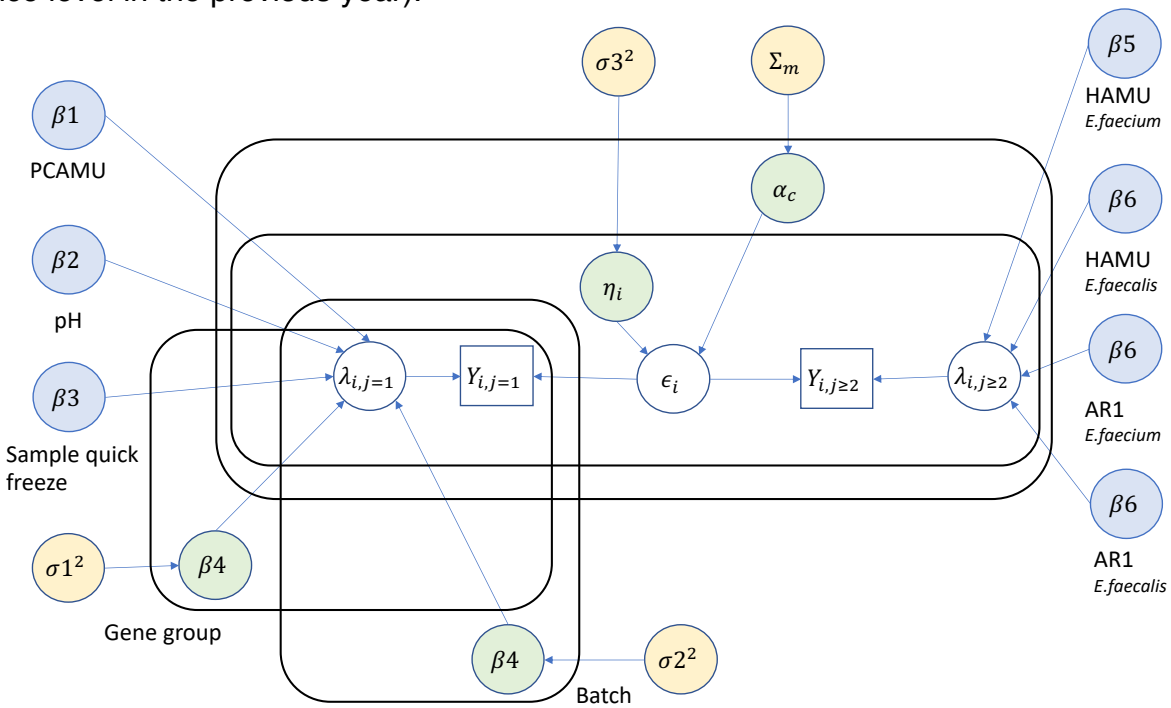
Appendix C: A multi-response model to combine sewage and hospital antibiotic resistance surveillance data

Appendix C Table 1: EARs-Net drug-bug combinations

Antimicrobial group	Organism	Specific antimicrobials
Fluoroquinolones	*E. coli*	Ciprofloxacin, ofloxacin, levofloxacin
	K. pneumoniae	Ciprofloxacin, ofloxacin, levofloxacin
	P. aeruginosa	Ciprofloxacin, levofloxacin
	Acinetobacter spp.	Ciprofloxacin, ofloxacin, levofloxacin
Aminopenicillins	*E. coli*	Amoxicillin, ampicillin
	E. faecalis	Amoxicillin, ampicillin
	E. faecium	Amoxicillin, ampicillin
Third generation cephalosporins	*E. coli*	Cefotaxime, ceftriaxone, ceftazidime
	K. pneumoniae	Cefotaxime, ceftriaxone, ceftazidime
	P. aeruginosa	Ceftazidime
Carbapenems	*E. coli*	Imipenem, meropenem
	K. pneumoniae	Imipenem, meropenem
	P. aeruginosa	Imipenem, meropenem
	Acinetobacter spp.	Imipenem, meropenem
Aminoglycosides	*E. coli*	Gentamicin, tobramycin, netilmicin
	K. pneumoniae	Gentamicin, tobramycin, netilmicin
	P. aeruginosa	Gentamicin, tobramycin, netilmicin
	Acinetobacter spp.	Gentamicin, tobramycin, netilmicin
	E. faecalis	High-level gentamicin
	E. faecium	High-level gentamicin
Vancomycin	*E. faecalis*	Vancomycin
	E. faecium	Vancomycin
Macrolides	*S. pneumoniae*	Erythromycin, clarithromycin, azithromycin

Appendix C Figure 1: Example directed acyclic graph of vancomycin resistance model structure

Showing variables in circles and constants in squares. Filled circles indicate estimated parameters (yellow for hyperpriors, green for random effects, blue for fixed effects). Arrows indicate dependency. Rounded 'plates' used to indicate aggregation levels in the data. HAMU: hospital antimicrobial usage. PCAMU: primary care antimicrobial usage. AR1: autoregression 1 (clinical resistance level in the previous year).



Appendix C Table 2: Priors used in specifying all models

I stands for the identity matrix and is of size $p \, x \, p$ where p is the number of parameters being estimated.

Parameter(s)	Prior
All fixed effects	Normal(0, 1)
Residual variance	Inverse-Wishart(I, 1)
Non-partitioned group variance	Inverse-Wishart(I, 1)
Partitioned group variance	Inverse-Wishart(I, p-1)

Example model code for vancomycin resistance model, including priors

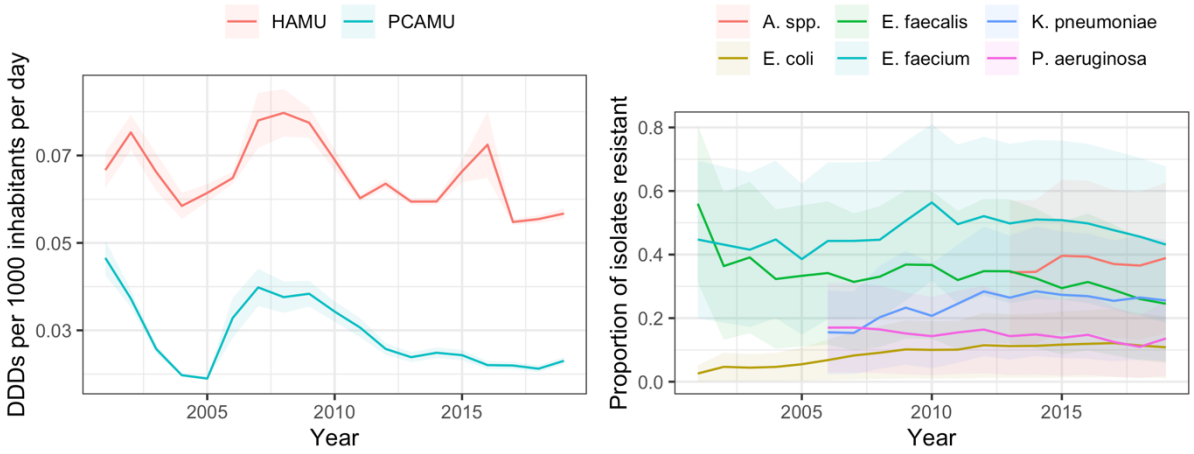
Written using R and the MCMCglmm package. Note that an adaptation to the MCMCglmm package (implemented by J. Hadfield) was required and family argument specification would not work for this data in the usual MCMCglmm package.

```
mvmod <- MCMCglmm2(cbind(n_success, n_fail) ~ ds-1 +
                         at.level(ds, 'gene'):PCAMUs +
                         at.level(ds, 'gene'):flow_cat +
                         at.level(ds, 'gene'):pH +
                         at.level(ds, 'gene'):storage_temp_cat +
                         at.level(ds, 'gene'):sum_lgl_ltb +
                         at.level(ds, 'isolate'):HAMUs +
                         at.level(ds, 'isolate'):AR1,
                   random = ~idh(at.level(ds, 'gene')):id +
                                idh(at.level(ds, 'gene')):batch +
                                us(ds):cca2,
                   rcov=~idh(ds):units,
                   data = fitting_dat, prior = prior,
                   family = "multinomial2", nitt = 100000,
                   burnin = 3000, verbose = FALSE, thin = 10)
```

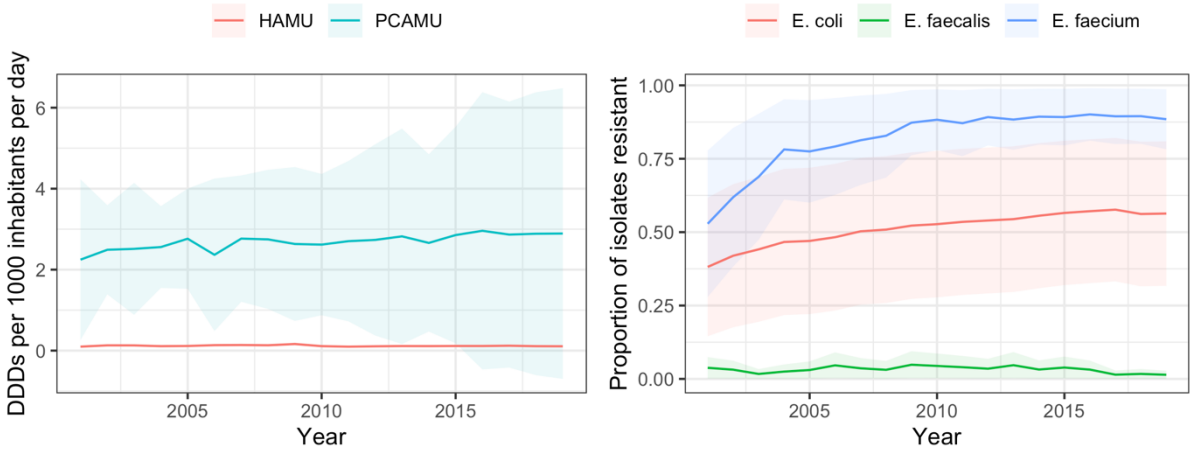
Appendix C Figure 2: EARS-Net and ESAC-Net time trends

Plots on left, the mean (points) and variance (lines) of usage rates of aminoglycosides in defined daily doses per 1000 people per day, in hospitals (HAMU) and primary care (PCAMU), for each year in the dataset. Plots on right, the mean (points) and variance (lines) of the proportion of isolates resistant to aminoglycosides in each year of the dataset, for each organism.

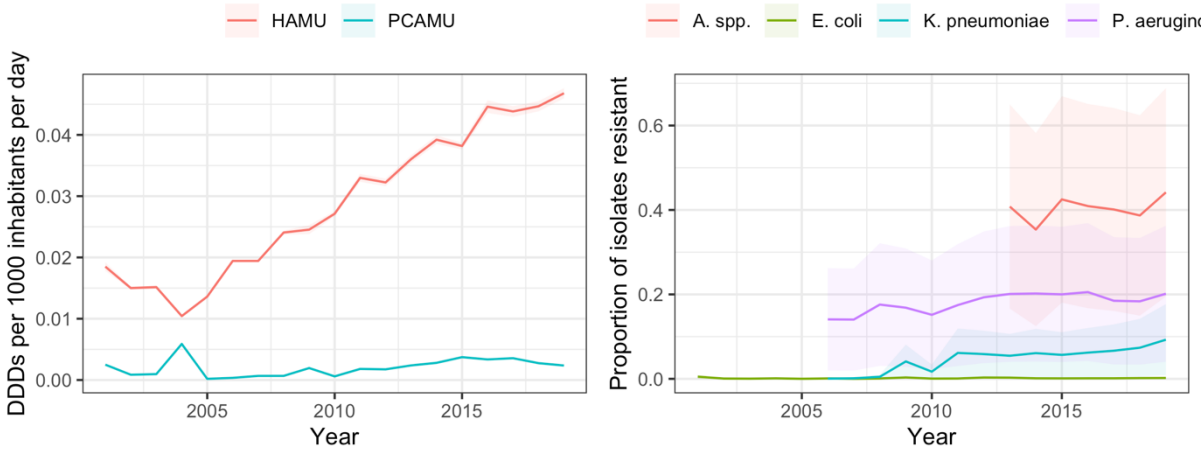
Aminoglycosides



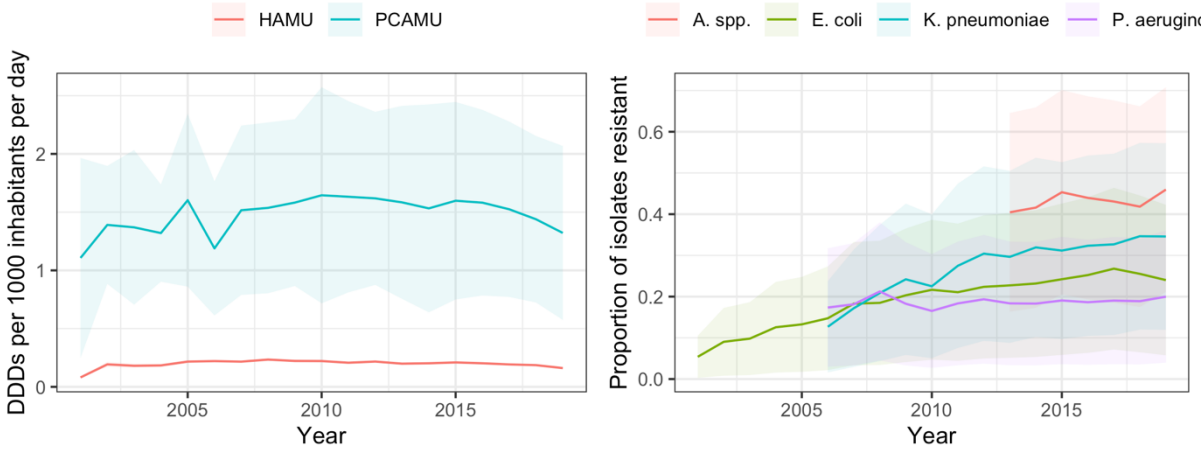
Aminopenicillins



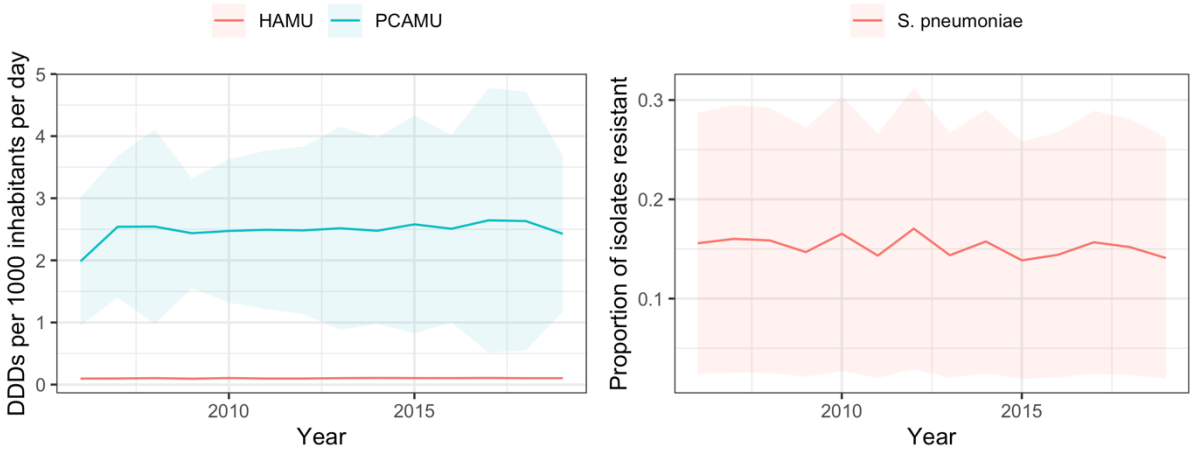
Carbapenems



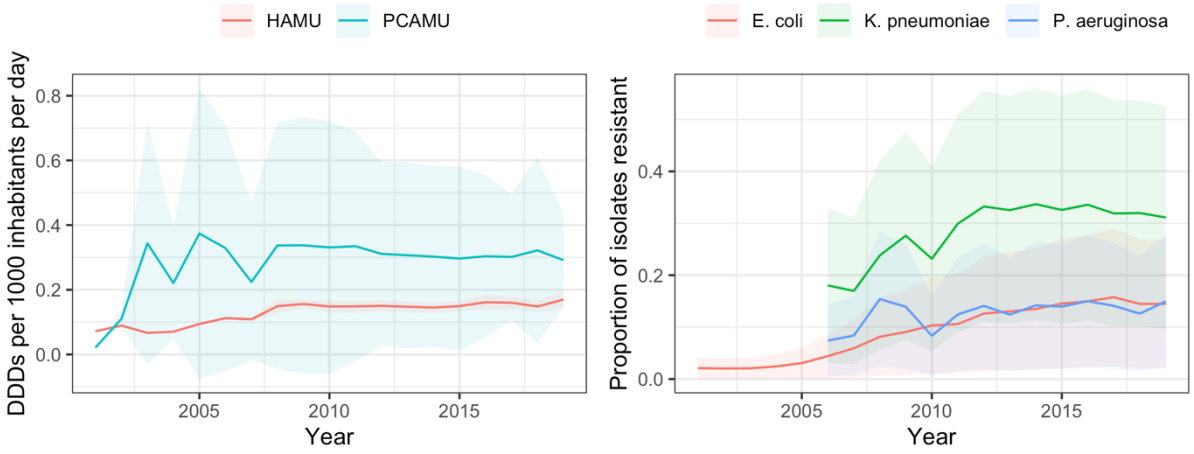
Fluoroquinolones



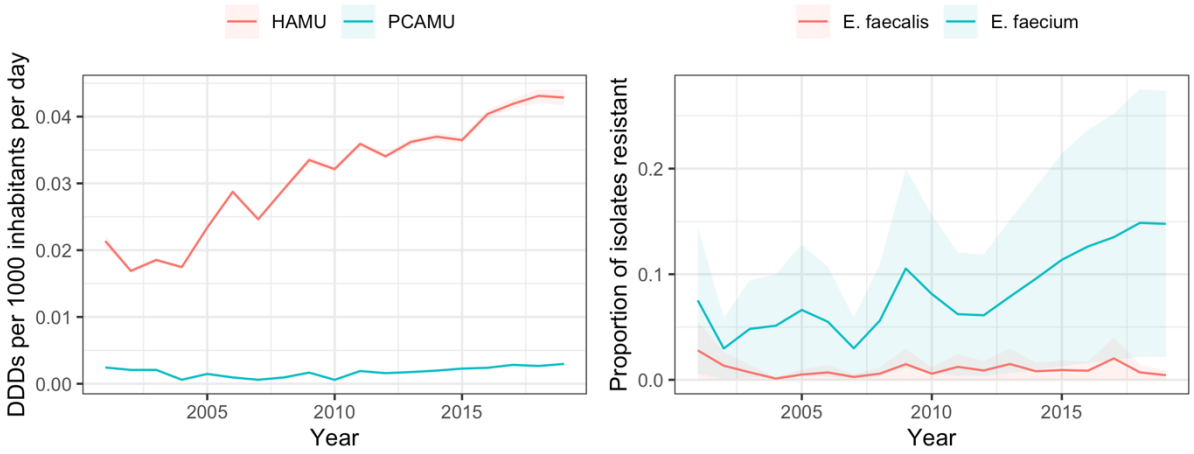
Macrolides



Third generation cephalosporins



Vancomycin



Appendix C Table 3: Odds ratios and uncertainty for the effect of hospital antimicrobial usage and resistance in the previous time point on clinical resistance

Posterior mode and 95% highest posterior density intervals given.

Aminoglycosides

Organism	Effect estimate (UI)	
	Antimicrobial usage (hospital)	AR1
Acinetobacter spp.	0.98 (0.92 – 1.04)	48.38 (17.86 – 84.46)
Enterococcus faecalis	1.03 (0.99 – 1.08)	29.17 (18.49 – 44.51)
Enterococcus faecium	1.03 (0.97 – 1.08)	35.06 (20.10 – 51.49)
Escherichia coli	0.98 (0.93 – 1.02)	501.70 (192.20 – 1278.31)
Klebsiella pneumoniae	1.05 (1.00 – 1.09)	29.52 (17.40 – 44.87)
Pseudomonas aeruginosa	1.04 (0.98 – 1.09)	14.14 (7.13 – 30.74)

Aminopenicillins

Organism	Effect estimate (UI)	
	Antimicrobial usage (hospital)	AR1
E. faecalis	1.33 (1.18 - 1.55)	40.71 (16.01 - 137.24)
E. faecium	0.84 (0.73 - 0.96)	76.22 (49.67 - 110.69)
E. coli	0.99 (0.95 - 1.03)	39.7 (25.55 - 53.20)

Carbapenems

Organism	Effect estimate (UI)	
	Antimicrobial usage (hospital)	AR1
Acinetobacter spp.	1.30 (1.00 - 1.70)	68.32 (15.76 - 131.36)
E. coli	1.29 (1.02 - 1.58)	1.00 (0.15 - 7.35)
K. pneumoniae	1.65 (1.28 - 2.07)	102.95 (25.85 - 292.31)
P. aeruginosa	1.09 (0.96 - 1.26)	10.30 (3.24 - 29.61)

Fluoroquinolones

Organism	Effect estimate (UI)	
	Antimicrobial usage (hospital)	AR1
Acinetobacter spp.	1.00 (0.85 - 1.14)	14.25 (5.60 - 30.41)
E. coli	1.00 (0.97 - 1.03)	160.15 (101.76 - 220.53)
K. pneumoniae	0.98 (0.93 - 1.04)	28.54 (20.48 - 45.06)
P. aeruginosa	1.07 (1.01 - 1.15)	8.14 (3.71 - 19.71)

Macrolides

Organism	Effect estimate (UI)	
	Antimicrobial usage (hospital)	AR1
S. pneumoniae	1.11 (0.98 - 1.27)	26.01 (11.71 - 52.98)

Third generation cephalosporins

Organism	Effect estimate (UI)	
	Antimicrobial usage (hospital)	**AR1**
E. coli	1.05 (0.97 - 1.17)	7435.43 (3369.36 - 17794.40)
K. pneumoniae	1.11 (0.99 - 1.27)	23.99 (10.97 - 43.71)
P. aeruginosa	1.21 (1.08 - 1.37)	5.54 (2.06 - 15.43)

Vancomycin

Organism	Effect estimate (UI)	
	Antimicrobial usage (hospital)	**AR1**
E. faecalis	0.93 (0.71 - 1.19)	2.61 (0.44 - 18.58)
E. faecium	1.11 (0.91 - 1.37)	359.30 (114.51 - 1190.02)

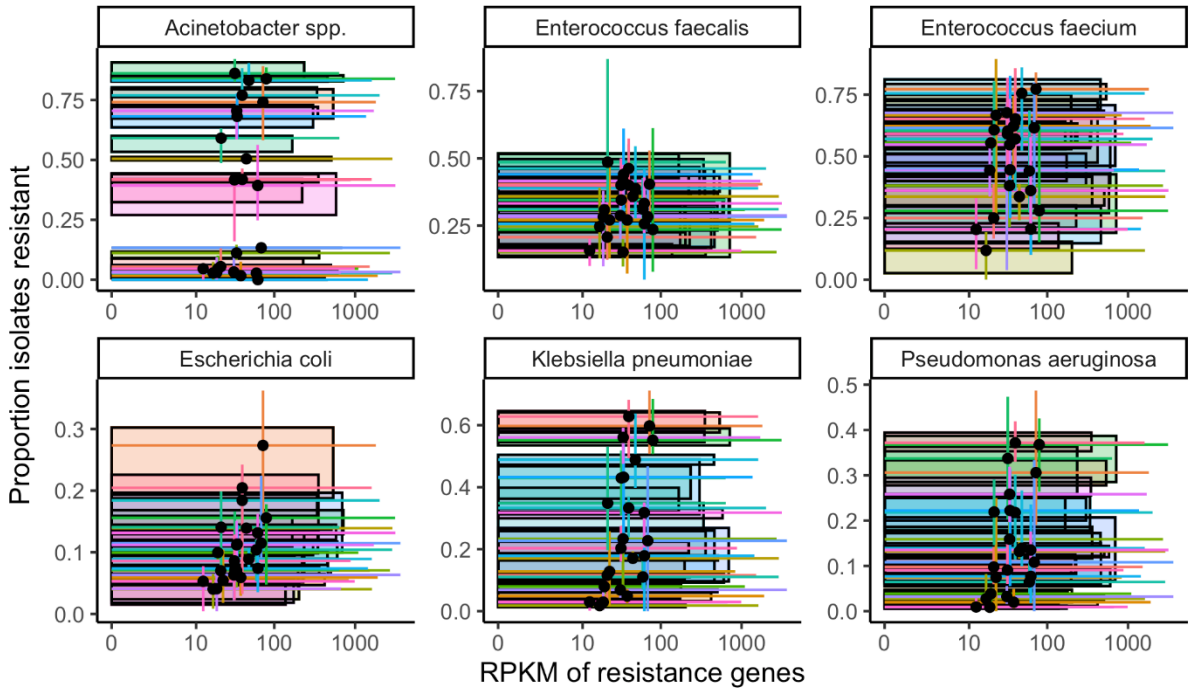
Appendix C Figure 3: Resistance measure country-level covariance

A) Double boxplots of proportion of isolates resistant against FPKM of genes conferring resistance to aminoglycosides for each country in dataset (indicated by box colour). For each country, box height indicates the upper and lower 95% quantiles of proportion of isolates resistance. Box width indicates 1.96 standard deviations around the mean of FPKM. Vertical lines indicate the range of proportion of isolates that were resistant. Horizontal lines indicate range of FPKM values. Points are the median and mean of proportion of isolates resistant and the mean of FPKM, respectively. B) For each organism, violin plots represent the posterior distribution for the correlation estimates from models with and without primary care antimicrobial prescriptions taken into account. Point and error bars for regression models are the posterior mode and 95% HPDs, and for Spearman's Rank are the estimate of correlation (rho), plus 95% confidence intervals. C) Posterior mode and 95% highest posterior density intervals in brackets for each element in the covariance matrix.

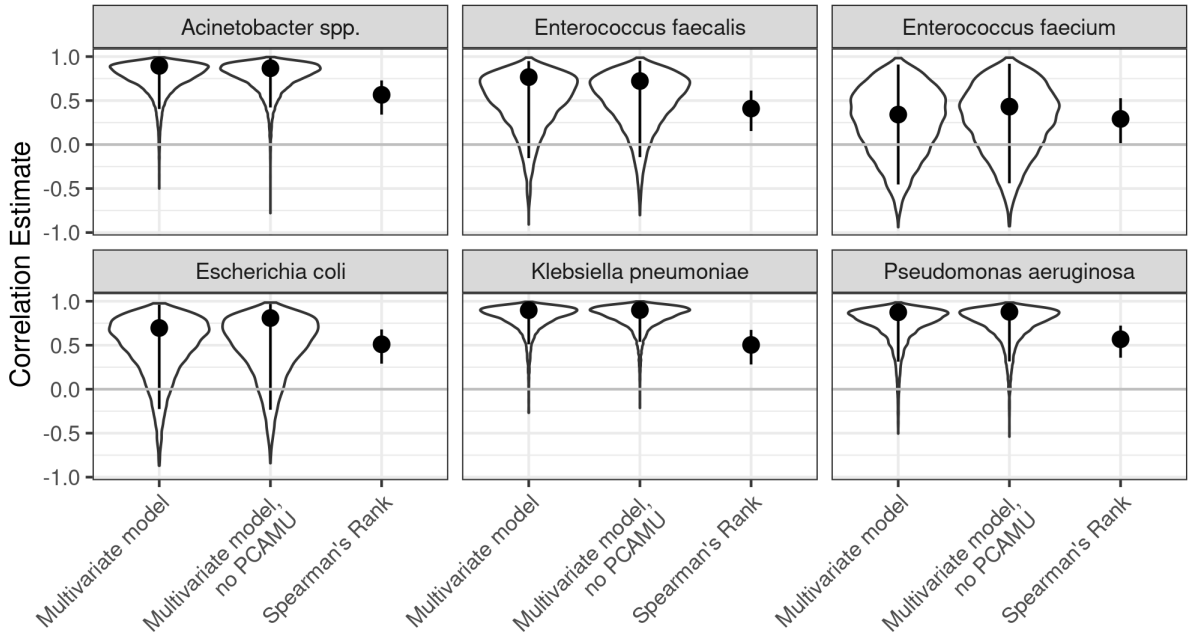
Figures begin on next page.

1. Aminoglycosides

A



B

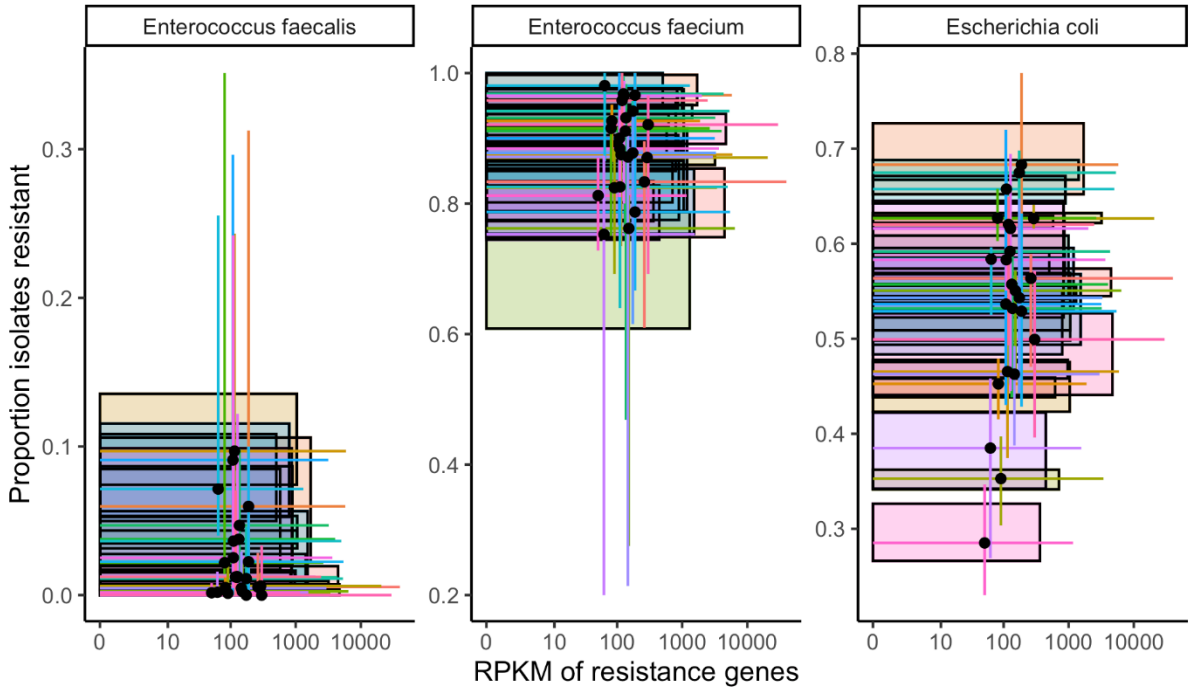


C

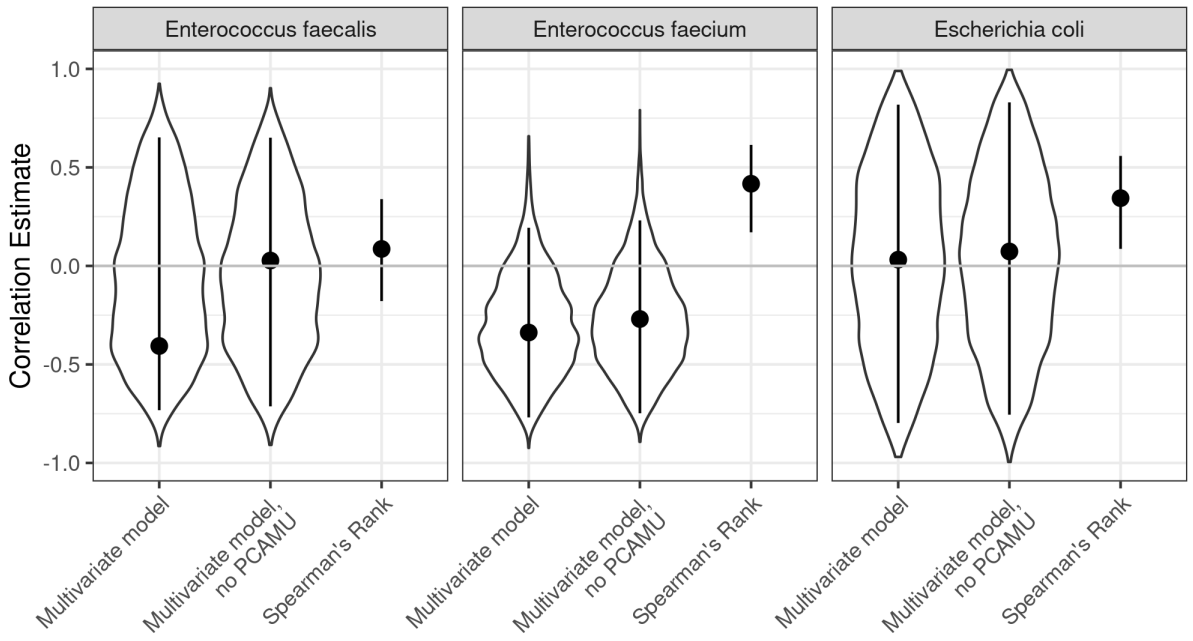
	Clinical data						Sewage data
	Acinetobacter spp.	E. faecalis	E. faecium	E. coli	K. pneumoniae	P. aeruginosa	
Acinetobacter spp.	0.89 (0.44 – 1.85)	0.06 (-0.05 – 0.23)	0.02 (-0.18 - 0.19)	0.49 (-0.04 - 1.45)	0.59 (0.30 - 1.24)	0.80 (0.44 - 1.75)	0.32 (0.10 - 0.91)
E. faecalis		0.02 (0.00. – 0.09)	0.01 (-0.02 - 0.1)	0.03 (-0.24 - 0.26)	0.05 (-0.02 - 0.18)	0.05 (-0.09 - 0.24)	0.03 (-0.04 - 0.19)
E. faecium			0.01 (0.00 – 0.26)	0 (-0.47 - 0.17)	0.02 (-0.1 - 0.15)	0.01 (-0.24 - 0.17)	0.02 (-0.13 - 0.17)
E. coli				1.06 (0.51 – 2.34)	0.44 (0.05 - 1.14)	0.7 (0.26 - 1.72)	0.3 (-0.24 - 1.04)
K. pneumoniae					0.48 (0.28 – 1.02)	0.69 (0.35 - 1.37)	0.3 (0.11 - 0.72)
P. aeruginosa						0.33 (0.06 - 1.03)	0.26 (0.1 - 0.75)
Sewage data							1.04 (0.58 - 2.07)

2. Aminopenicillins

A



B

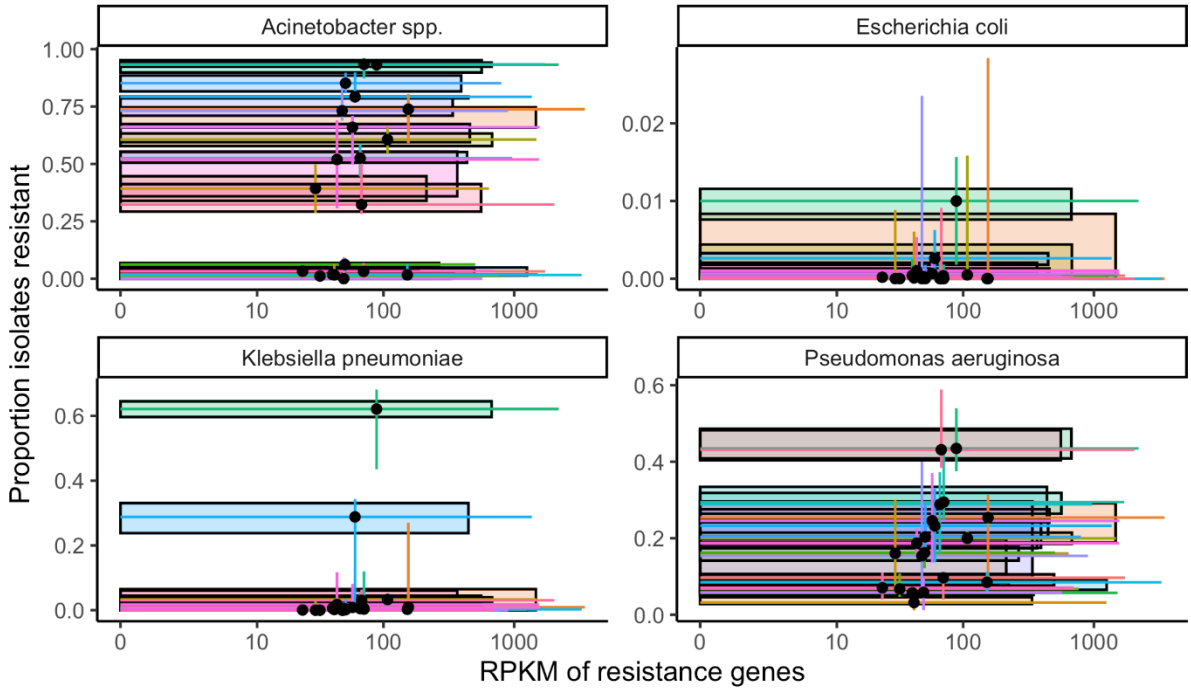


C

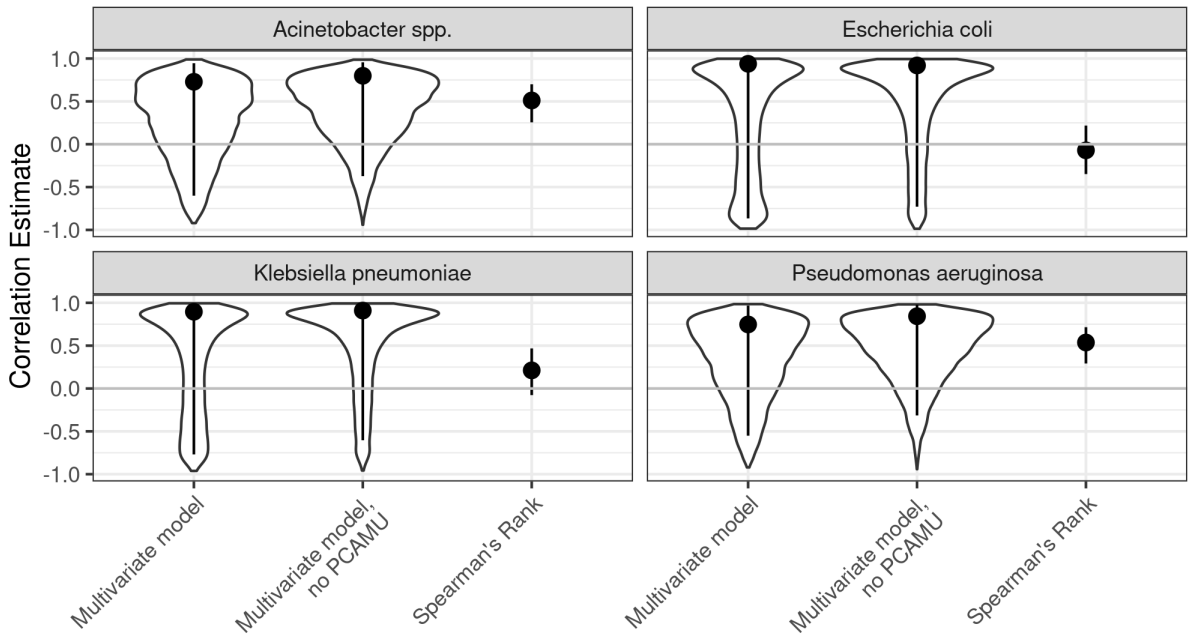
	Clinical data			Sewage data
	E. faecalis	E. faecium	E. coli	
E. faecalis	8 (3.7 - 111.75)	-0.36 (-2.01 - 1.51)	0 (-0.16 - 0.33)	0.05 (-3.06 - 2.04)
E. faecium		0.21 (0.09 - 0.61)	0 (-0.02 - 0.02)	-0.07 (-0.36 - 0.09)
E. coli			0 (0 - 0.01)	0 (-0.02 - 0.02)
Sewage data				0.26 (0.12 - 0.71)

3. Carbapenems

A



B

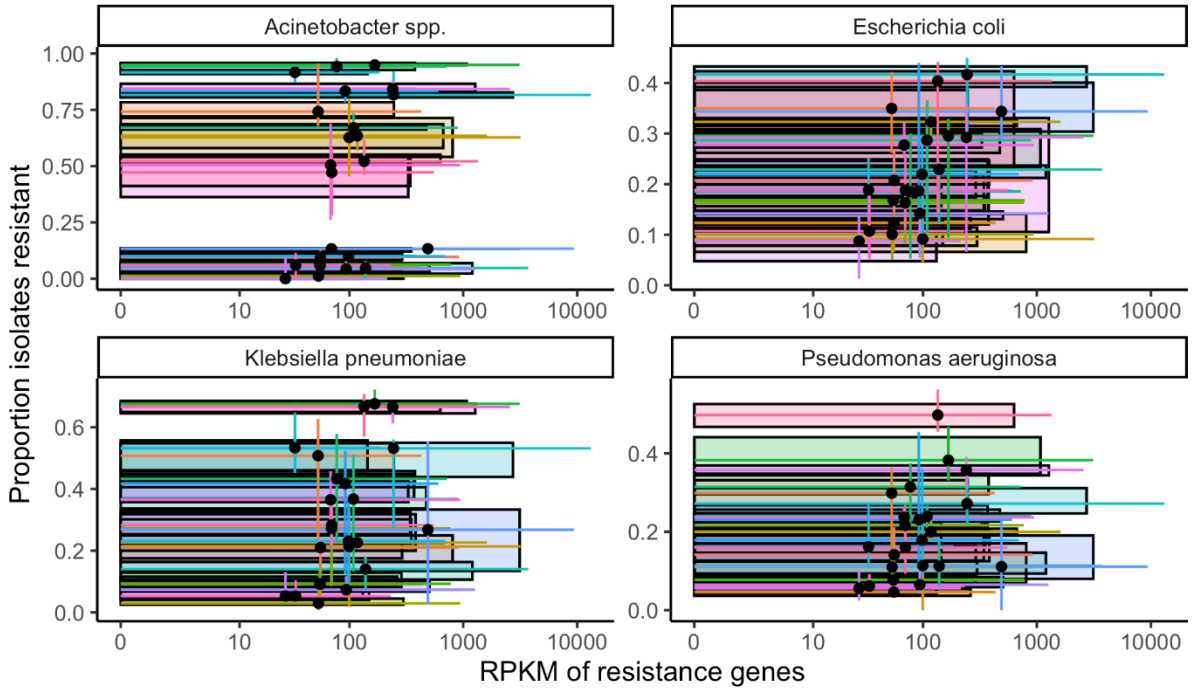


C

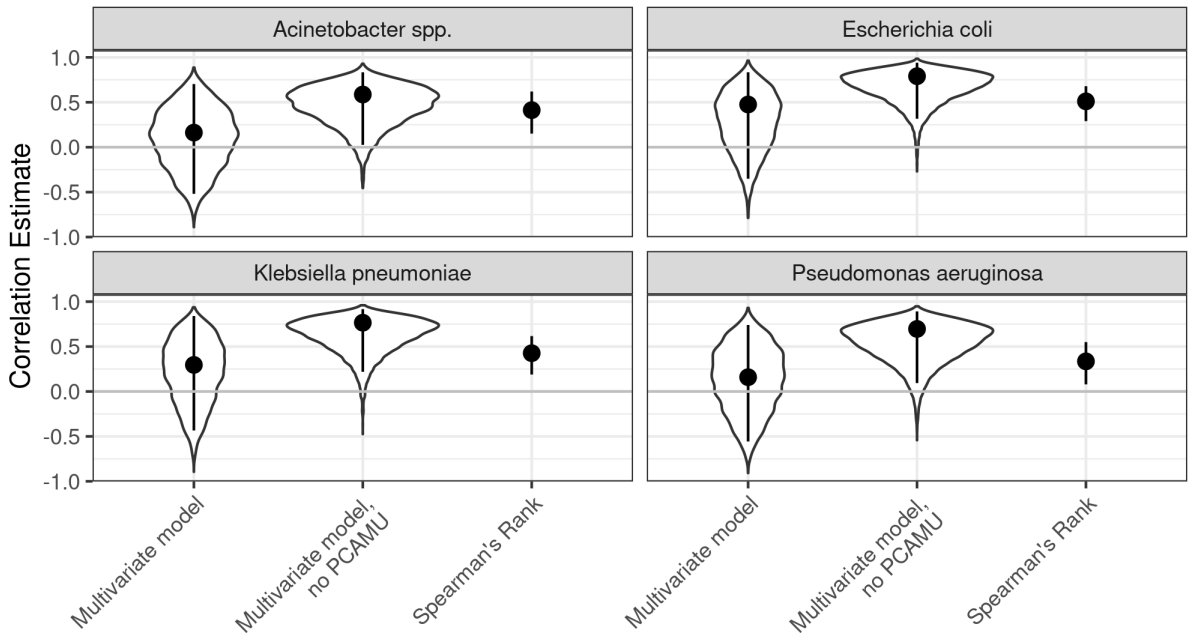
	Clinical data				Sewage data
	Acinetobacter spp.	E. coli	K. pneumoniae	P. aeruginosa	
Acinetobacter spp.	1.67 (0.59 - 5.95)	5.21 (-2.71 - 18.4)	3.31 (-0.23 - 11.38)	0.86 (0.21 - 3.76)	0.12 (-1.45 - 2.81)
E. coli		31.49 (13.52 - 80.98)	17.9 (6.93 - 41.02)	1.84 (-3.51 - 12.47)	1.25 (-7.78 - 12.08)
K. pneumoniae			10.43 (4.11 - 24.73)	1.29 (-1.19 - 7.54)	0.77 (-4.09 - 6.79)
P. aeruginosa				0.47 (0.13 - 2.63)	0.09 (-0.91 - 1.91)
Sewage data					0.24 (0.05 - 2.7)

4. Fluoroquinolones

A



B

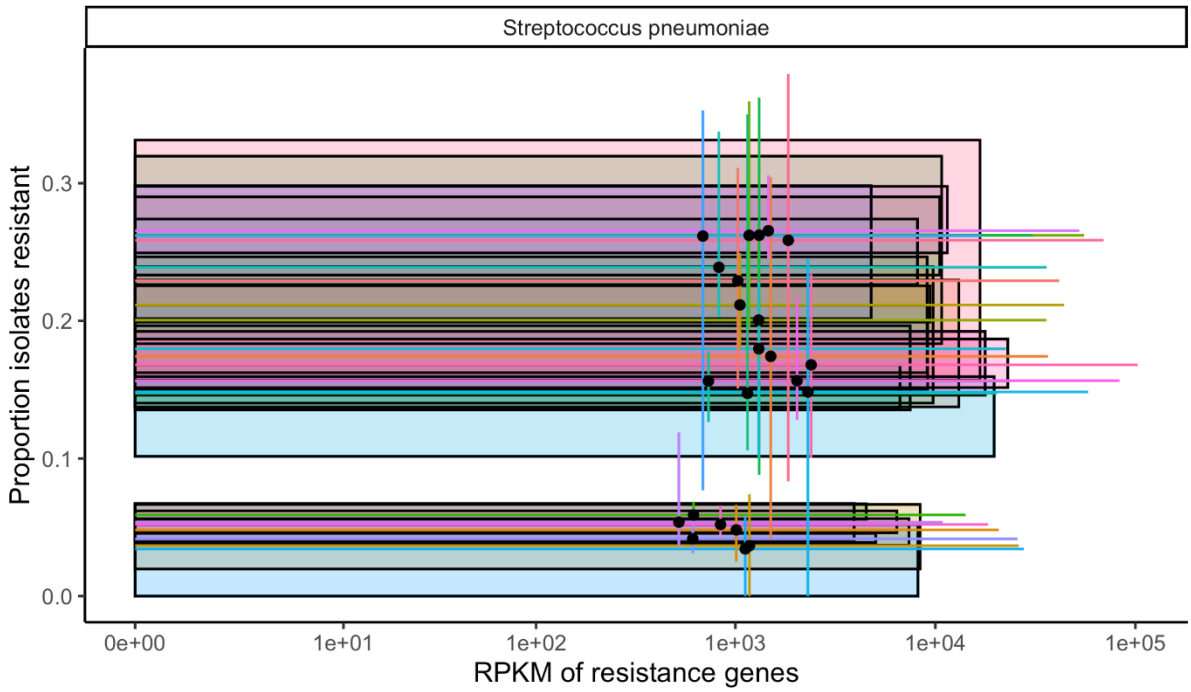


C

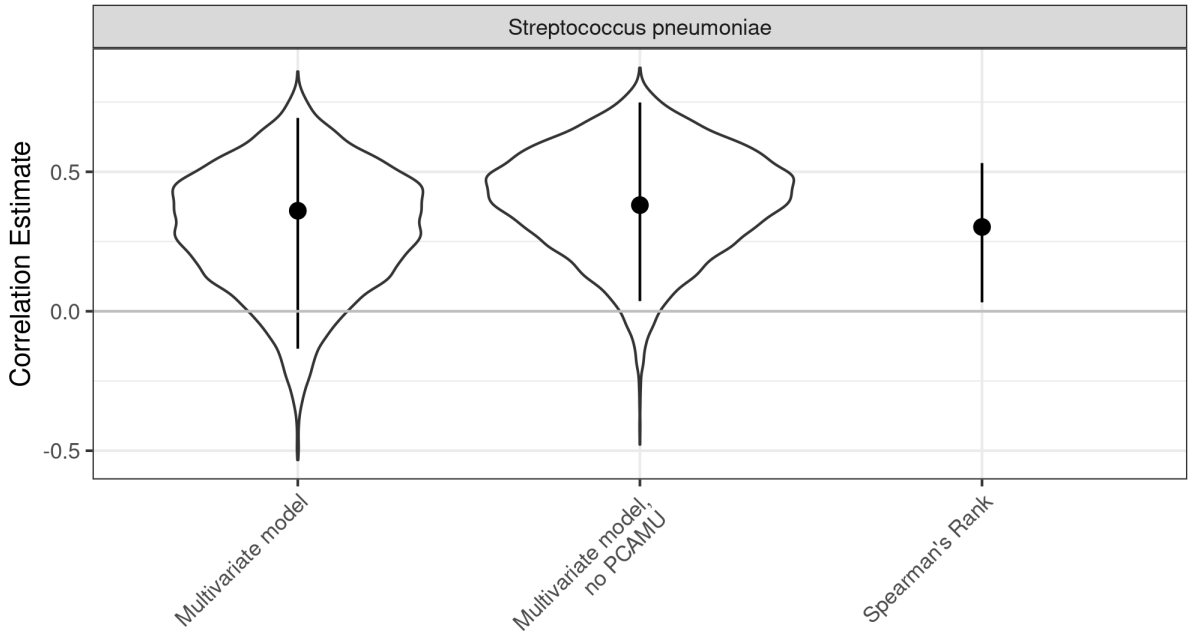
	Clinical data				Sewage data
	Acinetobacter spp.	E. coli	K. pneumoniae	P. aeruginosa	
Acinetobacter spp.	2.6 (0.95 - 5.41)	0.15 (0.03 - 0.4)	0.88 (0.41 - 2.06)	0.83 (0.41 - 2.14)	0.1 (-0.74 - 0.92)
E. coli		0.02 (0.01 - 0.06)	0.09 (0.02 - 0.2)	0.07 (0.02 - 0.2)	0.02 (-0.05 - 0.12)
K. pneumoniae			0.43 (0.19 - 1.01)	0.42 (0.18 - 0.97)	0.08 (-0.28 - 0.52)
P. aeruginosa				0.49 (0.21 - 1.24)	0.1 (-0.37 - 0.48)
Sewage data					0.27 (0.12 - 0.75)

5. Macrolides

A



B

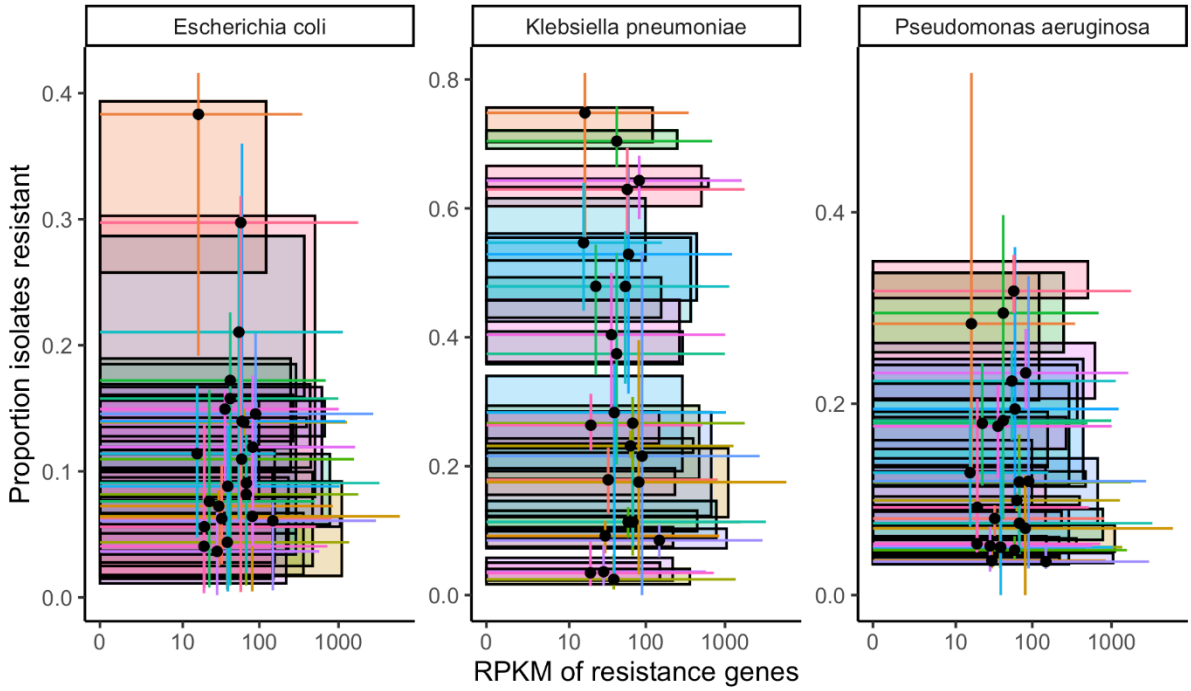


C

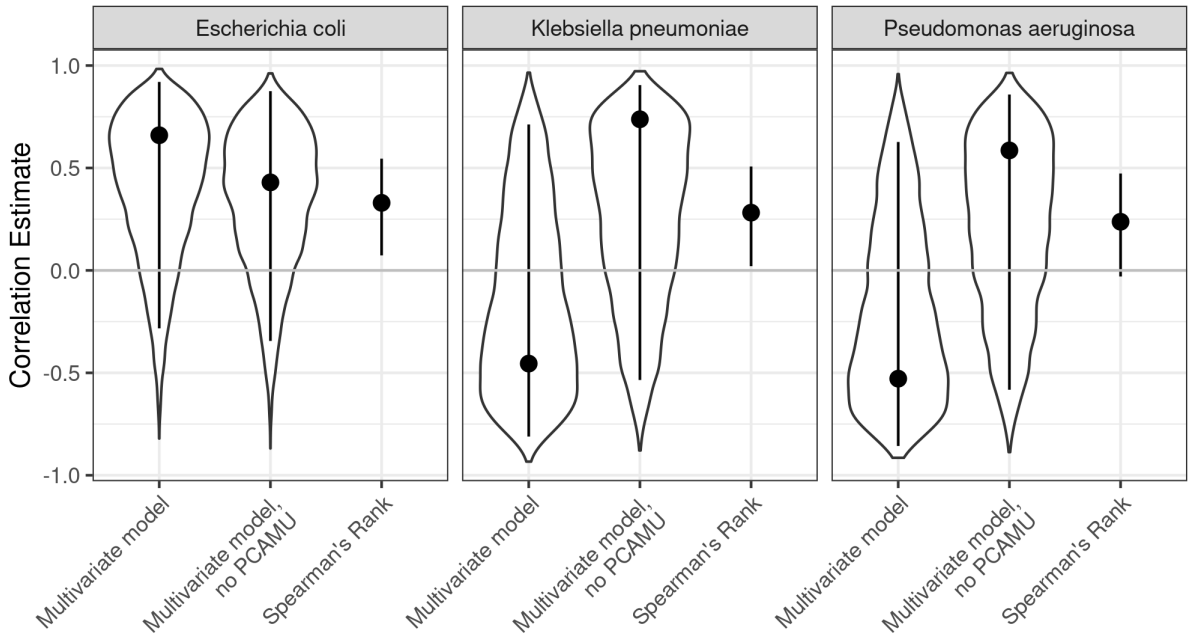
	Clinical data (S. pneumoniae)	Sewage data
Clinical data	0.24 (0.13 - 0.52)	0.09 (-0.07 - 0.25)
Sewage data		0.24 (0.14 - 0.56)

6. Third generation cephalosporins

A



B

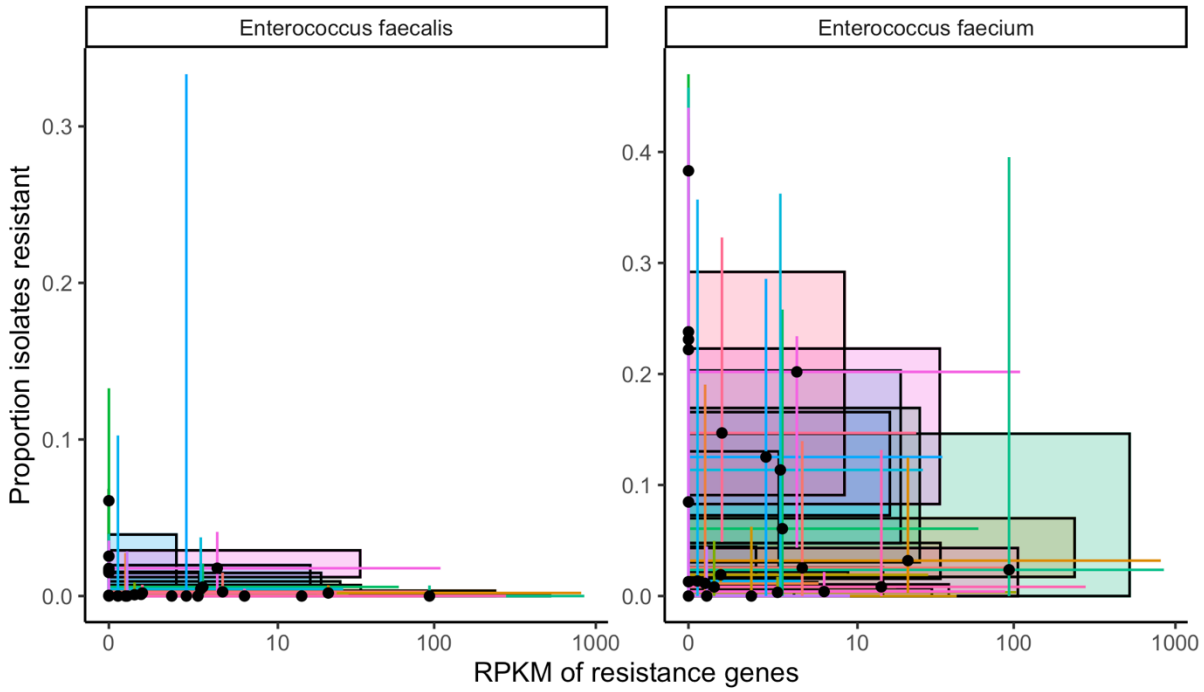


C

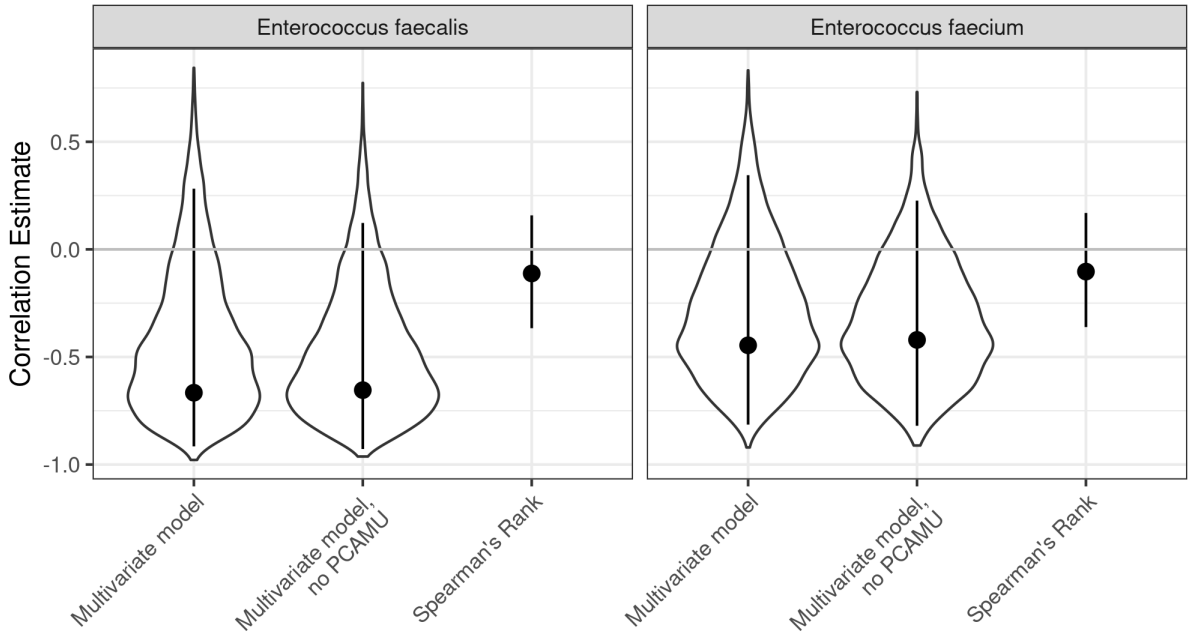
	Clinical data			Sewage data
	E. coli	K. pneumoniae	P. aeruginosa	
E. coli	0.11 (0.04 - 0.4)	-0.17 (-0.65 - 0.84)	-0.14 (-0.52 - 0.62)	0.08 (-0.16 - 0.58)
K. pneumoniae		1.83 (0.53 - 4.58)	1.32 (0.31 - 3.38)	-0.23 (-1.58 - 1.22)
P. aeruginosa			1.02 (0.28 - 2.8)	-0.25 (-1.34 - 0.85)
Sewage data				0.45 (0.16 - 1.72)

7. Vancomycin

A



B



C

	Clinical data		Sewage data
	E. faecalis	E. faecium	
E. faecalis	12.2 (5.95 - 26.07)	3.15 (0.89 - 9.46)	-6.47 (-22.81 - 6.69)
E. faecium		1.31 (0.51 - 3.9)	-0.74 (-7.04 - 2.51)
Sewage data			10.08 (2.83 - 40.11)

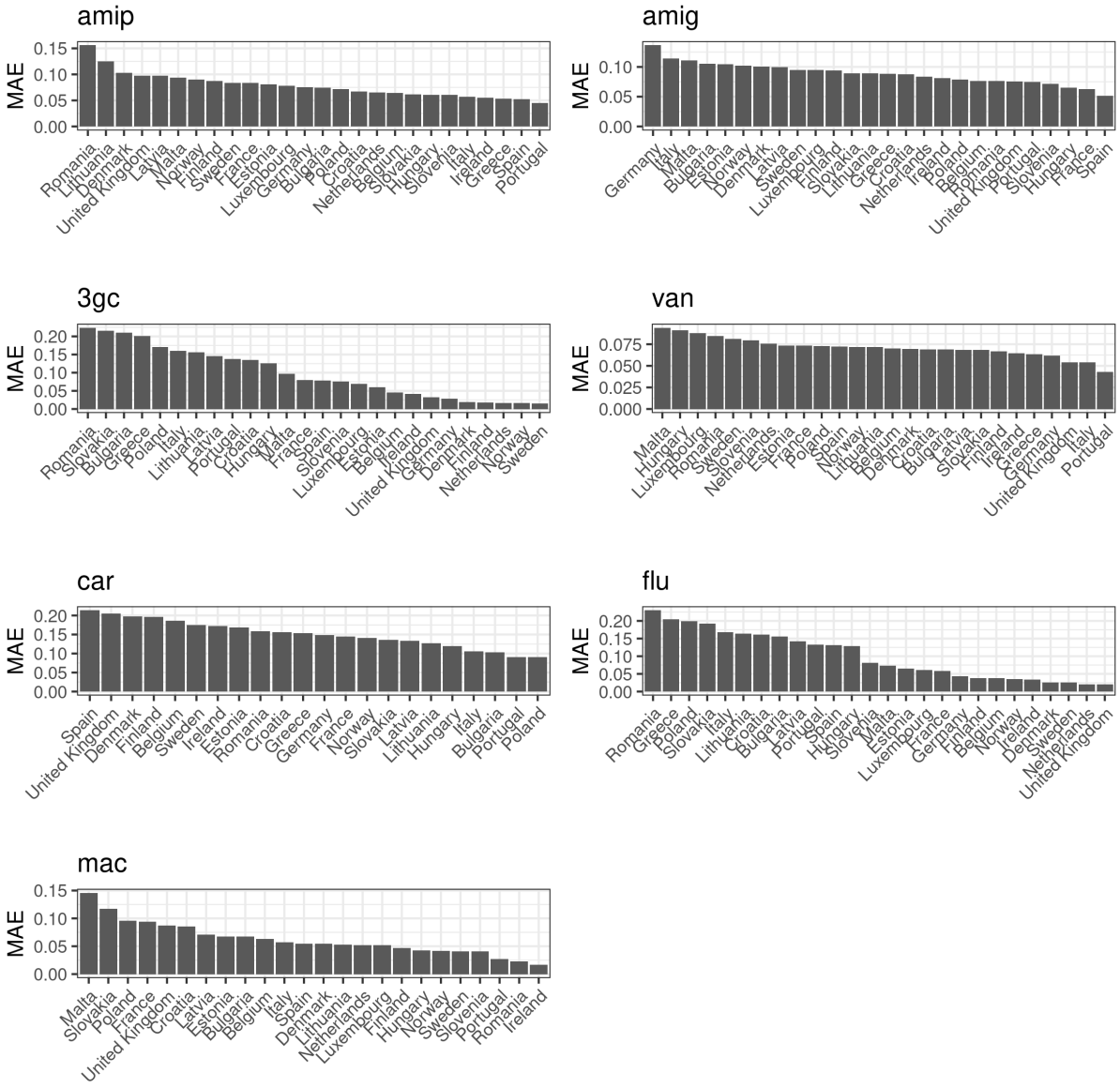
Appendix C Table 4: Other sources of non-partitioned variance in the model

Variance estimates from the multivariate model (posterior mode and 95% highest posterior density intervals).

Antimicrobial group	Residual variance (clinical data)	Residual variance (sewage data)	Resistance gene group variance	Sampling batch variance
Aminoglycosides	0.07 (0.07 – 0.08)	3.85 (3.46 – 4.22)	12.61 (8.72 – 21.26)	279.8 (83.72 – 1739.70)
Aminopenicillins	0.10 (0.08 - 0.12)	2.64 (2.42 - 2.89)	17.67 (11.07 - 40.13)	301.94 (84.66 - 1728.96)
Carbapenems	0.23 (0.20 - 0.29)	2.60 (2.18 - 3.04)	26.13 (11.77 - 68.28)	301.19 (93.73 - 1869.52)
Fluoroquinolones	0.04 (0.03 - 0.05)	2.07 (1.81 - 2.38)	26.13 (11.77 - 68.28)	213.31 (60.82 - 1410.88)
Macrolides	0.04 (0.03 - 0.06)	1.55 (1.45 - 1.72)	27.55 (14.95 - 48.44)	247 (65.05 - 1380.17)
Third generation cephalosporins	0.11 (0.10 - 0.13)	3.6 (3.17 - 4.09)	11.67 (6.64 - 24.26)	339.47 (82.39 - 1762.46)
Vancomycin	0.38 (0.30 - 0.51)	15.66 (10.08 - 24.08)	38.81 (3.59 - 1771.21)	69.62 (0.00 - 2733.24)

Appendix C Figure 4: Multivariate model prediction error for countries excluded from test set

Average absolute error in the predicted proportion of resistant isolates to the observed proportion for each country, after it was excluded from a test set for model fitting. MAE: mean absolute error; amip: aminopenicillins; amig: aminoglycosides; 3gc: third generation cephalosporins; van: vancomycin; car: carbapenems; flu: fluoroquinolones; mac: macrolides.



Appendix D: A cross-sectional metagenomics study of 8 wastewater networks in Scotland: hospital and community resistomes are distinct

Appendix D Table 1: Metagenomic characteristics

Site	Total read count	Human read count (%)	Bacterial read count (%)
Allers WWTP	40123778	0.83	96.86
East Kilbride community	71946727	0.94	96.85
Philipshill WWTP	55500121	0.87	96.89
Hairmyres Hospital	66206428	1.17	95.79
Hawick WWTP	40326498	0.36	98.83
Hawick community	18493700	1.17	95.36
Galashiels WWTP	40948895	0.40	97.88
Borders General Hospital	37034692	1.71	95.84
Stirling WWTP (March, 2021)	23399129	0.84	94.93
Stirling community (March, 2021)	39279013	0.23	98.69
Kirkcaldy WWTP (March, 2021)	43082063	0.28	98.66
Victoria Hospital (November, 2021)	76077131	1.21	96.02
Stirling WWTP (November, 2021)	32661326	0.92	96.43
Stirling community (November, 2021)	76950608	1.10	96.22
Kirkcaldy WWTP (November, 2021)	63501746	0.94	96.57
Victoria Hospital (November, 2021)	42202210	1.17	93.41
Kinneil Kerse WWTP	44129968	0.56	97.83
Grangemouth community	36553051	0.43	97.35
East Calder WWTP	36533779	1.03	96.95
St John’s Hospital	47766981	0.51	98.55

Appendix D Table 2: Sample Characteristics

Characteristics of the samples measured at the time of collection. *sample temperature could not be taken because autosampler was chilled. ** flow rate was not collected at the time, but estimated based on SW records later.

Catchment area	Sites	Sample temperature (°C)	Sample pH	Sample flow	Sample date and time	Sample colour
Galashiels	WWTP	12.9	7	54 L/s	17/11/21 09.40	Brown/grey
	Borders General Hospital	19.0	6	Category: 3	17/11/21 09.15	Brown/grey
Hawick	WWTP	12.6	7	62.1 L/s	17/11/21 10.20	Brown/grey
	Hawick community	11.7	7	Category: 2	17/11/21 12.45	Transparent
Kinneil Kerse	WWTP	-*	7	420 L/s	19/11/21 11.20	Grey
	Grangemouth community	11.5	7	Category: 2	19/11/21 11.55	Grey
East Calder	WWTP	12.6	7	652 L/s	19/11/21 10.12	Grey
	St John's Hospital	14.8	8	Category: 3	19/11/21 09.25	Transparent
Philipshill	WWTP	10.3	7	426 L/s	08/12/21 10.00	Grey
	Hairmyres University Hospital	17.0	6	Category: 2	08/12/21 08.30	Grey
Allers	WWTP	10.1	7	334 L/s	08/12/21 10.50	Grey
	East Kilbride community	10	7	Category: 3	08/12/21 10.25	Brown
Kirkcaldy (1st sample)	WWTP	-*	6	489.9 L/s **	17/03/21 10.05	Brown
	Victoria Hospital	13.4	6	Category: 1	17/03/21 09.13	Brown
Stirling (1st sample)	WWTP	-*	7	770 L/s **	17/03/21 11.20	Transparent
	Stirling community	7.8	6	Category: 2	17/03/21 11.50	Transparent
Kirkcaldy (2nd sample)	WWTP	-*	7	350 L/s	26/11/21 09.50	Grey
	Victoria Hospital	15.9	7	Category: 3	26/11/21 09.12	Brown
Stirling (2nd sample)	WWTP	12.1	7	368 L/s	26/11/21 12.00	Brown
	Stirling community	10.2	7	Category: 2	26/11/21 11.20	Transparent/grey

Appendix D Table 3: Resistance genes and bacterial species defining 'hospital' composition from PCoA analysis

Resistance gene groups and bacterial species with the strongest correlations with PCoA axes that discriminated between hospital and other sample type compositions (axis 1 for resistance genes, and axis 2 for microbiomes). Printed in order with strongest correlation value first.

Resistome PCoA			**Microbiome PCoA**	
Resistance gene	Correlation with PCoA1		Bacterial species	Correlation with PCoA2
ere(A)_5	-0.90		*Desulfovibrio desulfuricans*	-0.81
dfrA1_5	-0.90		*Roseburia inulinivorans*	-0.79
sul2_12	-0.89		*Phascolarctobacterium faecium*	-0.76
sul1_38	-0.88		*Simplicispira lacusdiani*	0.76
blaTEM-4_1	-0.86		*Klebsiella pneumoniae*	-0.76
erm(B)_9a	-0.85		*Klebsiella variicola*	-0.75
adA8b_1f	-0.84		*Comamonas jiangduensis*	0.74
osA_6	-0.83		*Acidovorax antarcticus*	0.74
mph(A)_1	-0.82		*Garciella* butyrate-producing bacterium SS3/4	-0.73
dfrA14_2	-0.82		*Acidovorax* sp. T1	0.73
tet(S/M)_2	-0.81		*Comamonas piscis*	0.73
blaVIM-42_1	-0.80		*Acidovorax soli*	0.73
mph(A)_2	-0.79		*Propionibacterium freudenreichii*	-0.73
erm(G)_2	-0.77		*Brevilactibacter flavus*	-0.73
aac(3)-Iia_4	-0.77		*Acidovorax carolinensis*	0.72
blaOXA-210_1	-0.77		*Acidovorax valerianellae*	0.72
blaDHA-18_1	-0.76		*Dorea formicigenerans*	-0.72
fosA_2	-0.76		*Bifidobacterium longum*	-0.72
oqxA_1	-0.76		*Comamonas composti*	0.72
erm(35)_1	-0.76		*Simplicispira hankyongi*	0.72

Appendix D Table 4: Model performance diagnostics

Reporting the lowest chain convergence ($\widehat{R}$) for a parameter as a measure of MCMC performance and Bayes R^2 (Gelman et al, 2019) as a measure of the proportion of the data variance explained by the model.

Model group	Model	Lowest $\widehat{R}$	Bayes R^2
Model structure 1	Impact of hospital on WWTP resistance gene abundance	1.00	0.43
Model structure 2	Association between upstream site and WWTP resistance gene abundance	1.00	0.52
Model structure 3	Association between sewage resistance levels and:		
	All community prescriptions	1.00	0.53
	Amoxicillin community prescriptions	1.00	0.50
	Vancomycin community prescriptions	1.00	0.20
	Association between sewage resistance levels and:		
	All hospital prescriptions	1.00	0.54
	Amoxicillin hospital prescriptions	1.00	0.45
	Carbapenem hospital prescriptions	1.00	0.34
	Vancomycin hospital prescriptions	1.00	0.29
Model structure 4	Association between sewage resistance levels and:		
	pH	1.00	0.52
	Sample temperature	1.00	0.50
	Sample collection time	1.00	0.52
	Weather	1.00	0.52
	Flow rate	1.00	0.30
	Wet/dry ground	1.00	0.52
Model structure 5	Association between net chance in sewage resistance levels and distance between sites	1.00	0.20

Appendix D Table 5: Number of observations and groups per model

Gene clusters were at the 70% homology level.

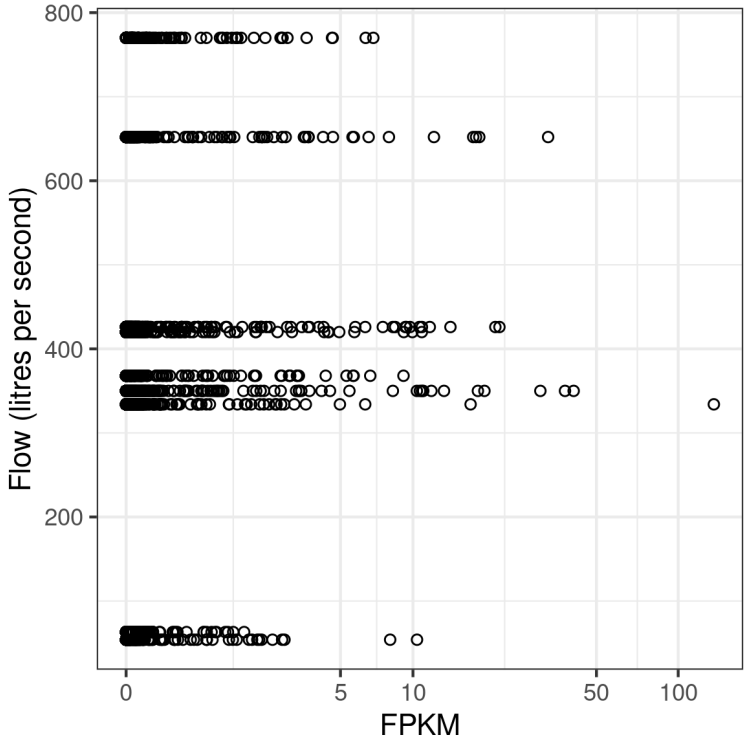
Model group	Model	Total observations	Number of samples	Number of resistance gene clusters
Model structure 1	Impact of hospital on WWTP resistance gene abundance	3240	9	300
Model structure 2	Association between upstream site and WWTP resistance gene abundance	3240	9	300
Model structure 3	Association between sewage resistance levels and:			
	All community prescriptions	4921	19	217
	Amoxicillin community prescriptions	760	19	34
	Vancomycin community prescriptions	154	14	7
	Association between sewage resistance levels and:			
	All hospital prescriptions	2331	9	217
	Amoxicillin hospital prescriptions	360	9	34
	Carbapenem hospital prescriptions	150	9	14
	Vancomycin hospital prescriptions	110	9	7
Model structure 4	Association between sewage resistance levels and:			
	pH	6840	19	300
	Sample temperature	5760	16	300
	Sample collection time	6840	19	300
	Weather	6840	19	300
	Flow rate	3240	9	300
	Wet/dry ground	6840	19	300
Model structure 4	Association between net chance in sewage resistance levels and distance between sites	3240	9	300

Appendix D Table 6: Correlations between environmental variables, using Spearman's Rank

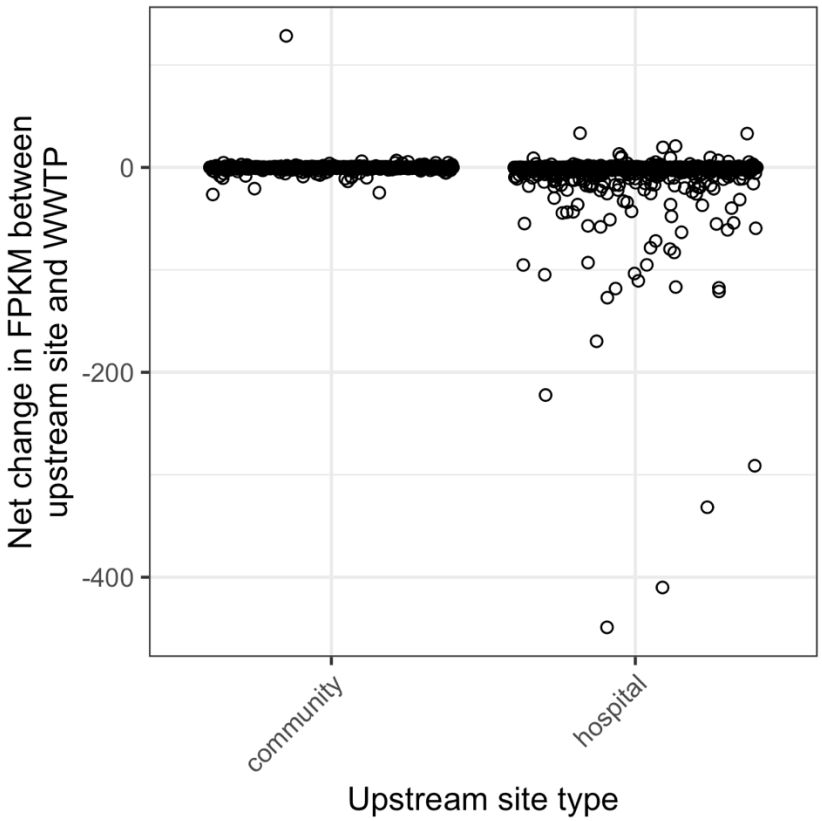
Weather was coded as rainy = 1, cloudy = 2, and clear/sunny = 3. A * indicates the correlation was significant ($p < 0.05$)

	Sample time	**Stream temperature**	**pH**	**Flow Rate**
Sample time	-	-	-	-
Stream temperature	-0.75 *	-	-	-
pH	0.25	-0.17	-	-
Flow rate	0.29	-0.38	-0.29	-
Weather	-0.06	0.37	0.30	0.26

Appendix D Figure 1: FPKM of resistance genes against the flow rate of the WWTP influent



Appendix D Figure 2: Net change in FPKM between upstream site and WWTP



Appendix E: Related preprints

The role of the environment in transmission of antimicrobial resistance between humans and animals: a modelling study

Authors

Hannah C. Lepper, Mark E.J. Woolhouse, and Bram A.D. van Bunnik.

Abstract

Background:

Antimicrobial resistance can be transmitted between animals and humans both directly or indirectly, through transmission via the environment (such as fomites or sewage). However, there is a lack of understanding of, and quantitative evidence about, the contribution of the environment to AMR epidemiology. In this study we incorporate the transmission of resistance via the environment into a mathematical model to study the potential importance of this form of transmission for human resistance levels and any effects of the impact of interventions to reduce antibiotic consumption in animals.

Methods:

We developed a compartmental model of human-animal AMR transmission with an additional environmental compartment. We compared the outcomes of this model under different human-animal-environment transmission scenarios, conducted a sensitivity analysis, and investigated the impact of curtailing antibiotic usage in animals on resistance levels in humans.

Results:

Our findings suggest that human resistance levels are most sensitive to both parameters associated with the human compartment (rate of loss of resistance from humans) and parameters associated with the environmental compartment (rate of loss of resistance from the environment and the transmission rate from the environment to humans). The impact of curtailing antibiotic consumption in animals on long term prevalence of AMR in humans was weaker when environmental transmission was assumed to be high.

Conclusions:

This study highlights that environment-human sharing of resistance can influence the epidemiology of resistant bacterial infections in humans and reduce the impact of interventions that curtail antibiotic consumption in animals. More data on the types and dynamics of resistance in the environment and frequency of human-environment transmission is crucial to understanding the population dynamics of antibiotic resistance.

The role of the environment in transmission of antimicrobial resistance between humans and animals: a modelling study

Introduction

Antimicrobial resistance (AMR) is a one-health issue, present in a variety of commensal and pathogenic bacteria found in humans, animals and the environment [1],[2]. The potential of the environment for dissemination of AMR has been increasingly recognised, with particular focus on the volume of resistance bacteria in human and agricultural wastewater effluent being discharged into natural waters and soils [3]–[5].

There are many potential routes for resistant bacteria into the environment. Several studies have demonstrated is it likely that resistant bacteria in humans can be transferred to the environment, including rivers[6], coastal waters[7], and soils[8]. In addition, studied have linked resistant bacteria in animals and their respective environments, such as between wild animals and human-impacted environments[9],[10], and between livestock and their environment, especially in aquaculture[11],[12]. However, the risk that the resistance in the environment poses to humans and animals remains poorly understood[13].

Mathematical models are an important tool to study complex dynamics inherent in the emergence and spread of resistance[14] and can therefore be used to improve our understanding and combat the spread of AMR in humans, animals and the environment. However, a lack of data and understanding about the burden, selection and transmission of resistant bacteria, especially in animals and the environment, presents a challenge to parameterising models of AMR from a one-health perspective. Consequently there are few models of resistant bacteria that connect humans, animals and the environment[15].

Some existing studies incorporate an environmental component into transmission models of resistant bacteria in hospitals or farms. Two compartmental models found that reducing or eradicating resistance in a hospital setting was harder to achieve when the environment was explicitly modelled[16],[17]. Studies taking the environment into account when modelling the spread of resistance in farms have found environmental parameters were key in dynamics of resistance in the farm [11],[18]. However, a recent modelling study found that interventions to reduce antibiotic consumption in animals would still be effective when the influence of resistance in animals and the environment is considered[19]. These findings indicate the need for further exploration of the role of the environment with fully dynamic transmission models.

In this study, we aimed to investigate the importance of the environment in the long term dynamics of resistant bacterial infections in humans, including how it might affect the impact of interventions to reduce resistance in humans. A compartmental of resistance transmission within and between humans,

animals and the environment was developed. We use a dynamic environmental compartment, improving on existing models by allowing us to assess the importance of within-environment processes. Our objectives were: 1) to investigate how adding an environmental compartment affects the long-term dynamics of resistance in humans, and the sensitivity of the model to its parameters; and 2) to investigate the impact of interventions to curtail antimicrobial usage in animals or environment to human transmission on the prevalence of resistance in humans in this model.

Methods

Model description

We extended the original model presented in van Bunnik and Woolhouse, 2017[20], to include an environmental compartment. Humans and animals gain resistant infection by exposure to antibiotics, or exposure to other humans, animals or environments carrying resistant bacteria. Resistance in the environmental compartment is increased by contact with humans or animals who carry resistant bacteria, or via exposure to antibiotics that have been excreted by humans or animals. The environment is not considered to be any one type of environment, such as water or soil, but rather a summation of these types.

We define the model using a system of coupled ordinary differential equations:

$$\frac{dR_H}{dt} = (1 - R_H) \cdot (\Lambda_H + \beta_{HH} \cdot R_H + \beta_{AH} \cdot R_A + \beta_{EH} \cdot R_E) - \mu_H \cdot R_H \quad (1)$$

$$\frac{dR_A}{dt} = (1 - R_A) \cdot (\Lambda_A + \beta_{AA} \cdot R_A + \beta_{HA} \cdot R_A + \beta_{EA} \cdot R_E) - \mu_A \cdot R_A \quad (2)$$

$$\frac{dR_E}{dt} = \gamma_H \Lambda_H + \gamma_A \Lambda_A + \beta_{HE} \cdot R_H + \beta_{AE} \cdot R_A - \mu_E \cdot R_E \quad (3)$$

R_H and R_A are the fractions of the human and animal population that are infected with resistant bacteria, respectively, and R_E is a measure of the amount of resistant infectious bacteria in the environment. Λ_H is the constant rate at which resistance is gained from exposure to antibiotics in humans, and Λ_A is the equivalent in animals. These are composite variables, taking into account both the amount of antibiotics consumed and the rate at which selection causes resistance in bacteria to arise. μ_H is the reversion rate of humans infected with resistant bacteria to having only sensitive bacteria, and μ_A is the equivalent in animals. This includes the rate of clearance of resistant infection and the rate of death in a fixed-size population. The parameters γ_H and γ_A are scaling parameters determining how much of the antibiotic exposure in humans (Λ_H) and animals (Λ_A) will result in excreted antibiotics selecting for an increase in resistant bacteria in the environment. μ_E is the rate of loss of resistant infectious bacteria from the environment. Transmission within and between the compartments is controlled by the β transmission coefficients, with the subscripts indicating the direction of transmission of each coefficient. For

example, β_{HH} is the transmission coefficient between humans, and β_{EH} is the transmission from the environment to humans.

Further details about parameter definitions, units and values ranges can be found in the Appendix Table 1. Fig. 1A shows a flow diagram representing the movement of infectious resistant material between and within the different compartments. All rates are per capita with respect to the human and animal populations, and per environmental unit with respect to the environment (see next section). We used the steady state solutions of this model, obtained numerically, as we were interested in long-term prevalence. The timestep of the model represents one month.

Capacity for resistance in the environment

Equation (3) represents the environment as an unbounded compartment, in which the amount of resistant infectious material in the environment is in the range 0 - ∞. We consider one "unit" of the environment to be the human infectious potential equivalent. This means that for a value of $R_E = 1$, if the transmission coefficients β_{EH} and β_{HH} were the same, each unit of the environment would transfer resistant material to humans at the same rate that an infected human would to another human. Although theoretically the environment may have some maximum capacity for resistant material, we do not have a way to determine this capacity, so we modelled the environment as an unbounded compartment. For comparison, we also explored a version of the model in which resistance levels in the environment cannot exceed 1. In this model the environmental compartment is specified:

$$\frac{dR_E}{dt} = (1 - R_E) \cdot (\gamma_H \Lambda_H + \gamma_A \Lambda_A + \beta_{HE} \cdot R_H + \beta_{AE} \cdot R_A) - \mu_E \cdot R_E \quad (4)$$

This model assumes that there is no growth or dissemination of resistant organisms within the environment. We also assume that the environment is only exposed to antibiotics that are excreted by humans or animals. The environment may be exposed to antibiotics directly through, for example, the effluent of pharmaceutical industries, but we do not consider this specific case here.

Impact of interventions on resistance in humans

We investigated the impact of two types of interventions on the levels of resistance in the human compartment. Firstly, we looked at interventions to remove antibiotic usage in livestock (reducing Λ_A to 0), and how changes to environmental parameters affect the effectiveness of this intervention. Secondly, we looked at interventions that would reduce the transmission of resistant bacteria from the environment to humans (reducing β_{EH} to 0).

We measured the impact of interventions as the percentage decrease in resistance levels in humans, following van Bunnik and Woolhouse (2017). We compare equilibrium values of R_H before (R_H^*) and after the intervention (RI_H^*), to obtain the impact, or percentage decrease in human resistance levels:

$$\omega = 1 - \frac{RI_H^*}{R_H^*} \tag{5}$$

We investigate the impact of reducing β_{EH} and of curtailing antibiotic usage in animals (Λ_A).

Sensitivity analysis

We use the extended version of the Fourier Amplitude Sensitivity Test (FAST)[21] to analyse the relative influence of each parameter on the value of R_H, the outcome measure of interest. A total sensitivity index for each parameter is calculated based on the variance of R_H over variation in all parameters. The R package fast was used for this analysis[22].

Parameterisation

Due to a paucity of data about many of the parameters in the model, we aimed to explore a wide range of parameter scenarios in this model. We chose the following transmission scenarios: 1) a baseline, with transmission parameter values similar to those of the original[20]; 2) a balanced transmission scenario, with all transmission coefficients equal; 3) human-driven transmission (i.e., if the subscript H denotes the humans and x denotes any other compartment $\beta_{Hx} > \beta_{xx}$); 4) animal-driven ($\beta_{Ax} > \beta_{xx}$); and finally 5) environment-driven ($\beta_{Ex} > \beta_{xx}$).

We also averaged our results across parameter sets generated randomly using sampling distributions for the three parameters R_H that was most sensitive to (viz. μ_H, μ_E, and Λ_H), to avoid over-reliance on model dynamics that are unusual to a particular combination of parameters rather than generally true of the system. All parameter values and sampling distributions can be found in the Appendix (Tables 2A and 2B), as well as the methods for obtaining transmission scenario parameters.

Software

Analyses were carried out using Wolfram Mathematica version 11.3[23], R 4.1[24], and Julia 1.7[25]. The code for the model, parameter set generation, and visualisations is available at https://github.com/hannahlepper/animal-human-env-model.

Results

All analyses were conducted in both bounded and unbounded environmental capacity versions of the model.

Long term dynamics of resistance in humans

Prevalence of resistance in humans

For all transmission scenarios, parameter sets were identified that corresponded to the intended target equilibrium human resistance prevalence of 71% in both the bounded and unbounded versions of the model (Appendix Fig. 1). Fig. 1B shows that the amount of resistance in the environment was influenced by the model structure and the transmission scenario. The highest level of resistance in the environment was in the environment-driven, unbounded version of the model, indicating that an implausibly high level of environmental contamination is not needed for observed human resistance levels.

Sensitivity analysis

Model sensitivity results are presented in Fig. 1C. In both bounded and unbounded models, human resistance prevalence was most sensitive to μ_H, the rate of loss of resistance from humans, but relatively insensitive to Λ_A, the antibiotic consumption in animals. The rate of transmission from the environment to humans, β_{EH}, was at least as important as β_{HH} and β_{AH}, rates of transmission to humans from other humans and from animals. Moreover, β_{EH} is more influential than any other transmission parameter in the unbounded model. The rate of loss of resistance from the environment, μ_E, was more important for human resistance levels in the unbounded than the bounded model.

Impact of interventions to reduce resistance in humans

Impact of curtailing antibiotic usage in animals

Curtailing antibiotic usage in animals had a small impact on human resistance levels, and the impact was lower when the environment was explicitly modelled or when animals contributed less to resistance transmission (Fig 2). The percentage decrease in human resistance levels achieved *without* an environmental compartment and using the parameters of the original model (the 'baseline transmission scenario) was 3.2%. Simply adding an environmental compartment and keeping other parameters reduced the percentage decrease to 2.8% in the unbounded and 2.9% in the bounded model. The animal-driven transmission scenario had the highest impacts (5.8% decrease in human prevalence), and the human-driven scenario had the lowest (0.064%). In the environment-driven transmission scenario, the environmental capacity was influential: when bounded, the impact was low (0.94%), and increased when unbound (3.2%). Both the environmental structure and the transmission parameters affected the impact of antibiotic usage reduction in animals.

Reducing Λ_A vs. reducing β_{EH}

We compared the impact (ω) of reducing either Λ_A (the antibiotic consumption in animals) or β_{EH} (the transmission of resistant material from the environment to humans) (Fig. 3). We considered pre-intervention values of 0.1 for each parameter, as well as the impacts in different transmission scenarios. This value was chosen so that the size of the intervention was consistent between transmission scenarios

in this model and with the previous model (van Bunnik and Woolhouse, 2017). Interventions to reduce β_{EH} had a greater impact than interventions to curtail Λ_A when transmission was human- or environment-driven, or when transmission was balanced. When livestock dominated transmission or for the baseline parameter set, the impacts of interventions to reduce β_{EH} or Λ_A were similar.

Effect of β_{EH} on impact of interventions to reduce antibiotic consumption in animals

We next identified the impact of reducing Λ_A across a range of values for β_{EH} (Fig. 4). Increasing β_{EH} decreased the size of the impact of curtailing antibiotic usage in animals in all transmission scenarios (Fig 4A). The peaked shape of the impact size in the environmental transmission scenario is caused by the increase in β_{EH} allowing increasing indirect transmission in animals and humans. This effect is only observed when there is little non-environmental transmission. Fig. 4B shows that the decrease in intervention impact was also observed across the range of pre-intervention values for Λ_A. These results indicate that increasing environmental transmission can reduce the impact of curtailing antibiotic usage in animals.

Discussion

Key findings

In this study we modelled the transmission of resistant bacteria between humans, livestock animals and the environment, and assessed the impact of interventions that reduce antibiotic consumption in animals or decrease the transmission of resistant bacteria from the environment to humans. We found that human resistance prevalence is sensitive to transmission between humans and the environment. Including an environmental compartment in the model decreased the impact of curtailing antibiotic resistance, and a more transmissible environmental reservoir of resistant bacteria further mitigated the impact of this intervention. Reducing the transmission of resistant bacteria from the environment to humans was found to be a more effective intervention than reducing antibiotic consumption in animals. Overall, these results indicate that resistant bacteria in the environment can influence the prevalence of resistance in humans. The size of environmental influence will depend on the amount and dynamics of resistant bacteria in the environment. Assessing the likelihood of observing these theoretical results in the real world is hindered by a lack of quantified, generalisable data on the types, amount, and degradation of resistance in the environment, and the transmission of resistance between humans, livestock and the environment.

Is curtailing antibiotic usage in animals an effective intervention to reduce human resistance levels?

The greatest observed impact of curtailing antibiotics in animals was a modest 10% decrease in human resistance level in a balanced transmission scenario, and the smallest impact was a <1% in the human-driven transmission scenario. This result provides little theoretical support that curtailment of antibiotics would appreciably decrease resistance in humans in many settings. In contrast, there is some empirical

evidence that curtailing antibiotics in livestock could reduce human resistance levels, although from a small set of observational studies[26]. A study of use of third-generation cephalosporin ceftiofur in broiler rearing in Canada found that resistance in *Salmonella* and *E. coli* was decreased in clinical isolates by 20% and 40%, respectively, after ceftiofur use decreased[27]. This real-world population-level effect is greater than our results would predict, and may indicate they are an underestimate, especially with respect to the degree of sharing of resistance between humans and animals. More data-based parameterisation will be crucial to improve the accuracy of one-health resistance transmission models.

The size of the effect of intervening to reduce antibiotic consumption in livestock varied by transmission scenario (balanced transmission, or transmission driven by either humans, livestock or the environment). Therefore, a key question for assessing the accuracy and relevance of the resulting intervention effect sizes is how realistic are the transmission scenarios? Although transmission of resistance between humans and animals is of great concern, evidence that conclusively demonstrates a case of direct transmission is rare[28],[29]. Accurately parameterising the relationship between resistance in humans and livestock is an ongoing area of research[30] which will be crucial for one-health modelling of resistance. It seems likely that on average across a large human population, human-human transmission is far more common than animal-human transmission and we suggest human-driven scenario to be most relevant for resistance dynamics in the human population.

As we increased the transmission rate from the environment to humans, the effectiveness of antibiotic curtailment was decreased. This suggests that the environment can provide a 'back door' transmission route from animals to humans that can reduce the effectiveness of antibiotic curtailment by adding to overall animal-human transmission rates. Using a two-pronged approach by intervening to reduce environmental transmission at the same time could therefore improve the impact of antibiotic usage curtailment. However, the effect of environmental transmission on antibiotic curtailment effectiveness was negligible in the human-dominated transmission scenario (Appendix Fig. 2.), again indicating the importance of transmission setting for this result. It remains unclear if non-human dominated transmission scenarios are realistic, and therefore what the real-world size of this back-door effect might be. There is some evidence that microbiomes in humans, animals and the environment become more shared as interactions become more frequent[31], suggesting that transmission scenarios in which humans do not dominate transmission (such as the balances and baseline scenarios) are possible. Further work to quantify environmental resistance concentrations and transmission could improve accuracy of outcome predictions of antibiotic usage interventions. As reducing antibiotic usage in livestock animals is a costly intervention, it is important to ensure optimal implementation.

Could the environment be an effective alternative intervention target?

The rate of transfer of resistant bacteria from environment to humans (β_{EH}) is also a potentially effective intervention target. Human resistance prevalence levels were sensitive to β_{EH} and μ_E, the rate of loss of resistant bacteria from the environment (sensitivity analysis, Fig. 1C), which suggests that interventions to reduce how much resistance humans gain from the environment would be effective. Indeed, the impact of reducing β_{EH} was more effective than antibiotic usage curtailment interventions, although the difference was small in the animal-dominated scenario (Fig. 2A). Interventions that improve sanitation have been proposed to reduce occurrences of transmission of resistance between humans and the environment in informal urban communities in LMICs where there is frequent exposure to resistance bacteria in the environment[32],[33]. Nadimpalli et al (2020) focus particularly on the potential benefits of improved water and wastewater infrastructure for controlling and preventing AMR transmission, but note that few studies have investigated the impacts of sanitation interventions on AMR.

Should the environment be included in AMR models?

In this model, the environment played an important role in the long-term dynamics of antibiotic resistance levels in humans. Mechanistically, the environment acts as a reservoir for antibiotic resistance from humans and animals in this model structure. Therefore, parameters that provide more opportunity for transmission to humans were influential in human resistance levels, especially the rate of loss or level of persistence of resistant bacteria in the environment (μ_E). Environmental parameters were also influential in the size of impact of interventions, and we show that it may be an effective intervention target itself. Existing models that incorporate an environmental component have also highlighted the potentially strong role the environment could play in increasing resistance levels in humans and undermining interventions[16]–[19]. Most models include environment as a constant rather than a dynamic compartment, with the exception of Booton et al, 2021. As we find comparable results to models with constant compartments, this may indicate that models incorporating the environment simply may be enough to account for this additional source of resistant bacteria. On the other hand, the model in Booton et al, 2021, assumes that transmission of resistance (including from the environment) is dependent on exposure to antibiotics and accordingly finds that human antibiotic usage is the most influential parameter for human resistance, downplaying the role of the environment. This contrasting result points to a need for further models that compare the contribution of the environment under different model structures and assumptions. Incorporating the environment into models of AMR spread may be important in understanding AMR prevalence and for evaluating intervention success.

Modelling the environment highlights data needs

The results highlight some key data needs for understanding the importance of AMR in the environment for humans. There are two influential parameters in the model which are difficult to parameterise from

existing data: the rate of transfer of AMR from the environment to humans, and the rate of loss of resistance in the human population.

How frequently humans gain resistant bacteria after exposure to an environmental source is unknown. There is evidence that humans can be exposed to resistant bacteria in the environment. For example, one study estimated that the amount of third-generation cephalosporin resistant *E. coli* that humans would ingest during recreational water use in coastal regions in England and Wales poses a risk of infection[7]. However it is not clear how often these exposures lead to infection or colonisation[34]. More research that demonstrates a close relationship and epidemiological link between resistant bacteria colonising the environment and humans is needed to understand the frequency of environment-human transmission events. Use of high resolution typing such as whole genome sequencing of, for example, isolates from hospital patients and the hospital environment in longitudinal studies would be ideal for this research.

Studies have provided data on the rate of clearance of resistant infections in humans. A systematic review on methicillin-resistant *S. aureus* (MRSA) and vancomycin-resistant *Enterococcus* (VRE) colonisation found that it takes a period of 88 and 26 weeks on average to clear MRSA and VRE infections, respectively[35]. However, they note that there is considerable methodological heterogeneity in studies of MRSA and VRE, including varying definitions of clearance, and length of follow-up[35]. The studies also focussed primarily on hospital-associated resistance. Data on resistant bacteria colonisation prevalence and clearance in the community, where the role of exposure to animals and the environment may play a greater role, appear to be rare. Parameterising generalisable one-health models will therefore be benefitted by more research into resistance in the community.

Limitations

There are some important limitations to this study that should be noted. Firstly, we make simplifying assumptions in the structure and parameterisation of the model. These are suitable to the questions posed in this study, but there are still many complexities in the spread and emergence of AMR in humans, animals and the environment to be explored. Further models should explore the importance of potential complexities, such as heterogeneity of transmission events, separate humans-specific and animal-specific environmental reservoirs, variation in the capacity for resistance in the environment, or the fitness costs to bacteria of carrying resistance in the three populations.

We do not model the dynamics of transmission of resistant bacteria and resistance genes separately, but assume that transmission parameters combine the transmission of both. This is in-keeping with the assumptions of the original model [20]. Resistance genes can be transferred between bacteria via plasmid transfer or bacteriophages, and can also be lost from bacterial lineages. The transmission rates of resistance genes in human population may therefore differ from resistant bacteria, and it is a

limitation that we do not capture this in the model. AMR epidemiology and surveillance is usually measured in resistant bacteria so there is little data on the prevalence and transmission rates of specific resistance genes.

Two further assumptions about resistance in the environment are that we assume that there is no growth of resistant material within the environment, and that all antibiotics secreted into the environment are from human and livestock usage. The dynamics of resistance genes and bacteria in the environment is a complex topic, and although there are potentially environments in which resistance may spread (especially in sewage) much more empirical and modelling research is needed[34],[36]. A recent review found that the sources of antibiotics in ground water include excretion from humans and animals (via sewage and manure) but also landfill, aquaculture and industrial sites[37], so not including these sources may limit the accuracy of the results of this model. However the relative contribution of each sources is not well known and may vary from one country to another[37].

Conclusions

This study illustrates the potentially important role of the environment in the epidemiology of resistant bacterial infections in humans. We highlight the need to consider the role of the environment in the design of AMR control strategies, as it can be influential in human prevalence of resistance, reduce the effectiveness of interventions that curtail antibiotic consumption in animals, and may be an effective intervention target itself via improved sanitation infrastructure. Incorporating the environment into a one-health model of antibiotic resistance as a dynamic compartment was useful for considering the role of the environment. However, assessing the uncertainty of model predictions is hindered by a lack of data on the types and frequency of resistance in the environment, and the frequency of environment-human transmission events.

The role of the environment in transmission of antimicrobial resistance between humans and animals: a modelling study

Figures

Figure 1. A: flow diagram indicating the model structure. B: R_E values in all transmission scenarios and both model structures. C: Fourier Amplitude Sensitivity Tests (FAST), indicating how much variation in R_H was explained by each model parameter. On the left, FAST for the version of the model in which R_E is bounded to 1. On the right, FAST for the version of the model in which R_E was unbounded.

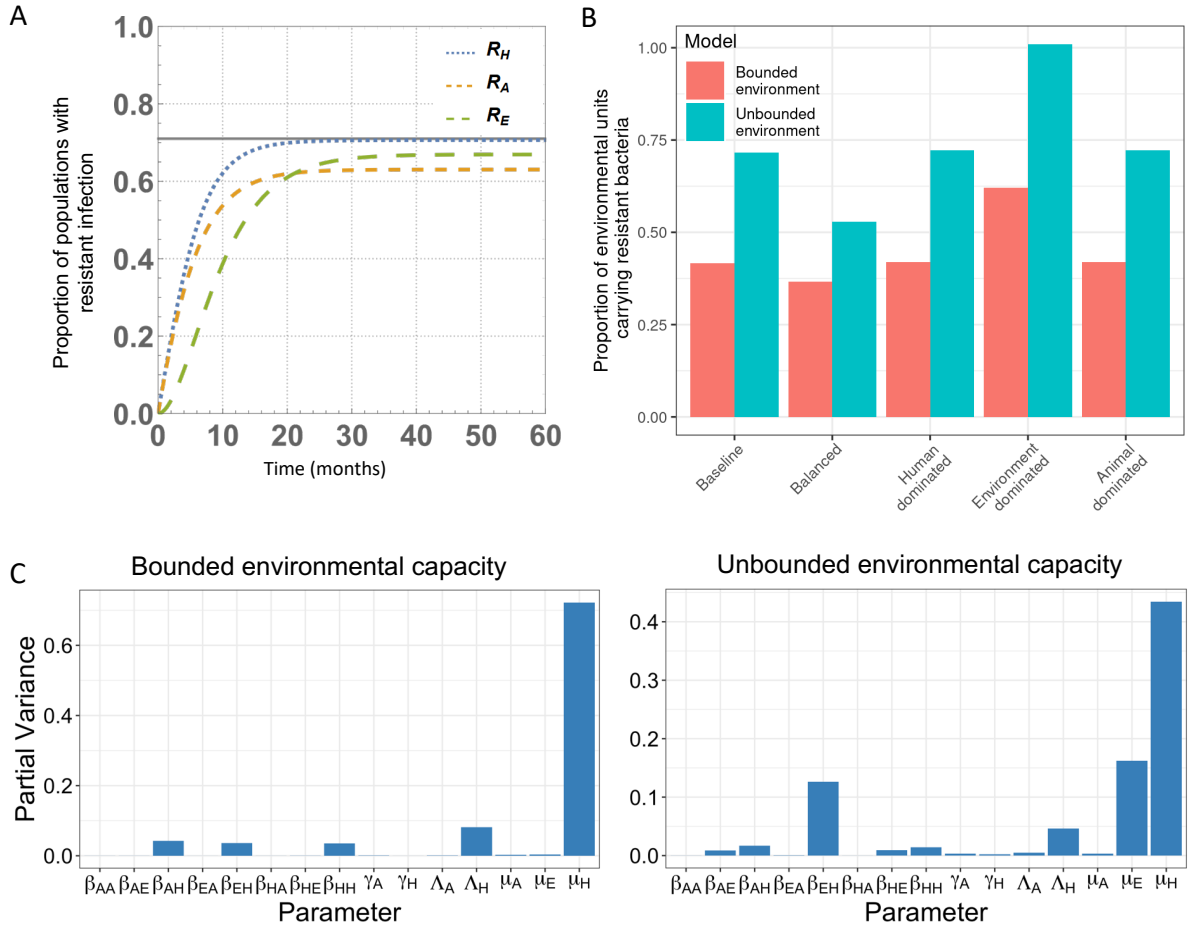


Fig 2: Mean impact of reducing Λ_A from 0.1 to 0 across transmission scenarios. The green point in the baseline transmission scenario group is the mean impact for the original van Bunnik and Woolhouse (2017) model, with no environmental compartment included. Results were averaged for parameter sets with μ_H, μ_E, and Λ_H varied, with error bars indicating standard deviation in results.

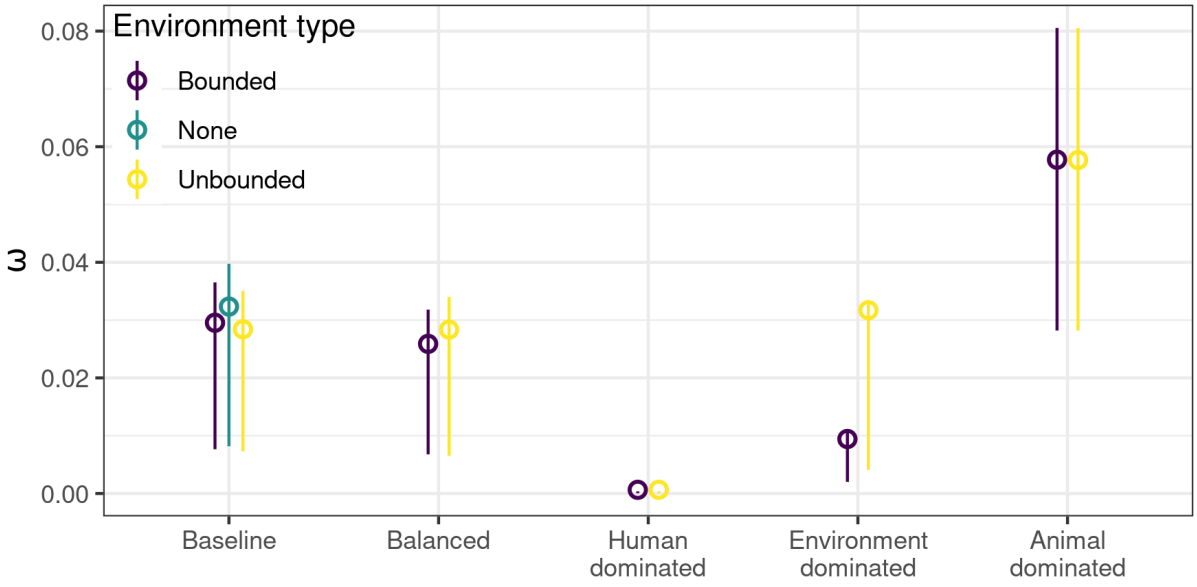


Fig 3. Violin plots of the impact (proportion decrease in R_H after the intervention) of reducing either β_{EH} or Λ_A in all transmission scenarios and for both model structures. The intervention target was reduced from 0.1 to 0 in each case for consistency.

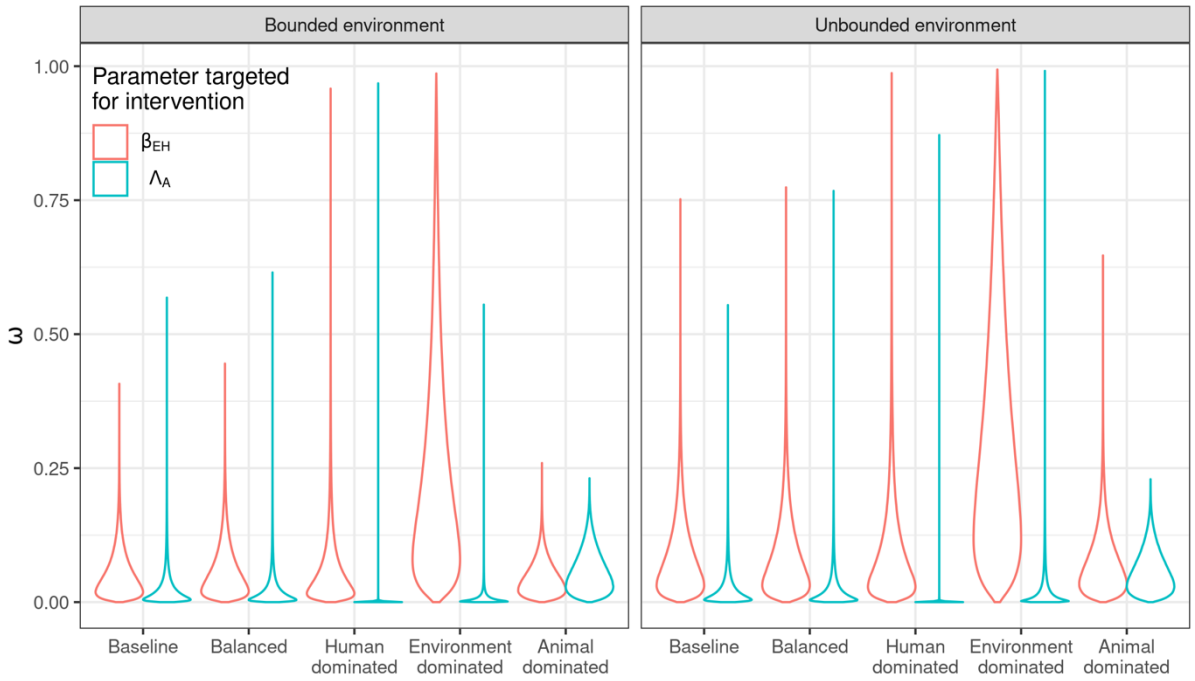
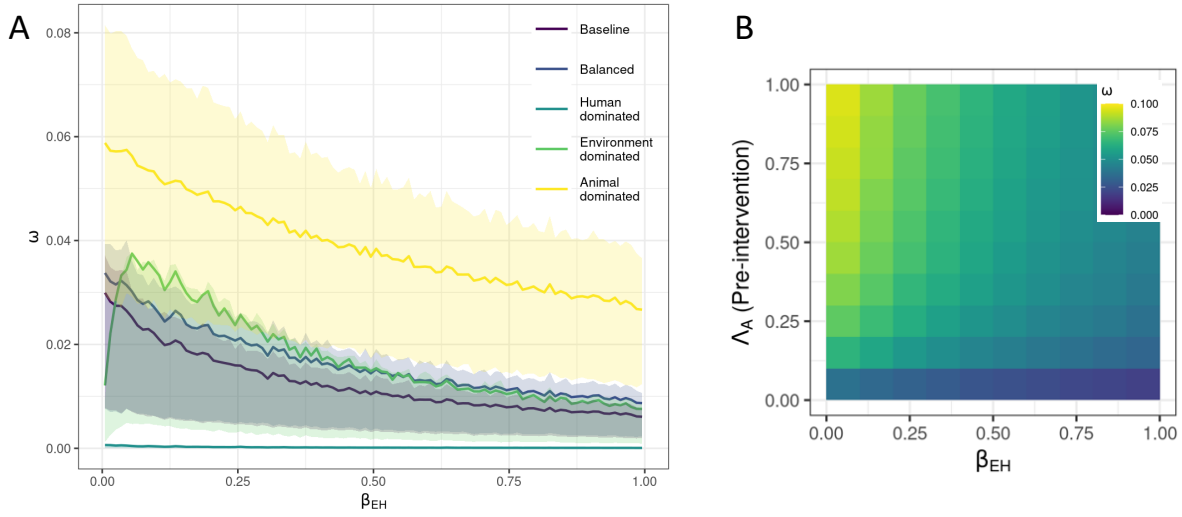


Fig 4: A) Mean impact of antibiotic decrease in animals on human resistance levels (proportion decrease in human resistance levels) for each transmission scenario with increasing rate of environment to human transmission (β_{EH}). Ribbons indicate 25% and 75% impact quantiles. B) Heatmap of the impact of different pre-intervention values of Λ_A (y axis) against different levels of environment to human transmission, β_{EH} (x axis), for the animal transmission scenario in the unbounded model. The colour of the tiles indicates the average value of the impact of the intervention from 17,000 parameter sets where μ_H, μ_E, and Λ_H were varied.



References

(1) Woolhouse, M.; Ward, M.; Van Bunnik, B.; Farrar, J. Antimicrobial Resistance in Humans, Livestock and the Wider Environment. *Philos. Trans. R. Soc. B Biol. Sci.* **2015**, *370* (1670). https://doi.org/10.1098/rstb.2014.0083.

(2) Robinson, T. P.; Bu, D. P.; Carrique-Mas, J.; Fèvre, E. M.; Gilbert, M.; Grace, D.; Hay, S. I.; Jiwakanon, J.; Kakkar, M.; Kariuki, S.; Laxminarayan, R.; Lubroth, J.; Magnusson, U.; Ngoc, P. T.; van Boeckel, T. P.; Woolhouse, M. E. J. Antibiotic Resistance Is the Quintessential One Health Issue. *Trans. R. Soc. Trop. Med. Hyg.* **2016**, *110* (7), 377–380.

(3) Larsson, D. G. J.; Andremont, A.; Bengtsson-Palme, J.; Brandt, K. K.; de Roda Husman, A. M.; Fagerstedt, P.; Fick, J.; Flach, C. F.; Gaze, W. H.; Kuroda, M.; Kvint, K.; Laxminarayan, R.; Manaia, C. M.; Nielsen, K. M.; Plant, L.; Ploy, M. C.; Segovia, C.; Simonet, P.; Smalla, K.; Snape, J.; Topp, E.; van Hengel, A. J.; Verner-Jeffreys, D. W.; Virta, M. P. J.; Wellington, E. M.; Wernersson, A. S. Critical Knowledge Gaps and Research Needs Related to the Environmental Dimensions of Antibiotic Resistance. *Environ. Int.* **2018**, *117* (May), 132–138. https://doi.org/10.1016/j.envint.2018.04.041.

(4) Bürgmann, H.; Frigon, D.; Gaze, W. H.; Manaia, C. M.; Pruden, A.; Singer, A. C.; Smets, B. F.; Zhang, T. Water and Sanitation: An Essential Battlefront in the War on Antimicrobial Resistance. *FEMS Microbiol. Ecol.* **2018**, *94* (9), fiy-101. https://doi.org/10.1093/femsec/fiy101.

(5) Manaia, C. M. Antibiotic Resistance in Wastewater: Origins, Fate, and Risks. *Prävention und Gesundheitsförderung* **2014**, *9* (3), 180–184. https://doi.org/10.1007/s11553-014-0452-3.

(6) Amos, G. C. A.; Gozzard, E.; Carter, C. E.; Mead, A.; Bowes, M. J.; Hawkey, P. M.; Zhang, L.; Singer, A. C.; Gaze, W. H.; Wellington, E. M. H. Validated Predictive Modelling of the Environmental Resistome. *ISME J.* **2015**, *9* (6), 1467–1476. https://doi.org/10.1038/ismej.2014.237.

(7) Leonard, A. F. C.; Zhang, L.; Balfour, A. J.; Garside, R.; Gaze, W. H. Human Recreational Exposure to Antibiotic Resistant Bacteria in Coastal Bathing Waters. *Environ. Int.* **2015**, *82*, 92–100. https://doi.org/10.1016/j.envint.2015.02.013.

(8) Palacios, O. A.; Contreras, C. A.; Muñoz-Castellanos, L. N.; González-Rangel, M. O.; Rubio-Arias, H.; Palacios-Espinosa, A.; Nevárez-Moorillón, G. V. Monitoring of Indicator and Multidrug Resistant Bacteria in Agricultural Soils under Different Irrigation Patterns. *Agric. Water Manag.* **2017**, *184*, 19–27. https://doi.org/10.1016/j.agwat.2017.01.001.

(9) Araujo, S.; Henriques, I. S.; Leandro, S. M.; Alves, A.; Pereira, A.; Correia, A. Gulls Identified as Major Source of Fecal Pollution in Coastal Waters: A Microbial Source Tracking Study. *Sci. Total Environ.* **2014**, *470*, 84–91. https://doi.org/10.1016/j.scitotenv.2013.09.075.

(10) Swift, B. M. C.; Bennett, M.; Waller, K.; Dodd, C.; Murray, A.; Gomes, R. L.; Humphreys, B.; Hobman, J. L.; Jones, M. A.; Whitlock, S. E.; Mitchell, L. J.; Lennon, R. J.; Arnold, K. E. Anthropogenic Environmental Drivers of Antimicrobial Resistance in Wildlife. *Sci. Total Environ.* **2019**, *649*, 12–20. https://doi.org/10.1016/j.scitotenv.2018.08.180.

(11) Call, D. R.; Matthews, L.; Subbiah, M.; Liu, J. Do Antibiotic Residues in Soils Play a Role in Amplification and Transmission of Antibiotic Resistant Bacteria in Cattle Populations? *Front. Microbiol.* **2013**, *4* (JUL), 1–8. https://doi.org/10.3389/fmicb.2013.00193.

(12) Cabello, F. C.; Godfrey, H. P.; Tomova, A.; Ivanova, L.; Dölz, H.; Millanao, A.; Buschmann, A. H. Antimicrobial Use in Aquaculture Re-Examined: Its Relevance to Antimicrobial Resistance and to Animal and Human Health. *Environ. Microbiol.* **2013**, *15* (7), 1917–1942. https://doi.org/10.1111/1462-2920.12134.

(13) Raven, K. E.; Ludden, C.; Gouliouris, T.; Blane, B.; Naydenova, P.; Brown, N. M.; Parkhill, J.; Peacock, S. J. Genomic Surveillance of Escherichia Coli in Municipal Wastewater Treatment Plants as an Indicator of Clinically Relevant Pathogens and Their Resistance Genes. *Microb. genomics* **2019**, *5* (5). https://doi.org/10.1099/mgen.0.000267.

(14) Knight, G. M.; Davies, N. G.; Colijn, C.; Coll, F.; Donker, T.; Gifford, D. R.; Glover, R. E.; Jit, M.; Klemm, E.; Lehtinen, S.; Lindsay, J. A.; Lipsitch, M.; Llewelyn, M. J.; Mateus, A. L. P.; Robotham, J. V.; Sharland, M.; Stekel, D.; Yakob, L.; Atkins, K. E. Mathematical Modelling for Antibiotic Resistance Control Policy: Do We Know Enough? *BMC Infect. Dis.* **2019**, *19* (1), 1011. https://doi.org/10.1186/s12879-019-4630-y.

(15) Niewiadomska, A. M.; Jayabalasingham, B.; Seidman, J. C.; Willem, L.; Grenfell, B.; Spiro, D.; Viboud, C. Population-Level Mathematical Modeling of Antimicrobial Resistance: A Systematic Review. *BMC Med.* **2019**, *17* (1), 1–20. https://doi.org/10.1186/s12916-019-1314-9.

(16) McBryde, E. S.; McElwain, D. L. S. A Mathematical Model Investigating the Impact of an Environmental Reservoir on the Prevalence and Control of Vancomycin- Resistant Enterococci. *J. Infect. Dis.* **2006**, *193*, 1473–1474.

(17) Kouyos, R. D.; zur Wiesch, P. A.; Bonhoeffer, S. On Being the Right Size: The Impact of Population Size and Stochastic Effects on the Evolution of Drug Resistance in Hospitals and the Community. *PLoS Pathog.* **2011**, *7* (4). https://doi.org/10.1371/journal.ppat.1001334.

(18) Græsbøll, K.; Nielsen, S. S.; Toft, N.; Christiansen, L. E. How Fitness Reduced, Antimicrobial Resistant Bacteria Survive and Spread: A Multiple Pig - Multiple Bacterial Strain Model. *PLoS One* **2014**, *9* (7). https://doi.org/10.1371/journal.pone.0100458.

(19) Booton, R. D.; Meeyai, A.; Alhusein, N.; Buller, H.; Feil, E.; Lambert, H.; Mongkolsuk, S.; Pitchforth, E.; Reyher, K. K.; Sakcamduang, W.; Satayavivad, J.; Singer, A. C.; Sringernyuang, L.; Thamlikitkul, V.; Vass, L.; Avison, M. B.; Turner, K. M. E.; Avison, M. B.; Alhusein, N.; Booton, R. D.; Buller, H.; Chantong, B.; Charoenlap, N.; Couto, N.; Dulyayangkul, P.; Feil, E.; Gibbon, M. J.; Gould, V. C.; Lambert, H.; Meeyai, A.; Mongkolsuk, S.; Montrivade, V.; Pitchforth, E.; Phoonsawad, K.; Rangkadilok, N.; Ratanakorn, P.; Reyher, K. K.; Sakcamduang, W.; Satayavivad, J.; Singer, A. C.; Sirikanchana, K.; Sringernyuang, L.; Suriyo, T.; Suwanpakdee, S.; Thamlikitkul, V.; Turner, K. M. E.; Vass, L.; Wichuwaranan, K.; Wiratsudakul, A. One Health Drivers of Antibacterial Resistance: Quantifying the Relative Impacts of Human, Animal and Environmental Use and Transmission. *One Heal.* **2021**, *12* (January), 100220. https://doi.org/10.1016/j.onehlt.2021.100220.

(20) van Bunnik, B. A. D.; Woolhouse, M. E. J. Modelling the Impact of Curtailing Antibiotic Usage in Food Animals on Antibiotic Resistance in Humans. *R. Soc. Open Sci.* **2017**, *4* (4). https://doi.org/10.1098/rsos.161067.

(21) Saltelli, A.; Tarantola, S.; Chan, K. P.-S. A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. *Technometrics* **1999**, *41* (1), 39–56. https://doi.org/10.1080/00401706.1999.10485594.

(22) Reusser, D. Fast: Implementation of the Fourier Amplitude Sensitivity Test (FAST). 2015.

(23) Wolfram Research Inc. Mathematica, Version 11.3. Champaign, IL 2018.

(24) R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria 2022.

(25) Bezanson, J.; Edelman, A.; Karpinski, S.; Shah, V. B. Julia: A Fresh Approach to Numerical Computing. *arXiv* **2017**, *59* (1), 65–98. https://doi.org/10.1137/141000671.

(26) Scott, A. M.; Beller, E.; Glasziou, P.; Clark, J.; Ranakusuma, R. W.; Byambasuren, O.; Bakhit, M.; Page, S. W.; Trott, D.; Mar, C. Del. Is Antimicrobial Administration to Food Animals a Direct Threat to Human Health? A Rapid Systematic Review. *Int. J. Antimicrob. Agents* **2018**, *52* (3), 316–323. https://doi.org/10.1016/j.ijantimicag.2018.04.005.

(27) Dutil, L.; Irwin, R.; Finley, R.; Ng, L. K.; Avery, B.; Boerlin, P.; Bourgault, A. M.; Cole, L.; Daignault, D.; Desruisseau, A.; Demczuk, W.; Hoang, L.; Horsman, G. B.; Ismail, J.; Jamieson, F.; Maki, A.; Pacagnella, A.; Pillai, D. R. Ceftiofur Resistance in Salmonella Enterica Serovar Heidelberg from Chicken Meat and Humans, Canada. *Emerg. Infect. Dis.* **2010**, *16* (1), 48–54. https://doi.org/10.3201/eid1601.090729.

(28) Muloi, D.; Ward, M. J.; Pedersen, A. B.; Fèvre, E. M.; Woolhouse, M. E. J.; van Bunnik, B. A. D. Are Food Animals Responsible for Transfer of Antimicrobial-Resistant Escherichia Coli or Their Resistance Determinants to Human Populations? A Systematic Review. *Foodborne Pathog. Dis.* **2018**, *15* (8), 467–474. https://doi.org/10.1089/fpd.2017.2411.

(29) Wee, B. A.; Muloi, D. M.; van Bunnik, B. A. D. Quantifying the Transmission of Antimicrobial Resistance at the Human and Livestock Interface with Genomics. *Clin. Microbiol. Infect.* **2020**, *26* (12), 1612–1616. https://doi.org/10.1016/j.cmi.2020.09.019.

(30) Thorpe, H.; Booton, R.; Kallonen, T.; Gibbon, M. J.; Couto, N.; Passet, V.; Sebastian Lopez Fernandez, J.; Rodrigues, C.; Matthews, L.; Mitchell, S.; Reeve, R.; David, S.; Merla, C.; Corbella, M.; Ferrari, C.; Comandatore, F.; Marone, P.; Brisse, S.; Sassera, D.; Corander, J.; Feil, E. J. One Health or Three? Transmission Modelling of Klebsiella Isolates Reveals Ecological Barriers to Transmission between Humans, Animals and the Environment. https://doi.org/10.1101/2021.08.05.455249.

(31) Pehrsson, E. C.; Tsukayama, P.; Patel, S.; Mejia-Bautista, M.; Sosa-Soto, G.; Navarrete, K. M.; Calderon, M.; Cabrera, L.; Hoyos-Arango, W.; Bertoli, M. T.; Berg, D. E.; Gilman, R. H.; Dantas, G. Interconnected Microbiomes and Resistomes in Low-Income Human Habitats. *Nature* **2016**, *533* (7602), 212–216. https://doi.org/10.1038/nature17672.

(32) Collignon, P.; Beggs, J. J.; Walsh, T. R.; Gandra, S.; Laxminarayan, R. Anthropological and Socioeconomic Factors Contributing to Global Antimicrobial Resistance: A Univariate and Multivariable Analysis. *Lancet Planet. Heal.* **2018**, *2* (9), e398–e405. https://doi.org/10.1016/S2542-5196(18)30186-4.

(33) Nadimpalli, M. L.; Marks, S. J.; Montealegre, M. C.; Gilman, R. H.; Pajuelo, M. J.; Saito, M.; Tsukayama, P.; Njenga, S. M.; Kiiru, J.; Swarthout, J.; Islam, M. A.; Julian, T. R.; Pickering, A. J. Urban Informal Settlements as Hotspots of Antimicrobial Resistance and the Need to Curb Environmental Transmission. *Nat. Microbiol.* **2020**, *5* (6), 787–795. https://doi.org/10.1038/s41564-020-0722-0.

(34) Berendonk, T. U.; Manaia, C. M.; Merlin, C.; Fatta-Kassinos, D.; Cytryn, E.; Walsh, F.; Buergmann, H.; Sorum, H.; Norstrom, M.; Pons, M.-N.; Kreuzinger, N.; Huovinen, P.; Stefani, S.; Schwartz, T.; Kisand, V.; Baquero, F.; Luis Martinez, J. Tackling Antibiotic Resistance: The Environmental Framework. *Nat. Rev. Microbiol.* **2015**, *13* (5), 310–317. https://doi.org/10.1038/nrmicro3439.

(35) Shenoy, E. S.; Paras, M. L.; Noubary, F.; Walensky, R. P.; Hooper, D. C. Natural History of Colonization with Methicillin-Resistant Staphylococcus Aureus (MRSA) and Vancomycin-Resistant Enterococcus (VRE): A Systematic Review. *BMC Infect. Dis.* **2014**, *14* (1). https://doi.org/10.1186/1471-2334-14-177.

(36) Bengtsson-Palme, J.; Kristiansson, E.; Larsson, D. G. J. Environmental Factors Influencing the Development and Spread of Antibiotic Resistance. *FEMS Microbiology Reviews*. Oxford University Press January 1, 2018, pp 68–80. https://doi.org/10.1093/femsre/fux053.

(37) Zainab, S. M.; Junaid, M.; Xu, N.; Malik, R. N. Antibiotics and Antibiotic Resistant Genes (ARGs) in Groundwater: A Global Review on Dissemination, Sources, Interactions, Environmental and Human Health Risks. *Water Research*. Elsevier Ltd December 15, 2020. https://doi.org/10.1016/j.watres.2020.116455.